Final version published in Journal of Law, Economics & Organization, 22 315-339

Gangs and Crime Deterrence

Abdala Mansour
CIRPÉE

Nicolas Marceau
Université du Québec à Montréal and CIRPÉE

Steeve Mongrain*
Simon Fraser University, RIIM, and CIRPÉE

A framework is developed in which the formation of gangs—the criminal market structure—is endogenous. As in standard models of crime, under a given gang structure, an increase in deterrence reduces criminal output. However, under identifiable circumstances, an increase in deterrence can lead to an increase in the number of competing criminal gangs and to an increase in total illegal output, possibly accompanied with a fall in the price. We show that an increase in demand can also modify the criminal market structure and can ultimately affect the output and the price in a similar way.

1. Introduction

Many governments have intensified their war on organized drug traffickers in the last decades, but those efforts appear to have been unsuccessful. This article offers an explanation for the ineffectiveness of these policies. Endogenizing the formation of gangs in a criminal market, we argue that an increase in deterrence can make the market more competitive and can lead to a higher output and lower prices.

Lee (1993) and Poret (2003) confirm the fact that in the United States, arrests and penalties for heroin and cocaine (including crack) drug violations rose significantly during the last twenty years. The U.S. Drug Enforcement Administration (DEA) saw its budget increase from 0.07 billion US$ to

*Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, BC, Canada, V5A 1S6. Email: mongrain@sfu.ca.

We are grateful to the three referees and to Vicky Barham, Claude Fluet, Anke Kessler, Steven Levitt, Gordon Myers, Stéphane Pallage, Krishna Pendakur, Sylvaine Poret, Said Souam, Marc Van Audenrode, and Jean-François Wen for their comments. We also thank seminar participants at Carleton University, Concordia University, McMaster University, Université de Sherbrooke, University of British Columbia, University of Maine, University of North Carolina, University of Washington and at the following conferences: Canadian Public Economics Study Group 2001, Journées du CIRPÉE 2001, and Journées de micro-économie appliquée 2001. Financial support from FCAR/FQRSC, RIIM, and SSHRCC is gratefully acknowledged.

1.55 billion US$ between 1973 and 2000, while the number of its special agents went from 1,470 to 4,561 over the same period.[1] The number of arrests by the DEA almost doubled during the 1990s, and the total value of drug seizures increased from 14 million US$ to 82 million US$ since 1986. Note that even if the DEA accounts for only part of the enforcement effort in the United States, similar trends can also be observed for other actors involved in drug enforcement. For example, Kuziemko and Levitt (2001) present evidence showing that in the United States, imprisonment for drug-related activities considerably increased during this period, while it had been relatively stable for other types of crime. In fact, over 80% of the increase in the federal prison population from 1985 to 1995 was due to drug convictions.[2]

More recently, governments have attempted to stiffen sanctions for organized drug traffickers. In the United States, the Foreign Narcotics Kingpin Designation Act (Kingpin Act) was signed by the president in December 1999 and became effective in June 2000. The purpose of the Kingpin Act is to deny access to the U.S. financial system to significant (i.e., large) foreign narcotics traffickers. The Kingpin Act provides for very stiff criminal penalties for individuals or businesses conducting financial or business transactions for designated foreign narcotics traffickers. In Canada, Bill C-24, modifying the Canadian criminal code to account for organized crime, was adopted in December 2001. Under Bill C-24, an individual who is simply a member of a criminal organization can be prosecuted and sanctioned.

Taken together, these facts lead to the conclusion that from the early 1980s to recent times:

- Efforts to detect criminal organizations involved in drug trafficking were intensified; and
- Effective sanctions for drug traffickers were stiffened, in particular for those belonging to large organizations (Kingpin Act).

While there was an increase in deterrence against drug consumers and producers, the market for drugs, far from shrinking, seems to have expanded:[3]

- *Estimated drug quantities:* Studies by Kuziemko and Levitt (2001) and by Everingham, Rydell, and Caulkins (1995) suggest that the consumption of cocaine in the United States increased significantly between 1980 and 1992. Possible estimates suggest that consumption of cocaine jumped from around 275 metric tons to around 450 metric tons over that period. In the 1990s, total cocaine production in Colombia, representing some 75% of world production, increased by a factor of two. However, a large share of this increase was due to the reallocation of cocaine production from

1. This information is available on the Web site of the U.S. DEA.
2. U.S. Bureau of Justice Statistics, Prisoners in 1996 (Washington, DC: U.S. Department of Justice, 1997).
3. Except where we refer to a specific source, the information reported here was obtained from the Web site of the DEA.

Bolivia and Peru to Colombia. Nevertheless, some DEA estimates show an increase of up to 20% in the potential cocaine production in the Andean region between 1991 to 2001, with most of the increase occurring in the early years of the period. Opium production in the western hemisphere, which constitutes about 90% of the U.S. heroin market, also increased over the same period. Over the last ten years, heroin-related deaths rose from 28 to 206 per year in Florida. In cities like Portland or Seattle, heroin overdose is now more important as a cause of death than murder. Since 1995, the estimated production of Canadian cannabis has tripled. Given the nature of the illegal sector in which drugs are traded, estimated quantities could be challenged. However, overall it would be extremely difficult to conclude that there was a decrease in production/consumption of different drugs from the start of the 1980s to the mid-1990s.

- *Prices:* While it may be difficult to estimate precisely the quantities of the various illegal drugs in a market, information on prices is known to be more reliable. This is because undercover agents can keep a record of the price of each transaction they undertake. According to Caulkins and Reuter (1998), the price of cocaine in the United States decreased from 500 US$ to 100 US$ per gram between 1980 and 1996, and the price of heroin has also been deflated by a factor of five.

To summarize, the increase in deterrence of the last decades was accompanied by an increase in output and a decrease in price for most illegal drugs. This article attempts to explain the simultaneous occurrence of these apparently irreconciliable facts.[4]

Our explanation is that deterrence affects the structure of the market for drugs. Our starting point is the assumption that the production and distribution of illegal drugs is controlled by well-organized criminal organizations.[5] As

4. In Figures 1 and 2 of their article, Kuziemko and Levitt (2001) present evidence on the evolution, for the last 20–30 years, of the number of arrests and state prison commitments (a proxy for the extent of deterrence), and of the price and estimated consumption of cocaine. These figures show very clearly that indeed, deterrence increased and cocaine output increased, while cocaine price went down. However, Kuziemko and Levitt argue that had it not been for the increase in deterrence, cocaine price would have declined even more, due to other factors such as increase in productivity in drug production and distribution. Our analysis states that for a given market structure, deterrence causes illegal output to fall and the price to increase, but that when accounting for possible changes in market structure, deterrence can cause the output to increase and the price to fall. The results of this study are perfectly consistent with Kuziemko and Levitt's if a change in market structure is interpreted as a factor contributing to a fall in price. Similarly to Kuziemko and Levitt's findings, an increase in deterrence can offset part of the decrease in price due to a change in market structure.

5. Some well-known examples of such organizations include the Hell's Angels and other motorcycle gangs for methamphetamine in North America, and the Medellin Colombian cartel for cocaine. No evidence can, however, confirm the degree to which those cartels are (were) successful at controlling prices. On the other hand, Levitt and Venkatesh (2000) argue that drug-selling gangs enjoy some monopoly power over their turf. Finally, as is surveyed in Poret (2003), it is standard in the literature to model criminal organizations as enjoying some market power.

would be expected, for a given market structure (i.e., for a given number of criminal organizations), an increase in deterrence will reduce total output and increase the price of drugs. However, it may not be so when the market structure reacts to deterrence. To endogenize the market structure, we use the theory of coalition-formation. Then, starting from a cartel, we show how an increase in deterrence can lead to a splintering of the cartel. From there, we simply make our own the observation of Buchanan (1973):[6] when more firms operate in an illegal good market (i.e., when the market is more competitive), the output is larger and the price is lower. Thus, by increasing the number of criminal organizations in the market, increased deterrence leads to an increase in output and to a fall in prices.[7]

Our story bears some resemblance to the events that took place in Colombia after increased deterrence efforts eventually lead to the dissolution of the Cali and Medellin drug cartels. Interestingly, after the dissolution of the cartels, the number of criminal organizations involved in the production of cocaine increased, and this was eventually accompanied by an increase in total production. With between 80 and 250 organizations involved in a market that was previously cartelized, it could be argued that the increase in production was due simply to the market being made more competitive (*The Economist*, September 11, 1999).

Of course, there are alternative explanations to the simultaneous occurrence of the above phenomena. One of them is that there simply was an increase in the demand for drugs in the last decades. But it is hard to reconcile the observed fall in prices with an increase in demand unless the market structure is endogenous.[8] Our model, in which market structure is endogenous, can generate a fall in prices following an increase in demand. Indeed, when demand increases, there is more of an incentive to exit the cartel for any of its members. And, as was pointed out above, the splintering of the cartel can lead to an increase in output and to a fall in prices.[9]

6. See also Neher (1978), in which organized criminal organizations attempt to extort money from some individual. This individual is viewed as a common pool of resources. Thus, and as in the tragedy of the commons, too much extortion takes place under competition, but a lower and efficient level of extortion is undertaken under a monopoly structure. Note that as in Buchanan (1973), Neher (1978) does not attempt to endogenize the market structure.

7. Clearly, it would be possible to design optimal deterrence policies taking this effect into account. The current article does not do so.

8. As pointed out in Kleiman (1993), greater sales can dilute enforcement (enforcement swamping) and consequently can create economies of scale at the industry level, leading to a downward-sloping supply curve. Given the fact that enforcement has increased, the effect of enforcement swamping would need to be significant for the price to go down after an increase in demand. Another potential reason for industry-level economies of scale would have to do with the network externalities in drug distribution, but such a story is not inconsistent with changes in the market structure.

9. Of course, prices could also increase. This will depend on the various elasticities and on the size of the shift in demand. Also note that other market structures could generate the observed phenomena. For example, an incumbent firm with market power could reduce its price to deter potential entrants.

There is a small literature that has tried to explain the puzzling inefficiency of the “war on drugs” in the 1980s. Among the possible explanations, Skott and Jepsen (2002) propose one based on switching costs. An increase in enforcement leads to higher switching costs, which in turn leads to more incentives for a drug seller who has market power to invest in the “marketing” of its product. Such an argument explains why consumption can increase in response to an increase in enforcement, but it cannot explain the observed reduction in prices, and therefore can resolve the puzzle only partially. Caulkins (1993) highlights the fact that zero-tolerance policies reduce marginal deterrence and can lead to an overall increase in drug consumption, but again this article restricts its attention solely to the quantity-related part of the puzzle. Poret (2003) shows that such a counterintuitive outcome can be replicated in a model with wholesalers (traffickers) and resellers when enforcement is focused on the wholesaler.[10] However, since the number of wholesalers is fixed, her argument cannot account for the apparent increase in the number of traffickers. Instead of focusing on the potentially counterintuitive effect of enforcement on consumption and prices, other studies focused on the potential impact of enforcement on the associated cost related to illicit consumption. Miron (1999) shows that enforcement of drug and alcohol prohibition raised the homicide rate. Property crime, as highlighted by Benson and Rasmussen (1991) or by Benson, Kim, Rasmussen, and Zuehlke (1992), and violence, as highlighted by Burrus (1999), can also rise with harsher enforcement. In their survey, Miron and Zwiebel (1995) argued again that more severe enforcement of prohibition can raise violence and property crime. They also argued that drug-related problems (overdose and addiction) can be made more severe because drug concentration may rise in response to stricter enforcement.

Other contributions are not directly related to our work but support the assumption that the drug market is somewhat non-competitive. Fiorentini and Peltzman (1995) have described environments in which organized crime is likely to flourish. First, they claim that organized crime is more likely to be prevalent when there are economies of scale and monopolistic power in the supply of some illegal good. In this article, we assume that gangs do exercise some market power, and so the environment we consider has the required feature. According to Fiorentini and Peltzman, another standard feature of organized crime is that gangs often exercise violence against other firms of the legal and illegal sectors. For example, Gambetta and Reuter (1995) show that criminal organizations can use violence to maintain their market power. In our model, we do not introduce violence explicitly. We note that the relationship between gang structure and violence is likely to be ambiguous. Indeed, an increase in the number of gangs means that more suppliers are sharing a given demand, and such an intensification in competition may lead to an increase in violence. On the other hand, a small number of gangs is likely to translate into larger rents for each organization and, because there is more at stake, violence may increase. Probably the most convincing argument for the presence of

10. Caulkins and Padman (1993) study a similar problem, but in an export-import context.

market power in the illicit drug sector was provided by Levitt and Venkatesh (2000), who used actual data on a Chicago gang to show the presence of market power.

This article is organized as follows. In the next section, we present an overview of the model and briefly discuss each of the two stages (gang formation and gang competition) of our overall game. The gang formation stage is the subject of Section 3. At this stage, individuals align themselves into gangs, anticipating the payoffs from various gang structures. In Section 4, we analyze the gang competition stage in which, given a gang structure, gangs compete in an illegal goods market. The equilibrium of the game as well as the impact of deterrence and of a change in demand are discussed in Section 5. In Section 6, we discuss how our results would change if we were to modify our assumption regarding the probability that gang activities are detected. We introduce the possibility for this probability to depend on the market share of a gang. We conclude in Section 7.

2. An Overview of the Model

Consider the following economy inhabited by two types of private agents and an authority (e.g., a government). First, there is a large number of potential consumers for a homogenous good that can only be purchased illegally. For simplicity, we assume that the aggregate inverse demand for this good is linear: $P = \beta - \gamma Q$, where Q is the quantity of the illegal good, P is its price, and $\beta, \gamma > 0$ are parameters. Note that any downward-sloping demand would generate qualitatively equivalent results. It is assumed that no individual can be arrested and sanctioned for consuming the good.[11] Second, there are three identical risk-neutral criminals,[12] denoted A, B, and C, which are to operate gangs in what follows. We denote by $E = \{A, B, C\}$ the set of criminals. Criminals will set up their businesses (gangs) to maximize their expected income. Note that using three criminals, rather than some larger arbitrary number of them, is mainly for ease of exposition. Indeed, as is shown in Appendix B, our main result below (i.e., that market structure reacts to changes in deterrence and in demand) would hold for any number of criminals.[13]

A criminal has two important decisions to make. The first decision, taken during what we call the *gang formation stage*, is to determine with whom he wants to operate a gang. We assume that criminals are members of one and only one gang. A *gang* is defined as a nonempty subset of E denoted G_j. A *gang structure* is defined as a partition of E and is denoted H, and the set of all

11. If we relaxed this assumption, the parameters of the demand would be affected, but not our results.

12. Risk neutrality is assumed for simplicity and is probably the most reasonable assumption. Indeed, risk neutrality is supported by Levitt and Venkatesh (2000), as they find that compensations (wages) received by Chicago gang members are low, despite the 7% annual death rates they face.

13. However, there are also some results, for example the precise path leading to the splintering of a large monopoly gang in Proposition 2, which may be maintained only by adding further restrictions on the parameters of the model.

possible gang structures is denoted $\boldsymbol{H}$. The process by which gangs merge or divide into smaller gangs acts in a similar fashion as entry and exit would. We choose to model the process this way because, as was highlighted by Levitt and Venkatesh (2000), takeover is an important aspect of the industry. Separation from an existing gang is at least as plausible as entry, because of the importance of specific human capital or on-hand experience in this industry.

In the second stage, the *gang competition stage*, the gangs will operate and make profits that will accrue to the criminals. We assume that the decisions of gang G_j are made cooperatively by its criminal members to maximize its total profits, $\Pi_{G_j}(H)$, from selling the illegal good. Note that profits depend the gang structure, which reflects the fact that other gangs affect the profits of gang G_j. We assume that the play across gangs is non-cooperative. Once the profits of gang G_j are realized, they are divided equally between its criminal members. We denote by $\pi_i(H) = \Pi_{G_j}(H)/|G_j|$, $i \in G_j$, the expected income of criminal i if he is a member of gang G_j, possibly with other criminals, the total number of them in gang G_j being given by $|G_j|$. Thus, $\pi_i(H)$ is the ultimate payoff of individual i, realized in the second stage, if gang structure H emerges in the first stage. Looking ahead from the first stage to the second stage, the three criminals will have a set of preferences (payoffs) over all possible gang structures, $\pi_i(H)$ for all $H \in \boldsymbol{H}$. Based on these preferences, criminals form gangs in the gang formation stage, which leads to an equilibrium gang structure, say, H^*, and thus to an equilibrium payoff for each player i, $\pi_i(H^*)$.

Once the gang structure has been determined, the gang competition stage begins, during which the gang members must decide how much to produce. It is assumed that to produce, a gang must pay a fixed cost F.[14] The output of gang G_j is denoted by n_j.

The authority attempts to reduce gangs' criminal activity by sanctioning the gang members it identifies. To identify those members, it invests α in detection. Let $p(\alpha, n)$ be the probability that the members of a gang will be detected, where in the case of a gang with multiple members, all of them are detected once one of them is. It is natural to assume that more effort α by the authority translates into a larger probability of detection, so $p_\alpha \geq 0$. It could also be argued that the probability of detection of a gang depends on its output, but this relationship could well be positive (larger gangs are easier to infiltrate) or negative (larger gangs have more at stake and invest more in avoidance activities), so for now, $p_n \gtreqless 0$.

14. According to Levitt and Venkatesh (2000), the greatest non-wage expenditure that accounts for 20% of the total revenue takes the form of tribute payments to central gang leadership. They mention in their article that street gangs pay this due and are residual claimants over the profits. Corruption of the authorities and setting up routes for the transportation of drugs are other forms of fixed costs that producers outside the United States face. Note that the scale economies generated by this fixed cost are not necessary, but allowed us to simplify the analysis. In order to exercise market power, criminals have an incentive to form a gang as large as possible. The fixed cost simply reinforces this incentive. It turns out that given the functional forms we have chosen, a positive fixed cost is required for the existence of the grand monopoly gang. However, a fixed cost would not be required if other functional forms were used.

When a gang is detected, each of its members is imposed an individual effective sanction $\hat{s}(\cdot)$. We assume that this effective sanction[15] depends on statutory sanction s and, possibly, on the output of the gang n, so $\hat{s}(s, n)$. In what follows, we assume that a larger statutory sanction translates into a larger effective sanction, so $\hat{s}_s > 0$. We also allow for the possibility that a criminal individual effective sanction, is increasing in the output of the gang to which this criminal belongs, so $\hat{s}_n \geq 0$. One could argue that the case where $\hat{s}_n > 0$ reflects laws similar to the Kingpin Act.

Because criminals are risk-neutral, we can restrict our attention to the expected punishment they face. Denote the expected punishment by $z(\cdot)$ and notice that $z(\alpha, s, n) = p(\alpha, n)\hat{s}(s, n)$. Given the assumptions above, expected punishment should be an increasing function of detection effort α and of the statutory sanction s, and a function of ambiguous sign of output n. To simplify our analysis in what follows, we assume a precise functional form: $z(\alpha, s, n) = k\alpha sn$, where $k > 0$ is a parameter. Note that in this functional form, expected punishment is increasing in output ($z_n = k\alpha s > 0$). We think this reflects the policies that were put in place from the early 1980s until recently, in particular the stiffening of effective sanctions for criminals involved in larger gangs (Kingpin Act). Note that below, we will consider the impact of increases in detection effort and in the statutory sanction. It should be clear by now that from the point of view of criminals, both are equivalent, as both lead to an increase in expected punishment.

We now proceed with the analysis of the two stages, starting with the first, the gang formation stage.

3. The Gang Formation Stage

Starting with the set of criminals E, we model how these players might choose to align themselves into gangs. We use the coalition formation approach of Burbidge et al. (1997), which was itself based on that of Hart and Kurz (1983) (see also Marceau and Myers, forthcoming).

Anticipating the gang competition stage, the players know $\pi_i(H)$ for all $H \in \boldsymbol{H}$, so they have a preference ordering over all possible gang structures. We use these preference orderings to construct a game in strategic form for this stage. We suppose that each criminal formulates a plan for joining partners to form a gang. A *strategy* of player i is a partnership plan in which i announces the gang to which he wants to belong. Formally, a strategy for player i is a subset of E, or G_i, with $i \in G_i$. A combination of one strategy for each player $g = (G_A, G_B, G_C)$ is a *strategy profile*. The set of all strategies for player i is denoted by $\boldsymbol{G_i}$, and $\boldsymbol{G} = \boldsymbol{G_A} \times \boldsymbol{G_B} \times \boldsymbol{G_C}$ will stand for the set of all strategy profiles.

15. There are many possible interpretations of what we are here labeling a sanction. As is standard, a sanction could be some amount of money to be paid or some non-monetary cost (i.e., pure utility loss) imposed on the members by the authority after detection. It could also be the opportunity cost of an illegal transaction that did not take place because of detection.

How a strategy profile $g \in \boldsymbol{G}$ is reconciled into a resultant gang structure is summarized by a function, $\psi{:}\boldsymbol{G} \rightarrow \boldsymbol{H}$, called the *coalition structure rule*, which assigns to any $g \in \boldsymbol{G}$ a unique gang structure $H = \psi(g)$. Several such rules exist, which are discussed in Burbidge et al. (1997). But for our purposes, we have selected the rule labeled the *similarity rule*. To save on notation, we assume that $\psi(\cdot)$ designates the similarity rule. Now, given any $i \in E$ and $g \in \boldsymbol{G}$, let $\psi_i(g)$ denote the gang to which i belongs in the gang structure $\psi(g)$ resulting from the profile g. Then, $\psi(\cdot)$ is the similarity rule if for any strategy profile $g \in \boldsymbol{G}$, and any $i \in \boldsymbol{E}$, we have:

$$\psi_i(g) = \{j \in E | G_i = G_j\}.$$

Thus, under the similarity rule, all players with the same partnership plan (i.e., an identical strategy) are in the same gang.[16]

The *gang formation game* is now well-defined. The coalitional players are the set E of criminals; the set of strategies available to each criminal $i \in E$ consists of all possible partnership plans, $\boldsymbol{G_i}$; every strategy profile g induces a gang structure $\psi(g)$ through the similarity rule, and thus a payoff for each criminal $i \in E$ of $\pi_i(\psi(g))$.

The equilibrium outcome of our game will be a gang structure $H^* = \psi(g^*)$, where g^* is the *equilibrium strategy profile*. We now wish to identify this equilibrium strategy profile. For reasons discussed in Burbidge et al. (1997), the equilibrium concept we use in this analysis is that of the *Coalition Proof Nash Equilibrium* (CPE), developed by Bernheim, Peleg, and Whinston (1987), which is a refinement of the Nash equilibrium. The concept of CPE is intended to deal with the possibility of a group of players coordinating on a joint deviation if such a deviation is to make each deviating player better off. Then, a strategy profile is a CPE if no set of players, taking the strategies of its complement as fixed, can fashion a profitable deviation for each of its members that is itself immune to further deviations by subsets of the deviating coalition.

4. The Gang Competition Stage

Various equilibrium outcomes can emerge from the first stage. When the equilibrium outcome is the grand gang $H^* = \{\{A, B, C\}\}$, this single gang will act as a monopoly in the illegal good market. But if H^* contains two gangs (i.e., if $H^* = \{\{A, B\}, \{C\}\}$, or if $H^* = \{\{A, C\}, \{B\}\}$, or if $H^* = \{\{A\}, \{B, C\}\}$), these will compete as a duopoly. Finally, if the outcome is one with three singleton gangs ($H^* = \{\{A\}, \{B\}, \{C\}\}$), they will compete as a triopoly. We

16. For example, in a game with a set of three players $\{A, B, C\}$, suppose that the players have the following strategy profile: $g = (G_A = \{A, B, C\}, G_B = \{B, C\}, G_C = \{B, C\})$. Then, the similarity rule leads to a duopoly gang structure in which A is alone while B and C are together: $\psi(g) = \{\{A\}, \{B, C\}\}$. Another example takes place in a world with four players $\{A, B, C, D\}$. Suppose the players have the following strategy profile: $g = (G_A = \{A, B, C\}, G_B = \{B, C, D\}$, $G_C = \{B, C, D\}, G_D = \{A, B, C, D\})$. Then, the similarity rule leads to a triopoly gang structure in which A is alone, B and C are together, and D is alone: $\psi(g) = \{\{A\}, \{B, C\}, \{D\}\}$. Note that there is a unique g, denoted g^m, which leads to the formation of the grand gang: $g^m = \{E, E, E\}$.

examine each of these in turn. Note that in all cases, we assume that deterrence cannot fully deter production; in other words, we assume that for each gang structure, expected punishment cannot be chosen so as to make it optimal for firms to stop producing.

4.1 Monopoly

The problem of the criminal members of the grand gang is to maximize total profits:

$$\max_n (\beta - \gamma n)n - F - 3k\alpha ns.$$

From the first-order condition, the solution to this problem is $n^m = (\beta - 3k\alpha s)/2\gamma$. Clearly, an increase in either the statutory sanction or detection effort leads to a decrease in production. After sharing, the profits for each member are $\pi_i^m = (\beta - 3k\alpha s)^2/12\gamma - F/3$, which are also decreasing with the statutory sanction and with detection effort.

4.2 Duopoly

When the equilibrium structure entails two gangs, one is necessarily a singleton gang and the other a doubleton gang. The two gangs compete as a duopoly. We denote by n^s and n^d the production of the singleton and doubleton gangs, respectively. The first step is to find the reaction functions for each gang. For the singleton gang, the problem to solve is

$$\max_n [\beta - \gamma(n + n^d)]n - F - k\alpha ns,$$

while for the doubleton gang, it is

$$\max_n [\beta - \gamma(n^s + n)]n - F - 2k\alpha ns.$$

Solving for the Nash equilibrium yields $n^d = (\beta - 3k\alpha s)/3\gamma$ and $n^s = \beta/3\gamma$. Output is decreasing in the statutory sanction and in detection effort for the doubleton gang, but not for the singleton gang. However, because the singleton gang faces a smaller total expected punishment, it produces more. The profits for each of the criminal members of the doubleton gang are $\pi_i^d = (\beta - 3k\alpha s)^2/18\gamma - F/2$, while the profits for each criminal member of the singleton gang are $\pi_i^s = \beta^2/9\gamma - F$. Again, the profits per member are decreasing in the statutory sanction and in detection effort for the doubleton gang, but not for the singleton gang.

4.3 Triopoly

The last possible gang structure is one with three singleton gangs. The three gangs compete as a triopoly. Let n_j^t denote the output of gang G_j. The problem of the criminal of gang G_i, $i \neq k, \ell$, is

$$\max_n [\beta - \gamma(n + n_k^t + n_\ell^t)]n - F - k\alpha ns.$$

Table 1. The Payoffs under Various Gang Structures

	π_A	π_B	π_C
$\{\{A,B,C\}\}$	$(\beta - 3k\alpha s)^2/12\gamma - F/3$	$(\beta - 3k\alpha s)^2/12\gamma - F/3$	$(\beta - 3k\alpha s)^2/12\gamma - F/3$
$\{\{A,B\},\{C\}\}$	$(\beta - 3k\alpha s)^2/18\gamma - F/2$	$(\beta - 3k\alpha s)^2/18\gamma - F/2$	$\beta^2/9\gamma - F$
$\{\{A\},\{B\},\{C\}\}$	$(\beta - k\alpha s)^2/16\gamma - F$	$(\beta - k\alpha s)^2/16\gamma - F$	$(\beta - k\alpha s)^2/16\gamma - F$

Solving for the Nash equilibrium and imposing symmetry so that $n_j^t = n^t, \forall j$, we find that $n^t = (\beta - k\alpha s)/4\gamma$, which is decreasing in the statutory sanction and in detection effort. The per criminal profits are $\pi_i^t = (\beta - k\alpha s)^2/16\gamma - F$, which are also decreasing in the statutory sanction and in detection effort.

4.4 Summary

In Table 1, we present a summary of the payoffs for each criminal for the three types of gang structures. Examination of this table reveals forcefully that without a theory of coalition formation, it would be impossible to identify the equilibrium gang structure.

We now turn to the analysis of the equilibrium of the overall game and present our results regarding the impact of an increase in expected punishment or of a shift in demand.

5. Equilibrium

Before turning to the description of the equilibrium, we first present a result that is in accord with conventional wisdom. Note that all proofs are in Appendix A.

Proposition 1. Given a gang structure, an increase in expected punishment (detection effort α or statutory sanction s) leads to a decrease in the illegal good total output and to an increase in its price. As for an increase in the demand parameter β given a gang structure, it leads to an increase in both the illegal good total output and in its price.

The important requirement for Proposition 1 to hold is that the gang structure *remains the same* after expected punishment or demand has increased. As we will now show, this is not guaranteed.

We now turn to a description of the equilibrium of the overall game. Recall that the model of coalition formation used in this article is that of Burbidge et al. (1997), itself based on that of Hart and Kurz (1983). An alternative approach on coalition formation has been developed in Ray and Vohra (1999). Each approach presents some advantages. For example, Ray and Vohra's (1999) approach is more complex (a cost), but it endogenizes both the division of a coalition's resources and the coalition structure rule (a benefit), which are exogenous in the current study. In principle, the equilibria under these two approaches could differ. However, in the language of Ray and Vohra (1999), the game considered here generates a symmetric partition function. In that case, applying Theorem 3.5 of Ray and Vohra (1999) yields the same equilibrium outcome as that described in our Lemma 1 (immediately below). The two approaches are therefore consistent in the specific model used here.

Lemma 1. There exists two critical levels of the fixed cost, denoted $\bar{F}$ and $\tilde{F}$, with

$$\bar{F} = \frac{\beta^2}{6\gamma} - \frac{(\beta - 3k\alpha s)^2}{8\gamma} > \tilde{F} = \frac{(\beta - k\alpha s)^2}{8\gamma} - \frac{\beta - 3k\alpha s)^2}{9\gamma}$$

and such that

a) If $F > \bar{F} > \tilde{F}$, then the equilibrium outcome is the grand gang;
b) If $\bar{F} > F > \tilde{F}$, then the equilibrium outcome is a duopoly gang structure; and
c) If $\bar{F} > \tilde{F} > F$, then the equilibrium outcome is a triopoly gang structure.

Thus, depending on the level of the fixed cost F relative to two critical levels, various equilibrium gang structures can obtain.[17]

The following result also obtains.

Lemma 2. For β large enough relative to $k\alpha s$, the critical levels $\bar{F}$ and $\tilde{F}$ are increasing in expected punishment (detection effort α or statutory sanction s). Similarly, critical levels $\bar{F}$ and $\tilde{F}$ are both increasing in the demand parameter β.

For the rest of this article, we assume that β is large enough so that Lemma 2 holds.[18] Then, combining Lemmas 1 and 2, we can show our main result:

Proposition 2. Starting from an equilibrium gang structure entailing the grand gang, increasing expected punishment (detection effort α or statutory sanction s) will eventually lead to a switch to a duopoly gang structure; further increases in expected punishment will lead to a switch to a triopoly gang structure. An increase in the demand parameter β has a similar impact.

Thus, the gang structure responds to deterrence and to shifts in demand.[19] There are factors that make a breakup of the gang structure more likely. Clearly, the larger the size of the change in expected punishment, the more likely the

17. As is shown in Part A of the Appendix, a given equilibrium outcome can be the result of several equilibrium strategy profiles.

18. The condition on β is only required to ensure that $\partial\tilde{F}/\partial\alpha > 0$ and $\partial\tilde{F}/\partial s > 0$. This implies that the results in Proposition 2 concerning the transition from the grand gang to a duopoly gang structure hold for any value of β.

19. This key element of Proposition 2 is extended to the case of an arbitrary large number m of criminals in Part B of the Appendix.

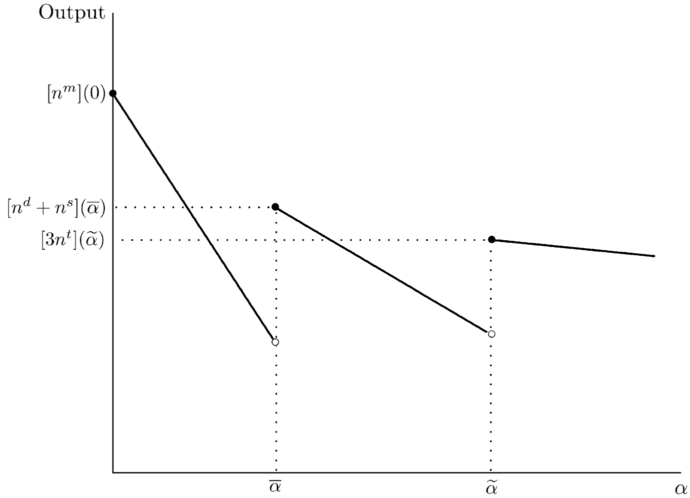

Figure 1a. Output as a Function of Detection Effort.

gang structure will respond.[20] Also, a larger demand for the illegal good (as measured by a larger β and/or a smaller γ) magnifies the response of the gang structure to expected punishment. Indeed, it is possible to show that starting from a monopoly, a given increase in expected punishment is more likely to induce a breakup of the monopoly when the demand is larger.[21]

Ultimately, an authority who would like to design an optimal deterrence policy should take the response of the gang structure into account. For example, if such an authority would like to minimize the illegal good total output, it would certainly benefit from understanding the following:

Corollary to Proposition 2. Starting from an equilibrium gang structure entailing the grand gang, increasing expected punishment (detection effort α or statutory sanction s) can lead to an increase in the illegal good total output and a reduction in its price. Also, starting from an equilibrium gang structure entailing the grand gang, an increase in the demand parameter β leads to an increase in the illegal good total output, and may translate into a reduction in its price.

This result contradicts conventional wisdom. It says that an increase in expected punishment can lead to more crime (as measured by output). This, of course, is due to the fact that by increasing expected punishment, a monopoly is broken and the illegal good market is made more competitive. And as is well known, more competition implies larger quantities.

In Figures 1a and 1b, we have depicted the relationship between total output and effort in detection (equivalent figures can be drawn with respect to the

20. This is because, for a given F, a larger change in α leads to a larger change in $\bar{F}$ and $\tilde{F}$, making it more likely that F will switch from being smaller than $\bar{F}$ (resp. $\tilde{F}$) to being larger than $\bar{F}$ (resp. $\tilde{F}$).

21. It can be shown that $\partial^2\bar{F}/\partial\alpha\partial\beta > 0$ and $\partial^2\bar{F}/\partial\alpha\partial\gamma < 0$.

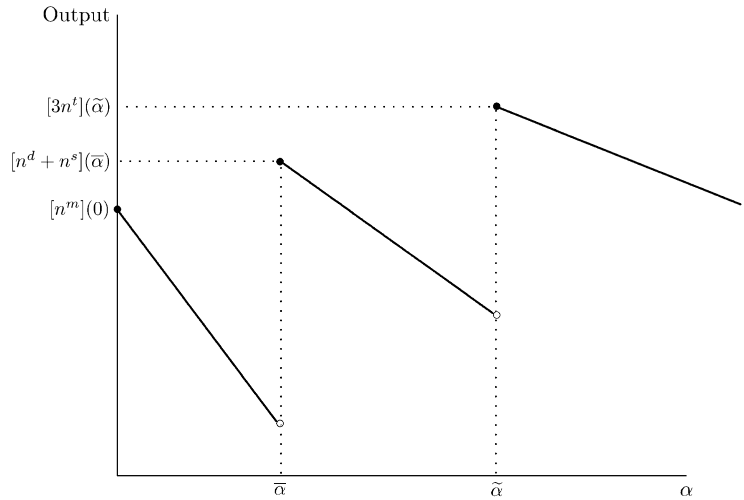

Figure 1b. Output as a Function of Detection Effort.

statutory sanction). In these figures, $\bar{\alpha}$ and $\tilde{\alpha}$ denote the levels of detection effort such that F equals $\bar{F}$ and $\tilde{F}$, respectively. Thus, for $\alpha < \bar{\alpha}$ (low detection effort), the equilibrium gang structure is a monopoly; for $\bar{\alpha} \leq \alpha < \tilde{\alpha}$ (medium detection effort), it is a duopoly; and for $\alpha \geq \tilde{\alpha}$ (high detection effort), it is a triopoly. In both figures, it can be noted that for a given gang structure, output declines with detection effort, but also that output jumps when the equilibrium gang structure changes. Overall, and as in Figure 1b, a breakup in the gang structure following an increase in detection effort may well lead to a level of output higher than any of those observed for lower levels of detection effort.

In Figure 2, the price increases when detection effort increases, as long as the market structure remains the same. This reflects the fact that for all gangs in a gang structure, an increase in detection effort is equivalent to an increase in their costs, which translates into a higher price. However, if the increase in detection effort triggers a splintering of the gang structure, the increased competition leads to an abrupt fall in the price. The overall impact on prices is generally ambiguous.[22]

One could argue that an increase in demand explains the increase in the production of drugs observed in the last decades. But, in a standard model with a fixed market structure, this potential explanation is unsatisfactory because it

22. Since the price does not necessarily fall, our model can also be made consistent with the view of Kuziemko and Levitt (2001), according to which prices would have fallen by more had it not been for increased deterrence.

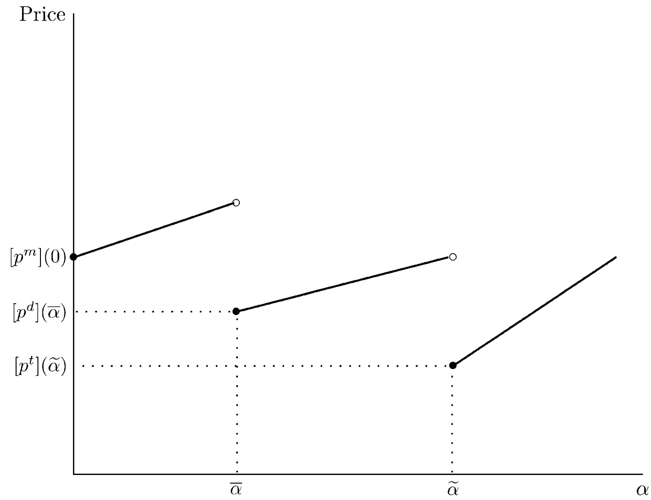

Figure 2. Price as a Function of Detection Effort.

fails to generate the reduction in prices that was also observed.[23] In our model, however, an increase in the demand parameter β can generate a splintering of the gang structure, and this can lead to a fall in prices. By staying with the grand gang, a criminal gets his share of monopoly profits, but by being the first to break away, he gets duopoly profits he does not have to share. As can be shown, for some size of demand, the incentive to break away comes to dominate that to stay with the grand gang. Note that by generating potentially lower prices for larger markets, our model is consistent with Caulkins (1995), which provides empirical evidence according to which larger markets (cities) have lower drug prices.

The next figures depict the relationship between the demand parameter β, output (Figure 3), and price (Figure 4). In these figures, the level of expected punishment is set at some arbitrary fixed level, and $\bar{\beta}$ and $\tilde{\beta}$ denote the levels of the demand parameter β such that F equals $\bar{F}$ and $\tilde{F}$, respectively. For $\beta < \bar{\beta}$ (low demand), the equilibrium gang structure is a monopoly; for $\bar{\beta} \leq \beta < \tilde{\beta}$ (medium demand), it is a duopoly; and for $\beta \geq \tilde{\beta}$ (high demand), it is a triopoly. In Figure 3, output rises with demand for a given gang structure, and output jumps up when the equilibrium gang structure is made more competitive. Overall, output unambiguously rises as demand rises. On the other hand, the impact on the price of an increase in demand is similar to that produced by an increase in expected punishment. The price increases when demand rises, as long as the gang structure remains constant. But when the market becomes sufficiently large (i.e., when the demand is large enough), a breakdown in the gang structure takes place. In such a case, the price falls abruptly.

23. Again, as was mentioned earlier, a model with entry deterrence could possibly generate a fall in the price.

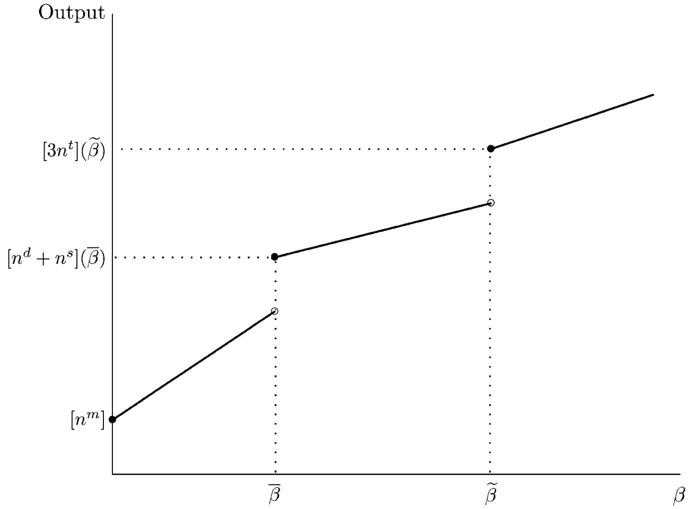

Figure 3. Output as a Function of Demand Size.

To summarize, our model can explain the increase in production and the fall in prices of several drugs in the last decades, whether these changes were due to an increase in deterrence or to an increase in demand.

6. Market Share and Detection

The probability that the activities of a gang will be detected is an important element of the preceding analysis. We argued that expected punishment could

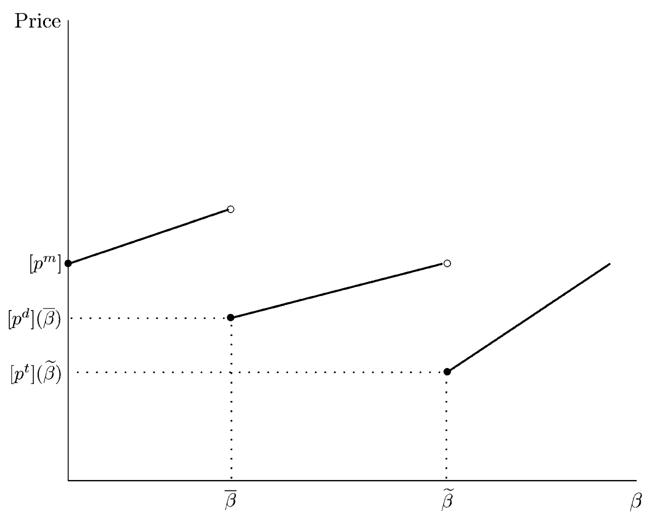

Figure 4. Price as a Function of Demand Size.

increase for larger gangs ($z_n > 0$) when the probability of detection is not affected by output ($p_n = 0$) but when effective sanction increases with output ($\hat{s}_n > 0$). An alternative view is simply that the probability of detection is an increasing function of the absolute level of output ($p_n > 0$), and that effective sanctions are not affected by output ($\hat{s}_n = 0$).[24] But in that case, for a given level of detection effort α, if each gangs double their output, they each double the probability that they are apprehended. This may be viewed as unsatisfactory when detection effort is fixed.

As was suggested to us by a referee, a possibly more satisfactory alternative is to suppose that the probability of detection is a function of market share. Further to conforming with intuition, such a specification also introduces interesting externalities between gangs: when a gang increases its output, it increases its market share and its probability of detection, but it also reduces the market share of other gangs and their probability of detection. We now turn to analyzing such a world for the particular case in which there are only two criminals.

If there are only two criminals, only two gang structures are possible: a two-member monopoly gang or a two singleton gangs duopoly. Assume that when apprehended, the effective sanction for each criminal belonging to a gang producing n units of output is $\hat{s}(s, n) = ns$, so that the effective sanction reflects laws similar to the Kingpin Act. To simplify further, assume that $s = 1$, so that $\hat{s} = n$. We will compare results in two cases.

a. *Detection independent of output*. To render meaningful the comparison with Case *b* below, we assume that detection effort α is either fully allocated to detecting the monopoly, or equally split in half to detect each gang in the case of a duopoly. Thus, $p(\alpha, n) = \alpha$ for a monopoly, and $p(\alpha, n) = \alpha/2$ for each gang in the case of a duopoly. Expected punishment is then $z(\alpha, s, n) = \alpha n$ for the monopoly, and $z(\alpha, s, n) = \alpha n/2$ for each gang in a duopoly.
b. *Detection depends on market share*. In this case, we assume that the probability of gang i producing n_i while gang j produces n_j is $p(\alpha, n_i, n_j) = \alpha n_i/(n_i + n_j)$. Thus, the larger the market share of i, the larger its probability of detection. Note that expected punishment is now $z(\alpha,s,n_i,n_j) = \alpha n_i^2/(n_i + n_j)$.

6.1 Detection Independent of Output

We use subscript I to identify equilibrium quantities in this case. When the two criminals operate together in a monopoly, they solve the following problem:

$$\max_n (\beta - \gamma n)n - F - 2\alpha n.$$

The solution to this problem is $n_I^m = (\beta - 2\alpha)/2\gamma$, and the profits for each member are $\pi_{I,i}^m = (\beta - 2\alpha)^2/8\gamma - F/2$. To ensure positive monopoly output, we impose that $\beta > 2\alpha$.

24. There are, of course, other combinations of p_n and $\hat{s}_n$ that can make $z_n > 0$.

In the case of a duopoly, each gang is a singleton. For each singleton gang, the problem to solve is

$$\max_n [\beta - \gamma(n + n')]n - F - \alpha\frac{n}{2}.$$

Solving for the Nash equilibrium yields $n_I^s = [\beta - (\alpha/2)]/3\gamma$. The profits for the criminal of each singleton gang are $\pi_{I,i}^s = [\beta - (\alpha/2)]^2/9\gamma - F$.

6.2 Detection Depends on Market Share

We use subscript M to identify equilibrium quantities in this case. When the two criminals operate together in a monopoly, because market share is unity, the problem they solve is identical to that of a monopoly in Case a. Thus, $n_M^m = (\beta - 2\alpha)/2\gamma$ and the profits for each member are $\pi_{M,i}^m = (\beta - 2\alpha)^2/8\gamma - F/2$.

In the case of a duopoly, the problem to solve for each singleton gang is

$$\max_n [\beta - \gamma(n + n')]n - F - \alpha\frac{n^2}{n + n'}.$$

Solving for the Nash equilibrium yields $n_M^s = [\beta - (3\alpha/4)]/3\gamma$. The profits for the criminal of each singleton gang are $\pi_{M,i}^s = \beta[\beta - (3\alpha/4)]/9\gamma - F$.

6.3 Comparison

The first point to note is that under a duopoly, output per gang is smaller when detection depends on market share than when it does not: $n_M^s < n_I^s$. The intuition is clear: each gang takes into account the fact that having a larger market share increases the probability of detection, so each gang reduces its output.

More interestingly, provided $\beta > 2\alpha$ (which we require above to ensure positive monopoly output), profits per gang (or per criminal) are larger when detection depends on market share than when it does not: $\pi_{M,i}^s > \pi_{I,i}^s$. The intuition for this result is that the non-cooperative duopoly total output is too large relative to the joint profit maximizing output (i.e., the monopoly output). The fact that detection depends on market share disciplines each gang, leading them to reduce output and thereby increasing profits.

As before, there is a level of the fixed cost $\hat{F}_r$ in regime $r = I, M$ such that if $F > \hat{F}_r$, then the equilibrium gang structure is the monopoly, while if $F < \hat{F}_r$, then the equilibrium gang structure is the duopoly. These critical levels of the fixed cost are

$$\hat{F}_I = 2\left[\frac{[\beta - (\alpha/2)]^2}{9\gamma}\right] - \frac{(\beta - 2\alpha)^2}{4\gamma}$$

$$\hat{F}_M = 2\left[\beta\frac{[\beta - (3\alpha/4)]}{9\gamma}\right] - \frac{(\beta - 2\alpha)^2}{4\gamma}.$$

It follows from simple manipulations that $\hat{F}_M > \hat{F}_I$. Thus, a monopoly gang structure is more likely to splinter after an increase in detection effort α or in demand β when detection depends on market share than when it does not. The intuition, of course, is that profits are larger when detection depends on market share, and this makes more attractive the unilateral deviation leading to a duopoly.

Thus, the main point of this article (i.e., that criminal market structure reacts to deterrence or to demand) is maintained in an environment in which detection depends on market share.

7. Conclusion

Previous studies of criminal gangs have assumed a fixed market (or gang) structure. This article's contribution has been to provide a framework in which gang structure is endogenous, thereby allowing for interesting and possibly counterintuitive phenomena. Our framework may, for example, provide an explanation for the failure of the "war on drugs" intensified in the 1980s under the Reagan administration. If market structure is fixed, more deterrence clearly leads to a reduction in criminal output. But as is shown in our analysis, when the gang structure responds to deterrence, more deterrence may lead to an increase in criminal output if it makes the market more competitive. Such a phenomenon was shown to possibly obtain in an oligopolistic market of the type we observe in the real world (e.g., cocaine or heroin markets). The failures of several wars on drugs may be explained by such a turn of events.

Appendix A: Proofs

Proof of Proposition 1. The grand gang produces n^m, which is decreasing in α, and the price is $P^m = \beta - \gamma n^m$, which is increasing with α. Under the duopoly structure, total output is $n^s + n^d$, which decreases with α, while the price $P^d = \beta - \gamma(n^d + n^s)$ is increasing with α. Finally, total output is $3n^t$ under a triopoly structure, and this is also decreasing in α; the price $P^t = \beta - \gamma 3n^t$ is increasing with α. The same analysis applies for changes in s.

A similar analysis can be undertaken with respect to changes in the demand parameter β. The grand gang produces n^m, which is increasing in β, and the price is $P^m = \beta - \gamma n^m$, which is increasing with β. Under the duopoly structure, total output is $n^s + n^d$, which increases with β, and the price $P^d = \beta - \gamma(n^d + n^s)$ is increasing with β. Finally, total output is $3n^t$ under a triopoly structure, which is also increasing in β, and the price $P^t = \beta - \gamma 3n^t$ is increasing with β. ■

Proof of Lemma 1. $\bar{F}$ is the fixed cost such that $\pi_i^m = \pi_i^s$, while $\tilde{F}$ is such that $\pi_i^d = \pi_i^t$. Note that $\bar{F} > \tilde{F}$ if

$$\frac{\beta^2}{6\gamma} - \frac{(\beta - 3k\alpha s)^2}{8\gamma} > \frac{(\beta - k\alpha s)^2}{8\gamma} - \frac{(\beta - 3k\alpha s)^2}{9\gamma}$$

Table A.1. Ranking of Payoffs when $F > \bar{F} > \tilde{F}$

	π_A	π_B	π_C
$\{\{A, B, C\}\}$	1	1	1
$\{\{A, B\}, \{C\}\}$	3	3	2
$\{\{A\}, \{B\}, \{C\}\}$	4	4	4

which, after some manipulations, reduces to $\beta^2 + 3k\alpha s(4\beta - 3k\alpha s) > 0$. This holds since $n^m = (\beta - 3k\alpha s)/2\gamma > 0$.

Next, to prove parts a), b), and c), consider the relative payoffs for the three possible levels of F.

a) $F > \bar{F} > \tilde{F}$ (Table A.1)

From the strategy profile $g^m = (E, E, E)$, with $E = \{A, B, C\}$ (i.e., all criminals want to be in a grand gang), there is no profitable deviation, as the payoffs in the first row (the equilibrium outcome for g^m) dominate those of the two other rows. Therefore, g^m is a CPE.

From any g that leads to the gang structures in rows 2 or 3, there are always profitable and credible deviations by the subset of players with $G_i \neq E$ to $G_i = E$. They are profitable (see Table A.1), and credibility is established by g^m being CPE. Therefore, the unique CPE is g^m, and the unique gang structure is the grand gang.

b) $\bar{F} > F > \tilde{F}$ (Table A.2)

From the strategy profile g^m, there is a profitable unilateral, and therefore credible deviation, by C, to $G_C = \{C\}$. From any g that leads to the triopoly gang structure $\{\{A\}, \{B\}, \{C\}\}$, there is at most one player with strategy E. If there is one player with strategy E, then this player, say A, and one of the other

Table A.2. Ranking of Payoffs when $\bar{F} > F > \tilde{F}$

	π_A	π_B	π_C
$\{\{A, B, C\}\}$	2	2	2
$\{\{A, B\}, \{C\}\}$	3	3	1
$\{\{A\}, \{B\}, \{C\}\}$	4	4	4

Table A.3. Ranking of Payoffs when $\bar{F} > \tilde{F} > F$

	π_A	π_B	π_C
$\{\{A, B, C\}\}$	2	2	2
$\{\{A, B\}, \{C\}\}$	4	4	1
$\{\{A\}, \{B\}, \{C\}\}$	3	3	3

two, say B (as in Table A.1), have a profitable deviation to $G_A = G_B = \{A, B\}$. This deviation is also credible, as there are no further profitable deviations for them (recall $G_C \neq E$). If no player has strategy E, then any two players, say A and B, have a profitable and credible deviation to $G_A = G_B = \{A, B\}$.

Therefore, if there is a CPE gang structure, it must be the duopoly gang structure. The profile $g = (\{A, B\}, \{A, B\}, \{C\})$ is immune to unilateral deviations, as they are not profitable. As for a joint deviation by A and B to E, it would have no effect because under the similarity rule, the equilibrium gang structure would not change. Thus, $g = (\{A, B\}, \{A, B\}, \{C\})$ is CPE, as are other profiles [e.g., $g' = (\{A, B\}, \{A, B\}, \{B, C\})$] leading to the duopoly gang structure. Thus, the unique CPE gang structure is the duopoly.

c) $\bar{F} > \tilde{F} > F$ (Table A.3)

From the strategy profile g^m, there is a profitable unilateral and therefore credible deviation, by C, to $G_C = \{C\}$. From any g that leads to the duopoly gang structure $\{\{A, B\}, \{C\}\}$, there are profitable unilateral and therefore credible deviations, by A to $G_A = \{A\}$, and by B to $G_B = \{B\}$.

Therefore, if there is a CPE gang structure, it must be the triopoly gang structure. The profile $g = (\{A\}, \{B\}, \{C\})$ is immune to unilateral deviations and to two-player deviations because they are not profitable. A joint deviation by all players to E is not credible as C would have an incentive to further deviate to $G_C = \{C\}$. Thus, $g = (\{A\}, \{B\}, \{C\})$ is CPE, as are other profiles [e.g., $g' = (\{A, B\}, \{B\}, \{C\})$] leading to the triopoly gang structure. Thus, the unique CPE gang structure is the triopoly. ■

Proof of Lemma 2. The result is obtained by differentiating $\bar{F}$ and $\tilde{F}$ with respect to α. First, $\partial\bar{F}/\partial\alpha = 6ks(\beta - 3k\alpha s)/8\gamma$, which is positive since $n^m = (\beta - 3k\alpha s)/2\gamma > 0$. Second, note that $\partial\tilde{F}/\partial\alpha = 2ks\{[(\beta - 3k\alpha s)/3\gamma] - [(\beta - k\alpha s)/8\gamma]\}$. Routine manipulations show that $\partial\tilde{F}/\partial\alpha > 0$ if $\beta > 4.2k\alpha s$, i.e., $\partial\tilde{F}/\partial\alpha > 0$ for β large enough relative to $k\alpha s$. The same analysis applies for changes in s.

Similarly, we need to differentiate $\bar{F}$ and $\tilde{F}$ with respect to β. First, $\partial\bar{F}/\partial\beta = (\beta + 9k\alpha s)/12\gamma$, which is positive. Second, note that $\partial\tilde{F}/\partial\beta = \{[(\beta + 15k\alpha s)/36\gamma]$, which is positive. ■

Proof of Proposition 2. Starting from $F > \bar{F} > \tilde{F}$, and by Lemma 2, increasing α or s will eventually lead to $\bar{F} > F > \tilde{F}$. Further increases will then lead to $\bar{F} > \tilde{F} > F$. That the gang structure changes then follows from Lemma 1. The same analysis is valid for an increase in β. ■

Proof of Corollary to Proposition 2. From Proposition 2, a marginal increase in α when $F = \bar{F}$ will induce a switch from the grand gang to the duopoly gang structure. Output will therefore go from $n^m = (\beta - 3k\alpha s)/2\gamma$ to $n^s + n^d = (2\beta - 3k\alpha s)/3\gamma > n^m$, and price will go from $P^m = (\beta + 3k\alpha s)/2$ to $P^d = (\beta + k\alpha s)/3 < P^m$. Also, an increase in α when $F = \tilde{F}$ will induce a switch from the duopoly to the triopoly gang structure. Output will therefore go from $n^s + n^d = (2\beta - 3k\alpha s)/3\gamma$ to $3n^t = 3(\beta - k\alpha s)/4\gamma > n^s + n^d$, and price will go from $P^d = (\beta + k\alpha s)/3$ to $P^t = (\beta + k\alpha s)/4 < P^d$. The same analysis applies for changes in s.

Also from Proposition 2, a marginal increase in β when $F = \bar{F}$ will induce a switch from the grand gang to the duopoly gang structure. Output will therefore go from $n^m = (\beta - 3k\alpha s)/2\gamma$ to $n^s + n^d = (2\beta - 3k\alpha s)/3\gamma > n^m$, and price will go from $P^m = (\beta + 3k\alpha s)/2$ to $P^d = (\beta + k\alpha s)/3 < P^m$. Also, an increase in β when $F = \tilde{F}$ will induce a switch from the duopoly to the triopoly gang structure. Output will therefore go from $n^s + n^d = (2\beta - 3k\alpha s)/3\gamma$ to $3n^t = 3(\beta - k\alpha s)/4\gamma > n^s + n^d$, and price will go from $P^d = (\beta + k\alpha s)/3$ to $P^t = (\beta + k\alpha s)/4 < P^d$. ■

Appendix B: Generalization to the Case of *m* Criminals

In the main text, we consider a simplified world in which there are only three criminals who have to choose with whom they want to operate a gang. In that context, Proposition 2 states that gang structure may be affected in a precise way by expected punishment and by shifts in demand.

We here extend the key element of Proposition 2 (i.e., that market structure reacts to changes in expected punishment or in demand) to the case in which there is an arbitrarily large number m of criminals. To make our point, we assume that initially, the m criminals have decided to operate together as a monopoly. The analysis below shows that when expected punishment increases (the equivalent holds for shifts in demand), a point will come where it will pay for one criminal (any of the m criminals, as they are all identical and profits are shared equally) to leave the grand monopoly gang and to start operating on his own, thereby creating a duopoly gang structure with a large gang of $(m - 1)$ criminals and a small gang with a single criminal.

Extending the model presented in the main text, consider first the problem of m criminals operating as a monopoly:

$$\max_n (\beta - \gamma n)n - F - mk\alpha ns.$$

From the first-order condition, the solution to this problem is $n^m = (\beta - mk\alpha s)/2\gamma$. After sharing, the profits for each member of the gang are $\pi_i^m = (\beta - mk\alpha s)^2/4m\gamma - F/m$.

Consider now the duopoly gang structure in which one is a singleton gang and the other an $(m-1)$-member gang. Denote by n^s and n^ℓ the production of the singleton and of the $(m-1)$-member gangs, respectively. The first step is to find the reaction functions for each gang. For the singleton gang, the problem to solve is

$$\max_n [\beta - \gamma(n + n^\ell)]n - F - k\alpha ns,$$

while for the $(m-1)$-member gang, it is

$$\max_n [\beta - \gamma(n^s + n)]n - F - (m-1)k\alpha ns.$$

Solving for the Nash equilibrium levels of output yields $n^\ell = [\beta + (3 - 2m)k\alpha s]/3\gamma$ and $n^s = [\beta - (3-m)k\alpha s]/3\gamma$. The profits for the criminal of the singleton gang are then $\pi_i^s = [\beta - (3-m)k\alpha s]^2/9\gamma - F$.

Our claim is that starting from the monopoly gang structure (with $\pi_i^s < \pi_i^m$), an increase in expected punishment or shifts in demand may alter the environment and make it attractive for any criminal to leave the grand monopoly gang $(\pi_i^s > \pi_i^m)$. Formally, the following holds:

Lemma 1A. There exists a critical level of the fixed cost, denoted $\check{F}$, with

$$\check{F} = \frac{m}{(m-1)\gamma}\left[\frac{[\beta - (3-m)k\alpha s]^2}{9} - \frac{[\beta - mk\alpha s]^2}{4m}\right]$$

such that if $F < \check{F}$, then the equilibrium outcome is not the monopoly gang structure.

Proof of Lemma 1A. $\check{F}$ is the fixed cost such that $\pi_i^m = \pi_i^s$. Suppose that initially, the grand monopoly gang is an equilibrium (it must then be that $\pi_i^m > \pi_i^s$). Simple manipulations show that if the environment becomes such that $F < \check{F}$, then $\pi_i^s > \pi_i^m$ and there is a profitable unilateral (and therefore credible) deviation: some criminal (any of the m criminals) can leave the grand monopoly gang to operate on his own and be better off. Therefore, the grand monopoly gang cannot be an equilibrium gang structure if $F < \check{F}$. ■

Further, note that $\check{F}$ is increasing in α, s, and β. Then, if those parameters are such that $F > \check{F}$ initially, an increase in any of those parameters will increase $\check{F}$ and will eventually reverse the inequality, making it impossible for the grand monopoly gang to remain an equilibrium.

Note that showing that the splintering of the monopoly gang structure would eventually happen is sufficient to make our point. However, we do not claim, as in Proposition 2, that following an increase in expected punishment or in demand, the first splintering will necessarily be from the monopoly to an $\{m-1, 1\}$ duopoly structure. In fact, the splintering could be from the monopoly to other duopoly structures, e.g., to an $\{m-m', m'\}$ duopoly structure with $0 < m' < m$, where m' criminals make a joint deviation. In fact, the first splintering could be from the monopoly to a large variety of gang structures involving an arbitrary number of gangs. Thus, the precise path in which splintering takes place when there are three players may not necessarily generalize to the case of *m* players. Finally, note that when the number of criminals increases, the potential number of gangs also increases. Clearly, when the number of gangs is large, a small increase in their number should have only a small impact on output and prices. However, a counterbalancing factor is that a given change in expected punishment (or a given shift in demand) could have a larger impact on the number of operating gangs when their number is large than when their number is small.

References

Benson, B. L., and D. W. Rasmussen. 1991. "The Relationship Between Illicit Drug Enforcement Policy and Property Crimes," 9 *Contemporary Economic Policy* 106–15.

Benson, B. L., I. Kim, D. W. Rasmussen, and T. W. Zuehlke. 1992. "Is Property Crime Caused by Drug Use or by Drug Enforcement Policy," 24 *Applied Economics* 679–92.

Bernheim, D., B. Peleg, and M. D. Whinston. 1987. "Coalition–Proof Nash Equilibrium 1. Concepts," 42 *Journal of Economic Theory* 1–12.

Buchanan, J. 1973. "A Defense of Organized Crime," in S. Rottenberg, ed., *The Economics Of Crime and Punishment*. Washington D.C.: American Enterprise Institute.

Burbidge, J. B., J. DePater, G. M. Myers, and A. Sengupta. 1997. "A Coalition Formation Approach to Equilibrium Federations and Trading Blocs," 87 *American Economic Review* 940–56.

Burrus, R. T. 1999. "Do Efforts to Reduce the Supply of Illicit Drugs Increase Turf War Violence? A Theoretical Analysis," 23 *Journal of Economics and Finance* 465–93.

Caulkins, J. P. 1993. "Zero-Tolerance Policies: Do they Inhibit or Stimulate Illicit Drug Consumption?," 39 *Management Science* 458–76.

Caulkins, J. P. 1995. "Domestic Geographic Variation in Illicit Drug Prices," 37 *Journal of Urban Economics* 38–56.

Caulkins, J. P., and R. Padman. 1993. "Interdiction's Impact on the Structure and Behavior of the Export-Impost Sector for Illicit Drugs," 37 *Zeitschrift fur Operations Research* 207–24.

Caulkins, J. P., and P. Reuter. 1998. "What Price Data Tell Us About Drug Markets," 28 *Journal of Drug Issues* 593–612.

Everingham, S. M., C. P. Rydell, and J. P. Caulkins. 1995. "Cocaine Consumption in the United States: Estimating Past Trends and Future Scenarios," 29 *Socio-Economic Planning Sciences* 305–14.

Fiorentini, G., and S. Peltzman. 1995. "Introduction," in G. Fiorentini and S. Peltzman, eds., *The Economics Of Organised Crime*. Stanford: Cambridge University Press.

Gambetta, D., and P. Reuter. 1995. "Conspiracy Among the Many: The Mafia in Legitimate Industries," in G. Fiorentini and S. Peltzman, eds., *The Economics Of Organised Crime*. Stanford: Cambridge University Press.

Hart, S., and M. Kurz. 1983. "Endogenous Formation of Coalitions," 51 *Econometrica* 1047–64.

Kleiman, M. A. 1993. "Enforcement Swamping: A Positive-Feedback Mechanism in Rates of Illicit Activity," 17 *Mathematical and Computer Modeling* 65–75.

Kuziemko, I., and S. D. Levitt. 2001. "An Empirical Analysis of Imprisoning Drug Offenders," NBER Working Paper 8489.

Lee, L. W. 1993. "Would Harassing Drug Users Work?," 101 *Journal of Political Economy* 939–61.

Levitt, S. D., and S. A. Venkatesh. 2000. "An Economic Analysis of Drug-Selling Gang's Finances," 115 *Quarterly Journal of Economics* 755–89.

Marceau, N., and G. Myers. Forthcoming. "On the Early Holocene: From Foraging to Early Agriculture," *Economic Journal*.

Miron, J. A. 1999. "Violence and the U.S. Prohibitions of Drugs and Alcohol," 1 *American Law and Economics Review* 78–114.

Miron, J. A., and J. Zwiebel. 1995. "The Economic Case Against Drug Prohibitions," 9 *Journal of Economic Perspectives* 175–72.

Neher, P. A. 1978. "The Pure Theory of Muggery," 68 *American Economic Review* 437–45.

Poret, S. 2003. "Paradoxical Effects of Law Enforcement Policies: The Case of Illicit Drug Market," 22 *International Review of Law and Economics* 465–93.

Ray, D., and R. Vohra. 1999. "A Theory of Endogenous Coalition Structures," 26 *Games and Economic Behavior* 286–336.

Skott, P., and G. T. Jepsen. 2002. "Paradoxical Effects of Drug Policy in a Model with Imperfect Competition and Switching Cost," 48 *Journal of Economic Behavior & Organization* 335–54.

The Economist. 1999. "A New Class of Trafficker," Volume 352, Issue 8136 (September 11, page 38).