

**INTERNATIONAL ECONOMIC REVIEW**  
Vol. 45, No. 3, August 2004

## **WHY DO GOOD COPS DEFEND BAD COPS?\***

BY JEAN-PIERRE BENOÎT AND JUAN DUBRA<sup>1</sup>

*Department of Economics and School of Law, New York University;  
Universidad de Montevideo*

Policemen are known to support colleagues who are the subject of criminal investigations. While we might expect guilty officers to defend each other, why do (otherwise) law-abiding policemen defend those who have broken the law? We investigate under what conditions it is in the interest of a group to defend its “bad” members.

### 1. INTRODUCTION

Investigations into unlawful police behavior inevitably confront the notorious *blue wall of silence*: Policemen and women are extremely reluctant to testify against their colleagues. In a similar vein, police officers accused of brutality or involved in questionable shootings are routinely staunchly defended by their fellow officers.<sup>2</sup> As has often been noted, this intragroup solidarity finds a counterpart on the other side of the law in the Mafia code of silence known as *omertà*. There is at least one important difference in these two protocols, however. Whereas the Mafia code of conduct involves lawbreakers protecting each other, the police code involves some honest policemen protecting criminal policemen, as well as competent policemen protecting incompetent ones.<sup>3</sup> This raises the question: Why do good cops defend bad cops?<sup>4</sup>

At the individual level, an immediate answer is that there are enforcement mechanisms in place to encourage police solidarity. Officers who testify against other officers face a variety of sanctions ranging from a social shunning to a lack

\* Manuscript received—; revised—.

<sup>1</sup> We thank Barry Adler, Yochai Benkler, Paul Chevigny, Federico Echenique, Barry Friedman, Néstor Gandelman, Lewis Kornhauser, John Ferejon, Yaw Nyarko, Ron Noble, Jerome Skolnick, and Alan Schwartz for helpful discussions; Stephanie Webster and James Perez for research assistance; and John Kleinig for kindly providing references. Benoît acknowledges the support of the C.V. Starr Center at NYU. Part of this work was done while Benoît was visiting at the Institut Analsi Economica, Barcelona. Please address correspondence to: Jean-Pierre Benoît, Department of Economics, New York University, 269 Mercer Street, 7th Floor, New York, NY 10003-6687. Phone: 212 998-8900. Fax: 212 995-3932. E-mail: [jpb1@nyu.edu](mailto:jpb1@nyu.edu).

<sup>2</sup> One exception is the police group 100 Blacks in Law Enforcement Who Care, which has been severely critical of the use of force in several controversial cases.

<sup>3</sup> The Mollen Commission writes, “The vast majority of honest officers still protect the minority of corrupt officers.” Mollen Commission (1994) as quoted by Kleinig (2000).

<sup>4</sup> In police parlance the term “good cop” often denotes a cop who respects the wall of silence. That is not, of course, how we are using the term.

Q1

of adequate “back-up” in dangerous situations.<sup>5</sup> However, this answer begs the question of why these sanctions are there in the first place. After all, one could just as easily imagine a police culture in which officers who testify against corrupt, incompetent, or overly violent colleagues are rewarded rather than sanctioned.

Policing is a stressful and dangerous job that engenders a “feeling of vulnerability and threat and need for protection, the sense that protection can and will come only from within . . . and . . . a need therefore not to be second guessed . . .” (Kleinig, 2000). These factors encourage a bonding that may find expression in an uncompromising support system. Although this explanation doubtlessly has much validity, it remains incomplete. Police unions, presumably acting on behalf of the general pool of officers, also strongly defend (almost) all officers accused of inappropriate behavior. The unions could adopt a more nuanced approach while line officers maintained a strict supportive attitude. In addition, other relatively impersonal institutional mechanisms could arise.

The reason for a universal defense of officers may seem obvious: An officer acting properly may be mistaken for one acting improperly, especially by outsiders but even by other officers. Defending all policemen, including justifiably accused ones, protects the falsely accused. This explanation, however, neglects the costs associated with a broad defense.

Hopefully, one reason that at least some people choose to become police officers is out of a desire to see the law upheld rather than broken. These people may feel a “moral” cost when dishonest or abusive officers are protected.<sup>6</sup> Even if such feelings are absent, there are costs of a purely selfish nature to the routine defense of inappropriate conduct. When police are defended indiscriminately, the fact that an officer is proclaimed innocent by his fellow officers and others acting on his behalf carries little, if any, weight. On the other hand, if only good policemen, or those perceived to be good by other officers, are defended, then the assertion that an officer is innocent of wrongdoing will be viewed as significant by outsiders. Good police are losing this potential benefit. In addition, any evaluation of a specific policeman will factor in the overall reputation of his or her department. To the extent that the defense of bad police officers is successful, these officers are not removed from the police force and the overall quality of the department is lowered. The reduced prior probability that a randomly selected individual is good works against anyone whose integrity or competence is questioned.<sup>7</sup>

Police are not the only group who protect each other. Although by law, serious hospital errors in Florida must be reported to the state, the Florida Agency for Health Care Administration estimates that less than 10% of such errors are actually reported.<sup>8</sup> A recent study by the Institute of Medicine (1999) puts the number of deaths in U.S. hospitals due to medical error<sup>9</sup> at between 44,000 and 98,000 per

Q2

<sup>5</sup> Noted whistle-blower Frank Serpico asserted that he was shot during a drug raid in which he did not receive proper support from other officers.

<sup>6</sup> The rare officers who breach the wall often cite such moral considerations.

<sup>7</sup> On the other hand, defending bad officers may be dynamically beneficial if the *perception* of the overall quality of officers depends upon the number who are convicted. We thank a seminar participant at the Universidad de Alicante for this observation.

<sup>8</sup> As reported by Anna Polk, an administrator in the Agency to CNN&TIME (2000).

<sup>9</sup> These errors include diagnostic errors, surgical errors, and medication errors.

year.<sup>10</sup> Despite this, “two thirds of the nation’s hospitals haven’t reported a single adverse incident involving a physician in the last eight years.”<sup>11</sup> According to a CNN&TIME (2000) investigation there exists a veritable *white wall of silence* in the medical profession whereby doctors do not report each other’s (or their own) mistakes. This wall has its own enforcement mechanism. One doctor explained that “other doctors could put him out of business by refusing to refer him patients.”

As with the blue wall of silence, the white wall of silence has its costs. There is the moral cost of seeing patients harmed. In addition, injured patients and their families may be more likely to sue for malpractice despite assurances that the treatment was appropriate, if such assurances are routinely given.

Not every group demands such loyalty. For instance, as far as we can ascertain, there is no comparable wall of silence among professors. Consider a professor accused of biased grading by a student. His or her colleagues might well initially react by saying that the charges are self-interested, that grading is a subtle process subject to misunderstanding by outsiders and that there is ample room for honest errors. These responses all mirror responses we might expect in the police and medical community to allegations of misconduct. Nevertheless, beyond this initial and superficial reaction most professors would not feel the compulsion to stand by a professor who appeared, in fact, to be in the wrong. To be clear, we are not suggesting that many professors spend their time rooting out irresponsible colleagues. Rather, our assertion is that if a professor chose to side with a student against another professor in a grading dispute, he or she would not face the opprobrium of uninvolved faculty members. There is no real academic wall of silence.

Of course, policemen and doctors are regularly engaged in activities that might result in the death or injury of another person, whereas professors are not. But what does this difference explain? One might well expect policemen and doctors to be more careful in distinguishing good colleagues from bad ones, not less careful.

Although there may be little pressure on professors to defend bad-acting colleagues, neither is there much pressure on them to condemn such colleagues. In contrast, the West Point Honor Code states that “A Cadet Will Not Lie, Cheat, or Steal, *nor Tolerate Those Who Do* (emphasis added).” Although this “anti-silence” stricture is, to a certain degree, dictated from above, it still requires the cooperation of the cadets themselves to function. All the aforementioned examples concern fixed protocols. On the other hand, the recent sex-abuse scandal involving priests has seen a shift in the Catholic church’s attitude from tight-lipped to open.

We propose a model to address the issue of why and when “good” group members defend “bad” ones.<sup>12</sup> Our explanatory variables include the difficulty that outsiders have in judging group members, the difficulty insiders have in judging each other, the overall quality of the group and the level of certainty needed to punish bad agents.

<sup>10</sup> In contrast, the number of deaths in automobile accidents in 1998 has been estimated at 43,458.

<sup>11</sup> This quote and the next one are from CNN&TIME (2000).

<sup>12</sup> Prendergast (2001) also considers the blue wall of silence, though this is not his primary concern. In his model (which is quite different from ours) society optimally chooses all police officers to be “bad,” and the blue wall consists exclusively of bad agents defending each other.

## 2. THE MODEL

A drug dealer accuses the officers arresting him of stealing money. Should these accusations be given credence? An unarmed man is shot and killed by policemen. Though the man posed no actual threat, was the shooting nonetheless justified? A patient dies during a minor operation. Was it poor luck or malpractice? Given the undesirable outcomes, a legal court may be called upon to determine the legitimacy of the actions, and group members may be asked to testify.

As the prisoner's dilemma illustrates, individual interests may deviate from the group interest to the detriment of all in the group. Discussion of codes of silence often focus on the mechanisms used to ensure individual compliance with group norms. In the present context, however, the appeal to the prisoner's dilemma is, at best, premature. When agents differ, it is not a priori clear what exactly is in the "group's interest." Identifying this group interest is our interest, rather than the mechanisms for achieving it. Accordingly, in our modeling we will abstract away entirely from the enforcement aspect of the problem.

We posit three different types of actors. At the bottom level there are the *agents* (e.g., police officers or doctors). The agents are of varying (action) types represented by a real number  $t$ , which we will take to be drawn from a normal distribution,  $N(a, A)$ . Lower numbers indicate worse types. Thus, a very high-type agent could be a police officer who uses the minimal amount of force needed in a given situation, a moderately low type an officer who uses more force than is desirable though still an "acceptable" amount, and a very low type one who uses an unwarranted and illegal amount. Only the agent knows his own type. Note that we make no distinction between the quality of the agent and the quality of the agent's action. Thus, when we speak of the agent's type the reader can understand this as the type of the agent's action.

At the top level there is the *court*, which is unconnected to the agents. Although the quality of the agents varies along a continuum the courts have a dichotomous view: Each agent, or each agent's action, is either bad (illegal) or good (legal).<sup>13</sup> A *bad* agent is one whose type is less than or equal to some value  $b$ . An agent of type  $t$  is *good* if  $t > b$ . The court's problem is to determine if the agent is bad or good, that is, to declare the agent guilty or innocent. The court has a distaste for convicting innocent people and will declare an agent guilty only if it believes the agent to be guilty with some minimal probability  $\bar{p} \geq \frac{1}{2}$ . We call  $\bar{p}$  the *standard of conviction*. The court believes that agent types are distributed  $N(\hat{a}, A)$ , and this belief is a common knowledge. Although we will mostly be concerned with the case where the court's beliefs are correct,  $\hat{a} = a$ , some interesting comparative statics arise when we allow for the court to misperceive the median agent quality.

At the intermediate level there is the group behavior, which we will refer to as the *union*. The union acts on behalf of the agents as a whole. For concreteness, we assume that the union, or group, acts in accord with majority sentiment, though for the most part the choice of this specific fraction is unimportant.

<sup>13</sup> In practice, for a single action an agent could be guilty of varying crimes of different seriousness. This still results in a (small) finite partition of the agents and we simplify by considering a partition of size two.

Given an agent, the court and the union conduct independent investigations and receive (conditionally) independent normally distributed signals  $c$  and  $u$  of the agent's type. For an agent of type  $t$ , the court's signal is distributed  $N(t, C)$ , whereas the union's signal is distributed  $N(t, U)$ . Thus it is possible, though relatively unlikely, for a good agent to send off a bad signal and for a bad agent to send off a good signal: Even a good doctor's patient might die (a bad signal), whereas a bad doctor's patient might recover (a good signal) despite a clumsy misdiagnosis.

The assumption that the signals and the agent type are normally distributed simplifies the analysis as the posteriors are then also normally distributed. The central limit theorem provides some justification for this assumption. However, the assumption of normals is not crucial. As will become clear, the essential property we use is that the signals can vary from informative (low variance) to uninformative (high variance).

The court's signal reflects publicly known information. The union's signal reflects two things: *insider* knowledge that fellow agents may possess by virtue of their intimate contact with each other and *expert* knowledge. Both these types of knowledge may be difficult to (credibly) convey to outsiders. Positive signals may be particularly difficult to convey. For instance, if a patient dies unexpectedly during an operation, a negative signal could be the administration of an incorrect drug or an improper incision, rather concrete signals. On the other hand, a positive signal could be the doctor's good reputation or an informed opinion that everything was done correctly, an opinion somewhat belied by the result. Or consider the fatal shooting of a suspect by a police officer. A weapon in the hands of the dead man, or "civilian" witnesses testifying that the suspect behaved threateningly are both positive indicators, but these would be part of the court's signal. A positive signal to the union might be the assertion from another officer that the suspect behaved threateningly for the officer's reputation in the department. In both the above examples, the positive union signals are difficult to transmit to outsiders, all the more so as there is a natural tendency to disbelieve positive assertions by the union that is, after all, acting on behalf of the agents.<sup>14</sup>

We make the simplifying stark assumption that in order to be believed the union must establish a reputation for honesty, and that it can do so only by always truthfully reporting its signal, be it good or bad. Thus, we restrict the union to two possible strategies. A *candid* union truthfully reports any signal it receives to the court. An *indiscriminate* union reports that it has received an arbitrarily high signal, regardless of the actual signal. In other words, an indiscriminate union defends all agents with equal vigor. This corresponds to the wall of silence in its purest form.

The court must decide whether or not the agent is bad ( $t \leq b$ ). In doing so the court considers the overall distribution of agents, the signal it has received and the union's report. The court has rational expectations and knows whether the union is following a candid strategy or an indiscriminate strategy. Hence, if the

<sup>14</sup> Of course, these signal examples are merely illustrative. Technically, the union's signal consists of that portion of its information that is not directly available to outsiders.

union is following a candid strategy the court has access to two signals, whereas if the union is following an indiscriminate strategy the court ignores the union's report, effectively leaving it with only its own signal.<sup>15</sup> As an example of the latter phenomenon, Chevigny (1995) notes that in 1989 the Civilian Complaint Review Board in New York city reported that it "had never had a case in which a police witness testified against another" and that, presumably as a consequence, "investigators routinely gave little weight to police officers' stories when they corroborated the account of another police officer."

To summarize—Each agent emits two signals, one to the union, one to the court. The union either candidly reports all signals or indiscriminately supports the agents. The court then updates its beliefs in a Bayesian fashion and convicts the agent if the probability that the agent is bad meets the standard of conviction.

Suppose the agent emits signals  $u$  and  $c$ . If the union is candid the court has a posterior belief about the agent's type, which is normally distributed (see Appendix)

$$N\left(\frac{uCA + cUA + \hat{a}UC}{CA + UA + UC}, \frac{UCA}{CA + UA + UC}\right)$$

The court convicts if  $\text{Prob}(t \leq b | u, c) \geq \bar{p}$ , i.e., if

$$\text{(Candid)} \quad \Phi\left(\frac{b - \frac{uCA + cUA + \hat{a}UC}{CA + UA + UC}}{\sqrt{\frac{UCA}{CA + UA + UC}}}\right) \geq \bar{p}$$

where  $\Phi$  is the cumulative distribution function of the standard normal.

If the union is indiscriminate the court has a posterior belief about the agent's type, which is normally distributed

$$N\left(\frac{cA + \hat{a}C}{A + C}, \frac{CA}{C + A}\right)$$

The court convicts if  $\text{Prob}(t \leq b | c) \geq \bar{p}$ , i.e., if

$$\text{(Indiscriminate)} \quad \Phi\left(\frac{b - \frac{cA + \hat{a}C}{C + A}}{\sqrt{\frac{CA}{C + A}}}\right) \geq \bar{p}$$

<sup>15</sup> Even with an indiscriminate strategy the union might have some success in transmitting some positive signals. However, this will be more difficult than with a candid strategy and we simplify by assuming that an indiscriminate strategy is effectively a babbling one.

Let  $x$  be such that  $\Phi(x) = \bar{p}$ . For a type  $t$  agent the (a priori) probabilities of conviction are

$$(1) \quad P_C(t) \equiv \Phi \left( \frac{b - \sqrt{\frac{UCA}{CA+UA+UC}}x - \left(\frac{UC}{CA+UA+UC}\right)\hat{a} - \left(\frac{CA+UA}{CA+UA+UC}\right)t}{\sqrt{\frac{C^2AU+U^2AC}{(CA+UA+UC)^2}}} \right)$$

with a candid union and

$$(2) \quad P_I(t) \equiv \Phi \left( \frac{b - \sqrt{\frac{CA}{C+A}}x - \left(\frac{C}{C+A}\right)\hat{a} - \frac{A}{C+A}t}{\frac{A}{C+A}\sqrt{C}} \right)$$

with an indiscriminate union.

Note that the model is a static one. A selfish type  $t$  agent prefers an indiscriminate union to a candid union if  $P_I(t) < P_C(t)$ .

- In principle, the median agent may be good or bad. The most interesting case is when most agents are good and so we assume that *the median agent is good* ( $a > b$ ). We also assume that the court believes the median agent to be good:  $\hat{a} > b$ . We remind the reader that in addition we are assuming that the standard of conviction  $\bar{p} \geq \frac{1}{2}$ . In Section 3 and Proposition 7 we relax these assumptions.
- Although we mostly use the terms “union” and “court,” the model should be interpreted more generally. Thus, in our leading interpretation the union is a proxy for individual behavior as dictated by the group culture. In addition to being a legal court, the court could be the “court” of public opinion,<sup>16</sup> or the police commissioner disciplining officers.
- We do not look for each agent’s optimal union strategy; instead we concentrate on two simple strategies, candid and indiscriminate as these two strategies effectively capture the trade-offs involved in deciding upon the group behavior. This last claim is confirmed in Section 4, where we consider the unrestricted optimal strategy in a model with only two types.

2.1. *Some Intuition.* Suppose each agent cares only about himself, and to simplify matters, assume that the court’s beliefs are correct,  $\hat{a} = a$ . At first blush it might seem that an above-average agent ( $t > a$ ) would always prefer a candid union to an indiscriminate one, since two signals provide more accuracy than one signal and the agent is better than the prior expectations about him. This would indeed be the case if this were a wage setting model in which risk-neutral agents were paid, say, their expected type. In such a setting, above-average agents, and

<sup>16</sup> The New York Times (2000) reports on the psychological cost to “street cops” of the negative public opinion of the police following a series of controversial police actions.

only above-average agents, prefer a candid union to an indiscriminate one.<sup>17</sup> The situation is different here, however. The agent's expected quality does not matter, only the probability that he is bad. We now provide some intuition as to why even a high-type agent may prefer an indiscriminate union. There are two effects at work, a "prior effect" and a "standard of conviction effect."

First suppose that  $\bar{p} = \frac{1}{2}$ . That is, an agent is declared guilty provided there is at least a 50% chance that he actually is guilty. This is, of course, far below the "reasonable doubt" standard of a criminal court but it corresponds to the "preponderance of evidence" standard of a civil court.

Suppose that  $a > b$  (the median agent is good), and consider an above average agent  $t > a$ . As an initial step, imagine that the court receives no signal of its own—essentially  $C = \infty$ . If the union is indiscriminate, the court is unable to update so that its beliefs about the agent are just its prior, which is centered at  $a$ . The court's median estimate of the agent's quality is lower than the agent's actual quality. Nevertheless, this estimate is still good enough for the court to acquit with certainty, since there is less than a 50% chance that the agent is bad. On the other hand, suppose the union is candid. Since the agent is above average, in all likelihood the court's median posterior belief about the agent will be better than when the union is indiscriminate. However, the (small) possibility also arises that the union's signal will be so poor that when the court updates the agent is now convicted. Hence, an above-average agent, indeed all agents, will prefer an indiscriminate union.

Now let us increase the accuracy of the court's signal by lowering  $C$  (from  $\infty$ ). For very large  $C$ , in calculating the posterior the court places much more weight on the "good" prior  $a$  than on the random signal  $c$ . Thus, it is virtually certain that the agent will not be convicted with an indiscriminate union and so the agent favors this type of union. As  $C$  decreases, more and more weight is shifted away from the prior toward the signal; eventually an above-average agent desires the increased accuracy that an extra signal provides and favors a candid union.

Now raise the conviction standard  $\bar{p}$  above  $\frac{1}{2}$ . Suppose that, as it happens, the signals  $u$  and  $c$  are such that the court's posterior distribution is centered at  $r$  whether the union is candid or indiscriminate. Although the two distributions have the same median, the indiscriminate distribution has a larger variance. If  $\bar{p}$  were equal to  $\frac{1}{2}$ , this difference in variances would be immaterial; with either union strategy the court would convict if  $r < b$  and acquit if  $r > b$ . When  $\bar{p} > \frac{1}{2}$ , however, the court is reluctant to convict an agent without accurate evidence. Now, even if  $r < b$  the court will not convict if the variance of the court's posterior is large enough. The fact that an indiscriminate union yields a larger variance than a candid union makes the indiscriminate union attractive to the agent. In general, when  $\bar{p} > \frac{1}{2}$  the imprecision of the indiscriminate union is a factor in its favor.

<sup>17</sup>The expected average posterior of a type  $t$  agent turns out to be  $\frac{CA+UA}{CA+UA+UC}t + \frac{UC}{CA+UA+UC}a$  with a candid union and  $\frac{A}{C+A}t + \frac{C}{C+A}a$  with an indiscriminate union. The former is greater than the latter if and only if  $t > a$ .

Note that whereas the court's utility is monotonically increasing in the number of signals,<sup>18</sup> the agent's utility is not necessarily monotonic. In particular, the agent's ranking may be no signals, followed by two signals, followed by one signal.

*2.2. Individual Agents' Preferences.* We will restrict our attention to the case of agents who are purely self-interested;<sup>19</sup> each agent seeks to minimize the probability that he is convicted. The first proposition indicates that, as one might expect, better agents favor greater candor.

**PROPOSITION 1.** *For all parameters, there is a unique type  $T$  such that all agents worse than  $T$  prefer an indiscriminate union to a candid union, whereas all agents better than  $T$  prefer a candid union to an indiscriminate one. The cutoff  $T$  is defined by the condition  $P_C(T) = P_I(T)$ .*

**PROOF.** All propositions are proved in the Appendix.

Thus, if an agent is good enough he favors a candid union. Nevertheless, as the next proposition shows, there are always some good agents who desire an indiscriminate union. The reason is that agents who are slightly good, but below (the believed) average, want the union to rely heavily on the relatively favorable prior. All bad agents also desire an indiscriminate union.

**PROPOSITION 2.** *For all parameters, all bad agents favor an indiscriminate union and some good agents favor an indiscriminate union.*

We see that quite generally some good *individual* agents want bad agents to be defended (by the group).<sup>20</sup> However, we have a *collective* action problem where the behavior of the entire group, or union, is determined by the will of the majority. Since the majority is defined by the median-type agent  $a$ , when  $T > a$  bad types will be defended, whereas when  $T < a$  bad types will not be defended. More precisely, if  $T > a$  agents who emit bad signals have these signals concealed, whereas if  $T < a$  these signals are revealed. Hence, it is important to understand under what conditions  $T > a$ . We consider this in the next section.

*2.3. The Union's Strategy.* We now investigate how individual preferences translate to the group strategy as embodied by the union. We assume that the union's strategy is determined by the desire of the majority (though most of our results do not depend upon this specific assumption). Hence, if the cutoff type  $T$

<sup>18</sup> This follows from Blackwell (1953); in Section 2.4 we explicitly model the court's utility.

<sup>19</sup> Thus, we ignore the "moral" costs mentioned in the Introduction. Since our model is static, we are ignoring the "dynamic" costs as well.

<sup>20</sup> In Holmstrom and Costa (1986) and Zwiebel (1995), managers do not take a "positive" action, despite the fact that it increases expected returns, because the action also leads to a diffuse posterior on the managers' types. In contrast, some good agents in our model shy away from the "positive" candid action because this leads to a tight prior. Thus, despite a surface similarity, the reasonings are opposite.

is greater than  $a$  the union is indiscriminate, whereas if  $T$  is less than  $a$  the union is candid. Recall that  $a > b$ . Therefore, when  $T > a$  the union defends bad agents as well as good agents despite the fact that bad agents form a minority.<sup>21</sup>

Criminal courts ask that a defendant be proven guilty “beyond a reasonable doubt.” This requirement for strong evidence makes it difficult to convict an agent with an uninformative indiscriminate union, which is beneficial to the agents. Thus we have:

**PROPOSITION 3.** *Ceteris paribus, if the standard of conviction  $\bar{p}$  is high enough then the union is indiscriminate. The fraction of agents favoring an indiscriminate union is increasing in  $\bar{p}$ , and as  $\bar{p}$  approaches 1 this fraction also approaches 1.*

Consider the interpretation of the court as an employer who wants to fire bad agents. An employer that uses a high  $\bar{p}$  is “employee-friendly” in that she is reluctant to decide against the employee. Somewhat perversely, the response to this friendliness is a lack of cooperation by the employees.

Medical doctors are trained for many years and must be accredited. One effect of this process is to temper the disparity in the quality of would-be doctors. As a consequence, one might expect the variance in the quality of actual doctors to be small. A small variance in the prior favors an indiscriminate union, as the court then places almost all its updating weight on the “good” prior.

**PROPOSITION 4.** *Ceteris paribus, if the prior is precise enough, that is if  $A$  is small enough, then the union is indiscriminate. Furthermore, as  $A$  approaches 0 all agents favor an indiscriminate union. If  $A$  is large enough the union is candid; in the limit as  $A$  approaches infinity only bad agents favor an indiscriminate union. The fraction of agents favoring an indiscriminate union is strictly decreasing in  $A$ .*

Policemen both good and bad may easily find themselves accused of improper conduct,<sup>22</sup> just as good and bad doctors may find themselves accused of incompetency. Furthermore, these accusations are often difficult for outsiders to properly assess. We can interpret these factors in our model as a large variance in the court’s signal. When  $C$  is large and the union is indiscriminate the court places almost all its weight on the prior and has relatively little information on which to go. Since  $\hat{a} > b$  and  $\bar{p} \geq \frac{1}{2}$  these last two facts are desirable.

**PROPOSITION 5.** *Ceteris paribus, if the court’s signal is imprecise enough, that is if  $C$  is large enough, then the union is indiscriminate. Furthermore, in the limit as  $C$  approaches infinity all agents favor an indiscriminate union. If the court’s signal is precise enough then the union is candid; in the limit as  $C$  approaches 0*

<sup>21</sup> The Fitzgerald Inquiry in Queensland, Australia, found that police extend “protection to corrupt officers, who may or may not be a *minority* of the department.” (emphasis added). Fitzgerald Inquiry (1989), as quoted by Kleinig (2000).

<sup>22</sup> In December 1998 alone, 108 San Francisco police officers faced citizen complaints (City and County of San Francisco, 1999). Of course, this figure is only suggestive as there is no way to accurately determine how many of these complaints were justified.

*only bad agents favor an indiscriminate union. The fraction of agents favoring an indiscriminate union is strictly increasing in  $C$ .*

If the union's information is accurate enough, then above-average agents will want the union's signal to be used. Thus,

*PROPOSITION 6. Ceteris paribus, if the union's signal is precise enough, that is if  $U$  is small enough, the union is candid. Furthermore, in the limit as  $U$  approaches 0 only bad agents favor an indiscriminate union. The fraction of agents favoring an indiscriminate union is strictly increasing in  $U$ .*

Notice that Proposition 6 does not include a converse statement. Large  $U$  may or may not lead to an indiscriminate union, depending upon the other parameters.

As the court's prior becomes better, agents want the court to rely more heavily upon this prior, leading to an indiscriminate union. This is the content of the next proposition (note that, exceptionally, the proposition allows for  $\hat{a}$  to fall below  $b$ ).

*PROPOSITION 7. Ceteris paribus, if the court's prior is high enough, that is if  $\hat{a}$  is large enough, then the union is indiscriminate. Furthermore, in the limit as  $\hat{a}$  approaches infinity all agents favor an indiscriminate union. If  $\hat{a}$  is small enough the union is candid; in the limit as  $\hat{a}$  approaches negative infinity all agents favor a candid union. The fraction of agents favoring an indiscriminate union is strictly increasing in  $\hat{a}$ .*

The previous proposition concerns the perceived agent quality. The following proposition relates to actual agent quality, and is a consequence of the fact that higher types tend to favor candid behavior and the fact that the cutoff  $T$  does not depend on  $a$ .

*PROPOSITION 8. Ceteris paribus, if the quality of the average agent,  $a$ , is high enough, then the union is candid. Furthermore, in the limit as  $a$  approaches infinity all agents favor a candid union. If  $a$  is low enough then the union is indiscriminate. The fraction of agents favoring an indiscriminate union is strictly decreasing in  $a$ .*

Notice that increasing  $\hat{a}$  has an effect opposite in sign to increasing  $a$ . The effect of increasing  $a$  and  $\hat{a}$  at the same time is analyzed in the next section.

**2.3.1. Downplaying the prior.** Arguably, the legal system demands that jurors accord greater weight to the current evidence at hand than to their prior beliefs about the defendant.<sup>23</sup> This can be captured in a Bayesian fashion by requiring that the court use an "effective" variance  $\lambda A$ , in its calculations (rather than the

<sup>23</sup> This seems to be the position articulated in *Smith v. Rapid Transit, Inc.* (1945), where the court ruled that Rapid Transit could not be held liable for an accident on the sole basis that it was a priori likely that the company owned a recklessly driven bus. At the same time, it is indisputable that jurors and others do put some weight on their priors. Indeed, the mere fact that a police officer is on trial, rather than an ordinary citizen, will affect the deliberations.

actual variance  $A$ ), where  $\lambda \geq 1$  is chosen such that  $\lambda A > C$ .<sup>24</sup> We say that the *court's signal weighs more heavily than the prior* when the effective prior variance the court uses is larger than its own signal variance.

Consider the effect of the overall quality of the agents when the court's beliefs are correct, i.e.,  $\hat{a} = a$ . As the median quality of the agents improves, the indiscriminate union looks better and better relative to the candid union, since the indiscriminate union places relatively more weight on the improving prior. Thus, more and more types prefer an indiscriminate union. There is a countervailing effect, however. As the median type increases there are more and more high quality agents and better agents tend to prefer a candid union. When the court's signal weighs more heavily than the prior, the second effect eventually dominates. To see why this is so, suppose that the variance of the prior is very large, say infinity. In that case, increasing the perceived median  $\hat{a}$  has no effect on the cut-off, since the court's beliefs do not depend on the prior, but just on the received signal. Therefore, only the second effect is present. More generally, when the effective prior variance is large enough the second effect is more important than the first.

As the next proposition indicates, the union is indiscriminate when the median agent is good, but not too good, and is candid when the median agent is very good.

**PROPOSITION 9.** *Suppose the court's signal weighs more heavily than the prior and  $\bar{p} > \frac{1}{2}$ . Ceteris paribus, there exists an  $a^* > b$  such that if  $\hat{a} \equiv a < a^*$  the union is indiscriminate, whereas if  $\hat{a} \equiv a > a^*$  the union is candid.*

The next proposition indicates that if there is an indiscriminate union when the court's signal weighs heavily, part of the impetus for it must come from the court using a high standard of conviction.

**PROPOSITION 10.** *Suppose the court's signal weighs more heavily than the prior and  $\hat{a} \equiv a$ . Ceteris paribus, there exists a  $\hat{p} > \frac{1}{2}$  such that if  $\bar{p} < \hat{p}$  the union is candid, whereas if  $\bar{p} > \hat{p}$  the union is indiscriminate.*

Suppose instead that the prior weighs more heavily than the court's signal. As Propositions 4 and 5 suggest, there is then a push toward an indiscriminate union. In contrast to Propositions 9 and 10, if  $A$  is small enough or  $C$  is large enough, the union is indiscriminate for all  $\hat{a} \equiv a > b$  and for all  $\bar{p} \geq \frac{1}{2}$ .<sup>25</sup>

2.3.2. "Only cops understand cops". A much-expressed sentiment among police officers is that the general public does not properly understand the nature of

<sup>24</sup> In the case of a candid union, we could also impose  $\lambda A > U$ . Adding this requirement does not affect the results of this section.

<sup>25</sup> This statement is a stronger version of Propositions 4 and 5, since it varies  $a$  and  $\bar{p}$  after  $C$  and  $A$ . The proof in the Appendix establishes this version.

police work. Doctors, too, believe that lay people are incapable of properly evaluating them. We can capture this by positing that the court misreads a candid signal  $u$  as  $u - m_u$ , and misreads its own signal  $c$  as  $c - m_c$ , where  $m_c, m_u \geq 0$ .

First suppose  $m_u \geq m_c$ . The court evaluates its own signal erroneously and with a candid union, the court has the opportunity to (mis)use the union's signal as well, compounding its error. This pushes toward an indiscriminate union. In fact, even if  $m_u < m_c$  the agents will prefer an indiscriminate union provided  $m_u$  is not too small. On the other hand, if  $m_u$  is sufficiently smaller than  $m_c$ , shifting weight away from the court's signal (and the prior) to the union's signal is desirable, pushing toward a candid union.

**PROPOSITION 11.** *There exists a number  $\beta \in (0, \frac{1}{2})$  such that if  $m_u > \beta m_c > 0$  more agents favor an indiscriminate union than when  $m_u = m_c = 0$ , whereas if  $m_u < \beta m_c$  more agents favor a candid union than when  $m_u = m_c = 0$ .*

**2.3.3. Different kinds of events.** On May 25, 1999, New York City police officer Justin Volpe admitted to brutalizing detainee Abner Louima by inserting a broken broomstick into his rectum. Volpe pleaded guilty after other officers testified against him. Although these officers breached the wall of silence, apparently they were not subject to the usual sanctions from their colleagues (New York Times 1999). Several commentators saw this relatively unpunished defection from the code of silence to be the result of the reprehensible nature of Volpe's crime. However, from a purely ethical standpoint, it is not clear that sodomizing a detainee with a broomstick is any worse than, say, choking a detainee to death. Yet when New York City police officer Daisy Boria testified against her partner Frank Livoti in the choke-hold death of Anthony Baez, she was shunned by her colleagues and threatened with death; when Mark Virginia died while handcuffed in Buffalo, New York, only one officer out of more than two dozen police witnesses testified that he had been choked by a policeman, and the testifying officer was harassed.<sup>26</sup> We next consider a nonethical distinction between the first case and the last two.

When a man dies in police custody, claims that the death was the unfortunate result of a justifiable response to the man's actions may be valid, even if the circumstances make this seem unlikely. On the other hand it seems virtually impossible to imagine a valid reason for using a broomstick as Volpe did.<sup>27</sup> In our terminology, the negative signals sent out by Volpe had an extremely small variance—they

<sup>26</sup> In the former case, Livoti was ultimately convicted of civil rights violations (New York Times, 1999); in the latter case the accused officer was found not guilty, despite the fact that the jurors believed an illegal choke hold had been applied (they were unconvinced that the choke was the cause of death) (Buffalo News, 1997). It is worth noting that neither of the victims was a criminal; the first victim was arrested after a thrown football accidentally struck a police car, while the second victim was mistakenly suspected of buying drugs, when he was, in fact, lending an employee some money.

<sup>27</sup> Indeed, rather than justifying the actions the defense initially implied that Louima had been injured in an act of consensual gay sex. However, there was no evidence that a consensual act would lead to the severe injuries Louima suffered or, for that matter, that Louima was in fact gay.

almost surely did not come from a good agent. Small  $C$  and  $U$  are conditions that favor candid reporting (Propositions 5 and 6).

The suggestion, then, is that the group may adopt a more sophisticated strategy, where exceptional events are treated differently—candidly instead of indiscriminately in the above instance. We emphasize that we do not mean that, say, candor would be used when it appeared that the signal in a *particular* case, as opposed to a *kind* of case, happened to be accurate. For instance, an otherwise indiscriminate union would not reveal a bad signal in a case involving the death of a suspect if it appeared that this signal was especially accurate. Allowing this element of discretion itself would inevitably introduce a variability, increasing the variance of the signal.

The following statement by a lieutenant commenting on the above-mentioned Louima case is instructive: “Cops want to believe that their fellow cops could not be capable of doing such a horrible thing . . . But with the Diallo case, the reaction was different. Everyone was, ‘Oh, that could have been me’” (New York Times, 1999). In the “Diallo case” to which the lieutenant referred, an unarmed and innocent suspect was killed when four plainclothes police officers shot at him 41 times. Presumably, by “Oh, that could have been me,” the lieutenant meant that a typical officer fears that he may be involved in what appears to be an unjustified killing, whereas he does not fear being accused of sodomizing a suspect. The statement emphasizes the difference in the two incidents. In particular, the death of a suspect involves the kind of noisy negative signal that any officer might well send, and where the wall of silence is most effective. Indeed, the fact that the jury refused to convict the officers of a wrongful death in the Diallo case, despite the apparent use of excessive force, confirms the noisiness of the signal. At the same time, it is not that one incident involved a death whereas the other did not which is dispositive—a police officer should not expect to be protected when accused of, say, murdering his or her spouse. Neither is the culpability of the officers pertinent, as reactions to the aforementioned choke cases demonstrate.

We note that a detective commenting on the fact that police officers offered incriminating testimony in the Louima case said, “In a situation like this, these guys coming forward and stating what they say, and being, honest, can only help the department,” (New York Times, 1999) showing that policemen are aware of the type of cost/benefit trade-off to the wall of silence that forms the basis of this article.

2.4. *Court Flexibility.* We have taken the standard  $\bar{p}$  used by the court as a primitive. Alternatively, this  $\bar{p}$  can be derived by positing that there are utility gains and losses from properly or improperly characterizing an agent. Specifically, we can write

- (i) Utility to convicting a guilty person =  $d$ .
- (ii) Utility to convicting an innocent person =  $-e$ .
- (iii) Utility to not convicting an innocent person =  $y$ .
- (iv) Utility to not convicting a guilty person =  $-z$ , where  $d, e, y, z \geq 0$ .

Let  $p_g$  be the probability that a given person is guilty. The court optimizes by convicting the person if

$$p_g d - (1 - p_g)e \geq (1 - p_g)y - p_g z$$

$$p_g \geq \frac{e + y}{d + e + y + z}$$

This then yields the standard  $\bar{p}$  as

$$(3) \quad \bar{p} \equiv \frac{e + y}{d + e + y + z}$$

Given a fixed strategy on the part of the union the court optimizes by using  $\bar{p}$  as defined by (3). However, the strategy of the union is not fixed; rather it is determined by the standard the court is using. Because of this, the court may prefer to use a different standard than  $\bar{p}$  in order to induce the union to be candid rather than indiscriminate. Presumably, the standard of conviction used by a legal court, be it criminal or civil, is exogenously given to it. Under other interpretations of the “court,” however, the standard of conviction used can be freely chosen. For instance, employers can decide how much evidence they need to fire their employees, and the district attorney has discretion in deciding which cases to prosecute (being prosecuted is costly to an agent regardless of the ultimate verdict). We now consider the selection of  $\bar{p}$ .

As an example, suppose that  $e = 4(d + z)$  and  $y = 0$ . Then the court would like to declare an agent bad only if there is at least an 80% chance that he is indeed bad. Nevertheless if  $U$  is small enough and  $C$  is large enough the court will have a higher expected utility by announcing a policy of convicting all agents who are bad with, say, at least a 50% chance. The reason is that an 80% standard induces an indiscriminate strategy and the court’s inaccurate signal means that it will not have enough information to convict many bad types. On the other hand, a 50% standard induces a candid strategy, and the fact that the union’s signal is accurate more than compensates for the too low standard.

In a similar manner, an employer who wishes to fire any employee whose type is below  $b$ , may benefit from using a different parameter. In particular, when  $C < A$  or  $U$  is small enough, lowering  $b$  increases the number of agents who favor a candid union. Under the opposite condition, increasing  $b$  increases the number of agents who favor a candid union.

### 3. THE CIVIL STANDARD

Doctors who err, or who may have erred, are usually subject to civil not criminal prosecution.<sup>28</sup> In a civil trial, a defendant loses if “the preponderance of evidence”

<sup>28</sup> While police action may result in civil suits, individual officers typically are not liable. Furthermore, police departments do not use the results of such suits to discipline officers. Thus, “a city may pay thousands, or millions, of dollars on behalf of a brutal officer, yet the officer pays no price whatsoever.” (Human Rights Watch, 1998)

is against her. This “preponderance” standard is generally interpreted to mean  $\bar{p} = 50\%$  (or just over 50%). According to Proposition 10, when  $\bar{p} = \frac{1}{2}$  and the court’s signal weighs more heavily than the prior, the group is candid, there is no wall of silence. This implication can be reconciled with the existence of a white wall of silence in at least three ways.

1. Jurors may, as a practical matter, require more confidence than the 50% implied by the term “preponderance.”
2. Jurors may, in fact, weigh the prior more heavily than the signal. This will be the case if jurors accept that doctors are generally well trained and jurors are hesitant to trust their ability to evaluate doctors. As we know, small  $A$  and large  $C$  yield a wall of silence.
3. Group behavior may be determined by a submajority. When the court’s signal weighs more heavily than the prior, a minority of agents (including some good agents) favor an indiscriminate union. If this minority determines group behavior, there will be a wall of silence. A minority might be controlling if worse agents, who face a greater probability of losing cases, are more intensive in their “lobbying” efforts. Proposition 10 does not then hold; there may be indiscriminate behavior even when  $\bar{p} = \frac{1}{2}$ .<sup>29</sup>

Of course, these considerations may also apply to groups other than doctors.

#### 4. EMPIRICAL ILLUSTRATIONS

In this section we discuss several empirical illustrations of the model.

Propositions 3 and 10 indicate that a low standard pushes toward candid behavior. Some support for this prediction is found in a research by Klockars et al. (2000). They find a disparity across agencies in the propensity for police officers to report their colleagues’ misdeeds. In particular, there seems to be a greater tendency toward candid behavior in agencies where officers feel there is a lower internal standard of conviction.<sup>30</sup>

Given that low standards of conviction favor candid behavior, it is interesting to note that the standard used for an accused West Point cadet is merely that it be “more likely than not” that the cadet is guilty. After all, the Honor Code injunction for (extremely) candid behavior on the part of cadets has the greatest chance of succeeding if this behavior is in the interest of the cadets themselves.

Proposition 7 indicates that low perceived overall agent quality leads to candid behavior. Consider recent developments involving the Catholic church. In the past a “clerical wall of silence,” both on the part of priests and church leaders, has protected accused priests;<sup>31</sup> now the church has changed its stance and vows

<sup>29</sup> Note that Proposition 10 is the only proposition that depends upon the assumption that the union behavior is determined specifically by a majority.

<sup>30</sup> Based on survey data, the authors report a more forthcoming attitude in departments where officers expect more severe punishments for misdeeds. In personal communication, Klockars (2003) has indicated that those officers feel that, for a given amount of evidence, it is more likely that they will in fact be punished.

<sup>31</sup> See, for instance, New York Times (2002).

to be forthcoming in investigations. How can we account for this change? New revelations have indicated that the number of abusive priests is far greater than had previously been suspected. In effect, the *perceived* distribution of priest types has shifted down, leading to candid behavior (the wall of silence is no longer an effective strategy: Presented with little (additional) information people will now believe a priest to be guilty where they would previously have believed him to be innocent).

Propositions 5 and 6 indicate that small signal variances push toward candid behavior. Let us return to the absence of an academic code of silence. Consider the grading process. There is a fair amount of arbitrariness in grading students. However, a typical response to a student's claim of unfair treatment would, at most, be a change in the grade with no penalty to the professor. Only repeated independent complaints would elicit any possibility of punitive action against a professor, and these are very unlikely to be made against an unbiased professor. In terms of the model, a signal is not a single draw (grading incident) but the average of  $n$  draws, resulting in effective court and union signals with small variances.<sup>32</sup> In contrast, with policemen and doctors even a single adverse incident may lead to costly legal proceedings (large  $C$ ).<sup>33</sup>

Policemen face both a high standard of conviction and a court signal with a high variance, while doctors face only the latter. Consistent with this difference, the blue wall of silence is more stringent than the white wall of silence.<sup>34</sup>

The model also has implications for settings quite different from the ones considered so far. At an abstract level, we have addressed the following question: When is it in the interest of an "object" that is being evaluated as acceptable or unacceptable not to have much information revealed about it, despite the fact that the object is, in fact, acceptable? The answer depends on the interaction between the quality of information and the rule the evaluator is using to decide on acceptability. One U.S. law school (not NYU) adopts the following policy with respect to potential employers. The rankings of students are not revealed to law firms (indiscriminate behavior) but are revealed to judges recruiting clerks (candid behavior). Our model applies to this situation as follows:

The agents are students and the court is the potential employer (either a judge or a law firm). The employer will not hire ("convict") the student if the probability that he or she is bad is greater than  $\bar{p}$ . Call  $(1 - \bar{p})$  the *hiring standard*. Presumably, the hiring standard is at least one-half, so that now  $\bar{p} \leq \frac{1}{2}$ . The union is the law school. A reasonable model of the law school is that it acts in the best interest of the students who are good (as determined by each employer).

<sup>32</sup> If  $x$  is distributed  $N(\cdot, X)$  then  $\sum \frac{x_i}{n}$  has variance  $\frac{X}{n}$ .

<sup>33</sup> Similarly, a single sexual harassment accusation against a professor may have dire consequences and it would be interesting to examine professorial attitudes toward this type of complaint.

<sup>34</sup> We note that while the morbidity and mortality conferences conducted at many hospitals would seem to reduce the variance of the union's signal for doctors, this reduction is mitigated by the fact that these conferences are explicitly intended *not* as peer review processes. Indeed, Hobgood et al. (2000) find that in almost 60% of emergency medicine resident conferences, cases are reviewed anonymously.

Judges are more selective than law firms in two respects: they have a higher hiring standard and a higher notion of what they consider a qualified student to be. In particular, for the judges the median student is unqualified ( $\hat{a} < b$ ), whereas for the law firms (as a group) the median student is qualified ( $\hat{a} > b$ ). The following result explains the behavior of the law school.

**PROPOSITION 12.** *If  $\hat{a} < b$ , all good students (and some bad students) favor candid behavior. If  $\hat{a} > b$  and the hiring standard is not too high then all good students (and some bad students) favor indiscriminate behavior.*

## 5. TWO TYPES

Although the court divides the agents into two classes, we have considered a model with a continuum of agent types. In this section we briefly describe a model with only two types of agents.

Suppose there are two types of agents, good ( $g$ ) and bad ( $b$ ), with a fraction  $p > \frac{1}{2}$  being good. The court receives a signal  $c_g$  or  $c_b$ , and the union receives a signal  $u_g$  or  $u_b$ , such that  $\Pr(c_g | g) = \Pr(c_b | b) \equiv c > \frac{1}{2}$  and  $\Pr(u_g | g) = \Pr(u_b | b) \equiv u > \frac{1}{2}$ . Note that when  $c$  or  $u$  is close to  $\frac{1}{2}$  the signal is very inaccurate, corresponding to high  $C$  or  $U$  in the previous model, whereas  $c$  or  $u$  close to 1 indicates a very accurate signal, corresponding to small  $C$  or  $U$ . A strategy for the union is a probability with which to truthfully report a signal  $u_g$  and a probability with which to truthfully report a signal  $u_b$ . A candid union always truthfully reports its signal whereas an indiscriminate union reports  $u_g$  regardless of its actual signal. Solving for the *optimal mixed* strategy on the part of the union (assuming rational expectations by the court) yields the following analogue to our previous results:

**PROPOSITION 13.** *If the court's signal is inaccurate enough then all agents favor an indiscriminate union; if the standard of conviction is high enough then all agents favor an indiscriminate union; if the fraction of good agents is large enough then all agents favor an indiscriminate union.*

The results of this two-type model reinforce those of the continuum model. While this two-type model has the advantage of permitting a derivation of the *overall* optimal union strategy, it has the disadvantage of allowing fewer comparative statics and of not revealing a sufficient heterogeneity of opinion among the agents,<sup>35</sup> hence we prefer the continuum model.<sup>36</sup>

**PROPOSITION 14.** *There exists a  $\bar{C}$  and  $\bar{b}$  such that if  $C > \bar{C}$  then a majority of agents find it overall optimal for the union never to reveal (a positive measure of) signals below  $\bar{b}$ . In particular, indiscriminate behavior is strictly better than revealing (a positive measure of) signals below  $\bar{b}$ .*

<sup>35</sup> Note that there is unanimity of opinion in Proposition 13.

<sup>36</sup> We note the following overall result in the continuum model:

6. CONCLUSION

We have provided an explanation for intragroup solidarity, and its limitations, within a selfish rational agent framework. This should not be taken to mean that we feel that bonding rituals, an esprit de corps, and the like are negligible factors. Rather, these may play a role in addition to the elements we have considered. Our discussion has focussed on people grouped by virtue of their profession, although groups are constituted on many other bases as well. We leave it to the reader to decide to what extent our arguments are applicable to these other groupings.

We have assumed that the behavior of the group is determined by the desire of the majority, although for the most part a sufficient assumption is that the probability that the group acts indiscriminately is an increasing function of the number of agents who favor such behavior. Even this latter assumption, though quite plausible, need not be true and one may imagine scenarios in which norms develop independently of the desires of group members. Nonetheless, it seems to us more surprising when a group develops norms that go against its interest than when it develops norms that go with its interest.

Recall from Section 2 the comment of the Civilian Complaint Review Board in New York city that police officers never testified against each other, leading investigators to ignore “the police officers’ stories when they corroborated the account of another police officer.” On the face of it, the quote reveals something a little puzzling, as there is an obvious sense in which the police strategy of not testifying is not working. Note, however, that it is one thing for the Review Board to ignore corroborating reports and quite another for it to rule against accused officers. In fact, as Chevigny (1995) notes, the Review Board used a high standard of proof and rarely sustained complaints against officers. Thus, despite the fact that the positive reports were being ignored, the indiscriminate strategy did “work.”

In our model, the agents’ actions are givens, precluding any incentive effects. Where the agent’s type is a reflection of his or her skill this approach may be proper; it is only a first step when the agent’s type reflects a conscious decision.

APPENDIX

In this section, we derive the results of the main body.

When the union is candid, the court’s posterior probability that an agent is of type  $t$  given the signals  $c$  and  $u$  is given by

$$\begin{aligned}
 p(t | c, u) &= \frac{\frac{1}{\sqrt{2\pi C}} e^{\{-\frac{1}{2C}(c-t)^2\}} \frac{1}{\sqrt{2\pi U}} e^{\{-\frac{1}{2U}(u-t)^2\}} \frac{1}{\sqrt{2\pi A}} e^{\{-\frac{1}{2A}(t-\hat{a})^2\}}}{\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi C}} e^{\{-\frac{1}{2C}(c-t)^2\}} \frac{1}{\sqrt{2\pi U}} e^{\{-\frac{1}{2U}(u-t)^2\}} \frac{1}{\sqrt{2\pi A}} e^{\{-\frac{1}{2A}(t-\hat{a})^2\}} dt} \\
 &= \frac{1}{\sqrt{2\pi} \frac{UCA}{CA+UA+UC}} e^{\{-\frac{1}{2} \frac{CA+UA+UC}{UCA} (t - \frac{uCA+cUA+\hat{a}UC}{CA+UA+UC})^2\}}
 \end{aligned}$$

Thus, the posterior is distributed

$$N\left(\frac{uCA + cUA + \hat{a}UC}{CA + UA + UC}, \frac{UCA}{CA + UA + UC}\right)$$

Similar reasoning establishes the distribution in the indiscriminate case.

Expressions (1) and (2) in Section 2 are easily derived, so that a type  $t$  agent strictly prefers an indiscriminate union if

$$\begin{aligned} & \Phi\left(\frac{b - \sqrt{\frac{UCA}{CA+UA+UC}}x - \frac{UC}{CA+UA+UC}(\hat{a} - t) - t}{\sqrt{\frac{C^2A^2U + U^2A^2C}{(CA+UA+UC)^2}}}\right) \\ & < \Phi\left(\frac{b - \sqrt{\frac{CA}{C+A}}x - \frac{C}{C+A}\hat{a} - \frac{A}{C+A}t}{\frac{A}{C+A}\sqrt{C}}\right) \end{aligned}$$

Solving for  $t$  yields that an agent strictly prefers an indiscriminate union if and only if his type is lower than  $T$ , where

$$\begin{aligned} \text{(A.1)} \quad T &= \left(1 - \frac{C\sqrt{U}}{A\sqrt{C+U}}\right)b + \frac{C\sqrt{U}}{A\sqrt{C+U}}\hat{a} \\ &+ \frac{[\sqrt{C+A} - \sqrt{C\frac{U}{C+U} + A}]\sqrt{C}\sqrt{U}}{\sqrt{A}(\sqrt{C+U} - \sqrt{U})}x \end{aligned}$$

This establishes Proposition 1.

Recall that  $x$  is implicitly defined by  $\Phi(x) = \bar{p}$ , where  $\Phi$  is the standard normal cumulative distribution function. Writing  $x(\bar{p})$  yields  $x(\frac{1}{2}) = 0$ ,  $x(\bar{p})$  is strictly increasing in  $\bar{p}$ , and  $\lim_{\bar{p} \rightarrow 1} x(\bar{p}) = \infty$ .

Since  $x \geq 0$  and  $\hat{a} > b$ , we have  $T > b$ , establishing Proposition 2. Moreover,  $T$  is linear and increasing in  $x$ , establishing Proposition 3. Proposition 12 follows from the fact that  $x(p) \leq 0$  in the hiring context.

Writing  $T = T(A)$ , the limit claims of Proposition 4 follow from the fact that  $\lim_{A \rightarrow 0} T = +\infty$ ,  $\lim_{A \rightarrow +\infty} T = b$ . In fact,  $A < \frac{C\sqrt{U}(b-\hat{a})}{\sqrt{C+U}(b-a)}$  implies  $T > a$  for all  $\hat{a} > b$ ,  $p \geq \frac{1}{2}$ . To see that  $T$  is decreasing in  $A$ , note that

$$\frac{dT}{dA} = -\frac{\sqrt{UC}}{A^2} \frac{\hat{a} - b}{\sqrt{C+U}} - \frac{\sqrt{U}\sqrt{C}^3 [\sqrt{C+U}\sqrt{\frac{CA}{C+A} + U} - U]x}{2\sqrt{A}^3 \sqrt{C+U}(\sqrt{C+U} - \sqrt{U})\sqrt{CA+UA+UC}}$$

Writing  $T = T(C)$ , the limit claims in Proposition 5 follow from the fact that  $\lim_{C \rightarrow 0} T = b$ , whereas  $\lim_{C \rightarrow +\infty} T = \infty$ . To see that  $T$  is increasing in  $C$ , note that the coefficient on  $x$  can be written as

$$\frac{\sqrt{U}\sqrt{C}}{\sqrt{A}\sqrt{C+U}} \frac{[\sqrt{C+U}\sqrt{C+A} - \sqrt{CA+UA+UC}]}{\sqrt{C+U} - \sqrt{U}}$$

The first term is clearly strictly increasing in  $C$ , and the second term is as well since its derivative

$$\frac{q \left[ \sqrt{\frac{U}{C+U}}(U+C)^2 - U(U+2C+A) \right] + U \left[ U \left( 1 - \sqrt{\frac{U}{C+U}} \right) + A \right]}{2\sqrt{UC+AC+AU}(C+U) \left( \sqrt{\frac{U}{C+U}} - 1 \right)^2 \sqrt{U}}$$

can be shown to be positive (where  $q = \sqrt{\frac{C}{C+A} \frac{U}{C+U} + \frac{A}{C+A}}$ ).

Writing  $T = T(U)$ , the limit claims of Proposition 6 follow from  $\lim_{U \rightarrow 0} T = b$  and

$$\lim_{U \rightarrow \infty} T = b + \frac{C(\hat{a} - b)}{A} + \frac{\sqrt{C+A}\sqrt{C}x}{\sqrt{A}}$$

To prove that  $T$  is increasing in  $U$  note that the coefficient on  $x$  can be written as

$$\sqrt{\frac{U}{C+U}} C \frac{C+A}{A} \frac{\sqrt{C+U} - \sqrt{\frac{CA}{C+A} + U}}{\sqrt{C+U} - \sqrt{U}}$$

where the first term is clearly strictly monotonically increasing in  $U$ . The second term is as well, since its derivative

$$\frac{\sqrt{U} - \sqrt{\frac{C+U}{\frac{CA}{C+A} + U}} \sqrt{U} + \sqrt{C+U} - \sqrt{\frac{CA}{C+A} + U}}{2\sqrt{U}\sqrt{C+U} (\sqrt{C+U} - \sqrt{U})}$$

is positive.

To prove Proposition 7, note that in Equation (A.1), the cutoff  $T$  is linear and increasing in  $\hat{a}$ . To prove Proposition 8 note that  $T$  does not depend on  $a$ , so that very large  $a$  implies that almost all agents favor candidness, and conversely for low  $a$ .

To prove Proposition 9, first replace  $A$  by  $\lambda A$  in (A.1) and write  $T(a)$ . Since  $x > 0$  when  $\bar{p} > \frac{1}{2}$ , we have that  $a = b \Rightarrow T(a) > a$ . Furthermore,

$$\frac{dT(a)}{da} = \frac{C\sqrt{U}}{\lambda A\sqrt{C+U}} < 1 \quad \text{for } C < \lambda A.$$

Thus, there exists a unique  $a^*$  for which  $T(a^*) = a^*$ , and  $T(a) \leq a$  according as  $a \geq a^*$ .

To establish Proposition 10 note that since  $x = 0$  at  $\bar{p} = \frac{1}{2}$ ,  $T < a$  when  $\bar{p} = \frac{1}{2}$  and  $C < \lambda A$ , and recall Proposition 3.

The distribution on the posterior when the court misreads the signals as  $u - m_u$  and  $c - m_c$  is given by

$$N\left(\frac{(u - m_u)CA + (c - m_c)UA + aUC}{CA + UA + UC}, \frac{UCA}{CA + UA + UC}\right)$$

in the candid case and by

$$N\left(\frac{(c - m_c)A + aC}{A + C}, \frac{CA}{C + A}\right)$$

in the indiscriminate case.

Letting  $T_{OLD}$  represent the indifferent agent with no misreading as defined by (A.1), the current indifferent agent is given by

$$T = T_{OLD} + m_c \frac{U + C \frac{m_u}{m_c} - \sqrt{U}\sqrt{C+U}}{U + C - \sqrt{U}\sqrt{C+U}}$$

Note that the numerator of the second term is negative if

$$m_u < m_c \frac{\sqrt{U}\sqrt{C+U} - U}{C}$$

Letting

$$\beta \equiv \frac{\sqrt{U}\sqrt{C+U} - U}{C} < \frac{1}{2}$$

we have that  $T < T_{OLD}$  if  $m_u < \beta m_c$  and  $T > T_{OLD}$  if  $m_u > \beta m_c$ , establishing Proposition 11.

#### REFERENCES

- BLACKWELL, D., "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics* 24 (1953), 265–72.
- BUFFALO NEWS (1997) "Jurors Believed Lukaszewski and Blosat."
- CHEVIGNY, P., *Edge of the Knife, Police Violence in the Americas* (New York: The New Press, 1995).
- CITY AND COUNTY OF SAN FRANCISCO, "1998 Annual Report of Office of Citizen Complaints," (1999), website.
- CNN&TIME (2000) "Thin White Line," April 30, in cnn.com.
- CRAWFORD, V., AND J. SOBEL, "Strategic Information Transmission," *Econometrica* 50 (1982), 1431–51.
- HOBGOOD, C., O. JOHN, MA, AND G. SWART, "Emergency Medicine Resident Errors: Identification and Educational Utilization," *Academic Emergency Medicine* 7 (2000), 1317–20.
- HOLMSTROM, B., AND J. R. I. COSTA, "Managerial Incentives and Capital Management," *Quarterly Journal of Economics* 101 (1986), 835–60.

- HUMAN RIGHTS WATCH, "Shielded from Justice: Police Brutality and Accountability in the United States," 1998. **Q4**
- INSTITUTE OF MEDICINE, *To Err is Human, Building a Better Safer Health Care System* (Washington DC: National Academy Press, 2000).
- KLEINIG, J., "The Blue Wall of Silence: An Ethical Analysis," Mimeo, 2000.
- KLOCKARS, C., S. IVKOVICH, W. HARVER, AND M. HABERFELD, "The Measurement of Police Integrity," National Institute of Justice publication, 2000.
- MILGROM, P., AND R. WEBER, "A Theory of Auctions and Competitive Bidding," *Econometrica* 50 (1982), 1089–122.
- NEW YORK TIMES "No One Cries 'Rat' as Officers Take Stand in Louima Trial," 1999.
- , "Branded By the Badge In a City Turned Hostile; Anger Over Shootings Permeates Job, Haunting A Veteran Police Officer," 2000.
- , "Scandals in the Church: The Struggle; Catholic Teachings Are Among Obstacles To Resolving Crisis," 2002.
- PRENDERGAST, C., "Selection and Oversight in the Public Sector, with the Los Angeles Police Department as an Example," Mimeo, 2001.
- SMITH V. RAPID TRANSIT, INC. 58 N.E. 2d 754, 1945.
- ZWIEBEL, J., "Corporate Conservatism and Relative Compensation," *Journal of Political Economy* 103 (1995), 1–25.

## QUERIES

- Q1** PE: Please provide article history.
- Q2** Author: Please check, Institute of Medicine, (1999) is not given in the list.
- Q3** Author: Please provide the web site in this references.
- Q4** Author: Please provide complete details of Ref. Human Rights Watch (1998).