

PAYMENT, PROTECTION AND PUNISHMENT

THE ROLE OF INFORMATION AND REPUTATION IN THE MAFIA

Alastair Smith and Federico Varese

ABSTRACT

A game theoretic model is used to examine the dynamics governing repeated interaction between Mafiosi running extortion rackets and entrepreneurs operating fixed establishments. We characterize the conditions under which violence occurs. Entrepreneurs pay protection money to the Mafia because they fear the Mafia's ability to punish. However, the entrepreneurs' willingness to pay encourages opportunistic criminals (fakers) to use the Mafia's reputation and also demand money. We show that two phenomena drive the repeated interaction between criminals and entrepreneurs: reputation-building and readiness to use violence on the part of the Mafiosi, and attempts to filter out fakers on the part of entrepreneurs. These two phenomena lead to turbulence: as entrepreneurs filter out fakers by not paying some of the times, some real Mafiosi are not paid and punish non-payment to establish their reputation. As Mafia reputation is re-established, fakers have again an incentive to emerge, setting in motion a spiral of never-ending filtering and violence. We also show how external shocks to this relationship, such as changes in policing practices, succession disputes within the Mafia or inflation, often lead to violence until beliefs are re-established. We conclude that a world where mafias operate is inherently turbulent. This conclusion goes against the widespread perception that racketeers are able to perfectly enforce territorial monopolies.

KEY WORDS • extortion • fakers • mafia • reputation • violence

Introduction

From afar, the world of the Mafia seems a simple one. A group controls a territory, everybody knows who the members of the group are, and shopkeepers pay regularly to the gang.¹ In this simplified

vision of a Mafia world, paying protection money amounts to paying a tax to a well-defined authority. Matters are not always so simple. Often, there is a great deal of confusion over who is a real Mafioso and who is not. Ladyzhenskii, a correspondent for the Russian monthly *Kriminal'naya Khronika*, gave the following account of the situation which occurred just at the time of the transition to the free market in Russia:

Perestroika was marked by the appearance of an embryonic form of free enterprise and ruthless and unregulated criminal racketeering. Everyone was involved in this racketeering: low-level gangs, students, sportsmen, former as well as current militiamen. As a result, it was necessary to have protection against such 'arbitrariness'. The state could not provide it, moreover it did not want to. . . . Only serious criminal structures could supply real help to businessmen. To pay them was expensive, but businessmen saw it as the lesser of two evils. (Ladyzhenskii 1994: 4)

How does one distinguish 'ruthless and unregulated criminal racketeering' from 'serious criminal structures'? A Russian businessman who testified to the US Senate on organized crime in Russia describes his uncertainty over whom to pay: 'As I said, things got out of control. I did not know which "roofs" to pay and which was safe to ignore. Men were showing up regularly asking for protection money' (Committee on Governmental Affairs 1996: 50).

In a world where there is an expectation of Mafia presence, impostors have an incentive to pass as real Mafiosi, rip-off the benefits of the Mafia reputation for violence, take the money and run. Joseph Serio, a security consultant who works in Moscow, has drawn attention to 'teenage wannabes' (*khuligany*): they 'are 17–20-year-olds who pass themselves as young toughs trying to take advantage of foreigners' well-known fear of "Mafia"'. Serio offers an example: 'An American firm was approached by three wannabes in search of easy money. They presented themselves as members of the Chechen criminal community (*obshchina*), knowing that the Chechens have a reputation for being particularly fierce' (Serio 1997: 97). The Moscow representative of the firm did not pay and the would-be Mafiosi failed to show up again (another instance of fly-by-night protectors is narrated in Serio 1997: 97).

This phenomenon occurs not just when Mafia groups are emerging for the first time.² Gambetta, a student of the Sicilian Mafia, was the first author to draw attention to the phenomenon of fakers: 'A local [Palermo] entrepreneur told me the (to him hilarious) story of a northern firm doing business in Sicily on a large contract. The firm was approached by a man making the vague sorts of

threats for which Mafiosi are renowned. So sure had they been that someone would at some point demand protection money in precisely this way that they took it for granted *this* was the person. They paid up for about two years before realizing they had been conned' (Gambetta 1993: 34). Another instance is recalled by Antonio Calderone, a prominent Catania Mafioso turned state witness. An employee of a northern firm operating in Catania started to make extortion calls to the manager, posing as a real Mafioso. It turned out that the same firm employed Calderone himself, precisely in order to be protected. The employee was duly killed (Testimony of A. Calderone, 1987–8, quoted in Gambetta 1993: 34. Another instance is narrated by Calderone in his memories. See Arlacchi 1992: 185). Most recently, Armando Giuliano has been arrested in Palermo. For several months he passed off as a member of the Borgo Vecchio Mafia family and extorted money from some antique dealers. Eventually, the victims themselves reported him to the police (*Giornale di Sicilia* 07/II/2001).

A name can be sufficient to induce payment: a group of Sicilians who had migrated to northern Italy started to make extortionary demands on entrepreneurs in a Piedmont village. According to a report by the Italian police, the man they used to collect money was Salvatore Badalamenti, whose last name happened to be the same as that of a well-known Sicilian Mafioso, Tano Badalamenti. Salvatore was neither a relative nor an accomplice of the true Mafioso (Tribunale di Torino 1995: 99, quoted in Sciarrone 1998: 252).

Mafia impostors have been reported also in Hong Kong. A senior officer of the Triad Society Bureau recalls:

Toward the end of 1973, a total of 57 people in good occupations were telephoned by a man who demanded money from them in the name of a triad society . . . all paid except one and as a result of this report to the police, an arrest was made and police learned about the other offences. The person concerned had made a very large sum of money and was never a member of a triad society at all. (*South China Morning Post*, 30/06/1975, quoted in Chu 1996: 98 and Chu 2000: 39)

Another attempt by bogus triad members to extract money from the manager of a construction site is reported in Chu 2000: 49).

These instances highlight a number of key questions that relate to the Mafia and to countries where Mafia groups flourish. First of all, reputation is a crucial asset for a Mafia group. Chechens have a reputation in Russia for being very cruel, and so do Triads in Hong Kong and the Sicilian Mafia in Palermo. A reputation for being tough enables them to save directly on the use of violent resources

when they interact with their victims (Gambetta 1994: 44; see also Reuter 1983). Few people would dare to challenge somebody who claims to be a Mafioso, enabling the latter to save time and effort in convincing their victims to pay.

As pointed out by Gambetta, the costs of probing whether the man asking for protection money is a real Mafioso are significant, hence signals that 'honestly' reveal one's type are of value. If one could be sure that dark skin in Russia or dark glasses in Palermo are signalling Mafia membership, an entrepreneur could make a reliable inference, and know whether to pay or not. However, too many people have dark skin in Russia and anybody can wear dark glasses in Palermo. Gambetta (1994) presents a simple game of incomplete information which tries to capture this situation. In this game 'if the victim does not know to which type the Mafioso belongs, only one equilibrium is possible, namely always complying' (1994: 356). Gambetta then suggests that real Mafiosi have an incentive to protect the signals associated with Mafia membership, and he shows how imaginative mafias can be. The limit of this argument is that the Mafia must, at the same time, preserve the secretness of the organization. There is a limit to the efforts of the Mafia to devise credible signals, such as a reliable Mafia ID. If they issued Mafia IDs, the police could easily identify them.³ A fuller picture of this dynamics also needs to incorporate violence. As pointed out by Calderone in his memoirs, 'Every Mafioso knows perfectly well, when all is said and done, where his power comes from. People are scared of being physically hurt, and, more than that, no one wants to risk being killed' (Arlacchi 1992: 191). We take the ability to use violence effectively against both non-payment of protection money and fakers as a credible signal of being a Mafioso. Punishment, however, must be constrained by other variables, such as level of policing in a given area. This is a variable absent from the original model by Gambetta.

Goals of the Paper

In this paper we look at how relevant actors—entrepreneurs who have to pay protection money, impostors, and real Mafiosi—solve the dilemmas outlined above. When is it more likely that both impostors and Mafiosi will request money? And what would the entrepreneur do when faced with a such a request? Entrepreneurs are likely to be punished for non-payments to the real Mafia, while

they would be better off by not paying impostors. Impostors are either unable to inflict punishment, for the lack of violent resources, or, if they do punish non-payers, they reveal their presence to the Mafia. Real Mafiosi are then likely to punish impostors, as in the instance narrated by Calderone. The model we present allows us to characterize the conditions under which violence occurs. A further element that we take into consideration is the expected level of policing, which affects the decisions of Mafiosi, impostors, and Mafia victims. Shocks to the Mafia/entrepreneur relationship, such as changes in policing as well as succession disputes or inflation, often lead to violence until beliefs are re-established.

The Game

We analyse the interaction between thugs and entrepreneurs. We consider the following scenario: a thug enters a business and demands protection money from its owner. The entrepreneur must decide whether or not to pay protection money. If he does, then the thug happily leaves. However, if the entrepreneur refuses to pay, the thug may opt to damage the business or harm the entrepreneur.

We examine the following questions: When do thugs demand money? Under what conditions do entrepreneurs pay protection money, and what influences whether the thug punishes non-payers?

In the first stage of the model, the thug decides whether to demand money from the entrepreneur. Empirically, focal points appear to govern the size of payments demanded.⁴ Indeed, it would be prohibitively costly for Mafiosi to carry out extensive inquiries into every business's accounts. Therefore, we assume that the size of the demand is fixed at size m .

For the purposes of this model it does not matter what the nature of the business is, beyond it being a fixed-location legal enterprise. The question under consideration is, when a thug demands money, does the entrepreneur pay? We assume that businesspeople want to avoid punishment. However, they also wish to avoid paying protection money. Avoiding punishment provides a powerful incentive for businesspeople to pay. If punishment is severe and certain, then the entrepreneur is certain to pay. Yet, this creates an incentive for other opportunistic thugs (non-Mafiosi). If entrepreneurs always pay, then anyone demanding money gets paid. Thus, non-Mafiosi join the protection racket, demanding money and pretending to be Mafiosi.

The entrepreneur is unsure whether or not the thug is Mafioso.

The thug could be a legitimate member of the local Mafia group. Alternatively, the thug could have no associations with the local Mafia group and instead be either a member of a rival group trying to poach business from another's turf or a fly-by-night opportunist. We describe these two types as *Mafioso* and *faker*, and where appropriate we represent them by the letters M and F, respectively.

When the thug first enters the entrepreneur's business, the entrepreneur does not know whether the thug is a *Mafioso* or a *faker*. Both *Mafioso* and *faker* look identical—like thugs. (As we shall see, as more thugs attempt to imitate *Mafiosi*, entrepreneurs become more reluctant to pay protection. This makes it harder for the Mafia to maintain its income.) Gambetta stresses that the Mafia has an incentive to develop reliable signals of Mafia membership. For instance, Yakuza members 'are identifiable by all-over tattoos and severed fingers (lopped off to punish themselves for their professional mistakes)' (Kaplan and Dubro 1986: 26 and 146, quoted in Gambetta 1994: 363). However, there is a limit to the manipulation of the signals by *Mafiosi*. A 'faker' could in principle be a member of a competing Mafia group, trying to poach into different territory. In this case, he is faking not Mafia membership in general, but just membership of a group. He would already have had his body tattooed and fingers chopped off.⁵

Given the above evidence, it should appear that it is not an easy matter to distinguish *fakers* from *Mafiosi* at first sight. Therefore, we assume that an entrepreneur cannot distinguish between *Mafioso* and *fakers* with certainty, but he will have certain beliefs based on his 'reading' of the available signals. We let θ represent the entrepreneur's beliefs about the type of thug. Thus, θ , a number between 0 and 1, is the probability that the thug is in the Mafia. If $\theta = 1$, then the entrepreneur is certain that the thug is a member of the Mafia. If the entrepreneur was certain that the thug had no Mafia connections, then his beliefs would be represented by $\theta = 0$.

As we shall show, the entrepreneur's beliefs are critical in his decision about whether to pay. Why does the entrepreneur want to pay *Mafiosi* but not *fakers*? The reason entrepreneurs pay up is not the identity of the thug, but rather the propensity of the different types to punish. *Mafiosi* are more likely to use violence against non-payers than *fakers*. This is so because *Mafiosi* and *fakers* face different incentives. *Mafiosi* and *fakers* fear official punishment, by which we mean the police and the judiciary. The Hong Kong impostor was indeed arrested by the police, after having extorted money from at

least 57 people. However, fakers face an additional risk, that of being discovered by the Mafia, as suggested by the instance narrated by Calderone. We start by considering the risk that Mafiosi face when using violence.

While one might think of the Mafia as above the law, in reality Mafiosi risk imprisonment if they are caught. Although their Mafia status might provide some insulation against the police, Mafiosi are not immune from prosecution. We assume that in certain situations the police are closely watching Mafia activities. On such occasions, punishing non-payers is risky, since the risk of arrest, prosecution, and conviction is high. On other occasions Mafiosi benefit from punishing non-payers. For example, when senior members of the Mafia watch the activities of a junior Mafioso, it is important for him to appear strong. Many other unpredictable factors influence the desirability of violence. For example, simple health considerations cannot be ignored. A wounded thug is less predisposed to violence than his healthy counterpart. So while in reality many factors influence the relative costs and benefits of using violence at any particular time, we model these through the simple device of police presence. We assume that the level of police presence varies across time. If the police presence is high, then Mafiosi face serious risks from damaging businesses and hence punishment is costly. If, on the other hand, the police presence is low, then there is little risk associated with violence and the Mafiosi benefit within the organization from using violence.

In addition to the risk faced by real Mafiosi, fakers face additional risks, since the Mafia wants to minimize competition. Therefore, if the Mafia discovers a faker's activities, then the faker's life is in jeopardy. This risk of discovery depends upon the faker's actions. If the faker does nothing, then he has little to fear from the Mafia. However, his risk of discovery goes up if he demands money. It is even more dangerous for fakers to use violence since the destruction of property is readily observable.

Having outlined the motivation for the model, we describe its components in detail.

The Model

Figure 1 is a pictorial representation of the game. At the first node, nature decides the type of thug. With probability θ the thug is a member of the Mafia, and with probability $(1 - \theta)$ the thug is a

faker. At node 2, the thug, whether Mafioso or faker, decides whether or not to demand money from the entrepreneur. If the faker demands money, then he risks discovery by the Mafia, which has expected cost r . If the thug makes no demands, then the game ends and both players, the thug and the entrepreneur, receive a payoff of zero.

If the thug makes a demand, then the entrepreneur decides whether or not to pay. When making this decision, the entrepreneur is not certain if he is dealing with a Mafioso or a faker. In Figure 1 this is represented by a dashed line. Although uncertain, the entrepreneur has beliefs about the type he is dealing with. These beliefs are determined by his prior beliefs, θ , and the choices made by the different types. If the entrepreneur pays, then the game ends with a transfer of m from the entrepreneur to the thug. Thus, the thug's payoff is m if he is a Mafioso, and $m - r$ if he is a faker. The entrepreneur's payoff is $-m$.⁶

If the entrepreneur refuses to pay, then the thug can use violence to punish him. As discussed above, the costs associated with violence differ between Mafiosi and fakers. If a faker uses violence,

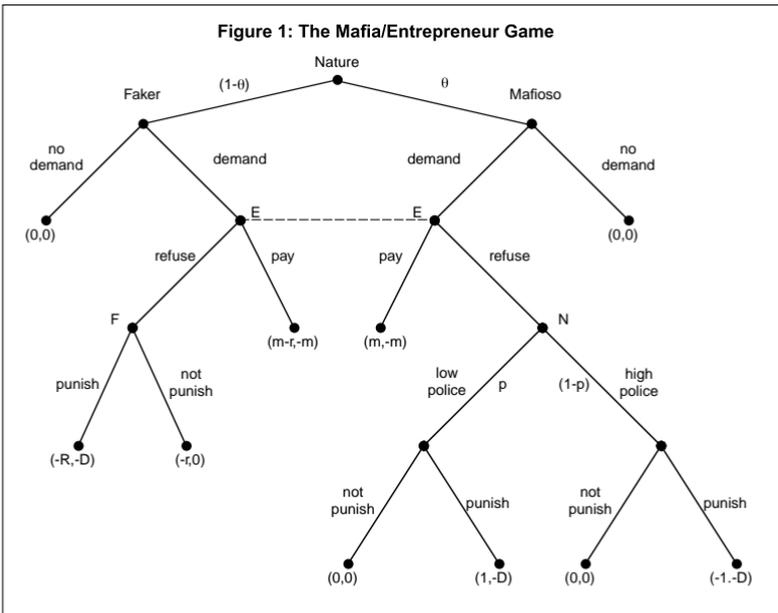


Figure 1. The Mafia/Entrepreneur Game

Table 1. Summary of variables

<i>Variable</i>	<i>Interpretation</i>
θ	Probability that the thug is Mafioso
m	Protection money
D	Damage inflicted if the the thug uses violence
r	Faker's risk for demanding money
R	Faker's risk for using violence
p	Probability of low police presence

then he faces an additional risk of discovery. Thus, if a faker uses violence, then his final payoff is $-R$. However, if a faker does not punish, he can avoid this additional risk. In this case his final payoff is $-r$ (where $R \gg r$). The Mafioso's costs depend upon many factors, but we shall model it as a variable, the level of police presence. If the level of policing is high, then using violence is costly. Yet, using violence benefits the Mafioso when the level of policing is low. We model this as follows: Immediately before the Mafioso contemplates violence he observes the level of policing, which is low with probability p , and high with probability $(1 - p)$. Variable p reflects the availability of police manpower, their level of training and the extent to which they can be trusted to uphold the law. Having observed the level of police presence, the Mafioso decides whether to use force. If he uses force and the level of policing is low, then his payoff is $+1$. Alternatively, if the level of policing is high, then his payoff for violence is -1 .⁷ The payoff for non-violence is 0.

Should the thug use violence to punish the entrepreneur, then the cost of the damage is D . Thus, if violence is used, the entrepreneur's payoff is $-D$. If no violence occurs, then his payoff is 0. We place some restrictions on these payoffs to keep the game plausible. Specifically, we assume that $D > m$ (otherwise the entrepreneur would prefer violence to payment and hence would never pay under any circumstance), and $m > 1$ (the Mafioso prefers payment to using violence).⁸

Results

Formally, we solve this game by characterizing the sequential equilibria (Kreps and Wilson 1982). The appendix contains all the

details of these equilibria. In the main text we explain the intuition behind these equilibria and discuss the results. Readers interested in technical aspects of the model should refer to the appendix.

We use backwards induction to explain the decisions of each actor in this game. In order to predict whether the thug demands money, we need to understand the likely consequences of this choice. In particular, the thug is more likely to demand money if he expects to get paid. Yet, the entrepreneur's decision depends upon the likelihood of punishment. Before we can understand the demand and payment decisions, we need to predict whether punishment will occur. The backwards induction algorithm starts with the last decision and works back up the tree, predicting each decision in light of what is anticipated to follow. In order to ask the demand and payment questions, we start by analysing the decision to punish.

Punishment. Suppose a faker demands money and the entrepreneur refuses. If the faker uses violence, then his payoff is $-R$, the risk associated with using violence. Without violence, the faker's payoff is $-r$. Therefore, fakers never punish non-payers. Mafiosi sometimes do. If there is a high police presence, then Mafiosi do not punish (the payoff from punishment, -1 , is worse than the payoff from not, 0). Yet, a low police presence leads to violent punishment (the payoff from punishment, $+1$, is better than the payoff from not, 0).

Payment. Fakers never punish and Mafiosi only punish when it is cheap to do so. Suppose the entrepreneur believes that the probability the thug is Mafioso is μ . As we shall show in a moment, the earlier behavior of the thug influences the entrepreneur's beliefs and so μ need not necessarily be equal to E's initial belief, θ . If the entrepreneur refuses to pay, then he is punished only when the thug is Mafioso and the police presence is low (when punished, the entrepreneur's payoff is $-D$). Thus, the probability of punishment is μp . Therefore, E's average payoff from refusing to pay is $-D\mu p$. This compares with a payoff of $-m$ if the entrepreneur complies with the thug's demands. Let σ_E represent the probability that the entrepreneur pays. If $\mu > \frac{m}{Dp}$ then E should pay, $\sigma_E = 1$. If $\mu < \frac{m}{Dp}$ then E should refuse to pay, $\sigma_E = 0$, and if $\mu = \frac{m}{Dp}$ then E is indifferent between paying and refusing and could do either; σ_E is any number between zero and one.

The Decision to Demand. Thugs should make demands only if they expect to gain from doing so. Mafiosi always expects to gain from demands; whether fakers also gain depends upon the reaction of the entrepreneur. Even when the entrepreneur refuses to pay, the Mafioso gets to use violence when the police presence is low. Formally, M's payoff is $U_M = m\sigma_E + p(1 - \sigma_E) + (1 - p)(1 - \sigma_E) 0 = m\sigma_E + p(1 - \sigma_E) > 0$. Thus, Mafioso always expects to do better by making demands. Thus, the probability that Mafioso make demands is $\sigma_M = 1$.

Fakers risk discovery by both the police and the Mafia when they make demands. Therefore, they will only risk making demands when they expect to get paid. If F makes a demand, then with probability σ_E he gets paid, which is worth $m - r$. Yet, with probability $(1 - \sigma_E)$ E refuses to pay. As we have already discovered above, F will not use violence, so his payoff is $-r$. On average, F's expected payoff is $U_F = \sigma_E(m - r) - r(1 - \sigma_E) = -r + \sigma_E m$. When this is positive, F demands money. If it is negative, then F makes no demand and when $-r + \sigma_E m = 0$ then F is indifferent.

Let σ_F represent the probability that fakers demand money. Note that F's demand decision depends upon his expectation about E's response. In turn, E's response depends upon his belief about the thug. In order to proceed we need to examine how E's beliefs change.

Initially, E's beliefs are θ . This means that the probability that the thug is a Mafioso is θ . Suppose that fakers decide not to demand money, $\sigma_F = 0$. Since fakers do not demand money, if the thug makes demands, then the thug must be a Mafioso, $\mu = 1$. Alternatively, consider the case where both Mafiosi and fakers always demand money, $\sigma_F = 1$. Upon being threatened, the entrepreneur's beliefs remain the same, $\mu = \theta$. In general, if a demand is made, then $\mu = \frac{\theta\sigma_M}{\theta\sigma_M + (1 - \theta)\sigma_F}$. Since $\sigma_M = 1$, this reduces to $\mu = \frac{\theta}{\theta + (1 - \theta)\sigma_F}$. This updating is known as Bayes' rule.

Predictions

Whether thugs demand money and whether entrepreneurs pay depends upon the entrepreneur's beliefs and expected level of policing. The results are summarized in Figure 2. There are three patterns of behavior, which are denoted regions 1, 2, and 3 in the figure. In region 1, only Mafioso demands money and the entrepreneur always refuses to pay. The Mafioso uses violence only

when the police presence is low. In region 2, both Mafioso and fakers demand money and the entrepreneur always pays. In region 3, the Mafioso always demands money, but the fakers only demand money sometimes. In response to demands, the entrepreneur sometimes but not always pays.

In this section, we describe each of these patterns of behavior and the conditions under which they occur in more detail. While we explain the intuition behind each case, the formal analysis is in the appendix. We start by examining region 1. Substantively, this region is the least interesting. Yet examining it helps to provide insight into the thug/entrepreneur relationship.

Region 1. The behavior in this region could be described as vandalism. This type of behavior occurs when the police presence is usually high. Entrepreneurs never pay the thug even if they are certain that the thug is a member of the Mafia. Since they never expected to get paid, only Mafiosi demand money. Yet despite being certain that the thug is Mafioso, the entrepreneur prefers to risk being punished rather than pay protection money.

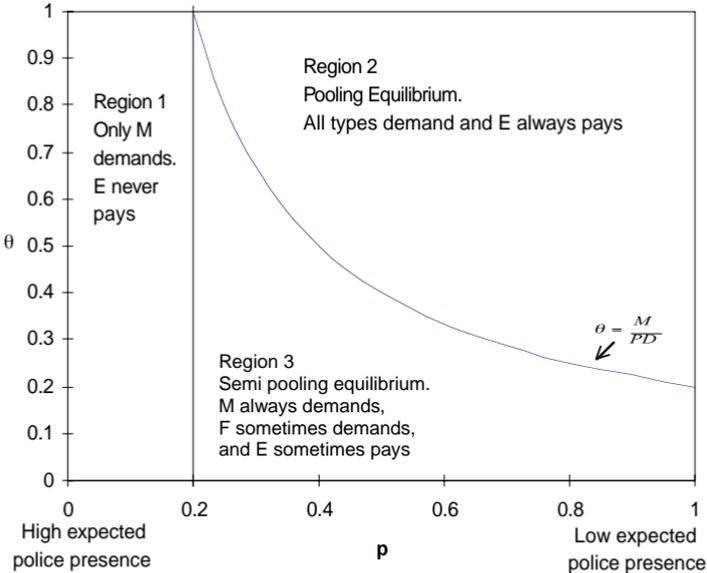


Figure 2. Outcomes in the Mafia Game

The reason for this behavior is as follows: The Mafioso only uses violence when the police presence is low. In region 1, the police presence tends to be high. Specifically, the probability of a low policing level is less than $\frac{m}{D}$. Even when the thug is certainly Mafioso, the probability of violence is still only p . Since $p < \frac{m}{D}$, the expected cost from refusing to pay, pD , is less than the protection money, m . Therefore, E never pays. Fakers have no incentive to demand money since it is never paid. In game theoretic terms, this pattern of behavior is called separating: different types behave differently.

Substantively, this region is uninteresting. Given the high average level of policing, entrepreneurs refuse to pay protection. It is impossible for organized crime to occur under these circumstances. It is interesting to note that as the level of policing increases (p goes down), the Mafia need to increase punishments (increase D) and reduce demands (decrease m) so as to stay out of region 1 and keep organized crime possible.

Having examined behavior when organized crime does not exist, we consider those situations where entrepreneurs pay if they are certain that they are dealing with a true Mafioso.

Region 2. In region 2, the entrepreneur believes that the thug is likely to be a legitimate member of the Mafia, $\theta > \frac{m}{Dp}$. Believing that the thug is likely to be Mafioso, the entrepreneur pays when threatened. Since the entrepreneur always pays, Fakers, as well as Mafiosi, demand money.

The entrepreneur knows that non-payment will only be punished by true Mafioso when the police presence is low. Yet, given his prior beliefs, $\theta > \frac{m}{Dp}$, the size of damages, D , and the probability of punishment, θp , outweigh the cost of paying protection, m . Since both M and F demand money, the entrepreneur cannot distinguish between them. Since he cannot update his beliefs about the thug's type and his prior suggestion that he should pay, the entrepreneur pays whenever threatened. This pattern of behavior is commonly referred to as pooling.

The pattern of behavior in region 2 only occurs when E is fairly certain that the thug is Mafioso, damages are large, and demands are small: $\theta > \frac{m}{Dp}$. Under these conditions, there is never any violence. If the Mafia establishes a reputation and can prevent fakers from entering the market for protection, then they never actually have to carry out threats, since entrepreneurs are compliant with their demands. In the dynamic setting, we might want to consider

the credibility of behavior in region 2. Under these conditions, entrepreneurs always pay and so it is extremely tempting for others to act as fakers. Hence, once established, the Mafia must work hard to exclude fakers, or other pretenders and rivals, from the protection market. The role of focal points in the size of demands is also important. Without such focal points, the Mafia would be tempted to bid the size of m up towards θDp . In effect, this moves the line ($\theta = \frac{m}{Dp}$) between regions 2 and 3 upwards.

At this point it is worth considering what happens on this line between regions 2 and 3 since it will be important when we return to the dynamic analysis later. On the line $\theta = \frac{m}{Dp}$, E is indifferent about whether to pay the demand: the expected punishment exactly equals the amount to pay in protection. Therefore, anything is a best response for the entrepreneur. He could pay, refuse to pay, or randomize whether he is going to pay. As discussed, whatever the entrepreneur does, the Mafioso always demands money. Fakers only demand money if the expected payoff from doing so is positive: $U_F = \sigma_E(m - r) - r(1 - \sigma_E) = -r + \sigma_E m$. Thus, if the entrepreneur often pays, $\sigma_E > \frac{r}{m}$, then fakers demand money ($\sigma_F = 1$). If the entrepreneur rarely pays, $\sigma_E < \frac{r}{m}$, then fakers do not demand money ($\sigma_F = 0$). If the entrepreneur pays with probability $\frac{r}{m}$, then fakers are indifferent about whether to demand. Therefore, providing that $\sigma_E \geq \frac{r}{m}$ we can support the pooling behavior characteristic of region 2 on the line $\theta = \frac{m}{Dp}$.

Region 3. In region 3, fakers sometimes demand money and entrepreneurs sometimes pay. This behavior occurs when the entrepreneur thinks the probability of the thug being a Mafia member is low, $\theta < \frac{m}{Dp}$. In this region, violence occurs because entrepreneurs sometimes refuse to pay Mafiosi.

To explain the logic behind this behavior, it is useful to consider why neither of the previous patterns can exist. Given the entrepreneur's ex ante beliefs, he should refuse to pay the thug. Since $\theta < \frac{m}{Dp}$, on average, it is cheaper to risk punishment than to pay. So suppose that the entrepreneur refuses to pay. In this situation only the true Mafiosi demand money. This is the separating behavior observed in region 1. However, under the conditions in region 3 this cannot be equilibrium behavior. In region 3, since $p > \frac{m}{D}$ the entrepreneur would pay if he were certain he were dealing with a Mafioso. Yet, this is not the situation that we have. Since only Mafiosi demand money, upon seeing a demand, E must infer that the thug is a legitimate Mafia member. Knowing this, E should pay.

However, if E pays demands then the fakers should also demand money.

Thus neither the separating behavior of region 1 nor the pooling behavior of region 2 is possible in region 3. If the thugs separate, as in region 1, then E should pay demands. But then fakers would also prefer to demand money. If the thugs pool, as in region 2, then E should refuse demands. But since E never pays, fakers do not want to make threats. In region 3, fakers sometimes demand money and entrepreneurs sometimes pay. We detail this behavior, often referred to as semi-pooling or semi-separating, below.

Fakers only sometimes demand money. Since sometimes fakers do not make demands, upon being threatened the entrepreneur is more likely to believe he is dealing with a Mafioso than he previously did. The intuition here is that initially there is a pool of thugs. Some of the fakers in this pool drop out (by not threatening). Since some fakers leave, the pool is richer in Mafiosi than before. In particular, if $\sigma_F = \frac{\theta(pD - m)}{m(1 - \theta)}$ then the entrepreneur is indifferent about whether to pay. In this situation, E randomizes whether to pay. Specifically, $\sigma_E = \frac{r}{m}$. Since E only pays with probability $\frac{r}{m}$ in expectation, F gets the same payoff whether or not he demands money. Hence randomizing his decision is optimal. Given this pattern of behavior, no player could do better by playing differently.

Having outlined the intuition behind the semi-pooling equilibria, it is worthwhile to examine it more closely. As the entrepreneur's beliefs change, then so does the probability that the thug demands money. When the thug is likely to be a faker (low θ), then the thug is unlikely to make demands. However, as the likelihood of Mafia membership increases, then fakers demand money more often. As these beliefs approach $\theta = \frac{m}{Dp}$, then all fakers demand money (the pooling equilibrium of region 2).

Comparative Statics. The behavior of thugs and entrepreneurs depends upon their circumstances. When the expected level of policing is high, entrepreneurs never pay even if they are certain they are dealing with a Mafioso. Under these conditions violence occurs, but infrequently enough that entrepreneurs are not coerced into paying protection. As such, high levels of policing prevent organized crime from becoming established. There is a further counter-intuitive conclusion that we can draw from the analysis of region 1, namely that the presence of a certain amount of violence

is not, as such, an indication of a significant presence of organized crime. Let's now consider a transition from low level of policing to higher level of policing. As policing increases, the Mafia must lower their demands (small m) and increase the damage (large D) if they ever expect to get paid. The level of violence observed by citizens actually increases precisely at the moment when the State is supplying better policing to the community. At the same time, the Mafia seems more reasonable in its demands.

When policing is sufficiently poor ($p > \frac{m}{D}$) and entrepreneurs think the thug is likely to be a Mafioso ($\theta > \frac{m}{Dp}$), then thugs always demand money and the entrepreneur always pays rather than risk punishment. Under these circumstances, violence never occurs. This is the symmetrically opposite situation to the one detailed above, as far as observed level of violence goes. The absence of violence actually means that the Mafia is fully in charge and nobody dares to challenge its monopoly over protection. When all is quiet, everything may be going wrong.

When the entrepreneur is less certain of the thug's type, $\theta < \frac{m}{Dp}$, he sometimes refuses to pay. This can lead to violence. In this region (region 3), fakers are more likely to make demands as E becomes more likely to believe they are Mafiosi (increasing θ), as policing levels fall (increasing p), as punishments increase (increasing D), and as demands decrease (decreasing m). The rate at which E pays increases as demands get smaller and as the risk to fakers increases. Thus, from the Mafia's point of view, the more they can identify and punish fakers the easier it is to receive payment from the entrepreneurs. It should not come as a surprise, therefore, that Mafiosi are very keen to punish fakers. The presence of the latter, in turn, increases the level of turbulence.

In region 3, the probability that violence occurs is $\theta p(1 - \frac{r}{m})$. There is no violence in region 2. As E's beliefs increase (increasing θ) or the level of policing falls (increasing p), the level of violence initially increases, and then drops to zero once $\theta > \frac{m}{Dp}$. Thus, the occurrence of violence is non-linear: initially increasing and then falling to zero.

The fact that the size of demands tends to be fixed by focal points has an interesting dynamic. Since inflation erodes the real value of m , we expect the following pattern to occur in settings which experience both high inflation and Mafia presence – such as Russia in the late 1980s and early 1990s. Fixing all the other parameters, as m becomes smaller over time then, fewer fakers make demands and E becomes more likely to pay when threatened. As the value of m falls, we then enter region 2, where E always pays. The declining

value of m also reduces the amount of violence. Although inflation increases compliance, it also rapidly erodes the value of supplying protection. The stickiness of focal points means that the size of m cannot be index linked. Thus, periodically the nominal level of demands will jump. Associated with each jump is an increase in non-payment and violence. In order to avoid this outcome, Mafia groups may prefer to be paid in hard currency or in kind. The available evidence (see e.g. Serio 1997: 97–101; Varese 1994 and 2001) seems to point to the fact that until the mid-1990s criminals demanded payments either in US dollars or in kind. Up to 1990 and after 1995 they resorted to using both rubles and US dollars (it is no coincidence that since mid-1995 the exchange rate of the ruble to the dollar has floated within a ‘corridor’ that has periodically been revised to allow the gradual depreciation of the ruble). The recent decision of the government and the Russian Central Bank (*RFE/RL*, 10/XI/1997) to support an average exchange rate of 6.1 rubles to \$1 during 1998 and an average rate of 6.2 rubles to the dollar from 1998 to 2000 should further stabilize the currency and make the ruble a currency used to pay the Mafia.

Repeated Interactions. Using the simple model above we have been able to tell several dynamic stories about how change affects the Mafia/entrepreneur relationship. However, we have not yet considered the most important dynamic effect: reputation (see Alt et al. 1988). If the entrepreneur believes that the thug is Mafioso, then he will always pay (region 2: $\theta > \frac{m}{Dp}$). This provides the thug with an incentive to build a reputation. If he takes actions that convince the entrepreneur that he is Mafioso, then in the future he is always paid. Thus, a thug might undertake myopically suboptimal actions today because the reputation this creates helps him collect tomorrow. The importance of reputation can hardly be exaggerated for the Mafia. Reputation is important on two counts. Like any other business, protection agencies thrive if they have a good name. More customers are attracted to the ‘family’ and competitors do not dare to enter the market. Furthermore, a Mafioso’s reputation enables him to save directly on ‘production costs’ (Gambetta 1993: 44). A reputation as a credible protector enables the Mafioso to save on the use of violence to convince reluctant victims to pay and competitors not to enter into the Mafia group’s territory.

The value of a good Mafia reputation can be appreciated indirectly by the fact that it may persist even if unfounded. Many Mafia families in the United States have lived off their reputation for a

considerable number of years after the Prohibition wars. If seriously challenged, they would not have been able to collect protection money and scare the competition off (Reuter 1986: ch. 6). The entrepreneur who pays protection money therefore has an incentive to find out if the thug is as strong as he claims to be. Ultimately, the test of being a genuine Mafia is the ability to punish non-payment. Rather than continually paying protection in the long run, the entrepreneur might find it worthwhile to risk being punished initially in order to weed out fakers demanding money. An Italian restaurant owner in Moscow waited for his car to be burned before starting paying protection money (Varese 2001: 68); also the representative of the American firm approached by supposedly Chechen Mafiosi refused to pay, but in this case no punishment ensued (see above p. 350). These are both instances of the same strategy: victims of the Mafia are testing out whether the source of the request is genuine or bogus.

How does the prospect of repeated interactions affect behavior? Under what conditions do thugs attempt to build a reputation and when are they successful? Given that the entrepreneur faces paying protection repeatedly, when does he test the credibility of thugs? In order to address these questions, we analyse the game in the repeated setting. Rather than playing the game only once, we examine what happens when the Mafia has a repeated opportunity to demand money. The game is played twice. Unlike the single period game where players wanted to maximize their immediate payoffs, in the repeated setting players must worry about how their actions affect future behavior. Although we repeat the game only twice, the model captures the tensions created by repeated play. As we have already observed, behavior and hence payoffs in the second period depend upon E's beliefs. In the first period, players are concerned not only with immediate rewards but also with the information that is revealed and how this will affect future interactions.

The mathematical analysis of the repeated game is considerably more complex than that for the single period game. For this reason, all the mathematics has been consigned to the appendix. In the main text, we concentrate on the intuition behind the results and use a series of pictures (Figures 3–5), rather than mathematics, to explain the logic.

For completeness' sake, we start by analysing behavior in region 1. When $p \leq \frac{m}{D}$, the entrepreneur never pays the thug even if he is convinced that the thug is Mafioso. When the level of policing is

high, the thug has too few opportunities to punish the entrepreneur. Repeating the game does not alter this situation. In the future, the entrepreneur will never pay, so the thug has no incentive to behave suboptimally today to build a reputation. The possibility of organized crime under these conditions is remote. Yet, as the level of policing falls, the threat of punishment encourages the entrepreneur to pay. We consider regions 2 and 3 next.

If $p > \frac{m}{D}$, then, myopically, the entrepreneur would sooner pay a Mafioso's request than risk punishment. In the short term, if the thug can convince the entrepreneur that he is Mafioso, then in the long run he expects to be paid. Thus, the thug has incentives to act tough today to ensure payment tomorrow. Yet, we should be wary of assuming thugs can easily build reputations. If all types have an incentive to act tough, then such toughness does not tell the entrepreneur anything about the thug's type. The entrepreneur also has an incentive to act tough in the short term. If fakers are unwilling to risk punishing him, then he can determine the type of thug he faces and avoid paying fakers in the future.

There are six equilibria: Ia, Ib, IIa, IIb, IIIa, and IIIb. We illustrate the conditions under which equilibrium occurs in Figures 3–5. When the faker's risk for using violence is large, then equilibria Ia, Ib, IIa, and IIb occur. When the contrary is true, $m \geq R$, then we observe equilibria IIIa and IIIb. In general, when E's initial beliefs are high, i.e. the thug is likely to be Mafioso, then both fakers and Mafiosi demand money. We denote these equilibria with an **a**. In the equilibria denoted **b**, the Mafiosi always demand money and fakers sometimes demand money. In these semi-pooling equilibria the entrepreneur sometimes refuses to pay. However, in the pooling equilibria (**a**), E always pays. In the situation where $R \geq m$, there are two patterns of behavior. When the expected police presence is low ($p \geq \frac{m}{r-m}(r+1-m)$), equilibria Ia and Ib, punishment only occurs when Mafioso observes a low police presence. In contrast, when the police presence is higher, $p \leq m-1$, Mafioso punishes regardless of police presence but fakers do not punish.⁹ In equilibrium IIIa, all types punish non-payment. In equilibrium IIIb, sometimes fakers punish non-payment, but Mafiosi always punish non-payment. Equilibria IIIa and IIIb only occur when payments are large relative to the risks that fakers face, i.e. $m \geq R$.

In the single period game, the demand decision acts as a signal of type. In the repeated game, the thug has two additional opportunities to signal his type: the first demand decision and the decision to punish. We discuss them in reverse order. Myopically, thugs only

punish non-payment when they are Mafiosi and the police presence is low. Since only Mafiosi should punish, using violence is a signal of Mafia status. Yet, it is a costly signal to use. Mafiosi risk imprisonment from using violence when the police presence is high. Fakers face the additional risk of Mafia detection. In the equilibria labelled Ia and Ib, thugs do not exploit this option, and non-payment is only punished by true Mafiosi when the police presence is low. For fakers the risk of detection by the real Mafia is too high. When the police presence is high, the Mafioso also regards it as too risky to punish.¹⁰ In the equilibria denoted II, M punishes whether or not the police are watching, but fakers never punish. In equilibria IIIa and IIIb all types punish (at least sometimes). The incentive to appear tough means thugs punish when myopically they should not.

Thugs' incentive to punish often means that their willingness to use force goes unobserved. For example, in equilibrium IIIa, for a large range of conditions non-payment is always punished. Therefore, as we might expect, entrepreneurs always pay up. It is important not to confuse the role of reputation in this result. In the pooling equilibria, Ia, IIa, and IIIa, in which all types make first round demands, no reputation is created. All types take the same action, so E's beliefs remain unchanged. However, it is the prospect that thugs want to build and maintain reputations that leads E to fear

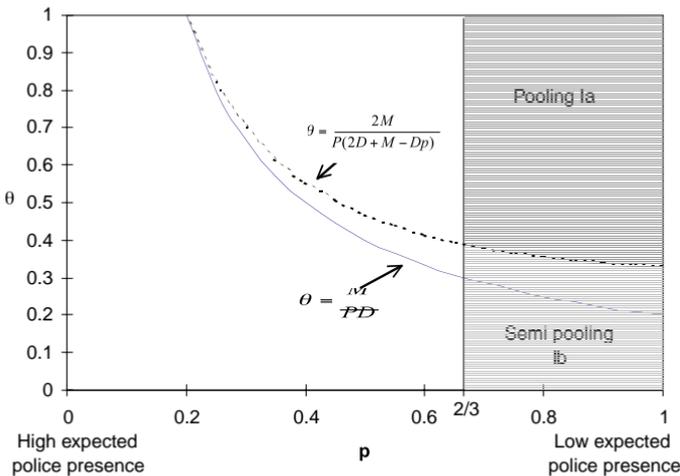


Figure 3. Equilibrium Ia and Ib in the Twice-Repeated Mafia/Entrepreneur Game

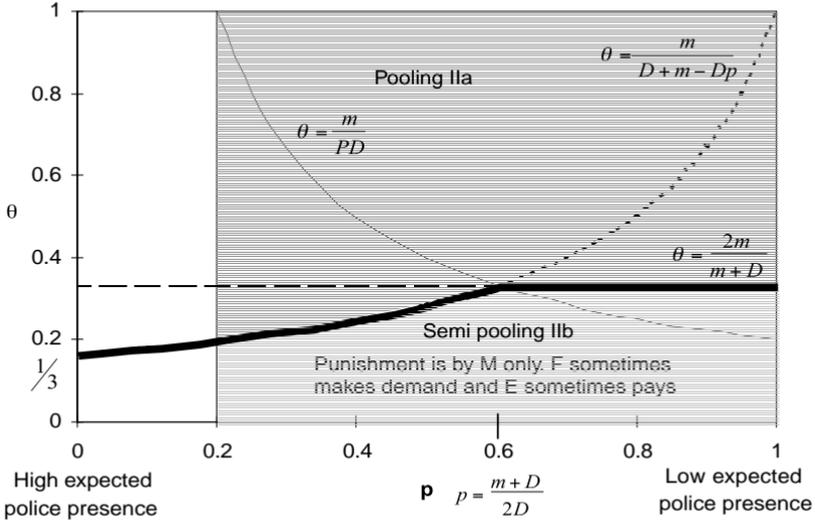


Figure 4. Equilibria IIa and IIb in the Twice-Repeated Mafia/Entrepreneur Game

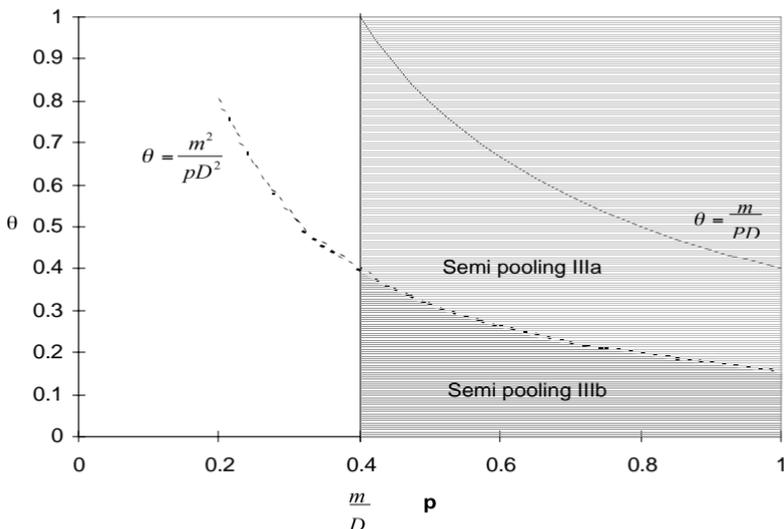


Figure 5. Equilibria IIIa and IIIb in the Twice-Repeated Mafia/Entrepreneur Game

punishment and hence to pay. When E's prior is above $\frac{m}{pD}$ (region 2), it is perhaps not unexpected that E pays. Given these priors, the probability of the thug being Mafioso leads the entrepreneur to want to pay. The entrepreneur is willing to pay demands in region 2 even in the single period game. Yet, E will sometimes pay even outside this range. Figure 4 shows that equilibrium IIa occurs in part of region 3. In this equilibrium the entrepreneur pays all demands. Myopically, E should not pay in this region. Yet, the incentives to build reputation mean that M punishes even if the police are watching. This bias towards violence compels E to pay. In summary, the desire to form a reputation means that thugs are more likely to punish than is the case in the single period game. This propensity to use violence can induce E to pay even when he would not normally do so in single period play (Figure 4: equilibrium IIa in region 3). Although thugs try to build a reputation, in the pooling equilibria E's beliefs are unchanged since the sincerity of demands is never tested.

Thus, in conclusion, thugs want to build a reputation. To do so they prefer to use force when myopically they should not. This propensity to use force deters E from testing the thug's resolve. Hence entrepreneurs pay even though myopically they should not. This behavior occurs in those regions of equilibria IIa and IIIa that lie in region 3. For real Mafiosi, the prospect of few opportunities to punish in the future causes them to risk using force to establish a reputation ($p \leq m - 1$). As payments increase, both M and F become more likely to punish non-payment since the reputation this generates is worth more.

The long run interests of reputation-building make thugs overly aggressive. Long-run benefits also cause entrepreneurs to act against their short-term interests. As we saw in the single period play, myopically E should pay if his beliefs are $\theta \geq \frac{m}{Dp}$ (region 2). Given these beliefs, the risk of punishment is larger than the protection money demanded. Yet the entrepreneur may refuse to pay the thug under these conditions. While it is better to pay once to avoid the risk of punishment, the risk does not justify paying twice. In a repeated setting, entrepreneurs use the early interactions to gauge whether or not they should be paying the thug.

We introduced the idea of filtering out the fakers in the semi-pooling equilibria of region 3 in the single play game. In these equilibria, the entrepreneur sometimes refuses to pay and this sometimes results in E being punished. Yet, the advantage of refusing to pay is that some of the fakers drop out, either by making no

demands or by refusing to punish. These actions help E identify the fakers and so avoid paying them in the future.

In single period play we saw that if $\theta \geq \frac{m}{Dp}$ (region 2), then E prefers to pay rather than risk punishment. In repeated play, the entrepreneur has an incentive to refuse payment when myopically he should pay. To illustrate, consider the region of equilibrium I**b** that lies in region 2: $\theta < \frac{2m}{m+D}$, $R \geq m$, and $p \geq \frac{m+D}{2D}$ (Figure 4).

Suppose that in this zone all types make demands (as in single period play). If E pays in the first period, then he learns nothing about the thug and, hence, given his prior belief, he also pays in the second round. Instead of paying in both rounds, E can risk taking a lottery. Given the large risks of being identified ($R \geq m$), fakers will not punish; but Mafiosi punish regardless of the police presence. If the entrepreneur refuses to pay, then he will be punished by Mafioso but not by fakers. This enables E to identify the fakers and so refuse to pay them in the second round. Thus, with probability θ , E's payoff is $-D - m$, but with probability $(1 - \theta)$ E identifies the faker and avoids having to pay protection. Thus, if $\theta < \frac{2m}{m+D}$, then E prefers to risk punishment in order to filter out the fakers and avoid paying them in the long run.

While it risks punishment, refusing to pay in the first round enables the entrepreneur to identify fakers and avoid repeatedly paying them. This prevents fakers from always demanding money in the first place. Thus, even when on the basis of his prior beliefs E should pay, fakers will not always demand money. When E's beliefs are low, i.e. the thug is likely to be a faker, all the equilibria are semi-pooling: fakers sometimes demand money and the entrepreneur sometimes risks punishment to filter fakers out.

The desire of E to filter out the fakers can also be seen in Figure 3. In equilibria I**a** and I**b**, M only punishes when the police presence is low. For $\frac{2m}{p(2D+m-Dp)} > \theta > \frac{m}{pD}$, E sometimes refuses to pay. Again the filtering works by two mechanisms. First it discourages fakers from demanding money in the first place and second E partially identifies fakers by their refusal to punish.

In the semi-pooling equilibria (I**b**, I**ib**, and III**b**) E learns about the thug. Thugs that make no demands are identified as fakers. If demands are made, then E's updated beliefs make him indifferent about paying. Although reputation and information drive equilibrium behavior they do not do so in a direct manner. Specifically, in the pooling equilibria the desire of thugs to further their reputation

prevents the entrepreneur from testing the thug's resolve. This can be seen in the section of equilibria IIa and IIIa that occur in region 3. In these equilibria, E pays protection money even though he thinks the thug is likely to be a faker. He does so because of the thug's propensity to punish in order to build a reputation. At the same time, precisely because E pays, he fails to learn about the thug. Although in the first period E pays because of the incentive to build reputation, in the second period E does not always pay because no learning occurs. Learning and reputation-building occur in the semi-pooling equilibria. The entrepreneur occasionally risks punishment to filter out fakers and avoid long-term payments.

Although we repeated the game only twice, we can gain insight into repeated interaction between thugs and entrepreneurs. The desire to build and maintain a reputation encourages thugs to punish when myopically they should not. Once the entrepreneur's beliefs are suitably high, the probability that thugs will punish to maintain their reputation deters the entrepreneur from refusing demands. The result is no violence and no additional learning. Dynamically the game becomes frozen with beliefs remaining unchanged. As the game is repeated, then either the thug is identified as a faker or E's beliefs increase until he believes that the thug is most likely Mafioso. At this point, E always pays and no additional information gets revealed.

The entrepreneur filters out fakers by sometimes refusing to pay. Each time he does so, some fakers drop out of the pool and the entrepreneur becomes more convinced that he is dealing with a real Mafioso. Once he is suitably convinced, then he always pays in the future. The length of the filter process depends upon E's beliefs and other conditions. For example, in equilibrium Ib filtering typically only takes a single period. Either the faker drops out and is identified by E or E's beliefs increase ($\mu(\text{pay}) > \frac{m}{pD}$) such that E pays in the future. Yet if the thug refuses to punish in the first period, then E filters again in the second period. Under other circumstances the filtering process typically takes two periods. For example, in equilibrium IIb if $p < \frac{m+D}{2D}$, then filtering takes two periods because after the first round E's beliefs still lie in region 3.¹¹

When the entrepreneur's initial beliefs are low, then fakers are more reluctant to demand money since the entrepreneur sometimes refuses to pay in order to partially filter out the fakers. Yet, after the initial filtering E's beliefs tend to become fixed. It is during

the initial period that E filters. This remains true even if the number of interactions is increased. To see why, consider the motivation for filtering. By refusing to pay, E tests whether the thug is really Mafioso and also discourages fakers from demanding money in the first place. Once a faker is identified, the entrepreneur never has to pay him again. However, it is costly to filter because sometimes E refuses to pay the legitimate Mafia and is punished as a consequence. If the game were 10 periods long there would be no point trying to identify fakers in the penultimate period. Suppose the thug was a faker. The entrepreneur uselessly paid him for eight periods and carried the same risk of punishment. Entrepreneurs partially filter in their initial interaction. In subsequent periods E's beliefs remain fixed and violence does not occur.

Despite the perpetual threat of violence, after the initial interactions beliefs become fixed and dynamically the game is frozen. Once the Mafia has established itself the amount of violence is minimal. However, exogenous shocks to the system may restore violence. Changes in the size of demands, the level of policing, or confusion over succession may lead to a new period of violence as entrepreneurs partially filter and re-establish their beliefs that they should pay in the long run. Since we have already discussed inflation effect on demands, we first examine changes in policing and then confusion over succession. To avoid having to discuss pre-emption, we consider unanticipated changes.

We illustrate possible effects in changes in policing in Figure 6, which is an example of equilibria Ia and Ib. Following the initial interaction, E's beliefs become frozen on the line $\theta = \frac{2m}{p(2D+m-D\beta)}$. Once at these beliefs, little violence occurs because E realizes that it is in his interest to pay. To start with, consider a shock that reduces the level of policing. Following the change, E strictly wants to pay when threatened and observationally there is no increase in violence. The beliefs that E held prior to the shock were on the line $\theta = \frac{2m}{p(2D+m-D\beta)}$. Yet, following the shock, E's beliefs are in the interior of region 3. At this point, it is never worth testing the thug's resolve and the entrepreneur pays in all future interactions. Paradoxically, one can reduce Mafia violence by taking police off the streets. Yet, the absence of violence is a sign of Mafia entrenchment and policy failure.

Policies that lead to an increase in policing levels lead to a short-term increase in violence. Consider an increase in the expected

level of policing from p to p' . E's beliefs now lie below the line $\theta = \frac{2m}{p(2D+m-Dp)}$. Given these beliefs, the entrepreneur attempts to partially filter out fakers. Thus, immediately following an increase in policing, entrepreneurs sometimes refuse to continue paying. This inevitably leads to an increase in violence, since E's refusal to pay is punished by real Mafioso. The analysis suggests that although increasing the level of policing reduces the overall number of thugs that engage in organized crime, the initial impact of the policy is to increase the amount of violence. The increase in violence is only a short-term phenomenon while entrepreneurs attempt to filter out fakers. Violence is an unavoidable consequence of combating organized crime.

Similarly, confusion over succession in a Mafia family may cause uncertainty over whom to pay. When a boss goes to prison, is deposed, or becomes unavailable for some other reason, new arrangements emerge. For instance, when Joe Bonanno, a prominent American Mafioso, left the US in 1957, he was 'careful to make preparations to avoid confusion and to ensure continuity in his absence' (Bonanno 1983: 195, quoted in Gambetta 1993: 62). In a world where information flows perfectly, customers would be informed immediately and would have no doubts over whom to pay in the next stage of the interaction with the Mafioso. The Mafia, however, is not this ideal world. For instance, confusion among the Mafia's customers ensued when Mariano Marsala, the boss of the small Sicilian town of Vicari, was deposed as a consequence of the Mafia war of the late 1970s and 1980s. Some of his old customers continued to ask Don Mariano for protection, even though he was no longer in a position to supply it. As far as our model is concerned, he had become a faker. Don Mariano kept supplying protection behind the back of the new boss until 1983, when he was exposed and killed (Vicenzo Marsala's testimony, 1985, quoted in Gambetta 1993: 63).¹² The issue of succession can be addressed within the framework of this model. Our game models behavior with respect to beliefs. Hence if news of a succession is completely confined within the Mafia, an entrepreneur's beliefs would not be altered and he will continue to pay the previous Mafioso, as in the case of Mariano Marsala's customers. If information flows perfectly, entrepreneurs would update their beliefs and switch to paying the new Mafioso. If information flows imperfectly, an entrepreneur would be unsure about whether or not he is dealing with a real Mafioso. Depending on the level of his uncertainty, he would

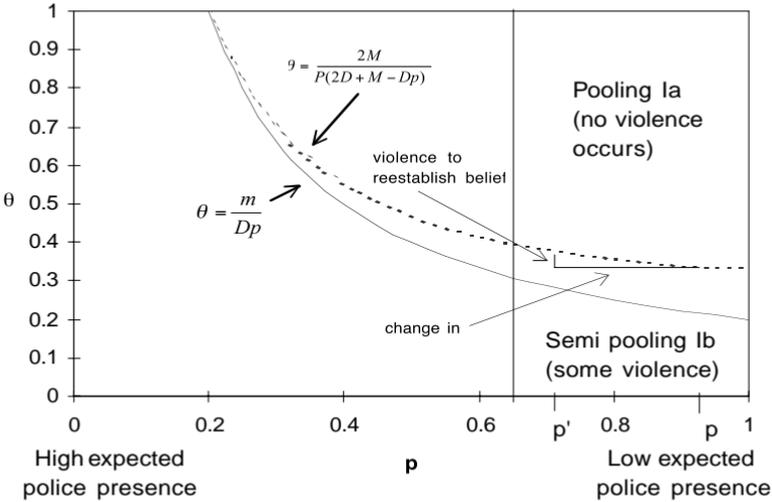


Figure 6. The Effect of Changing the Level of Policing on the Occurrence of Violence in the Dynamic Setting

act accordingly. This model can account for the effect of differing levels of information dissemination on the entrepreneur’s decision to pay.

Once entrepreneurs are convinced that they are dealing with the Mafia, they pay. Outwardly there is no sign of Mafia involvement. Mafia activity becomes visible only when its position is challenged. Increased policing, succession disputes, and outside rivals all lead to increased violence as entrepreneurs test credibility and filter out the fakers. The Mafia punish to restore their position as the only supplier of private protection.

Conclusions

The model presented in this paper allows us to characterize the behavior of fakers, Mafiosi running extortion rackets, and their victims. When the expected police presence is high, entrepreneurs know that those making demands are Mafiosi, but the expected level of damage they can inflict is not high enough for the entrepreneur to comply (region 1 of the one-shot game). Paradoxically, in such well-policed areas there may be sporadic violence. Hence, this

setting can more fruitfully be thought of as a situation when the level of policing has recently increased and hence a previously existing Mafia organization exists. In the long run, Mafia organizations do not survive in this setting, since the high police presence prevents them from collecting revenues, but during their demise violence occurs sporadically. Yet, from a public policy perspective, these instances of violence are signs of success, rather than failure.

A transition from low to higher levels of policing leads Mafiosi to use harsher punishment (big D) and make smaller demands (small m) in order to continue their racketeering activities. In this scenario, more violence is an indicator of an increased effort of the state to stamp out the Mafia.

The absence of violence does not indicate that the Mafia is not operating in a certain area. Our model shows that no violence occurs when entrepreneurs always pay the Mafia. When all is quiet, everything might be wrong. A dynamic pattern emerges from a setting where the Mafia is in charge and the entrepreneurs always pay. In this situation (region 2), fakers have an incentive to enter the market. These entrants produce a dynamic which makes entrepreneurs less likely to believe thugs asking for money are true Mafiosi. As more fakers enter the protection market, entrepreneurs are no longer sufficiently certain that they are dealing with real Mafiosi and refuse to pay (region 3). This refusal to pay is sometimes punished with acts of violence, but in other cases allows the entrepreneur to filter out fakers. As this dynamic further unfolds, fakers drop out of the protection business since payment is not certain any more. This in turn helps entrepreneurs to restore the belief that thugs asking for money are real Mafiosi and as such should be paid.

A Mafia world is plagued not only by internally generated turbulence, but also by external shocks. External economic shocks often increase Mafia violence. For example, if protection money is paid in the local currency, then high inflation erodes the real value of Mafia payments. The size of the payments is typically a focal denomination; for example, payment in Russia in the late 1980s was typically 500 or 1000 rubles. Eroding value makes the entrepreneur more willing to pay when asked, even if he is not entirely certain of the identity of the thug. The same decline in the value of payments causes the Mafia to decide on sudden jumps in the size of demands, for example to 2000 rubles. These jumps in demands are accompanied by violence. Since the entrepreneur's beliefs remain static, the increased demands lead entrepreneurs to question the identity of the thug and attempt to filter out fakers by refusing payment. In

conclusion, high inflation leads to jumps in the level of violence associated with each discrete price jump and a consolidation of mob control as fakers are progressively weeded out at each price jump. Precisely to avoid this scenario, many Mafia groups prefer to be paid in hard currency or in kind.

Although the one-shot game yielded a number of insights, it could not account for a fundamental aspect of this social situation, namely reputation-building. When the Mafioso expects to interact a second time with the entrepreneur, he has an incentive to build up his reputation in order to save on the production of violence. This might lead to an 'excessive' use of violence in the first interaction to create or preserve a reputation. Yet, precisely because of a thug's incentive to use violence, even if it is costly (in our model, a high police presence), entrepreneurs know a refusal to pay will be punished and so they are more likely to pay up. In other words, thugs are more likely to use violence in the first interaction in order to establish their reputation, but since entrepreneurs know this, victims of racketeers are more likely to pay, hence little violence is actually observed (see Chin 1996: 57 for a similar conclusion). Despite the propensity to use violence to establish reputation, beliefs remain unchanged because entrepreneurs always pay. As mafias become more entrenched in society (i.e. as entrepreneurs expect the thug to return) the display of violence is reduced.

The equilibrium analysis of the repeated game still confirms that filtering takes place among entrepreneurs, as predicted by the one-shot game. Entrepreneurs who anticipate multiple interactions sometimes refuse payment, even though they are sufficiently certain that a thug is Mafioso and that from a myopic viewpoint they should pay. By refusing to pay in the initial interaction the entrepreneur can sometimes filter out fakers. Although this creates a high risk of punishment in the first period, the entrepreneur might learn the identity of the thug, which provides the potential long-term benefit of avoiding paying a faker in all future interactions. The incentive to filter means that the majority of instances of payment refusal and punishment occur early in the relationship between entrepreneurs and thugs.

We have now established that both overly aggressive behavior to create and maintain a reputation and filtering exist when actors expect to interact more than once. We have also pointed out an important aspect of reputation-building which has been overlooked by other authors: the expectation that violence will be used to build reputation in the first interaction leads the victim to comply before

the violence is actually displayed. So when do we observe filtering (and the ensuring potential for violence), and when do we observe the acquiescence of the victims who pay right away to avoid punishment?

Filtering behavior is more prevalent as the faker's risk from using violence increases ($R > m$, equilibria Ia, Ib, IIa, and IIb). In contrast, the acquiescence of the victims caused by the fact that Mafiosi are motivated to use excessive violence in order to build their reputation is more prevalent when the faker's risk from using violence is low ($m \geq R$, equilibria IIIa and IIIb). In conclusion, even in this more complex setting, there is an inherent turbulence in the system, which mirrors what we observed in the one-shot game. Once entrepreneurs start to believe they are paying true Mafiosi, fakers still have an incentive to emerge, making some entrepreneurs refuse payment to real Mafiosi, which in turns leads to violence in a never-ending spiral.

External shocks can lead to an increase in violence even when actors expect to interact more than once. We analysed two potential external shocks, changes in the level of policing and changes in the Mafia leadership. If there is a shock that reduces the level of policing, the entrepreneur strictly wants to pay when threatened and we observe no increase in violence. Yet, counter-intuitively, policies that increase policing lead to a short-term increase in violence. Changes in Mafia leadership may also cause entrepreneurs to question whom they should be paying. If entrepreneurs sense that there have been changes in Mafia leadership but they are not quite sure about their nature, they will be inclined to test whether they are paying the right person. In conclusion, our analysis suggests that a world where mafias operate is inherently turbulent, with fakers emerging and disappearing, and violence being used to re-establish identities. This conclusion goes against the widespread perception that racketeers are able to perfectly enforce protection monopolies.

NOTES

We are grateful to Diego Gambetta, two anonymous referees and the editor of *Rationality and Society* for their useful comments. Special thanks to Galina Kravtchenko-Varese for proofreading this article. A version of this paper appeared as Sociology Working Paper no. 3, Department of Sociology, Oxford University, May 2000.

1. By 'Mafia' we refer to criminal rackets that force their service of 'protection' on customers operating in the legal economy. A Mafioso (plu. Mafiosi) is a member

- of a Mafia group. This definition covers a number of cases, such as the Sicilian Mafia (Gambetta 1993) the Triads in Hong Kong and abroad (Chu 2000; Chin, 1996), the Yakuza (Hill 2000) and the Russian mafia (Varese 1994; 2001). Some urban gangs in America also fall under this definition (see Jankowski 1991: 121–3, 187–8 and 305–6). Although these organizations are engaged in other criminal activities as well, they are all territorially-based rackets which extract resources from legal entrepreneurs. This definition is restrictive: for instance, it does not apply to entrepreneurs trafficking in illegal commodities, such as drugs.
2. For an argument as to why the Mafia has emerged in Russia at the time of the transition to the market economy, see Varese (1994; 2001).
 3. Even perspective members of the Sicilian Mafia did not know with certainty who belonged. Antonino Calderone, narrating his initiation ritual in Catania in 1962, recalls that he was surprised to see at the gathering people he knew and he never suspected were in the Mafia, and not to see people he thought were in the organization (Arlacchi 1992: 191).
 4. For example, in 1988, there were more than 6,000 reported cases of racketeering in the Soviet Union; of these, in almost half the cases (2,800) the demand was for 500 rubles, in 535 it was for 1,000 rubles; in 928 instances, criminals tried to charge more than 1,000 rubles (*Trud* 19/V/1990). Racketeers tried to extort 1,000 rubles in protection money from a food cooperative in Moscow in 1990 (*Novoe russkoe slovo* 15/I/1990). Kiosk owners in a district of the Russian city of Perm are asked to pay a fixed sum at regular intervals, usually monthly. The sum paid is 10,000 rubles per day (Varese 2001: 107).
 5. In most countries, tattoos are a feature of prison life (or, indeed, of closed male communities, such as the navy and the army), and do not distinguish Mafiosi from fakers. According to conservative estimates, twenty-eight to thirty million of the Russian inmates were tattooed (data refer to the 35 million people who went through the prison system between mid-60s to the 1980s. Bronnikov, 1993: 50). Bronnikov reports that 'as convictions increase and the terms of incarceration become more severe, the tattoos multiply'. In the Russian criminal world, tattoos signal the inmate's standing in the prison hierarchy, rather than his membership to a Mafia group that operates outside.
 6. This is not the place to discuss whether the Mafia provide a real service, which has economic value (Goldstock 1990), or simply protection against a danger the Mafia itself creates (Schelling 1971; Chin 1996). The payoff values assigned in the context of this game capture the fact that entrepreneurs who operate fixed establishments would rather not pay the money requested by the Mafioso. Whether Mafia protection turns out to be beneficial in the long run, we assume here that entrepreneurs do not search for Mafia protection. In other words, the first interaction between entrepreneurs and mafiosi is governed by (the fear of) violence and attempts to avoid payments. To the extent that it is not, the dynamics studied by this model does not apply. Such situation would occur when the entrepreneur is himself a Mafioso or is connected to the Mafia to start with.
 7. Since Fakers face the risk of discovery by the police and the Mafia we should think of the risk to the faker as always being much greater than that of the Mafioso. Hence we set $R \geq 2$.
 8. The level of damage done to a business is partly at the discretion of the thug. However, certain bounds are likely to exist. First, for thugs the amount demanded should be less than the value of the damage, otherwise the entrepreneur would prefer to suffer the violence rather than pay. Second, if the size of the damage is

too large then the entrepreneur goes out of business, and cannot pay in the future. Within these limits we consider the comparative statics of how the amount of damage affects behavior. There is a wealth of evidence that confirms the above reasoning. Racketeers usually start with minor threats and only in the face of stubborn opposition do they increase the amount of damage (see, e.g. Arlacchi 1986: 26; Chin 1996: 43–46; Chu 2000: 47 and 51–2; Jankowski 1991: 122; Varese 2001: 68).

9. There are values of p for which both constraints hold. Thus, if $m - 1 \geq p \geq \frac{m}{r-m}(r+1-m)$ then there are two equilibrium predictions.
10. Paradoxically, it is when the police presence is likely to be low that M does not punish. The reason is as follows: Since p is high (relative to m) then tomorrow M expects to enjoy punishing E at low cost. Although, M could persuade E to pay tomorrow by using force, the risk does not justify the benefit.
11. The regions of equilibria IIa and IIIa that lie below $\theta = \frac{m}{pD}$ suggest that filtering is delayed. In these equilibria, pooling occurs in the first period and filter occurs in the semipooling equilibrium in the second period. However, this late filter is only an artifact of the two period structure. In these equilibria the thug's desire to create a reputation means that he punishes even when myopically he should not. This deters E from refusing to pay. This incentive to create a reputation disappears in the last period. However, as long as the game continues then thugs want to create reputation and E continues to pay.
12. In situation where the Mafia is more stable, succession is announced almost publicly. After the death of Carmelo Colletti, boss in the town of Ribera (Agrigento, Sicily), in 1983, a meeting was held to appoint a successor. At the end of the meeting the participants paraded to the town's main bar, in full public view. According to investigators, this parade served to inform the public of the identity of the new boss, Gennaro Sortino, who marched at the head of the procession (Gambetta 1993: 60). In the Japanese Yakuza, which enjoys greater official tolerance than the Sicillian Mafia, 'succession is announced with some formality in the relevant social segments of the underworld' (Iwai 1986: 217). In one case, a leaflet was produced with the name of the new boss and his supporters (Iwai 1986: 217).
13. Note that if $m = pD$ then $\sigma_M = 1$, and $\sigma_F = 0$ and $\sigma_E \leq \frac{r}{m}$ is also a sequential equilibrium. This is a special knife edge case that occurs in the line separating regions 1 and 3.
14. The only exception is if $\theta = \frac{m}{pD}$ in which case there are multiple equilibria, each with different payoffs for the thug.

REFERENCES

- Alt, J.E., R.L. Calvert and B.D. Humes. 1988. 'Reputation and Hegemonic Stability: A Game Theoretic Analysis.' *American Political Science Review* 82: 445–66.
- Arlacchi, P. 1986. *Mafia Businesses: The Mafia and the Spirit of Capitalism*. London: Verso.
- Arlacchi, P. 1992. *Men of Dishonor: Inside the Sicilian Mafia: an Account of Antonino Calderone*. New York: William Morrow & Co.
- Banks, Jeffrey J.S. 1991. *Signaling Games in Political Science*. Chur, Switzerland: Harwood Academic Publishers GmbH.

- Bonanno, J. 1983. *A Man of Honour: the Autobiography of a Godfather*. London: Unwin Press.
- Bronnikov, A.G. 1993. 'Telltale Tattoos in Russian Prisons.' *Natural History* 11: 50-9.
- Chin, K. 1996. *Chinatown Gangs: Extortion, Enterprise, Ethnicity*. Oxford: Oxford University Press.
- Chu, Y.K. 1996. 'Hong Kong Triads: The Core Business of Organized Crime.' PhD. Thesis, University of Exeter.
- Chu, Y.K. 2000. *The Triads as Business*. London: Routledge.
- Committee on Governmental Affairs. 1996. *Russian Organized Crime in the United States. Hearing before the Permanent Sub-Committee on Investigation*. Washington, DC.
- Gambetta, D. 1993. *The Sicilian Mafia*. Cambridge, MA: Harvard University Press.
- Gambetta, D. 1994. 'Inscrutable Markets.' *Rationality and Society*, 6: 353-68.
- Goldstock, R. 1996. *Corruption and Racketeering in the New York City Construction Industry*. New York: New York University Press.
- Hill, P. 2000. 'Botaiho: Japanese Organised Crime Under the Boyokudan Countermeasures Law'. Ph.D. Thesis, University of Stirling.
- Iwai, H. 1986. 'Organized Crime in Japan.' In *Organized Crime: A Global Perspective*, ed. R.J. Kelly, pp. 208-33 Totowa, NJ: Rowman and Littlefield.
- Jankowski, M.S. 1991. *Islands in the Street. Gangs and American Urban Society*. Berkeley and Los Angeles, CA: University of California Press.
- Kaplan D. and A. Dubro. 1986. *Yakuza: The Explosive Account of Japan's Criminal Underworld*. New York: Macmillan.
- Kreps, David M. and Robert Wilson. 1982. 'Sequential Equilibria.' *Econometrica* 50: 863-94.
- Ladyzhenskii, A. 1994. 'Krysha.' *Kriminal'naya Khronika* 7: 4.
- Reuter, P. 1983. *Racketeering in Legitimate Industries: Two Case Studies*. Washington DC: National Institute of Justice.
- Reuter, P. 1986. *Disorganized Crime. Illegal Markets and the Mafia*. London/Cambridge, MA: MIT Press.
- Serio, J. 1997. 'Threats to the Foreign Business Community in Moscow.' In *Russian Organized Crime*, ed. P. Williams, pp. 88-108. London: Frank Cass.
- Schelling, T.C. 1971. 'What is the Business of Organized Crime?', In *Choice and Consequence*, pp. 179-94. Cambridge, MA and London: Harvard University Press, [1971]1984.
- Sciarrone, R. 1998. *Mafie Vecchie, Mafie Nuove*. Rome: Donzelli.
- Tribunale di Torino. 1995. 'Sentenza nella Causa Penale contro Adorno Filippo + 12'. Quinta sezione penale. Torino.
- Varese, F. 1994. 'Is Sicily the Future of Russia? Private Protection and the Rise of the Russian Mafia.' *European Journal of Sociology* 35: 224-58.
- Varese, F. 2001. *The Russian Mafia. Private Protection in a New Market Economy*. Oxford: Oxford University Press.

ALASTAIR SMITH is Assistant Professor of Political Science at Yale University. He studies the role of political institutions in determining relations between states and how the timing of

parliamentary elections affects electoral outcomes. He has published papers in *American Political Science Review*, *American Journal of Political Science*, *Economics and Politics*, *Journal of Conflict Resolution*, *International Studies Quarterly*, *International Interactions*.

FEDERICO VARESE is William H. Orrick Visiting Professor at Yale University. His main research interests are the study of organized crime, Soviet criminal history, and the dynamics of altruistic behavior. He is author of *The Russian Mafia* (Oxford University Press, 2001) and has published papers in *Archives Européennes de Sociologie*, *Low Intensity Conflict and Law Enforcement*, *Political Studies*, *Cahiers du Monde Russe* and *Rationality & Society*.

Appendix

In the main text we have characterized the best responses for all players. We know that $\sigma_M = 1$. Rather than repeat the exact analysis in the text more formally, we cut straight to characterizing the equilibria. Since $\sigma_M = 1$, there are three possible cases: separating ($\sigma_F = 0$), pooling ($\sigma_F = 1$), and semi-pooling ($\sigma_F \in (0,1)$). These correspond to regions 1, 2, and 3, respectively, in Figure 2.

Separating Equilibria. We start by characterizing the conditions under which separating equilibria can occur: $\sigma_M = 1$ and $\sigma_F = 0$. Examining the best response function for the fakers, we see that F does not make demands when $\sigma_E m - r \leq 0$. Therefore, if a separating equilibrium exists then $\sigma_E \leq \frac{r}{m}$. Since E does not always pay, $U_E(\text{pay}) = -m \leq -\mu pD = U_E(\text{not pay})$. If M and F play separating strategies, then by Bayes' rule $\mu = 1$. This implies that $m \geq pD$.

Therefore, $m > pD$ then $\sigma_M = 1$, and $\sigma_F = 0$ and $\sigma_E = 0$ is a sequential equilibrium in which E's beliefs upon being threatened are $\mu = 1$.¹³ This is the behavior in region 1 of Figure 2.

Pooling Equilibria. Suppose $\sigma_M = 1$, and $\sigma_F = 1$. In order that $\sigma_F = 1$, $U_F(\text{demand}) = \sigma_E m - r \geq 0 = U_F(\text{no demand})$. Therefore, $\sigma_E \geq \frac{r}{m}$. This implies that $U_E(\text{pay}) = -m \geq -\mu pD = U_E(\text{not pay})$. Since both types pool on the same message, $\mu = \theta$. Therefore, pooling equilibria only exist if $\theta \geq \frac{m}{pD}$. We can now state the equilibria formally.

If $\theta = \frac{m}{pD}$ then $\sigma_M = 1$, $\sigma_F = 1$, and $\sigma_E = 1$, and the entrepreneur's beliefs are $\mu = \theta$. This is the equilibrium that occurs in region 2 of Figure 2.

When $\theta = \frac{m}{pD}$ there is a range of equilibria in which $\sigma_M = 1$, $\sigma_F = 1$, and $\sigma_E \geq \frac{r}{m}$ and the entrepreneur's beliefs are $\mu = \theta$. These equilibria occur on the line separating regions 2 and 3 in Figure 2.

Semi-pooling equilibria. Suppose $\sigma_M = 1$, and $\sigma_F \in (0,1)$. In order that the faker randomizes the demand decision, F must be indifferent about whether or not to make demands: $U_F(\text{demand}) = \sigma_E m - r = 0 = U_F(\text{no demand})$. This implies that $\sigma_E = \frac{r}{m}$. In any equilibrium in which F randomizes, it must be the case that $\sigma_E = \frac{r}{m}$. However, since this requires that the entrepreneur is randomizing, E must be indifferent between paying and not: $U_E(\text{pay}) = -m - \mu pD = U_E(\text{not pay})$. Therefore, $\mu = \frac{m}{Dp}$. By Bayes' rule, $\mu = \frac{\theta}{\theta + (1 - \theta)\sigma_F}$. Therefore, if E is indifferent, then $\sigma_F = \frac{\theta(Dp - m)}{m(1 - \theta)}$. We can now formally state the equilibrium.

If $\theta \in (0, \frac{m}{Dp})$ and $m < Dp$ then $\sigma_M = 1$, $\sigma_F = \frac{\theta(Dp - m)}{m(1 - \theta)}$, and $\sigma_E = \frac{r}{m}$, and E's beliefs are $\mu = \frac{m}{Dp}$. This is region 3 in Figure 2.

Since we have exclusively searched all the possible strategy profiles for thugs, the equilibria represent all the possible equilibria.

Size of demands. So far we have assumed that the size of m is fixed. However, what, if any, are the consequences of allowing the thugs to endogenously choose m ? The first thing to note is that in equilibrium M and F must demand the same amount. (Note: we do not regard $m = 0$ as equivalent to making no demand.) If M and F were to send different signals then the entrepreneur could instantly recognize the fakers and refuse to pay them. Thus, F would prefer either to demand the same as M or to make no demand.

So M and F must demand the same amount. However, can this amount vary? Game-theoretically there are equilibria where only single size demands are made. Suppose in equilibrium, when they make demands, both M and F ask for m . In equilibrium, other size demands are not asked for. If, for example, a thug demanded $m' \neq m$, then what should the entrepreneur believe about the thug's type. Bayes' rule tells us nothing about this situation, as it is a zero probability event (the denominator in the Bayes' formula is zero and hence our beliefs are $\frac{0}{0}$, which is undefined). Therefore, consistent with sequential equilibrium, we are free to fix any beliefs. If these beliefs are that fakers are the type that demand m' then no type would ever want to ask for m' . Hence in equilibrium all types ask for m . If we relax the assumption that m is exogenously fixed, then

observationally equivalent equilibria exist. As we discussed informally above, the Mafia may have preferences over the set of possible equilibria. In terms of equilibrium selection, we should probably concentrate on the most preferred equilibria from the Mafioso point of view. Since, in practice, focal points seem central to the size of demands we treat m as fixed.

Twice-Repeated Game

In order to examine the effect of reputation, we examine the twice-repeated game. Additional notation is required to analyse this game. We index the notation by the time period in which the decision takes place: $t = 1, 2$. In addition, the strategies of players may also depend upon the previous outcomes. Let h represent the history of play, i.e. the previous decisions of the players. For convenience, we represent the histories in as short a form as possible where this leads to no ambiguity. For example, suppose in the first round the thug demands money, E refuses and the thug punishes, then we would represent E's beliefs as $\mu(\textit{violence})$, since the outcome violence can only be reached via demand and refusal to pay. In the single period game, we did not introduce any notation for the thug's decision to punish. Let $s_F^{t=1}$ be the probability that the faker punishes in the first period. Similarly, $s_M^{t=1}(\textit{high})$ and $s_M^{t=1}(\textit{low})$ represent the probabilities that Mafioso punishes in the first round given high and low police presence, respectively.

The solution concept, sequential equilibrium, requires that strategies be sequentially rational. This means that when making choices, players must optimize at every point and they cannot commit to take suboptimal actions in the future. Therefore, the behavior in the second period of the game is given in the earlier analysis. Given this, for any set of beliefs held by E we can predict behavior in the second round. Hence, for any given belief, we know the expected payoff for each player in the second period.¹⁴ The question then becomes, given that we know how beliefs affect payoffs in the second round, how do players act in the first round to affect these beliefs.

Before proceeding, it is worthwhile summarizing the second round expected payoff for each player for any given set of beliefs. First consider region 2: $p \geq \frac{m}{D}$ and $\mu(h) \geq \frac{m}{pD}$. Suppose that given the first round history of play h , that $\mu(h) > \frac{m}{pD}$. In the second period, all types of thug demand money and E always pays. Hence

E's second round payoff, $U_E^{t=2}(h)$, is $-m$. Similarly, $U_F^{t=2}(h) = m - r$ and $U_M^{t=2}(h) = m$. If the history of play is such that $\theta(h) = \frac{m}{pD}$, then $U_E^{t=2}(h) = -m$, $U_F^{t=2}(h) = \sigma_E^{t=2}(h) m - r$ and $U_M^{t=2}(h) = \sigma_E^{t=2}(h) m + (1 - \sigma_E^{t=2}(h))p$, where $\sigma_E^{t=2}(h)$ is the probability that E pays in the second round. If the beliefs generated in the first round are in region 3, then in the second round we observe semi-pooling equilibrium behavior. Hence if $\mu(h) < \frac{m}{pD}$, then $U_E^{t=2}(h) = -\mu(h)Dp > -m$, $U_F^{t=2}(h) = 0$, and $U_M^{t=2}(h) = \sigma_E^{t=2}(h) m + (1 - \sigma_E^{t=2}(h))p = r + p - \frac{p\mu}{m} > 0$.

With these preliminaries over, we analyse the equilibria. Details, such as out of equilibrium belief refinements, are introduced as required.

Region 1: $p \leq \frac{m}{D}$. In region 1, whatever the entrepreneur's beliefs, he refuses to pay. In single period play, this meant that only Mafioso demands money, and that despite being convinced of the thug's Mafia status, E refused to pay. The Mafioso punished only when the police presence was low. In this case the extension to repeated play is trivial since E's beliefs do not affect behavior. Players have no incentive to manipulate their play in early rounds to affect E's beliefs because E's play is independent of such information. In every period, play is identical to that in the single period game.

Regions 2 and 3: $p \geq \frac{m}{D}$. In the main text we informally discussed the equilibria. In this section, our intention is to formally characterize the equilibria and to provide some intuition as to why other patterns of behavior cannot be equilibria. During these characterizations we make several assumptions about out of equilibrium beliefs. To illustrate, suppose that in equilibrium all types demand money in the first round. Since all types threaten, what should E believe if the thug makes no first round threat. The solution concept provides no constraints in this situation since Bayes' rule is undefined in such situations. There are numerous equilibrium refinements (see Banks 1991 for discussion). The spirit of these refinements is to find the type that would gain most from sending the out of equilibrium message. The refinements assume that it is this type that sends the out of equilibrium message. Therefore, in this situation, if not demand is the out of equilibrium message, then E should infer that the thug is a faker, $\mu = 0$. Similarly, an out of

equilibrium failure to punish implies that the thug is a faker. Alternatively, an out of equilibria threat or use of violence punishment should lead E to believe that the thug is Mafioso, $\mu = 1$.

Before characterizing the equilibria we demonstrate why other forms of behavior cannot be equilibrium. In the first period, M and F never separate on making demands. To demonstrate why, suppose they do separate: $\sigma_M^{t=1} = 1$, and $\sigma_F^{t=1} = 0$. E's beliefs are $\mu(\text{demand}) = 1$ and $\mu(\text{no demand}) = 0$. Since they are completely identified, there is no subsequent benefit from punishing unless it is cheap: $s_F^{t=1} = 0$, $s_M^{t=1}(\text{low}) = 1$ and $s_M^{t=1}(\text{high}) = 0$. Given this, E's best response to a threat is to pay, since he knows he is dealing with M who will punish with probability p : $\sigma_E^{t=1} = 1$. However, if E always pays then F also demands money. This contradicts the original premise that the types separate. Hence, there are no separating equilibria. Thus, we need only consider strategy profiles in which types pool or semi-pool in their decision to demand money.

The two-period game offers the thug an alternative opportunity to signal his type. If E refuses to pay then the thug decides whether or not to punish. Given our out of equilibrium belief refinements, M always punishes when it is cheap. We can place an additional restriction on thugs' decisions to punish. Specifically, $s_F^{t=1} = 1$ and $s_M^{t=1}(\text{high}) < 1$ can never be part of an equilibrium profile. Suppose it were, $\mu(\text{no violence}) = 1$. In the second period E would always pay: $\sigma_E^{t=2}(\text{no violence}) = 1$. But then in the first period $U_F(\text{not punish}) = -r + m - r$. If F punishes then $U_F(\text{punish}) = -R + \zeta$, where the maximum value for ζ is $m - r$. But this contradicts the original premise that F punishes.

In order to prove the strategy profiles are indeed equilibria, we analyse each profile separately and find the conditions under which it is a sequential equilibrium. Since equilibrium behavior in the second period is completely described by the original analysis, we do not repeat these results unless ambiguity arises.

Equilibrium Ia.

If $p \geq \left(\frac{m}{r-m}\right)(r+1-m)$, $R \geq m$ and $\theta \geq 2 \frac{m}{p(2D+m-Dp)}$ then the strategy profile $\sigma_F^{t=1} = 1$, $\sigma_M^{t=1} = 1$, $\sigma_E^{t=1} = 1$, $s_F^{t=1} = 0$, $s_M^{t=1}(\text{high}) = 0$, $s_M^{t=1}(\text{low}) = 1$ is a sequential equilibrium with beliefs $\mu(\text{demand}) = \theta$, $\mu(\text{no demand}) = 0$, $\mu(\text{violence}) = 1$, and $\mu(\text{no violence}) = \frac{\theta(1-p)}{\theta(1-p) + (1-\theta)} < \theta$.

Proof: Note that the beliefs are all consistent with Bayes' rule. Therefore, we simply need to check that every player is utility maximizing given their type, their beliefs and the strategy of the other players.

Consider E's decision to pay in the first period. $U_E(\text{pay}) = -2m$ since given these beliefs E also pays in the second period (region 2). If E refuses to pay then there are two cases: (i) $\mu(\text{no violence}) \geq \frac{m}{Dp}$, $U_E(\text{no pay}) = \mu(\text{demand}) p (-D - m) + (1 - \mu(\text{demand})) p (0 - m)$. Thus, $\sigma_E^{t=1} = 1$.

(ii) $\mu(\text{no violence}) < \frac{m}{Dp}$, $U_E(\text{no pay}) = \mu(\text{demand}) p (-D - m) + (1 - \mu(\text{demand})) p (0 - Dp\mu(\text{no violence}))$, where $\mu(\text{demand}) = \theta$, and $\mu(\text{no violence}) = \frac{\theta(1-p)}{\theta(1-p) + (1-\theta)}$. Hence, $U_E(\text{no pay}) = \theta p (-D - m) + (1 - \theta p) (0 - Dp \frac{\theta(1-p)}{\theta(1-p) + (1-\theta)}) = - (2D + m - Dp)\theta p$.

Thus, if $\theta \geq 2 \frac{m}{p(2D + m - Dp)}$ then E's best response is to pay given that $\mu(\text{demand}) = \theta \geq 2 \frac{m}{p(2D + m - Dp)}$. Therefore, $\sigma_E^{t=1} = 1$.

Next consider F's decision to punish: $U_F(\text{punish}) = -R + m - r$ and $U_F(\text{no punish}) \geq -r + 0$ (F's payoff in the second period is 0 if $\mu(\text{no violence}) = \frac{\theta(1-p)}{\theta(1-p) + (1-\theta)} < \frac{m}{pD}$, region 3). If $R \geq m$ then F prefers not to punish. Obviously, F prefers not to punish if $\mu(\text{no violence}) \geq \frac{m}{pD}$.

Does M punishes when the police presence is high? $U_M(\text{punish}|\text{high}) = -1 + m$ and $U_M(\text{no punish}|\text{high}) = 0 + r + p - \frac{pr}{m}$. ($r + p - \frac{pr}{m}$ is M's expected payoff in the 2nd period for region 3). Thus, if $p \geq \left(\frac{m}{r-m}\right) (r + 1 - m)$ then M does not punish during a high police presence.

Finally, consider the initial decision to demand money. Since E always pays demands, both types prefer to make demands. QED.

Equilibrium Ib

If $p \geq \left(\frac{m}{r-m}\right) (r + 1 - m)$, $R \geq m$ and $\theta < \frac{2m}{p(2D + m - Dp)}$ then the strategy profile $\sigma_F^{t=1} = x = \frac{\theta(p-2)(m-pD)}{2m(1-\theta)}$, $\sigma_M^{t=1} = 1$, $\sigma_E^{t=1} = \frac{r}{(-r+2m)}$, $s_F^{t=1} = 0$, $s_M^{t=1}(\text{high}) = 0$, and $s_M^{t=1}(\text{low}) = 1$ is a sequential equilibrium with beliefs $\mu(\text{demand}) = \frac{\theta}{\theta + (1-\theta)x} =$

$$\frac{2m}{p(2D+m-Dp)}, \mu(\text{no demand}) = 0, \mu(\text{violence}) = 1, \text{ and } \mu(\text{no violence}) = \frac{2}{p} m \frac{-1+p}{m-2D+Dp} < \frac{m}{pD}.$$

Proof: Note that the beliefs are all consistent with Bayes' rule. Therefore, we simply need to check that every player is utility maximizing given their type, their beliefs, and the strategy of the other players.

Consider E's decision to pay in the first period. Given F's strategy, $\mu(\text{demand}) = 2\frac{m}{p(2D+m-Dp)} > \frac{m}{Dp}$ and $\mu(\text{no violence}) < \frac{m}{Dp}$. Therefore, $U_E(\text{pay}) = -2m$. $U_E(\text{no pay}) = \mu(\text{demand})p(-D-m) + (1-\mu(\text{demand})p)(0-Dp\mu(\text{no violence}))$, where $\mu(\text{no violence}) = \frac{\mu(1-p)}{\mu(1-p)+(1-\mu)} < \frac{m}{Dp}$. Hence, $U_E(\text{no pay}) = \mu p(-D-m) + (1-\mu p)(0-Dp\frac{\mu(1-p)}{\mu(1-p)+(1-\mu)}) = -(2D+m-Dp)\mu p$. Hence, E should pay if $\mu(\text{demand}) \geq (2\frac{m}{p(2D+m-Dp)})$. Hence, randomizing is a best response for E.

$$\text{Therefore, } \sigma_F^{t-1} = x = \frac{\theta(p-2)(m-pD)}{2m(1-\theta)}.$$

Next consider F's decision to punish: $U_F(\text{punish}) = -R + m - r$ and $U_F(\text{no punish}) \geq -r + 0$ (F's payoff in the second period is 0 since $\mu(\text{no violence}) < \frac{m}{Dp}$, region 3). If $R \geq m$ then F prefers not to punish.

Does M punish when the police presence is high? $U_M(\text{punish|high}) = -1 + m$ and $U_M(\text{no punish|high}) = 0 + r + p - \frac{pr}{m}$ ($r + p - \frac{pr}{m}$ is M's expected payoff in the 2nd period for region 2). Thus, if $p \geq (\frac{m}{r-m})(r+1-m)$ then M does not punish during a high police presence.

Finally, consider the initial decision to demand money. $U_F(\text{demand}) = -r + \sigma_E^{t-1}(m-r+m)$ and $U_F(\text{no demand}) = 0$. Thus, providing $\sigma_E^{t-1} = \frac{r}{(-r+2m)}$ randomizing in the first period is optimal. QED.

Equilibrium IIa

If $p \leq m-1$, $R \geq m$ and if either (i) $\theta \geq \frac{2m}{m+D}$ and $\theta \geq \frac{m}{Dp}$, or (ii) $\theta \geq \frac{m}{(D+m-Dp)}$ and $\theta < \frac{m}{Dp}$ then the strategy profile $\sigma_F^{t-1} = 1, \sigma_M^{t-1}$

$= 1, \sigma_E^{t=1} = 1, s_F^{t=1} = 0, s_M^{t=1}(\text{high}) = 1, s_M^{t=1}(\text{low}) = 1$ is a sequential equilibrium with beliefs $\mu(\text{demand}) = \theta, \mu(\text{no demand}) = 0, \mu(\text{violence}) = 1,$ and $\mu(\text{no violence}) = 0.$

Proof: Note that the beliefs are all consistent with Bayes' rule. Therefore, we simply need to check that every player is utility maximizing given their type, their beliefs, and the strategy of the other players.

Consider E's decision to pay in the first period. There are two cases (i) $\theta \geq \frac{m}{Dp}$. In this case, if E pays in the first period he would also pay in the second period. $U_E(\text{pay}) = -2m.$ If E refuses to pay then $U_E(\text{no pay}) = \mu(-D - m) + (1 - \mu)(0 + 0) = \mu(-D - m),$ where $\mu = \mu(\text{demand}) = \theta.$ If $\theta \geq \frac{2m}{m+D}$ then E should always pay: $\sigma_E^{t=1} = 1.$

(ii) $\theta < \frac{m}{Dp}$. In this case, E would be in region 3 in the second period if he pays in the first. Hence E's second period payoff is $-Dp\mu(\text{demand}) = -Dp\theta.$ Therefore, $U_E(\text{pay}) = -m - Dp\theta,$ and $U_E(\text{no pay}) = \mu(-D - m) + (1 - \mu)(0 + 0) = \theta(-D - m).$ Therefore, E should pay if $\theta \geq \frac{m}{(D+m-Dp)}.$ Note that lines $\theta = \frac{m}{(D+m-Dp)},$ $\theta = \frac{m}{pD},$ and $\theta = \frac{2m}{m+D}$ all intersect at the same point: $p = \frac{m+D}{2D}$ and $\theta = \frac{2m}{m+D}.$ Therefore, E should pay if either (i) $\theta \geq \frac{2m}{m+D}$ and $\theta \geq \frac{m}{Dp}$ or (ii) $\theta \geq \frac{m}{(D+m-Dp)}$ and $\theta < \frac{m}{pD}.$

Next consider F's decision to punish: $U_F(\text{punish}) = -R + m - r$ and $U_F(\text{no punish}) \geq -r + 0$ (F's payoff in the second period is 0 since $\mu(\text{no violence}) = 0 < \frac{m}{pD},$ region 3). If $R \geq m$ then F prefers not to punish.

Does M punish when the police presence is high? $U_M(\text{punish|high}) = -1 + m$ and $U_M(\text{no punish|high}) = 0 + p$ (p is M's expected payoff in the 2nd period since E's beliefs $\mu(\text{no violence})$ mean that E never pays). Thus, if $p \leq m - 1$ then M punishes during a high police presence.

Finally, consider the initial decision to demand money. $U_F(\text{demand}) = m - r + \zeta,$ where ζ is $m - r$ if $\theta \geq \frac{m}{pD}$ (region 2) and $\zeta = 0$ else (region 3). $U_F(\text{no demand}) = 0.$ Therefore, F always demands money. QED.

Equilibrium IIb

If $p \leq m - 1, R \geq m$ and $\theta < \min \left\{ \frac{2m}{m+D}, \frac{m}{m+D-Dp} \right\}$ then the strategy profile $\sigma_M^{t=1} = 1,$

$$s_F^{t=1} = 0, s_M^{t=1}(\text{high}) = 1, s_M^{t=1}(\text{low}) = 1, \sigma_F^{t=1} =$$

$$x = \begin{cases} \frac{\theta(D-m)}{2m(1-\theta)} & \text{if } p \geq \frac{m+D}{2D} \\ \frac{\theta D(1-p)}{m(1-\theta)} & \text{if } p \leq \frac{m+D}{2D} \end{cases}, \text{ and}$$

$$\sigma_E^{t=1} = \begin{cases} \frac{r}{2m-r} & \text{if } p \geq \frac{m+D}{2D} \\ \frac{r}{m} & \text{if } p < \frac{m+D}{2D} \end{cases} \text{ is a sequential equilibrium with}$$

beliefs $\mu(\text{demand}) = \frac{\theta}{\theta + (1-\theta)x} = \frac{2m}{m+D}$, $\mu(\text{no demand}) = 0$, $\mu(\text{violence}) = 1$, and $\mu(\text{no violence}) = 0$.

Proof: Note that the beliefs are all consistent with Bayes' rule. Therefore, we simply need to check that every player is utility maximizing given their type, their beliefs and the strategy of the other players.

Consider E's decision to pay in the first period. There are two cases: (i) $p \geq \frac{m+D}{2D}$. In this case $U_E(\text{pay}) = -2m$ if $\mu(\text{demand}) \geq \frac{m}{pD}$ and $U_E(\text{pay}) = -m - pD\mu(\text{demand})$ if $\mu(\text{demand}) < \frac{m}{pD}$. If E refuses to pay then $U_E(\text{no pay}) = \mu(-D - m) + (1 - \mu)(0 + 0) = \mu(-D - m)$, where $\mu = \mu(\text{demand}) = \frac{\theta}{\theta + (1-\theta)x}$. If $\mu(\text{demand}) = \frac{2m}{D+m}$ and hence $x = \frac{\theta(D-m)}{2m(1-\theta)}$ then E is indifference pay: $\sigma_E^{t=1} \in [0,1]$. (Note that we require that $\theta < \frac{2m}{m+D}$ in order that $x < 1$.)

(ii) $p < \frac{m+D}{2D}$. In this case $U_E(\text{pay}) = -2m$ if $\mu(\text{demand}) \geq \frac{m}{pD}$ and $U_E(\text{pay}) = -m - pD\mu(\text{demand})$ if $\mu(\text{demand}) < \frac{m}{pD}$. If E refuses to pay then $U_E(\text{no pay}) = \mu(-D - m) + (1 - \mu)(0 + 0) = \mu(-D - m)$, where $\mu = \mu(\text{demand}) = \frac{\theta}{\theta + (1-\theta)x}$. First note that if $\mu(\text{demand}) \geq \frac{m}{pD}$ then E strictly prefers to pay. But then all types of thugs demand money, which is a contradiction. Therefore, $\mu(\text{demand}) < \frac{m}{pD}$. E is indifferent between paying and not if $\mu(\text{demand}) = \frac{m}{m+D-Dp} = \frac{\theta}{\theta + (1-\theta)x}$. Thus, if $\sigma_F^{t=1} = x = \frac{\theta D(1-p)}{m(1-\theta)}$ then E randomizes its choice.

Next consider F's decision to punish: $U_F(\text{punish}) = -R + m - r$ and $U_F(\text{no punish}) \geq -r + 0$. If $R \geq m$ then F prefers not to punish.

Does M punish when the police presence is high? $U_M(\text{punish}|\text{high})$

= -1 + m and $U_M(\text{no punish}|\text{high}) = 0 + p$ (p is M's expected payoff in the 2nd period since E's beliefs, $\mu(\text{no violence})$, mean that E never pays). Thus, if $p \leq m - 1$ then M punishes during a high police presence.

Finally, consider the initial decision to demand money.

Case (i) $p \geq \frac{m+D}{2D}$ and $\theta < \frac{2m}{m+D}$. In order that F randomizes, F must be indifferent. $U_F(\text{demand}) = -r + \sigma_E^{t=1}(m + m - r)$ and $U_F(\text{no demand}) = 0$. Therefore, if $\sigma_E^{t=1} = \frac{r}{(2m-r)}$ then F is indifferent and randomizing is optimal.

Case (ii) $p \leq \frac{m+D}{2D}$ and $\theta < \frac{m}{m+D-Dp}$. In order that F randomizes, F must be indifferent. $U_F(\text{demand}) = -r + \sigma_E^{t=1}(m)$ and $U_F(\text{no demand}) = 0$. Therefore, if $\sigma_E^{t=1} = \frac{r}{m}$ then F is indifferent and randomizing is optimal. QED.

Equilibrium IIIa

If $R \leq m, p \leq m - 1$ and $\theta \geq \frac{m^2}{pD^2}$ then the strategy profile $\sigma_F^{t=1} = 1, \sigma_M^{t=1} = 1, \sigma_E^{t=1} = 1,$

$$s_F^{t=1} = \begin{cases} 1 & \text{if } \theta \geq \frac{m}{pD} \\ y = \theta \frac{-m + Dp}{m(1 - \theta)} & \text{if } \theta < \frac{m}{pD} \end{cases}, s_M^{t=1}(\text{high}) = 1, s_M^{t=1}(\text{low}) = 1, \text{ and}$$

$$\sigma_E^{t=2}(\text{violence}) \begin{cases} 1 & \text{if } \theta \geq \frac{m}{pD} \\ \frac{R}{m} & \text{if } \theta < \frac{m}{pD} \end{cases} \text{ is a sequential equilibrium with}$$

beliefs $\mu(\text{demand}) = \theta, \mu(\text{no demand}) = 0, \mu(\text{violence}) =$

$$\begin{cases} \theta & \text{if } \theta \geq \frac{m}{pD} \\ \frac{m}{pD} & \text{if } \theta < \frac{m}{pD} \end{cases}, \text{ and } \mu(\text{no violence}) = 0.$$

Proof: Note that the beliefs are all consistent with Bayes' rule. Therefore, we simply need to check that every player is utility maximizing given their type, their beliefs, and the strategy of the other players.

Consider E's decision to pay in the first period. There are two cases: (i) if $\theta \geq \frac{m}{pD}$ then $U_E(\text{pay}) = -2m$ (E would also pay in the second period given these beliefs). Since all types punish, E learns nothing from punishment and so also pays in the second period. If E refuses to pay then $U_E(\text{no pay}) = -D - m$. Therefore, $\sigma_E^{t=1} = 1$.

(ii) if $\theta < \frac{m}{pD}$ then $U_E(\text{pay}) = -m - \mu(\text{demand})pD$. If E refuses to pay then $U_E(\text{no pay}) = -D - m + (1 - \mu(\text{demand}))(1 - y)(m + D)$. Since $y = \theta \frac{-m + Dp}{m(1 - \theta)}$, $U_E(\text{no pay}) = -D - m + (1 - \theta)(1 - \theta \frac{-m + Dp}{m(1 - \theta)})(m + D) \leq U_E(\text{pay})$ if $\theta \geq \frac{m^2}{pD^2}$.

Therefore, $\sigma_E^{t=1} = 1$ is a best response.

Next consider F's decision to punish. There are two cases: (i) $\theta \geq \frac{m}{pD}$. $U_F(\text{punish}) = -R + m - r$ and $U_F(\text{no punish}) \geq -r + 0$. If $R \leq m$ then F prefers to punish.

(ii) $\theta < \frac{m}{pD}$. $U_F(\text{punish}) = -R + \sigma_E^{t=2}(\text{violence})m - r$ and $U_F(\text{no punish}) \geq -r + 0$. Since $\sigma_E^{t=2}(\text{violence}) = \frac{R}{m}$ then F is indifferent and $s_F^{t=1} = \theta \frac{-m + Dp}{m(1 - \theta)}$ is a best response. Note that given $s_F^{t=1} = \theta \frac{-m + Dp}{m(1 - \theta)}$, $\mu(\text{violence}) = \frac{m}{pD}$ and E is indifferent about paying in the second period.

Does M punish when the police presence is high? There are two cases: (i) $\theta \geq \frac{m}{pD}$. $U_M(\text{punish|high}) = -1 + m$ and $U_M(\text{no punish|high}) = 0 + p$ (p is M's expected payoff in the 2nd period since E's beliefs $\mu(\text{no violence})$ mean that E never pays). Thus, if $p \leq m - 1$ then M punishes during a high police presence.

(ii) $\theta < \frac{m}{pD}$. $U_M(\text{punish|high}) = -1 + \sigma_E^{t=2}(\text{violence})m + (1 - \sigma_E^{t=2}(\text{violence}))p = -1 + \frac{R}{m}m + (1 - \frac{R}{m})p = \frac{-m + Rm + pm - pR}{m}$ and $U_M(\text{no punish|high}) = 0 + p$. Since $p \leq \frac{m}{R}(R - 1)$ then punishing is a best response.

The punishment decision depends upon whether or not $\theta \geq \frac{m}{pD}$. However, the equilibria are operationally equivalent since E always pays.

Finally, consider the initial decision to demand money. Since E always pays it is always optimal to demand money. QED.

Equilibrium IIIb

If $R \leq m$, $p \leq m - 1$ and $\theta < \frac{m^2}{pD^2}$ then the strategy profile $\sigma_F^{t=1} = x = \theta \frac{pD^2 - m^2}{m^2(1 - \theta)}$, $\sigma_M^{t=1} = 1$, $\sigma_E^{t=1} = \frac{r}{m}$, $\sigma_E^{t=2}(\text{violence}) = \frac{R}{m}$, $s_F^{t=1} = y = \frac{-m + Dp}{m(1 - \theta)}$, $s_M^{t=1}(\text{high}) = 1$, $s_M^{t=1}(\text{low}) = 1$ is a sequential equilibrium

with beliefs $\mu(\text{demand}) = \frac{\theta}{\theta + (1 - \theta)x} = \frac{m^2}{pD^2}$, $\mu(\text{no demand}) = 0$, $\mu(\text{violence}) = \frac{\mu}{\mu + (1 - \mu)y} = \frac{m}{pD}$ (where $\mu = \mu(\text{demand})$), and $\mu(\text{no violence}) = 0$.

Proof: Note that the beliefs are all consistent with Bayes' rule. Therefore, we simply need to check that every player is utility maximizing given their type, their beliefs, and the strategy of the other players.

Consider E's decision to pay in the first period. $U_E(\text{pay}) = -m - \mu(\text{demand})Dp$, since $\mu(\text{demand}) = \frac{\theta}{\theta + (1 - \theta)x} < \frac{m}{pD}$. E plays the semi-pooling equilibrium in the second period. If E refuses to pay then $U_E(\text{no pay}) = (-D - m)(\mu + (1 - \mu)y) + (1 - \mu)(1 - y)0$, where $\mu = \mu(\text{demand}) = \frac{\theta}{\theta + (1 - \theta)x}$. Thus, if $-m - \mu Dp = (-D - m)(\mu + (1 - \mu)y)$ then $y = \frac{-\mu D - \mu m + m + Dp\mu}{(1 - \mu)(D + m)}$ makes E indifferent: $\sigma_E^{t=1} \in [0, 1]$.

Next consider F's decision to punish: $U_F(\text{punish}) = -R + \sigma_E^{t=2}(\text{violence})m - r$ and $U_F(\text{no punish}) = -r + 0$. If $\sigma_E^{t=2}(\text{violence}) = 1$ then F always punishes; if $\sigma_E^{t=2}(\text{violence}) = 0$ then F never punishes; and if $\sigma_E^{t=2}(\text{violence}) = \frac{R}{m}$ then F is indifferent. In order that E randomizes in the second period we require that $\mu(\text{violence}) = \frac{m}{pD} = \frac{\mu}{\mu + (1 - \mu)y}$. Hence, $y = \mu \frac{pD - m}{m(1 - \mu)} = \frac{-\mu D - \mu m + m + Dp\mu}{(1 - \mu)(D + m)}$. This implies that $\mu(\text{demand}) = \frac{m^2}{pD^2} = \frac{\theta}{\theta + (1 - \theta)x}$ and hence that $x = \theta \frac{pD^2 - m^2}{m^2(1 - \theta)}$ and $y = m \frac{-m + Dp}{pD^2 - m^2}$.

Does M punish when the police presence is high? $U_M(\text{punish}|\text{high}) = -1 + \sigma_E^{t=2}m + (1 - \sigma_E^{t=2})p = -1 + R + (1 - \frac{R}{m})p$ and $U_M(\text{no punish}|\text{high}) = 0 + p$. Thus, if $p \leq \frac{Rm - m}{R} = \frac{m}{R}(R - 1) > 1$ then M punishes during a high police presence.

Finally, consider F's decision to demand money. $U_F(\text{demand}) = \sigma_E^{t=1}(m - r + 0) + (1 - \sigma_E^{t=1})(-R + \sigma_E^{t=2}(\text{violence})m - r)$. $U_F(\text{no demand}) = 0$. Therefore, if F randomizes his demand decision, $\sigma_E^{t=1} = \frac{r}{m}$, F is indifferent and randomizing is a best response. QED.