

Witness Intimidation*

Brendan O’Flaherty[†]

Rajiv Sethi[‡]

Forthcoming in: *Journal of Legal Studies*

Abstract

Witness intimidation involves strategic complexity and two-sided uncertainty: criminals cannot know whether threats will deter witnesses, and witnesses cannot know whether threats will be carried out. We model this interaction and explore the manner in which rates of intimidation, testimony, and conviction respond to changes in the value of testimony, police-community relations, and witness protection programs. If the value of testimony rises, criminals face stronger incentives to threaten, but threats are less credible. The increase in threats may be large enough to offset the greater value of testimony, with the paradoxical outcome that fewer criminals are convicted. Counterintuitive results are most likely when witness intimidation is a severe problem: few witnesses testify although prosecutors are competent. When the harm faced by witnesses depends on whether or not the criminal is convicted, communities can be trapped in equilibria with collective silence: no witness testifies because none expects others to testify.

*We thank the Santa Fe Institute and the Institute for Advanced Study for financial support and hospitality, and are especially grateful to Megan Ambrosio, members of the Newark Police Department Homicide Squad, two anonymous referees, and the Editor, Omri Ben-Shahar, for suggestions that have considerably improved the paper.

[†]Department of Economics, Columbia University (bo2@columbia.edu).

[‡]Department of Economics, Barnard College, Columbia University and the Institute for Advanced Study (rs328@columbia.edu).

1 Introduction

If you commit a crime, you will often escape punishment if no one testifies against you. So you have an interest in keeping witnesses from testifying. If criminals routinely succeed in deterring testimony, however, the criminal justice system withers, and laws can be broken with impunity. Witness intimidation is a fundamental threat to the rule of law.

Empirical data on intimidation are hard to come by. When the practice is successful, law enforcement officials seldom hear about it, and the police may not even learn about the underlying crime. A survey in the Bronx courts in 1988 found astonishing levels of intimidation: 36% of victims and witnesses said they had been threatened, 57% of those who had not been explicitly threatened feared reprisals, and 71% of witnesses said they would feel threatened if the defendants were released on bail (Finn and Healey, 1996). Since these respondents were in court, this survey provides a snapshot of *unsuccessful* intimidation; it tells us nothing about the extent of successful intimidation. A British survey (Maynard, 1994) found much smaller rates of (unsuccessful) intimidation in the general population, but considerably more activity in high-crime housing estates, where victims, offenders, and witnesses all tended to know each other and live in close proximity. Later British reports (Tarling *et al.*, 2000) relied on household surveys and so could detect some successful intimidation, but they found much smaller rates of occurrence than the Bronx survey did. Both British surveys are mainly about crimes like vandalism and car theft; little is known about murder and other more serious crimes.

While quantitative data are scarce, anecdotal evidence is plentiful. Police in many cities, including Baltimore, New Orleans, Boston, and Newark, describe chronic difficulties with witnesses who refuse to step forward, and witnesses who change their testimony at the last minute. And certain spectacular instances of intimidation have occasionally been reported in the national press (Kocieniewski, 2007):

A woman who was standing 10 feet away when a stray bullet from a gang fight struck 7-year-old Tajahnique Lee in the face told the police she had been too distracted by her young son to see who fired the shots. A man who was also in the courtyard when that .45-caliber round blew Tajahnique off her bicycle told detectives he had been engrossed in conversation with neighbors and ducked too quickly to notice what had happened. Indeed, at least 20 people were within sight of the gunfight among well-known members of the Sex Money Murder subset of the Bloods gang 15 months ago, but the case remains unsolved because not a single one will testify or even describe what they saw to investigators. The witnesses include Vera Lee, Tajahnique’s grandmother, who declined to be interviewed for this article. People who have spoken to her about the shooting said she would not talk to the police for fear she would “have to move out

of the country.”

In some cases, witnesses fear a loss of reputation rather than physical harm. Hip-hop artist Lil’ Kim served 10 months in prison for perjury after testifying before a federal grand jury that she had not seen her manager and bodyguard at the scene of a 2001 shootout in New York City. Prior to her incarceration, the artist was featured on a television reality show “Lil’ Kim: Countdown to Lockdown”, promoted with the slogan “She’s going to prison with her mouth shut and her head held high” (Kennedy, 2006). Less prominent witnesses fear a different kind of reputation loss; even if there is no danger of injury, exclusion from social networks can be costly enough to deter cooperation with police. Acceptance in such networks can be a matter of survival, especially if the witness finds himself in prison on some other matter.

Threats to witnesses and their families need not be explicit in order to be effective. Symbolic displays work as well as (or better than) verbal threats as long as they are understood by witnesses. Witness intimidators, for instance, can make phone calls in the middle of the night and hang up, puncture tires, or leave dead fish in the mail box. More diffusely, “acting like a gangster (or gangsta)” can be an implicit threat in a society where gangsters are known to harm witnesses. In this context, “acting like a gangster” can mean not only wearing gangster clothing, but also committing crimes as a gangster would commit them—brazenly, fearlessly, without subterfuge. Thus, ironically, a person who tries to shoot someone surreptitiously and flee quickly may end up with more witnesses testifying against him than someone who shoots his victim in broad daylight on a busy street and strides away unperturbed. The latter *modus operandi* is a non-verbal threat to bystanders. While such behavior is often associated with urban gangs today, it has a long history in the annals of organized crime (Adler, 2006):

The behavior identified with the Black Hand, in short, obscured the motives of many Italian murderers and stymied police investigations, fueling the Black Hand mystique and encouraging killers to hide behind its shroud. This strategy succeeded. Between 1910 and 1920 Chicago policemen and prosecutors secured convictions for 21 percent of the city’s homicides. Among homicides labeled Black Hand killings, only 4 percent of killers were convicted—providing another incentive for non-gangsters to kill in public (ideally near Death Corner), use sawed-off shotguns, and whisper about extortion letters, *omerta*, and the code of personal vengeance. Similarly, police made arrests in only 21 percent of Black Hand killings, compared with 71 percent for all of the city’s homicides. During the trial in one of the few cases in which law enforcers made an arrest and secured an indictment, a stranger walked into the courtroom and waved a red handkerchief. Seeing this, the Italian immigrant on the witness stand instantly fell silent and “refused to answer any questions by the prosecutor”...

Neither law enforcers nor newspaper reporters were able to distinguish Black Hand murders from murders merely committed in the Black Hand style. Thus, the Black Hand provided a kind of default explanation for Italian homicides. Italian killers eager to disguise their motives invoked the name and imagery of the secret society of assassins, but so, too did law enforcers unable to solve Italian crimes and journalists itching to write about shotgun-packing Sicilians.

As the example of the Black Hand illustrates, witnesses can never be entirely certain that the threatening signals they observe come from authentic sources or from less dangerous imitators. A witness therefore faces a complex inference problem in deciding how seriously to take a threat. Furthermore, the manner in which witnesses respond to threats will influence the extent to which criminals find it worthwhile to make threats in the first place. Hence the extent of intimidation, the credibility of threats, and the likelihood of witness testimony are all jointly determined, and depend in complex ways on such factors as the effectiveness with which the prosecutor's office can win convictions once a witness comes forward.

This paper is an attempt to explore the manner in which rates of intimidation, testimony, and conviction respond to changes in the value of testimony, witness protection policies, and police-community relations. We develop a strategic model of criminal and witness behavior that allows for uncertainty on the part of both groups: some witnesses testify despite being threatened, and some criminals make threats that they have no intention of carrying out. We focus on two key probabilities: the likelihood of witness testimony conditional on a threat, and the likelihood that a witness will be harmed conditional on testifying. Both of these are endogenously determined in equilibrium. Under empirically plausible conditions, an increase in prosecutorial effectiveness that raises the value of testimony also raises the incentive for criminals to threaten witnesses. This makes these threats less credible. The reason is that the marginal criminal (who is indifferent between threatening the witness and not doing so) is bluffing and has no intention of carrying out the threat. A shift in the margin which raises the incidence of threats also disproportionately raises the incidence of bluffing. Since witnesses cannot distinguish between genuine and empty threats, this makes the average threat less credible and raises the likelihood of witness testimony conditional on a threat.

Despite this effect, the overall rate of witness testimony may decline, since more witnesses are threatened. Greater value of testimony through prosecutorial effectiveness reduces the unconditional rate of witness testimony when the distribution of witness willingness to testify satisfies a condition we call 'declining average density.' In certain cases, even the conviction rate may decline: a prosecutor's office that is better able to obtain a conviction once a witness has come forward may end up with fewer convictions per unit of crime. Prosecutors who are better in this sense will sometimes convict fewer criminals because of witness intimidation. Prosecutor effectiveness tends

to work in this counter-intuitive direction when the problem of witness intimidation is most severe.

Witness protection policies, which make it more costly to harm witnesses, have effects that are similar to increases in the value of testimony. Threats become less credible and witnesses are accordingly more likely to ignore them. If increases in the value of testimony raise the rate of testimony, then witness protection increases the rates of testimony and conviction. However, a narrowly tailored policy that simply makes threatening behavior more costly can have precisely the opposite effect. Fewer criminals threaten witnesses, but those that do are less likely to be bluffing. Threats therefore provide a stronger deterrent to witnesses, and the rates of testimony and conviction may therefore decline. On the other hand, certain kinds of improvements in the underlying willingness of witnesses to testify, which we call courage-increasing policies, unambiguously raise rates of testimony and convictions.

We extend this basic model to allow for the possibility that a conviction makes it less likely that a threat will be carried out, and for multiple witnesses. When there are two witnesses, criminals must decide whether to threaten one, both or none of them. Since witnesses are unaware of whether or not other witnesses have been threatened, this can give rise to tipping phenomena and multiple equilibria. Some of these equilibria involve *collective silence*: witnesses fear that others have been threatened and will not come forward, which makes conviction less likely, and hence makes it more likely that they will be harmed if they testify. Communities can be trapped in an equilibrium in which no witness testifies because each expects no one else to testify. Breaking this cycle of reticence is a critically important component of any crime reduction policy.

The analysis in this paper begins at the stage when a crime has already been committed in the presence of one or more witnesses. This restriction means we cannot deal with two important questions: what effects witness intimidation and attempts to combat it have on the volume of crime, and how they affect the probability that an innocent person will be convicted. The first restriction puts us squarely in the truth-seeking camp, as opposed to the incentive-creating camp (Sanchirico, 2004). The danger of convicting innocent people rules out one otherwise obvious solution to the problem of witness intimidation: prosecutors cannot pay witnesses who testify (at least overtly). Purchased testimony is not credible testimony. Authorities can and do pay for tips rather than testimony in the investigative phase of a case, but every successful case has an adjudicative phase in which testimony cannot be purchased.

We also rule out cases where the criminal kills the witness to keep her from testifying. Marvell and Moody (2001) show that three-strikes laws increase the number of such murders, and numerous specific murders have been attributed to this motive. (For instance, a famous 2004 quadruple homicide in which the victims' bodies were dumped next to St. Thomas Aquinas Church in Newark was alleged to have been driven by this concern.) The strategic considerations in these circumstances are very different from the ones we consider here. Accordingly, we deal only with situations in which the criminal is either unwilling or unable to kill the witness preemptively.

Our work may be viewed as a contribution to the economics of law enforcement (see Polinsky and Shavell, 2000, for an extensive survey), and more specifically to the theory of *detection avoidance* (Sanchirico, 2006). Despite its prevalence and importance, witness intimidation has received little theoretical attention in the literature on the economic analysis of law. As Sanchirico notes, our “theories of crime and regulation view evidence too much as something that investigators uncover, and not enough as something that violators cover up.” A notable exception is the work of Malik (1990), who argues that criminals make greater investments in avoiding detection when they face larger penalties, and uses this to show that maximal fines may not be socially optimal. This is related to our finding that intimidation rises with prosecutor effectiveness and rates of testimony may therefore fall. What is missing in Malik’s framework is the strategic nature of the interaction between criminal and witness, and the complexities arising from incomplete information. This is the gap we seek to fill.

Also related to our work is that of Akerlof and Yellen (1994), who model the interactions between a community, a gang, and a government, and explore the determinants of the equilibrium level of crime and community cooperation. They allow for the possibility that a fear of reprisals by the gang can affect the community’s willingness to cooperate with police. However, their model does not explain why the community believes it has to fear reprisals, or why the gang would carry them out, or what steps authorities could take to make the community less fearful. We concentrate on a relatively small part of the world that Akerlof and Yellen discuss, but derive our results from first principles, and explore in detail the effects of public policy on the extent of witness intimidation.

2 The Model

Consider an individual who has already committed a crime but has yet to be convicted. The case against him involves the testimony of a single witness (we consider the situation with multiple witnesses below). If the witness does not testify, the probability of conviction is e_0 , which we call the base conviction probability. The base conviction probability depends on both the nature of the crime and the skills of the prosecutor. A paper trail, for instance, may be key for a white collar crime, while a crime of violence may require witnesses. Prosecutors may be more or less capable of marshalling and explaining this evidence.

If the witness testifies, the probability of conviction is greater than e_0 . Let e denote the *increase* in the probability of conviction that the witness testimony causes. We call this the *value of testimony*. Like the base conviction probability, the value of testimony depends both on the nature of the crime and skill of the prosecutor. The value of testimony may be low in the white collar crime and high in the violent crime if the prosecutor is competent, but if the prosecutor cannot ask the right questions, make the right inferences, or tie the testimony to other parts of the case, the value of testimony will be low no matter what the circumstances are. We are concerned in this paper

with crimes for which the value of testimony is potentially significant.

Changes and differences in the value of testimony e will be among the key independent variables in our analysis. Such differences can arise because of the nature of different cases, the knowledge and credibility of different witnesses, or the skill and diligence of different prosecutors. The last class of differences will interest us the most, since public authorities have the greatest degree of control over it. For example, a prosecutor who spends more time preparing a cooperative witness, who asks better questions during the trial, who better anticipates questions and objections of defense counsel, and who marshals physical and verbal evidence into a more coherent and convincing summary is more likely to prevail at trial than a prosecutor who does not accomplish these tasks as well. Hence we would say that the value of testimony is greater with the more accomplished prosecutor, and that the difference in value of testimony is the difference in conviction probabilities—given the willingness of the witness to cooperate. (We will ask whether this difference in conviction probabilities between the accomplished prosecutor and the unaccomplished one remains positive *unconditionally*.)

Notice that we have assumed that testimony is a single act by the witness, and so we have not distinguished between the investigative and adjudicative phases of an investigation. Intimidation can occur at either phase. Although we will usually use language appropriate to the adjudicative phase, the impact of testimony may in many cases be greatest at the investigative stage, because authorities may not know where to look, even for physical evidence, unless someone steps forward to point them in the right direction. In such cases, we would set $e_0 = 0$. The value of testimony would be high for a prosecutor who could use these tips to find the perpetrator and make a good case; it would be low for a prosecutor who could not.

If the criminal is convicted, he suffers punishment $\rho > 0$. We normalize the criminal's payoff in the event he is not convicted to zero. Before the witness has the opportunity to testify, the criminal may send a costly signal, which we shall refer to as a *threat*. The signal itself does not harm the witness but (for reasons given below) may deter him from testifying. If a witness chooses to testify despite having received a threat, the criminal may attack the witness. We assume that no criminal wishes to attack a witness who does *not* testify, since there is no revenge motive to outweigh the costs of doing so, and we have ruled out preemptive murder. If the criminal does attack he incurs a cost while the witness suffers damages δ . While δ is exogenously given and commonly known, γ is private information, known only to the criminal, and drawn from a distribution $F(\gamma)$ with corresponding density $f(\gamma)$. We assume that $\delta > 0$ and $0 \leq \gamma \leq \gamma_{\max}$.

The cost γ of attacking includes not only the physical exertion the criminal must expend to hurt the witness and the associated cost of time, but also the expected value of any repercussions he might suffer as a result of his attack: he might feel remorse, the witness or his associates might retaliate, or he might suffer legal punishment for making the attack. Our assumption is that these costs vary among criminals, and that witnesses do not know what they are.

Making threats is costly to criminals in two respects: there is a sunk cost $\lambda \geq 0$ which is paid regardless of whether the threat is carried out, and reputational cost $\mu > 0$ that is paid if a threat is made but not carried out. Both costs are taken to be commonly known and constant across all offenders. A criminal who has made a threat will choose to carry it out if $\gamma < \mu$, but will not do so if $\gamma \geq \mu$. (We are assuming here that, when indifferent, a criminal will not carry out the threat.) We say that a criminal is *dangerous* if and only if $\gamma < \mu$, since a threat from such a criminal, if ignored, will result in harm to the witness. The proportion of criminals who are dangerous is $F(\mu)$, and we assume that $F(\mu) < 1$ (not all criminals are dangerous). We show below that (under certain restrictions on out-of-equilibrium beliefs) dangerous criminals will always make threats in equilibrium, and also that some threats will come from criminals who are not dangerous. Since γ is unobservable, witness behavior will depend on expectations of γ conditional on a threat, which in turn will depend on the equilibrium behavior of criminals of all types.

If the witness does not testify his payoff is 0 (regardless of whether or not he was threatened). If he does testify, he obtains a payoff β which is private information and drawn from a distribution $G(\beta)$ with corresponding density $g(\beta)$. We assume that $0 \leq \beta \leq \beta_{\max}$. The payoff from testifying can be interpreted as the subjective benefit of doing one's duty as a citizen net of the private costs of doing so (for instance, time lost at work or the distress of being surrounded and questioned by lawyers), or as the satisfaction of gaining revenge against the offender, or as the instrumental advantage of gaining the friendship or trust of the police or prosecutor. Different witnesses have different motives and gain different amounts from testifying. The offender does not know how much the witness he is thinking about intimidating would gain from testifying. The distribution of β may be different in different communities: for instance, in communities where the criminal justice system is not viewed as legitimate, β may be low for most people most of the time. Similarly, the distribution of β would be different for a criminal who is genuinely well-liked than for one who is despised.

We assume for the moment that the willingness and ability of a criminal to retaliate against a witness who testifies is independent of the eventual outcome of the trial; the possibility that retaliation may be contingent on conviction is explored in Section 8.

We make two key assumptions. First, the cost of making threats is not as great as the expected cost (to the criminal) of witness testimony:

$$\lambda < \epsilon\rho. \tag{1}$$

If this were not the case, then threats would never be made, even if it were believed that a threat was certain to deter a witness from testifying. In other words, we are assuming here that the crime is serious enough to carry significant penalties relative to the costs associated with making threats. For trivial crimes with minor penalties, threats would not generally be worth making. In addition,

we assume:

$$F(\mu)\delta < \beta_{\max} < \delta. \tag{2}$$

The latter inequality implies that even the most civic minded witness would not testify if he were threatened by a criminal known to be dangerous. The assumption is made to simplify notation, and all our results continue to hold without modification if the proportion of witnesses willing to testify when faced with certain harm is positive but sufficiently small. The first inequality implies that if threats are uninformative (so the witness belief about the likelihood that the criminal is dangerous is not affected by the threat), then some witnesses will testify. That is, the proportion of dangerous criminals in the population is not so large as to deter *all* witnesses from testifying when threats are uninformative.

3 Equilibrium Behavior

Since γ and β are both positive, a witness who has not been threatened will not fear harm, and will therefore testify. Furthermore, the decision to harm a witness who has testified despite having received a threat is made mechanically: the witness is harmed if and only if the criminal is dangerous. Hence the only strategic decisions involve the making of threats and the response to them. We represent this strategic interaction as a game of incomplete information with two stages: the threat stage and the testimony stage. At the first stage the criminal decides (based on his privately observed cost γ) whether or not to threaten the witness, and at the second stage the witness decides whether or not to testify. The latter decision depends both on the privately observed benefit β , and on whether or not the witness has been threatened at the first stage. Our focus is on equilibria of this game.¹

Let q denote the equilibrium likelihood that a witness will testify conditional on being threatened, and p the likelihood that a witness will be attacked conditional on testifying after being threatened.² That is, p denotes the likelihood that the threat comes from a dangerous criminal. For any given p , the set of witnesses who testify is composed of those for whom $\beta \geq p\delta$. (We are

¹Formally, we analyze a Bayesian game in which a strategy for a criminal is a function that associates with each possible γ a probability of making a threat, and a strategy for a witness is a function that associates with each possible β a probability of testifying conditional on having received a threat. A (perfect Bayesian) equilibrium of this game is a pair of strategies (one each for the criminal and witness respectively) with the following properties: (i) the belief of a witness who has been threatened regarding the likelihood that the criminal is dangerous is derived from the equilibrium strategy of criminals using Bayes' rule, (ii) the strategy of the witness is a best response to these beliefs, and (iii) the strategy of a criminal is a best response to the strategy of the witness.

²In the special case of equilibria in which no threats are made, p represents the belief held by witnesses regarding the likelihood that an (out-of-equilibrium) threat comes from a dangerous criminal, and q the likelihood that a witness holding this belief will testify. We discuss this case separately below.

assuming here that indifferent witnesses choose to testify.) Hence

$$q = 1 - G(p\delta). \quad (3)$$

If a witness is not threatened, he will not fear harm since γ is positive. All unthreatened witnesses will therefore testify, since β is positive. Hence a criminal who does not threaten will have expected payoff $-(e_0 + e)\rho$. The expected payoff to a criminal who does make a threat is

$$-e_0\rho - q(e\rho + \min\{\gamma, \mu\}) - \lambda. \quad (4)$$

A criminal will threaten in equilibrium if and only if his cost γ is such that the expected payoff of making a threat exceeds that of not doing so, given his equilibrium belief q .

We are interested in the manner in which changes in policy affect equilibrium rates of witness intimidation, testimony and conviction, defined as follows. Let r_i denote probability that the witness will receive a threat, which we refer to as the *rate of witness intimidation*. Recall that p is the likelihood that a threat comes from a dangerous criminal, $F(\mu)$ is the proportion of criminals who are dangerous, and all dangerous criminals make threats. Hence $F(\mu) = pr_i$ and so

$$r_i = \frac{F(\mu)}{p} \quad (5)$$

The rate of witness intimidation is inversely related to p : the greater the proportion of criminals who threaten witnesses, the smaller the likelihood that a threatened witness who testifies will subsequently be attacked. This is because of a pool composition effect: the increase in r_i comes about through a greater willingness to threaten on the part of those criminals with $\gamma > \mu$, who would not subsequently attack a testifying witness.³

Recall that an unthreatened witnesses testifies with certainty, while a threatened witness testifies with probability q . The unconditional probability that a witness will testify is therefore

$$r_t = 1 - r_i + qr_i. \quad (6)$$

We shall refer to r_t as the (unconditional) *rate of witness testimony*. The likelihood that the criminal will be convicted, which we call the *conviction rate* r_c , is then

$$r_c = e_0 + er_t. \quad (7)$$

Our focus here is on the manner in which equilibrium rates of intimidation, testimony, and conviction respond to changes in the value of testimony, witness protection programs, and police-community relations.

There are two classes of equilibria in the model, those in which all criminals take the same action (so their behavior is uninformative to witnesses) and those in which criminal behavior is

³This is similar to the effect identified in O’Flaherty and Sethi (2009) to explain why robberies have become less frequent but more violent.

heterogeneous (so witnesses can make inferences about a criminal based on whether or not they have been threatened). We refer to these as *informative* and *uninformative equilibria* respectively. While our focus will be on the former, we begin with the latter, which have a much simpler structure.

3.1 Uninformative Equilibria

Uninformative equilibria may be of two types: those in which criminals of all types make threats, and those in which none do. (There are no equilibria in which all criminals randomize.)

When all criminals make threats, $p = F(\mu)$ and $q = 1 - G(\delta F(\mu))$. In order for such an equilibrium to exist, it must be the case that even criminals who are *not* dangerous find it in their interest to make threats at these values of p and q . This can only be the case if $e\rho \geq q(e\rho + \mu) + \lambda$, which, given the equilibrium value of q , simplifies to yield

$$\frac{\mu + \lambda}{\mu + e\rho} \leq G(\delta F(\mu)). \quad (8)$$

Note that $G(\delta F(\mu))$ is the proportion of witnesses who are successfully deterred from testifying when they receive an uninformative threat. This proportion cannot be too low if all criminals are to threaten in equilibrium.

A very different type of uninformative equilibrium arises when no criminal threatens. In this case witness beliefs conditional on receiving a threat are unconstrained, since a threat is necessarily an out-of-equilibrium event. Such equilibria will exist only if witnesses who observe a threat are sufficiently sure that it comes from a criminal who is *not* dangerous. (If this were not the case then the witness would be deterred by such a threat, inducing some criminals to issue threats, contradicting the hypothesis that there are no threats in equilibrium.) Suppose, for instance, that witnesses who receive a threat believe that it comes from a randomly drawn criminal, and hence that the likelihood of violence conditional on testimony is simply $F(\mu)$. In this case, the likelihood of witness testimony conditional on a threat would be $q = 1 - G(\delta F(\mu))$. Given this, in order for no threats to be issued in equilibrium, even a criminal with $\gamma = 0$ must be unwilling to bear the costs of making a threat. For this to be the case, we require $e\rho \leq qe\rho + \lambda$ which, given that $q = 1 - G(\delta F(\mu))$, simplifies to

$$G(\delta F(\mu)) \leq \frac{\lambda}{e\rho}. \quad (9)$$

In this case, the proportion of witnesses who are successfully deterred from testifying when they receive an uninformative threat cannot be too *large*. If it were, then some criminals would be tempted to threaten witnesses, thus deviating from the equilibrium strategy.

Comparing (8) and (9), we see that the two types of uninformative equilibria cannot coexist. In fact, for some parameter values, neither type of uninformative equilibrium exists.

A special case of an uninformative equilibrium arises when threats are merely implicit, in which case $\lambda = 0$. Setting $\lambda = 0$ in (8), we obtain the following condition for the existence of an equilibrium

with implicit threats:

$$\frac{\mu}{\mu + e\rho} \leq G(\delta F(\mu)). \quad (10)$$

Such an equilibrium will exist if μ is sufficiently small (so there is very little reputational cost to making a threat that is not subsequently carried out). The condition (10) is more likely to be satisfied for any given μ if $e\rho$ is large (so criminals have much to fear from testimony) or $\delta F(\mu)$ is large (so dangerous criminals are plentiful and capable of inflicting considerable harm).

3.2 Informative Equilibria

Suppose that (8) fails to hold, so we have

$$\frac{\mu + \lambda}{\mu + e\rho} > G(\delta F(\mu)). \quad (11)$$

The following result establishes that this condition is necessary and sufficient for the existence of an informative equilibrium and characterizes some of the properties that any such equilibrium must satisfy.

Proposition 1. *There exists an informative equilibrium if and only if (11) holds. Equilibrium values of p and q are uniquely determined. In equilibrium, all dangerous criminals make threats, some threats come from harmless criminals, and some (but not all) threatened witnesses testify.*

Note that informative equilibria cannot coexist with uninformative equilibria in which all criminals threaten, but they can coexist with uninformative equilibria in which none threaten. Also, although p and q are uniquely determined in an informative equilibrium, there may be some indeterminacy at the level of individual behavior. Specifically, among the set of criminals who are not dangerous, some make threats while others do not. While the overall incidence of bluffing is uniquely determined in equilibrium, the precise identity of those who bluff is not.

Uninformative equilibria are generally unaffected by small changes in policy, since such changes leave p and q unaffected. That is, neither the likelihood of testimony conditional on a threat, nor the likelihood of harm conditional on testimony are affected by policy changes as long as one remains at an uninformative equilibrium. This is not the case at informative equilibria, the characteristics of which are responsive to changes in underlying parameters. In the remainder of this paper we focus exclusively on informative equilibria.

4 The Value of Testimony

The properties of the distribution $G(\beta)$ are the key to understanding some of the effects of changes in the value of testimony. Define the *average density* between 0 and x as follows:

$$a(x) = \frac{G(x)}{x}$$

Since $G(0) = 0$, this is simply the mean of the density $g(x)$ over the interval $[0, x]$. We say that the distribution G has *declining average density* at x if $a(x) > g(x)$, *rising average density* at x if $a(x) < g(x)$, and *constant average density* otherwise. Clearly, a distribution with declining (rising) density between 0 and x has declining (rising) average density, but distributions with non-monotonic densities can also have declining or rising average densities.

The following result identifies the effects of changes in the value of testimony on equilibrium behavior.

Proposition 2. *An increase in e increases the rate of witness intimidation and the likelihood of testimony conditional on a threat, and reduces the likelihood of attack conditional on testimony. It reduces (increases) the rate of witness testimony if and only if G has declining (rising) average density at $p\delta$.*

A rise in e lowers criminal payoffs regardless of the actions they choose, but raises the relative payoff of making a threat for the marginal criminal. Hence the rate of witness intimidation r_i increases. Since the marginal criminal is bluffing, the credibility of threats p falls. And since threats are less credible, q rises.

Although an increase in the value of testimony raises the likelihood that a *threatened* witness will testify, it may lead to a decline in the overall rate of witness testimony since the proportion of witnesses who are threatened also rises. Hence the effect of changes in e on r_t is theoretically indeterminate. Proposition 2 states that increases in e raise the rate of witness testimony if and only if G has rising average density at $p\delta$. The intuition is as follows. Decreasing p increases testimony by decreasing the proportion of witnesses $G(p\delta)$ who are vulnerable to threats (the witness effect), but it reduces testimony by raising the proportion of criminals r_i who threaten (the criminal effect). The witness effect operates on the margin $g(p\delta)$ while the criminal effect works on the average $a(p\delta)$. The criminal effect decreasing testimony is more powerful than the witness effect increasing testimony if and only if the marginal density is smaller than the average density, which corresponds to the case of declining average density.

Since declining average density can give rise to considerable difficulties in the effectiveness of law enforcement policies, it is useful to consider what this property means in practical terms. An example of a family of probability distributions that has declining average density at all points is the exponential, which is defined as follows:

$$G(\beta) = 1 - e^{-c\beta},$$

where c is some positive constant. The corresponding density function is $g(\beta) = ce^{-c\beta}$, which is decreasing in β ; hence the average density $a(\beta)$ is also declining for all values of β . This means that witnesses with high values of β (who are not easily deterred from coming forward to testify) are scarce relative to witnesses with low values of β (who are easily deterred). This would be a

typical situation in communities where the criminal justice system was not seen as legitimate. In such communities, the greater incidence of threatening behavior induced by a rise in the value of testimony would be highly effective in making witnesses more reticent, causing an overall decline in the rate of testimony (despite the fact that any given threat is now less likely to come from a dangerous criminal). This clearly presents a serious potential obstacle to increasing the effectiveness of law enforcement initiatives.

Declining average density does not even require a monotonically declining density function. Suppose, for instance, that the distribution is unimodal with a positive mode: witnesses who gain some positive amount m from testifying are the most common, and witnesses who gain a lot more or a lot less are rare. If $p\delta \leq m$, then the distribution has rising average density at $p\delta$, but if $p\delta$ is considerably larger than m , it has declining average density. Thus if the marginal witness is a “fringe personality”—either in the sense that the modal witness is willing to accept almost no hardship or in the sense that the threat of retaliation is considerably greater than what the modal witness would accept—declining average density applies and an increase in the value of testimony e reduces the rate of witness testimony. When the threat is large or the community’s willingness to testify is low, better prosecutors see fewer witnesses.

What can we say about rates of conviction? Clearly increases in the value of testimony raise the conviction rate if they raise the testimony rate. So declining average density at $p\delta$ is a necessary (but not sufficient) condition for the counterintuitive effect on conviction. Since $r_c = e_0 + er_t$,

$$\frac{dr_c}{de} = r_t + e \frac{dr_t}{dp} \frac{dp}{de}$$

where $dp/de < 0$, and $dr_t/dp > 0$ only in the case of declining average density. Given declining average density at $p\delta$, greater value of testimony is likely to reduce conviction rates if e is large and r_t is small. Thus, when witness intimidation is the most serious problem—few witnesses testify and testimony is valuable—improvements in prosecutorial performance that raise the value of testimony are most likely to be counterproductive. Prosecutors who are better at using witness testimony at either the investigative or adjudicative phase to obtain convictions may not in fact obtain more convictions.⁴

⁴This reasoning might help account for a recent empirical finding that laws mandating arrests in domestic violence cases increase intimate partner homicide (Iyengar, 2006). With intimate partner violence, the witness is often the victim, and the decision to call the police is similar to the decision to testify. The laws that Iyengar studies require “police to arrest a suspect without a warrant, if there is probable cause to suspect that an individual has committed some form of assault (either misdemeanor or felonious) against an intimate partner or family member” (p. 4). We can think of these laws as increasing probability that an offender will be punished, conditional on the witness testifying against him. Intimate partner threats are highly credible, and retaliation very costly to victims. Under these circumstances, an increased likelihood of conviction (mandatory arrest laws) could reduce the rate of testimony (the probability of calling the police). If the fraction of domestic violence incidents that escalate into murder is roughly constant, then the incidence of murder will rise.

Of course, increases in prosecutor effectiveness that raise base conviction probabilities—better handling of physical evidence, for instance—will always increase convictions (as long as they are not contingent on receiving tips from witnesses). Moreover, even increases in e that drive down conviction rates force criminals to make more threats, and these threats are costly to them. Better prosecutors in this sense raise the cost of crime, even if they do not obtain more convictions. This is a stark example of the phenomenon that Sanchirico (2004) discusses.

Note that better prosecutors may also raise the costs of making threats or harming witnesses, which we have assumed here to be constant. We deal with these effects separately below.

5 Witness Protection

As noted above, the value of testimony e refers to the original crime and not the likelihood of conviction on charges related to witness intimidation. However, law enforcement officials can also use resources to make intimidation more costly, by prosecuting threats to witnesses or attacks on them, or by protecting witnesses from threatening or harmful behavior.

Most discussion in the criminal justice literature about witness intimidation is about raising the cost of making or carrying out threats (see, for instance, Finn and Healey, 1996). Thus police and prosecutors strive to keep witnesses and defendants apart in precincts and court houses, hide the identity of witnesses from defendants as much as the law will permit, separate incarcerated witnesses from the general population, seek quick trials so that intimidation schemes have less opportunity to work, request high bail so that defendants have to stay in jail, relocate witnesses temporarily or permanently, and prosecute offenders under a panoply of different laws about witness tampering, accomplices after the fact, and obstruction of justice.⁵

In the context of the model, witness protection can take one of two forms: an increase in λ , which raises the costs of making threats, or a shift to the right in the distribution $F(\gamma)$, which raises the costs of carrying out the threatened retaliation. We refer to the former as a *threat reduction policy* and the latter as a *retaliation reduction policy*. Examples of threat reduction policies are penalties for making threats (whether or not they are carried out), isolation of offenders or witnesses (so communicating the threat is hard), and prohibitions on gang paraphernalia or colors. Examples of retaliation reduction policies are laws against retaliation and vigorous enforcement of those laws, and relocation of witnesses. While any given witness protection initiative is likely to affect both types of cost, the two types of policy have very different effects and it is analytically useful to consider each in isolation.

Consider first a policy which raises the costs of symbolic or literal threats, and hence affects

⁵ “Vertical prosecution” might also be viewed as a tool for reducing witness tampering. In vertical prosecution, a large proportion of a gang’s hierarchy is indicted simultaneously and held on high bail, so that few members are left on the street to intimidate witnesses.

only λ . Specifically, suppose that λ_1 and λ_2 satisfy $\lambda_1 < \lambda_2$. Then we say that the change from λ_1 to λ_2 is a threat reduction policy. Such policies are most effective precisely when increases in prosecutor resources are *ineffective* in raising the level of testimony:

Proposition 3. *A threat reduction policy lowers the rate of witness intimidation and the likelihood of testimony conditional on a threat, and increases the likelihood of attack conditional on testimony. It increases (reduces) the rates of witness testimony and conviction if and only if G has declining (rising) average density at $p\delta$.*

The logic underlying this result is as follows. When the costs of making threats rise, fewer people make threats, but relatively more of them are people who will carry out those threats. Threats become more credible. This is the opposite of what happens when the value of testimony increases. Hence p rises and q falls relative to the initial equilibrium. The effect of such a policy on the rate of testimony is therefore precisely the opposite of the effect of increases in the value of testimony.

This raises an interesting possibility: rates of testimony and conviction could be increased by making symbolic threats *more* rather than less common. This dilutes the information that they carry, and makes them less effective deterrents to witness testimony. When threats are symbols like gangster clothing, there may be gains to subsidizing such activities. Encouraging people to adopt the insignia of gangsterhood erodes the value of these symbols; p falls for any q , and in the new equilibrium, “threats” are more frequent but less credible and more often resisted. If increases in the value of testimony raise the rate of testimony, then so does paying people to acquire MS-13 tattoos.⁶

Next consider policies that shift the cost distribution $F(\gamma)$ to the right, raising the costs of retaliating against witnesses for criminals of every type. Specifically, let F_1 and F_2 be two distributions of γ , and suppose that F_2 stochastically dominates F . Then we say that the change from F_1 to F_2 , everything else being equal, is a *retaliation reduction policy*.

Proposition 4. *A retaliation reduction policy lowers the rate of witness intimidation, and increases the rates of testimony and conviction.*

To summarize, retaliation reduction policies are universally effective, threat reduction policies are effective only if increases in prosecutor resources are not. As in the case of changes in e , there are circumstances under which threat reduction policies can lower rates of testimony and conviction.

What about adopting both witness protection and threat reduction policies simultaneously? This question is important because many of the traditional tactics for reducing witness intimidation—holding defendants on high bail, for instance—do both. In this case the rate of witness intimidation

⁶MS-13, or Mara Salvatrucha, is “a fast-growing, violent, and increasingly mobile street gang” whose members “often wear clothing or sport tattoos incorporating MS-13 or the number 13.” (Federal Bureau of Investigation, 2005).

declines, threats are more credible, and threatened witnesses testify less often. Rates of testimony and conviction rise when G has declining average density at $p\delta$ but if not, then the effect on these rates is indeterminate.⁷

6 Police-Community Relations

Akerlof and Yellen (1994) have emphasized the importance of public cooperation with police. Such cooperation can be diminished if law enforcement agencies are perceived as being unjust or discriminatory, or if codes of silence become fashionable through the behavior of trend setters. The stop snitchin’ phenomenon exemplifies the spread of such norms (CBS News, 2007):

Reluctance to talk to police has always been a problem in poor, predominantly African-American communities, but cops and criminologists say in recent years something has changed: fueled by hip-hop music, promoted by major corporations, what was once a backroom code of silence among criminals, is now being marketed like never before. The message appears in hip-hop videos, on T-shirts, Web sites, album covers and street murals. Well-known rappers talk about it endlessly on DVDs. It is a simple message heard in African-American communities across the country: don’t talk to the police.

In our model, willingness to assist the police is represented by the distribution $G(\beta)$ of gains from testifying. What happens when this distribution shifts to the left—people become more reluctant to testify—either because of alienating actions by law enforcement, or because of the spread of oppositional culture like Gangsta Rap?

Among witnesses who receive a threat, those with $\beta = \delta p$ are marginal. A shift to the right in $G(\beta)$ makes these witnesses more willing to testify: we call such changes *courage increasing*.

Proposition 5. *A courage increasing policy lowers the rate of witness intimidation, increases the likelihood of harm conditional on testimony, and increases the rates of testimony and conviction.*

The criminal justice literature (Finn and Healey, 1996) outlines a number of tactics police and prosecutors might use to induce more witnesses to testify. Most of these appear to be courage-increasing, and so unambiguously raise rates of testimony and conviction (but may also increase the number of witnesses who are harmed). Simply being kind and generous to witnesses can help

⁷Note, however, that these results arise in models where criminals can only intimidate witnesses; they cannot kill them. Witness protection and threat reduction policies raise the relative incentive to kill rather than intimidate, and so a more complete model could produce more counterintuitive results.

prosecutors, since most people feel bad about disappointing those whom they like. Community policing and active engagement with community organizations can help potential witnesses understand how their testimony is protecting their neighbors, not feeding youths into the jaws of a racist system. (This is the reason why many police do not want to enforce immigration laws, since doing so is likely to make illegal aliens their enemies.) Some prosecutors have even established “neighborhood support groups” for witnesses—neighbors who sit in the courtroom when the witness testifies so she sees friendly faces as well as hostile ones, and can immediately experience the admiration and gratitude of her friends. Other authors have identified legitimacy as the key to changing the distribution $G(\beta)$ and emphasized how the actions of criminal justice authorities can raise or lower legitimacy (Tyler 1990, Tyler and Fagan 2008).

7 Multiple Witnesses

Crimes often have more than one witness, and the presence of multiple witness can significantly increase the likelihood of conviction. In such cases an offender may choose to threaten only a subset of witnesses, and witnesses must consider the possibility that others may also testify. This makes the decisions of all parties more strategically complex.

Suppose that there are precisely two witnesses to a crime and they are *ex ante* identical. In this case the offender may threaten none, one, or both of them, and each of the threatened witnesses may or may not testify. We assume that witnesses make their decisions independently, and each is unaware of whether or not the other has been threatened. As before, let p denote the likelihood that a threatened witness who testifies will be harmed. Then a witness will testify if $\beta > p\delta$, and hence the likelihood that a threatened witness will testify is $q = 1 - G(p\delta)$ as in the case of a single witness. But in identifying the likelihood of attack p conditional on a threat witnesses face a more complicated inference problem.

For notational simplicity, we assume that the base conviction probability $e_0 = 0$ (so a conviction cannot be obtained without witness testimony). Let $e_2 > e_1$ denote the likelihood of conviction if both witnesses testify, where e_1 is the likelihood of conviction if only one witness testifies. The difference reflects the extent to which a second witness raises the effectiveness with which a case can be brought. We shall focus on the case in which adding a second witness more than doubles the likelihood of obtaining a conviction:

$$e_2 > 2e_1. \tag{12}$$

This is commonly the case, and some jurisdictions will not even attempt a prosecution without a corroborating witness. Note that the importance of a second witness raises the possibility that even dangerous criminals may choose to intimidate only one witness, bringing the likelihood of conviction down to acceptable levels.⁸

⁸If $e_2 < 2e_1$, it is never optimal for a criminal to threaten just one of the witnesses. Criminals will threaten either

Let $\pi_0(\gamma)$, $\pi_1(\gamma)$, and $\pi_2(\gamma)$ denote the expected payoffs to a criminal of type γ from threatening zero, one and two witnesses respectively. These payoffs are:

$$\pi_0(\gamma) = -e_2\rho \tag{13}$$

$$\pi_1(\gamma) = -q(e_2\rho + \min\{\gamma, \mu\}) - (1-q)e_1\rho - \lambda \tag{14}$$

$$\begin{aligned} \pi_2(\gamma) &= -q^2(e_2\rho + 2\min\{\gamma, \mu\}) - 2q(1-q)(e_1\rho + \min\{\gamma, \mu\}) - 2\lambda \\ &= -q^2e_2\rho - 2q(1-q)e_1\rho - 2q\min\{\gamma, \mu\} - 2\lambda \end{aligned} \tag{15}$$

Note that we are assuming that both the sunk cost and the reputational cost are linear in the number of the witnesses threatened. The sufficient condition (11) for the existence of an informative equilibrium in the single witness case needs to be modified for the multiple witness case as follows:

$$\frac{\mu + \lambda}{\mu + \rho(e_2 - e_1)} > G\left(\frac{2\delta F(\mu)}{1 + F(\mu)}\right). \tag{16}$$

As in the case of a single witness, this requires that not too many witnesses testify when threats are uninformative.

Proposition 6. *Suppose (12) and (16) both hold. Then, if λ is sufficiently small, there exists an informative equilibrium with the following structure: some dangerous criminals threaten both witnesses while the remainder threaten one, and some harmless criminals threaten one witness while the remainder threaten none.*

As in the single witness case, the receipt of the threat signals to the witness that the criminal is more likely to be dangerous. In the multiple witness case, however, the likelihood of conviction depends on the collective behavior of witnesses: the marginal impact of one person's testimony depends on whether or not the other also testifies. If a conviction were to substantially diminish the ability of a criminal to harm a witness, then the response of a witness to a threat will depend on his expectations regarding the manner in which other witnesses respond to threats. We consider this case next.

8 Collective Silence

To this point we have assumed that the damages δ inflicted on witnesses are contingent only on whether the witness testifies, and not on the outcome of the trial. However, a criminal who is convicted and incarcerated may be unable to harm a witness even if he wishes to do so. Witnesses in the 1988 Bronx survey thought that incarceration was conducive to their safety, as the overwhelming majority felt safer with the defendant in jail. In the British surveys analyzed by Tarling et al. (2000), the original offender was the person doing the intimidating in three quarters of the cases both or none, and the resulting analysis is then similar to the single witness case.

in which intimidation occurred. This is especially true of offenders who are not affiliated with a gang or criminal organization, and who cannot therefore rely on proxies to act on their behalf, and witnesses who are not likely to find themselves in prison later.

If the conviction of an offender reduces the likelihood or extent of damage to the witness, then an increase in the value of testimony could increase the likelihood of witness testimony conditional on a threat being made. This in turn could affect the behavior of criminals and result in a change in the likelihood that a witness will be threatened. From the perspective of a witness who has been threatened, the expected loss from testifying depends on whether or not the other witness testifies. The interaction here is characterized by strategic complementarity: the greater the likelihood that the other witness testifies, the lower the expected damages, and hence the greater the incentive to testify oneself. This can give rise to multiple equilibria with varying levels of witness testimony.⁹

In order to explore the conditions under which multiple equilibria arise, we assume that

$$\lambda = e_1 = e_0 = 0.$$

That is, we assume that threats do not have an explicit cost (although empty threats do have a reputational cost), and two witnesses are required to obtain a conviction. We also assume that a criminal who is convicted does not have the capacity to harm a witness, and that

$$\mu < e_2\rho.$$

While these assumptions are not essential to the argument, they considerably simplify notation.

The costs to a criminal of harming a witness (or the reputational costs of not doing so) are paid only if the criminal is not convicted. Hence the expected payoffs from threatening none, one or both witnesses are respectively as follows:¹⁰

$$\pi_0(\gamma) = -e_2\rho \tag{17}$$

$$\pi_1(\gamma) = -q(e_2\rho + (1 - e_2)\min\{\gamma, \mu\}), \tag{18}$$

$$\begin{aligned} \pi_2(\gamma) &= -q^2(e_2\rho + 2(1 - e_2)\min\{\gamma, \mu\}) - 2q(1 - q)\min\{\gamma, \mu\} \\ &= -q^2e_2\rho - 2q(1 - qe_2)\min\{\gamma, \mu\} \end{aligned} \tag{19}$$

Now consider a witness who has received a threat, and let τ denote the (endogenously determined) probability that the *other* witness has also been threatened. As before, let p denote the likelihood that the threat comes from a dangerous criminal, and let q denote the likelihood that a witness

⁹We maintain the assumption that witness decisions are made independently, although when damages are contingent on conviction, sequential witness behavior could raise interesting new issues. For instance, the eventual outcome could depend on the sequence in which decisions to testify are made, which prosecutors could manipulate to their advantage.

¹⁰We are assuming here that the costs of threatening and harming witnesses exhibits constant returns to scale. This assumption could be relaxed to some degree without affecting the basic structure of equilibrium.

will testify conditional on being threatened. Then, if this witness testifies, the likelihood that the criminal will be convicted is $(\tau q + 1 - \tau) e_2$. The witness will testify if the expected cost of doing so based on this likelihood is no greater than β . Hence, in equilibrium, we must have

$$q = \varphi(q) \equiv 1 - G((1 - (\tau q + 1 - \tau) e_2) p \delta). \quad (20)$$

Since $\varphi(q)$ is increasing in q (for any given τ), (20) could have multiple solutions with different levels of witness testimony. However, since τ is itself endogenously determined, the conditions under which multiple equilibria arise are not immediately transparent. We demonstrate that multiplicity is indeed possible by identifying two types of equilibria: those with a high incidence of intimidation (including bluffing) and low rates of testimony, and those in which threats are quite rare and highly credible but also likely to be ignored.

Consider a candidate equilibrium in which all criminals threaten both witnesses. In order for such widespread threatening behavior to be consistent with equilibrium, it must be the case that even those criminals who are not dangerous find it in their interest to threaten both witnesses. A sufficient condition for this is $\pi_2(\mu) \geq \pi_1(\mu) \geq \pi_0(\mu)$. Using (17-19), it may be verified that this condition is equivalent to $q \leq q_l$, where

$$q_l = \min \left\{ \frac{e_2 \rho - \mu(1 + e_2)}{e_2 \rho - 2\mu e_2}, \frac{e_2 \rho}{e_2 \rho + (1 - e_2)\mu} \right\}.$$

If the likelihood q that a threatened witness testifies is small enough ($q \leq q_l$), then all criminals will threaten both witnesses. In this case $p = F(\mu)$ and $\tau = 1$. That is, the likelihood that a threat comes from a dangerous criminal is simply equal to the proportion of such types in the population, and a witness who has received a threat is certain that the other witness has also been threatened.

In order for a low testimony equilibrium of this type to exist, there must be a solution to (20) that satisfies $q \leq q_l$ when $p = F(\mu)$ and $\tau = 1$. Since $\varphi(0) \geq 0$ and φ is continuous, a sufficient condition for this is $q_l \geq \varphi(q_l)$ or

$$G((1 - q_l e_2) F(\mu) \delta) \geq 1 - q_l. \quad (21)$$

If (21) holds, then there exists an equilibrium in which threats are widespread and very effective in silencing witnesses. In such an equilibrium threats are common and lack credibility but are nevertheless highly effective in deterring testimony since each witness expect other to be threatened with certainty and to testify in the face of such threats with low probability. The expectation that others will be threatened and are likely to be silenced by such threats makes conviction unlikely, and hence makes testimony risky. This expectation is therefore self-fulfilling.

Consider next a candidate equilibrium in which threats are rare, and come only from dangerous criminals. The condition $\pi_0(\mu) \geq \pi_1(\mu) \geq \pi_2(\mu)$ is sufficient to ensure that harmless criminals choose not to make threats. Using (17-19), it may be verified that this condition is equivalent to

$q \geq q_h$, where

$$q_h = \max \left\{ \frac{e_2 \rho - \mu(1 + e_2)}{e_2 \rho - 2\mu e_2}, \frac{e_2 \rho}{e_2 \rho + (1 - e_2)\mu} \right\}.$$

If witnesses ignore threats with sufficiently great frequency ($q \geq q_h$) then only dangerous criminals will make threats. Furthermore, within the set of dangerous criminals, some will threaten both witnesses, some with threaten just one, and the remainder will threaten none. Again using (17-19) it may be verified that for any given $q \geq q_h$, there exist numbers $\gamma_l(q)$ and $\gamma_h(q)$ such that $0 \leq \gamma_l \leq \gamma_h \leq \mu$ and the following holds: criminals with $\gamma < \gamma_l$ threaten both witnesses, those with $\gamma \in (\gamma_l, \gamma_h)$ threaten one witness, and the remainder threaten none. Using Bayes' rule, a witness who receives a threat believes that the likelihood $\tau(q)$ that the other witness has also received a threat is given by

$$\tau(q) = \frac{F(\gamma_l)}{F(\gamma_l) + \frac{1}{2}(F(\gamma_h) - F(\gamma_l))} = \frac{2F(\gamma_l)}{F(\gamma_l) + F(\gamma_h)}. \quad (22)$$

Since $\gamma_h \leq \mu$, the probability that the threat comes from a dangerous criminal is $p = 1$. Let τ_h denote $\tau(q_h)$. Then, since $\varphi(1) \leq 1$ and φ is continuous, a sufficient condition for the existence of an equilibrium with high rates of testimony is $q_h \leq \varphi(q_h)$ or

$$G((1 - (1 - \tau_h(1 - q_h))e_2)\delta) \leq 1 - q_h \quad (23)$$

In this equilibrium threats are rare and come only from dangerous criminals. They are not very successful in deterring testimony because they are expected to be ignored, and because testimony results in high rates of conviction. The expectation that others are unlikely to be threatened, and are likely to ignore any threats that they do receive, is self-fulfilling, since it lowers the expected costs of testifying.

It is not clear that the conditions (21) and (23) can be satisfied simultaneously. We show that they can using the following simple numerical example, which illustrates that the two types of equilibria can coexist.

Example. Suppose that $\delta = 12$, $\rho = 3$, $\mu = 1$, $e_2 = \frac{3}{4}$, β is normally distributed with mean 4 and variance $\frac{1}{2}$, and γ is normally distributed with mean $\frac{9}{10}$ and variance $\frac{1}{30}$. Then there exist two equilibria: (i) a low testimony equilibrium with $p = 0.71$, $q = 0.00$ and $\tau = 1.00$, and (ii) a high testimony equilibrium with $p = 1.00$, $q = 0.91$ and $\tau = 0.08$.

In a high testimony equilibrium, all threats come from dangerous criminals so $p = 1$. Nevertheless, there is considerable heterogeneity among dangerous criminals: some threaten both witnesses, some threaten just one, and the remainder threaten none. Because of this, the value of τ is quite low. Despite the fact that a threatened witness is certain that the criminal making the threat is dangerous, the vast majority of threatened witnesses are willing to testify because they are confident that others will also testify (either because the others have not been threatened, or because

they are expected to ignore the threat). Hence conviction is likely. This dramatically reduces the danger entailed in testifying, and justifies the decision to do so.

In a low testimony equilibrium, all criminals (including those who are not dangerous) threaten both witnesses. Hence witnesses are certain that all other witnesses have been threatened and $\tau = 1$. Almost all witnesses refuse to testify, expecting (correctly) that almost all *other* witnesses will do the same. Since witness testimony is unlikely to be corroborated, the rate of conviction is low and witnesses face considerable danger. As a result, bluffs are rarely called, and very little retaliation takes place. The expectation that few witnesses will testify becomes self-fulfilling, resulting in a state of collective silence. Under these circumstances even criminals who are not dangerous are therefore able to preserve their reputations. This equilibrium can be interpreted as one with purely implicit threats that are taken very seriously by witnesses.

9 Conclusions

Empirical data on witness intimidation are scarce. It is easy to understand why: witnesses who never come forward are very hard for outside observers to detect. Given the paucity of empirical information, we have relied largely on theoretical arguments in our attempt to understand witness intimidation. We have confirmed the result of Akerlof and Yellen (1994) that community respect for law enforcement and willingness to cooperate are indispensable. For multiple witness cases, moreover, multiple equilibria are possible, and so small changes in community relations or perceptions might trigger major improvements in outcomes.

For many other policies, we identify considerable ambiguity in the direction of their effects. This ambiguity indicates that empirical work is likely to be valuable. Data may be available on potential witnesses who are killed and on witnesses who recant, but these phenomena are not at all direct measures of witness intimidation (the first is a measure of witnesses who were incapacitated, not intimidated, and the latter is a measure of witnesses who came forward enough at some point that they could go back). Perhaps the best way to measure witness intimidation is the conviction rate on murders, with appropriate lags to reflect the length of time it takes to bring a murder case to trial (since the payoff from witness intimidation is usually greatest in homicide cases, and these cases don't rely on victims to file a complaint). State laws and policies on witness tampering, and average time to trial could be used to proxy for differences in witness protection and threat reduction strategies. There appears to be considerable variation in these laws and policies. Empirical work in this vein could probably tell us something about witness reaction curves.

This paper, as an exercise in theory, has shown that witness intimidation is indeed a challenging problem, with few obvious solutions and considerable danger of perverse outcomes. As Cook (2007, p.30) notes: "In essence the citizens who become involved in crime are invited to make a charitable contribution of their time and possibly their safety, in exchange for knowing they have done a good

deed for their community.” The rule of law and well-functioning society rely much more on civic altruism than we often care to acknowledge.

Appendix

Proof of Proposition 1. First we show that $q > 0$ at any informative equilibrium. Suppose, by way of contradiction, that $q = 0$. Then, from (4), the payoff to each criminal who makes a threat is $-\lambda$. From (1), $\lambda < e\rho$. Hence if $q = 0$, all criminals will make threats contradicting the hypothesis that the equilibrium is informative. Hence $q > 0$.

Note that (2) implies $G(\delta) = 1$, so from (3) and $q > 0$ we have $p < 1$. Since $q > 0$, (4) is strictly decreasing in γ for all $\gamma \leq \mu$ and is independent of γ thereafter. Hence if *any* criminal with $\gamma > \mu$ threatens then *all* criminals with $\gamma \leq \mu$ will also threaten. Since $p < 1$, some threats must come from harmless criminals, and so all dangerous criminals must threaten. Since the equilibrium is informative, not all harmless criminals threaten, and $p > F(\mu)$. Since all harmless criminals have the same expected payoff, they must all be indifferent between making and not making threats. This requires: $(e_0 + e)\rho = e_0\rho + q(e\rho + \mu) + \lambda$, or

$$q = \frac{e\rho - \lambda}{e\rho + \mu} = 1 - \frac{\mu + \lambda}{e\rho + \mu}. \quad (24)$$

In order for such an equilibrium to exist, we must have $p \in (F(\mu), 1)$, where

$$\frac{\mu + \lambda}{e\rho + \mu} = G(p\delta).$$

So we must have

$$\frac{\mu + \lambda}{e\rho + \mu} < G(\delta)$$

and

$$\frac{\mu + \lambda}{e\rho + \mu} > G(\delta F(\mu)).$$

The first of these inequalities is trivially satisfied since $G(\delta) = 1$ from (2). The second inequality is equivalent to (11); which is therefore both necessary and sufficient for an informative equilibrium to exist. The remaining claims follow from the fact that $q \in (0, 1)$ and uniquely determined from (24) and $p \in (F(\mu), 1)$ and uniquely determined from (3).

Proof of Proposition 2. From (6),

$$\frac{dr_t}{de} = -\frac{dr_i}{de}(1 - q) + r_i \frac{dq}{de}$$

From (24),

$$q = 1 - \frac{\mu + \lambda}{e\rho + \mu}$$

so

$$\frac{dq}{de} = \frac{\rho(\mu + \lambda)}{(e\rho + \mu)^2} = \frac{\rho(1 - q)}{e\rho + \mu}$$

Using (3),

$$\frac{dq}{de} = -\delta G'(p\delta) \frac{dp}{de}$$

and using (5),

$$\frac{dr_i}{de} = -\frac{F(\mu)}{p^2} \frac{dp}{de}$$

Hence

$$\begin{aligned} \frac{dr_t}{de} &= -\frac{dr_i}{de} (1 - q) + r_i \frac{dq}{de} \\ &= F(\mu) (1 - q) p^{-2} \frac{dp}{de} - F(\mu) p^{-1} \delta G'(p\delta) \frac{dp}{de} \\ &= -F(\mu) \delta p^{-1} \frac{dp}{de} \left(G'(p\delta) - \frac{1 - q}{p\delta} \right) \end{aligned}$$

Using (3) and the definitions of marginal and average density,

$$\frac{dr_t}{de} = -F(\mu) \delta p^{-1} \frac{dp}{de} (g(p\delta) - a(p\delta))$$

Since p is decreasing in e , the sign of dr_t/de is equal to the sign of $g(p\delta) - a(p\delta)$.

Proof of Proposition 3. From (24),

$$1 - q = \frac{\lambda + \mu}{e\rho + \mu}$$

so a rise in λ lowers q . From (3), this means that p increases. From (5), r_i declines. From (6), $r_t = 1 - r_i + qr_i$.

$$\begin{aligned} \frac{dr_t}{d\lambda} &= r_i \frac{dq}{d\lambda} - \frac{dr_i}{d\lambda} (1 - q) \\ &= \frac{F(\mu)}{p} \frac{dq}{d\lambda} - \frac{dr_i}{d\lambda} (1 - q) \\ &= F(\mu) p^{-1} \frac{dq}{d\lambda} + F(\mu) p^{-2} (1 - q) \frac{dp}{d\lambda} \\ &= -F(\mu) p^{-1} \delta G'(p\delta) \frac{dp}{d\lambda} + F(\mu) p^{-2} (1 - q) \frac{dp}{d\lambda} \\ &= \frac{dp}{d\lambda} F(\mu) \delta p^{-1} (a(p\delta) - g(p\delta)) \end{aligned}$$

Since p is increasing in λ , the sign of $dr_t/d\lambda$ is equal to the sign of $a(p\delta) - g(p\delta)$. Furthermore, since e is held constant, the rates of testimony and conviction change in the same direction.

Proof of Proposition 4. A shift in $F(\gamma)$ has no effect on q from (24), and hence no effect on p from (3). Since $F(\mu)$ declines, the rate of witness intimidation falls from (5). Since $r_t = 1 - r_i(1 - q)$ and r_i falls while q is unchanged, r_t must increase. Since e is unchanged, r_c must also increase.

Proof of Proposition 5. A shift in $G(\beta)$ has no effect on q from (24). Since $1 - q = G(p\delta)$ from (3), a decline in G must be offset by an increase in p . The rate of witness intimidation falls from (5). Since $r_t = 1 - r_i(1 - q)$ and r_i falls while q is unchanged, r_t must increase. Since e is unchanged, r_c must also increase.

Proof of Proposition 6. Consider a candidate equilibrium in which dangerous types threaten at least one witness, with the lowest γ types threatening both, and some harmless criminals threatening one while others threaten none. Harmless types must therefore be indifferent between making threats and not doing so, which implies $\pi_0(\mu) = \pi_1(\mu)$. Using (13-14), this simplifies to yield

$$q = \frac{\rho(e_2 - e_1) - \lambda}{\rho(e_2 - e_1) + \mu}. \quad (25)$$

For this to be consistent with equilibrium, we need to verify that $\pi_2(\mu) < \pi_0(\mu) = \pi_1(\mu)$. Using (13) and (15), this implies

$$\begin{aligned} q^2(e_2\rho + 2\mu) + 2q(1-q)(e_1\rho + \mu) + 2\lambda &> e_2\rho \\ -e_2\rho + q^2e_2\rho + 2q^2\mu + 2q(1-q)e_1\rho + 2q(1-q)\mu + 2\lambda &> 0 \\ -e_2\rho(1-q^2) + 2q(1-q)e_1\rho + 2q\mu + 2\lambda &> 0 \\ (1-q)\rho(q(2e_1 - e_2) - e_2) + 2q\mu + 2\lambda &> 0 \\ (1-q)\rho(q(e_2 - 2e_1) + e_2) &< 2(q\mu + \lambda) \\ \frac{\lambda + \mu}{\rho(e_2 - e_1) + \mu}\rho\left(\frac{\rho(e_2 - e_1) - \lambda}{\rho(e_2 - e_1) + \mu}(e_2 - 2e_1) + e_2\right) &< 2\left(\frac{\rho(e_2 - e_1) - \lambda}{\rho(e_2 - e_1) + \mu}\mu + \lambda\right) \\ (\rho(e_2 - e_1) - \lambda)(e_2 - 2e_1) + \rho e_2(e_2 - e_1) + \mu e_2 &< 2(e_2 - e_1)(\rho(e_2 - e_1) + \mu) \\ (\lambda + \mu)(e_2 - 2e_1) &> 0 \end{aligned}$$

So condition (12) is sufficient for this. Hence when q is given by (25) and $e_2 > 2e_1$, we have

$$\pi_2(\mu) < \pi_0(\mu) = \pi_1(\mu).$$

For $\varepsilon > 0$ sufficiently small, since π_1 is strictly decreasing in γ for $\gamma < \mu$ when $q > 0$ and π_0 does not depend on γ , we therefore have

$$\pi_2(\mu - \varepsilon) < \pi_0(\mu - \varepsilon) < \pi_1(\mu - \varepsilon).$$

So dangerous criminals of type γ sufficiently close to μ will threaten one witness. Note from (14-15) that

$$\frac{d(\pi_2 - \pi_1)}{d\gamma} = -q < 0, \quad (26)$$

so as γ declines, $\pi_2 - \pi_1$ increases, while π_0 (which is independent of γ) remains unchanged. Note also that

$$\begin{aligned} \pi_2(0) - \pi_1(0) &= qe_2\rho - q^2e_2\rho + (1-q)e_1\rho - 2q(1-q)e_1\rho - \lambda, \\ &= (1-q)(q\rho(e_2 - e_1) + e_1\rho(1-q)) - \lambda. \end{aligned} \quad (27)$$

If $\lambda = 0$ then $\pi_2(0) - \pi_1(0) > 0$, and the same is true if λ is sufficiently small. Hence for λ sufficiently small, there exists $\tilde{\gamma} \in (0, \mu)$ such that criminals with $\gamma < \tilde{\gamma}$ threaten both witnesses and criminals with $\gamma \in (\tilde{\gamma}, \mu)$ threaten one.

To complete the proof we need to verify that if q is given by (25) and p by (3), then p is consistent with the beliefs of witnesses regarding the likelihood of harm. It is clear from (3) that $p < 1$ is satisfied for any q . The other bound for p is obtained by considering witness beliefs when all harmless criminals threaten one witness. In this case, using Bayes' rule, the likelihood of harm conditional on a threat is

$$p = \frac{F(\tilde{\gamma}) + \frac{1}{2}(F(\mu) - F(\tilde{\gamma}))}{F(\tilde{\gamma}) + \frac{1}{2}(1 - F(\tilde{\gamma}))} = \frac{F(\tilde{\gamma}) + F(\mu)}{F(\tilde{\gamma}) + 1}.$$

Hence for feasibility of p we require

$$p > \frac{F(\tilde{\gamma}) + F(\mu)}{F(\tilde{\gamma}) + 1}.$$

Since $\tilde{\gamma} < \mu$, a sufficient condition for this is

$$p > \frac{2F(\mu)}{F(\mu) + 1}. \tag{28}$$

If (16) is satisfied, then when q is given by (25) and p by (3)

$$G(p\delta) > G\left(\frac{2\delta F(\mu)}{1 + F(\mu)}\right),$$

so (28) is necessarily satisfied. Hence there exists an equilibrium in which criminals with $\gamma < \tilde{\gamma}$ threaten both witnesses, those with $\gamma \in (\tilde{\gamma}, \mu)$ threaten one, some of those with $\gamma \geq \mu$ threaten one, and the remainder threaten none.

References

- [1] Adler, Jeffrey S., 2006, *First in Violence, Deepest in Dirt: Homicide in Chicago, 1875-1920*, Cambridge: Harvard University Press.
- [2] Akerlof, George A. and Janet L. Yellen, 1994, "Gang behavior, law enforcement, and community values," pp.173-209 in Henry Aaron, Thomas Mann, and Timothy Taylor, eds., *Values and Public Policy*. Washington, D.C.: Brookings.
- [3] CBS News, 2007, "Stop Snitchin'," 60 Minutes, August 12, 2007, accessed at <http://www.cbsnews.com/stories/2007/04/19/60minutes/main2704565.shtml>
- [4] Philip J. Cook, 2007, "Crime in the City." Paper for conference on Unraveling the Urban Enigma; Wharton School, May 4.
- [5] Federal Bureau of Investigation, 2005, "How we're ganging up on MS-13 and what you can do to help," *Headline Archives*, July 13, 2005, accessed at <http://www.fbi.gov/page2/july05/ms071305.htm>
- [6] Finn, Peter, and Kerry Murphy Healey, 1996, *Preventing Gang- and Drug-Related Witness Intimidation*. Washington, DC: National Institute of Justice.
- [7] Iyengar, Radha, 2006, "Does the certainty of arrest reduce domestic violence? Evidence from mandatory and recommended arrest laws," *National Bureau of Economic Research Working Paper* 13186.
- [8] Kennedy, David, 2006, "The Neighborhood War Zone", *Washington Post*, August 13, 2006.
- [9] Kocieniewski, David, 2007, "A Little Girl Shot, and a Crowd That Didn't See," *New York Times*, July 9, 2007.
- [10] Malik, Arun S., 1990, "Avoidance, Screening and Optimum Enforcement," *RAND Journal of Economics* 21: 341-353.
- [11] Marvell, Thomas B. and Carlisle E. Moody, 2001, "The lethal effects of three-strikes laws," *Journal of Legal Studies* 30:89-106.
- [12] Maynard, Warwick, 1994, *Witness Intimidation: Strategies for Prevention*. London: Home Office Police Department, Crime Detection and Prevention Series Number 55. Accessed on October 3, 2007 at <http://www.homeoffice.gov.uk/rds/prgpdfs/fcdps55.pdf>.
- [13] O'Flaherty, Brendan and Rajiv Sethi, 2009, "Why Have Robberies become Less Frequent but More Violent?" *Journal of Law, Economics and Organization*, in press.

- [14] Polinsky, A Mitchell and Steven Shavell, 2000, “The Economic Theory of Public Enforcement of Law.” *Journal of Economic Literature* 38: 45–76.
- [15] Sanchirico, Chris William, 2004, Evidence tampering, *Duke Law Journal* 53 (4): 1215-1336.
- [16] Sanchirico, Chris William, 2006, “Detection Avoidance.” *New York University Law Review*, 81: 1331-1399.
- [17] Tarling, Roger, Lizanne Dowds, and Tracey Budd, 2000, Victim and Witness Intimidation: Findings from the British Crime Survey. London: Home Office. Accessed on October 7, 2007 at <http://www.homeoffice.gov.uk/rds/pdfs/occ-victandwit.pdf>.
- [18] Tyler, Tom R., 1990, *Why People Obey the Law*. New Haven: Yale University Press.
- [19] Tyler, Tom R. and Jeffrey Fagan, 2008, Legitimacy and cooperation: Why do people help the police fight crime in their communities? *Ohio State Journal of Criminal Law* 6 (1): 231-75.