The London School of Economics and Political Science

Essays on the Economics of Crime

Iain William Long

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy, London, June 2012.

Declaration

I certify that the thesis I have presented for examination for the Ph.D. degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 33,025 words.

Statement of inclusion of previous work

I can confirm that chapter 1 was based upon the result of previous study (for a Master of Research in Economics) I undertook at the London School of Economics and Political Science.

Abstract

I present three essays on the economics of crime. The first considers an activity associated with 55% of all criminal offences in the UK: binge drinking. One group inextricably linked with such behaviour is the sports team. Members regularly engage in post-match drinking, where the team's reputation is at stake. Teams often apply peer pressure (the threat of punishment for refusal to compete) to ensure each member gets involved. Chapter 1 presents a simple model of competitive drinking, and evaluates the amount of peer pressure a team needs to apply when multiple equilibria exist.

The thesis then turns attention towards criminal organisations. Chapter 2 discusses the use of initiation by protection rackets. Such rituals are widely used, and serve several purposes. Firstly, they allow initiates' skills to be assessed. Secondly, they act as an incentive to invest in skills. Thirdly, they signal to the racket's customers. The chapter derives conditions on the underlying distribution of abilities such that a racket can adjust initiation difficulty to improve its reputation. It then discusses these conditions in light of "key player" policies, suggesting they may be more effective than previously thought.

Chapter 3 evaluates the impact of a variety of anti-crime policies on how a criminal gang recruits. Gangs counteract policy effects by adjusting the wage they offer and the intensity of violence they require their members to inflict. This can lead to policies backfiring; increasing the social cost of the gang. A policy which reduces the youths' incentive to join a gang leaves only hardened criminals as recruits. If gang size and violence are weak revenue complements, this causes the gang to substitute towards more violence. Policies are therefore most effective when they not only reduce the incentive to join the gang, but also increase youths' sensitivity towards inflicting violence.

Contents

Declaration **2**

1 Peer Pressure and Binge Drinking **8**

1.1 Introduction 8

1.2 Literature Review 12

1.2.1 Empirical Work 13

1.2.2 Theoretical Work 15

1.3 The Model 18

1.4 Equilibria 23

1.4.1 Acquired Signal Equilibria 24

1.4.2 No Signal Equilibria 27

1.5 Peer Pressure 29

1.6 Conclusions 34

1.A Proof of Proposition 1 36

1.B Proof of Proposition 2 39

1.C Proof of Proposition 3 40

1.D Proof of Proposition 4 41

1.E Proof of Proposition 5 45

1.F Proof of Corollary 1 46

1.F.1 All players choose action C 47

1.F.2 *"Strong"* players choose action C, whereas *"Weak"* players choose action D 48

1.F.3 *"Strong"* players choose action D, whereas *"Weak"* players choose action C 49

1.F.4 All players choose action D 49

2 Initiation and Protection Rackets **51**

2.1 Introduction 51

2.2 Stylised Facts About Protection Rackets 55

2.2.1 Reputation in Protection Markets 55

2.2.2 Initiation as a Determinant of Reputation 57

2.2.3 Acquisition of Criminal Skills 59

2.3 The Model 61

2.4 Equilibrium 65

2.4.1 Youth Decisions with Initiation 66

2.4.2 Youth Decisions without Initiation 69

2.4.3 Mafia Decisions 72

2.5 Initiation as a Signal 73

2.6 Discussion 78

2.7 Conclusions 81

2.A Proof of Proposition 6 83

2.A.1 $\frac{\partial \bar{c}_1}{\partial I} > 0 \Rightarrow \frac{\partial \sigma_H}{\partial I} > 0$ and $\frac{\partial \sigma_L}{\partial I} > 0$ 83

2.A.2 $\frac{\partial \sigma_H}{\partial I} > 0$ and $\frac{\partial \sigma_L}{\partial I} > 0 \Rightarrow \frac{\partial \bar{c}_1}{\partial I} > 0$ 83

2.B Proof of Proposition 7 84

3 Recruitment to Organised Crime **88**

3.1 Introduction 88

3.1.1 Criminal Organisations as Firms 91

3.1.2 Opportunities in a Criminal Neighbourhood 94
3.1.3 Tackling Organised Crime 96
3.2 A Model of Recruitment to Organised Crime 100
3.3 The Simple Gang 103
3.3.1 Youth Decisions 103
3.3.2 Gang Leader Decisions 106
3.4 Policy with a Simple Gang 111
3.4.1 Results for a Generic Policy 112
3.4.2 Results for Specific Policies 117
3.5 The Separating Gang 125
3.5.1 Youth Decisions 125
3.5.2 Gang Leader Decisions 127
3.6 Policy with a Separating Gang 131
3.6.1 Results for a Generic Policy 132
3.6.2 Results for Specific Policies 136
3.7 Conclusions 141
3.A Proof of Proposition 8 143
3.B Proof of Proposition 11 145
3.C Proof of Proposition 12 147
3.D Proof of Proposition 13 148

List of Figures

2.1 A simple model of protection 56

2.2 Equilibria with high and low recruits, only low recruits and no recruits. 70

2.3 A situation in which Proposition 7 holds. 78

2.4 A situation in which Proposition 7 does not hold. 79

3.1 A σ-ability youth's choice of criminal skill. 104

3.2 Youths' payoffs from gang membership versus the primary labour market. 105

3.3 The range of values over which Assumption 4 and Assumption 5 hold. 109

3.4 Restricted factor demands with complements and substitutes. . . 111

3.5 Policy effects with complements. 114

3.6 Policy effects with substitutes. 115

3.7 Effect of an increase in the severity of punishment (f) or the market wage (w) 118

3.8 Effect of an increase in the arrest and conviction rate (p) or in the prevention of juvenile crime (k). 121

3.9 An example of an implementable contract 128

Chapter 1

Peer Pressure and Binge Drinking

1.1 Introduction

Binge drinking is one of the most high profile issues facing modern British society. A recent government report suggests that alcohol abuse costs the National Health Service around £2.7 billion per year in England (Department of Health 2008). This figure has seen a significant increase over recent years, growing from £1.7 billion per year only five years earlier (Cabinet Office 2003). The upsurge in concern in policy, media and academic circles is therefore unsurprising.

Binge drinking is also the source of a wide variety of negative externalities. Increased morbidity, absenteeism and premature death resulting from alcohol abuse are estimated to cost England around £6.4 billion per year in lost productivity (Cabinet Office 2003). Moreover, the cost of alcohol-fuelled crime (including victimisation costs) was estimated at around £11.9 billion per year. Matthews, Shepherd, and Sivarajasingham 2006 show that alcohol consumption is a significant risk factor in admission to hospital as a result of being injured due to violence. Moreover, they suggest that this relationship is causal. They found that increases in the price of beer in the UK significantly reduced the number of hospital admissions, even when controlling for numerous socioeconomic and seasonal variables. It is this relationship to crime that often captures the headlines. Thirty percent of all recorded crime in the UK in 2003 was committed by young binge drinkers (compared with 17% for other young people), with binge drinkers of all ages accounting for 55% of all offences (Home Office 2005). Whilst it is important to differentiate between correlation and causation, 63% of young binge drinkers surveyed claimed to have been involved in criminal or disorderly behaviour whilst drunk in the last twelve months. Twenty-five percent claimed to have been involved in a fight.

What is the extent of the problem? According to a survey conducted by the Home Office, 62% of young men[1] and 47% of young women are classed as binge drinkers, claiming to feel very drunk at least once a month for the last twelve months (Home Office 2006). Perhaps more worryingly, 28% of boys, and 32% of girls aged 10-15 reported binge drinking. The issue is not restricted to the UK. In the US, many universities and colleges have adopted an alcohol policy in response to the effects of binge drinking. In a recent survey of online policies (Faden and Baskin 2002), 73% of colleges sampled did not allow consumption of alcohol in public places. Moreover, 36% of institutions explicitly forbade the use of alcohol references in advertising by student groups. In 78% of cases, breaches of the policy could result in suspension, and in 82% even expulsion.

Particularly amongst young people, peer pressure may play an important role in shaping the decision to binge drink. Youths are forced to spend time with the same group of individuals every day at school. This has additional strength during the teenage years, as socialising and establishing identity as part of a group is particularly important. Individuals may be willing to engage in personally costly actions to protect this. I define peer pressure as the (implicit

[1] Defined as those between eighteen and twenty-five years old.

or explicit) cost a group imposes on its members for not engaging in a particular activity.

The specific example of peer pressure that I shall discuss in this paper is that of binge drinking amongst members of sports teams. After sporting competitions, particularly when representing a university, it is common for clubs to engage in drinking games. These games are highly competitive, and involve copious amounts of alcohol. Usually, the aim is to consume your drink faster than your opponent. For example, consider a "boat race". In a boat race, each university fields a team of four players. The two teams then stand in line with a drink each. Upon hearing the command to begin, the first member of each team drinks their drink. As soon as they have finished, the next member of their team begins consuming their drink. The winners are the team who finish all four drinks in the shortest time. Such activities, although perhaps fun to begin with, soon lose their appeal. When several universities are gathered together, it is not uncommon to turn the boat race into a knock-out tournament, with finalists perhaps drinking as many as four beverages within fifteen minutes prior to reaching the final round. However, despite the inevitable feelings of illness, students still line up against members of the opposing university in the final, so as not to bring shame upon the team. Key features of this scenario are that, when presented with a member of an opposing team, a player does not know how well they can hold their alcohol. Moreover, as the pride of their team is at stake, they are likely to feel pressured into engaging in the game, as to back out immediately would look bad, and may result in some form of group punishment. The model described in the following section aims to include these features. Note that, under the hypothesis that drinking games are individually costly, nobody should engage in them in an efficient setting. Thus peer pressure introduces an inefficiency, by forcing players to get involved.

Empirical research suggests that a large part of measured peer effects may be due to self-selection of peers. Consider, for example, a person with a higher than average propensity to use drugs. It may be the case that they seek out other with similar propensities and form a group. When observing the group, an outsider would see a lot of people using drugs and may mistakenly attribute it to peer pressure. This has led to a major discrepancy between the theoretical literature on peer pressure and empirical observation, which I hope to rectify in this paper. The model I present restricts attention to situations where peer pressure is likely to be pervasive. Specifically, I model circumstances in which there are two opposing peer groups, competing against each other. I propose that under such conditions, each individual will feel pressured by his/her compatriots to engage in (individually costly) activities so as not to 'let the side down'. Competition seems to play a key role in determining the identity of the group. Although self-selection into peer groups is still an issue, I would suggest that peer effects are likely to be an important factor. In many such situations, there are numerous benefits to being a member of a particular peer group, e.g. a sense of group identity, friendship etc. The costly activity I model should be considered to be a necessary evil. The key result of the paper is a testable hypothesis regarding the minimum (and hence optimum, if the application of peer pressure is itself costly) amount of peer pressure that needs to be applied to agents in such a situation, given a particular prior distribution of players and cost of acquiring information. I also show that peer pressure results in inefficiencies related to this cost.

Although I do not claim that all binge drinking is a result of sports teams engaging in drinking games, I would suggest that a proportion of it is. Moreover, it seems plausible that competitive behaviour amongst young drinkers extends beyond drinking games. For example, individuals may be chastised, or even bullied, for not keeping up with their colleagues, or for consuming drinks with lower alcohol content. This could also be considered a form of peer pressure. However self-selection is likely to be more of an issue in this case, which is why I restrict attention to circumstances in which there is an explicit competitive aspect to drinking behaviour.

The remainder of this paper is set out as follows. In section 1.2 I review the relevant empirical and theoretical literature. In section 1.3, I describe peer pressure as a simple dynamic game of incomplete information. In section 1.4, I begin to analyse the model. under the assumption that each team applies sufficient peer pressure to ensure all its members engage in drinking games. I outline various equilibria in which signals are acquired (or not), and give conditions for their existence. I return to the idea of peer pressure in section 1.5 and attempt to derive a threshold level of peer pressure that ensures players choose to participate in the drinking game, given that they may be unaware of which equilibrium they will end up in at the moment they decide to enter. I make some concluding comments in section 1.6.

1.2 Literature Review

Economists have long recognised the importance of reference points in economic decisions. As early as the 1930s, John Maynard Keynes (Keynes 1936) suggested that workers reference around their nominal wages. He further postulated that they would not accept falls in their nominal wages, resulting in a permanent disequilibrium in the labour market. The inclusion of reference points in both macroeconomics and, more importantly for this paper, microeconomics has since flourished. Habit formation has long been an established part of the literature on intertemporal choice, helping to explain why agents would choose to consume more and save less. In a seminal paper, Kahneman and Tversky (1979) adopted a psychological approach (which later became known as behavioural economics) in order to consolidate the Allais paradox (Allais 1953) with expected utility theory, and thereby enshrined the importance of reference points. Prospect theory takes a far more flexible approach to decisions under uncertainty, emphasising that agents respond differently to lotteries depending on how they are proposed. For example, phrasing a given lottery in terms of money won may result in a completely different decision compared to when the same lottery is phrased in terms of money lost relative to the maximum possible payoff.

Peer pressure is a particularly pertinent example of a reference point. It has long been recognised by other social sciences that individuals often choose to make similar decisions to those with whom they associate. I begin this section by reviewing attempts to measure the effects of peer pressure, emphasising the problems of endogeneity arising from self-selection, and commenting on the results. I then turn attention to three main theoretical approaches explaining why agents may choose to take similar actions to their peers.

1.2.1 Empirical Work

Several attempts have been made over the last twenty years to measure the effects of peer pressure using panel data. The first major study was conducted by Evans, Oates, and Schwab 1992. In this study the authors identified potential endogeneity problems that would likely occur when attempting to come up with a measure of peer pressure. These issues arise from the fact that agents have a certain amount of choice with regard to their peers. Consequently in observing, for example, a group of drug users, it may be the case that each member of the group has a higher than average propensity for drug use. When choosing a group of people to associate with, it is very likely that these individuals will seek out peers with a similar high propensity for drug use It is therefore often very difficult to disentangle the peer effects from these individual propensities when conducting econometric analysis. What, in a simple analysis may appear to be a strong peer effect, may in fact simply be the result of individuals with a higher propensity for an activity grouping together. In their study, Evans *et al* consider a teenager's propensity to become pregnant, and estimate peer effects in a probit setting. They run two regressions, first based upon a single equation model, and then a set of simultaneous equations (where the teenager's peer group is considered to be endogenous[2]). They find that the peer group effect is significant and positive in the single equation model. However, when they endogenise the peer group, it not only becomes insignificant, but it changes sign as well. They conclude that peer effects can be attributed almost entirely to family choices in this case. They finally derive the same result for teenagers' propensity to drop out of school, again suggesting that peer effects are less prevalent than previously imagined.

Several more recent papers have found significant peer effects, even when controlling for endogeneity. Gaviria and Raphael (Gaviria and Raphael 2001) derived estimates of peer effects on teenagers decision to use drugs, drink alcohol, smoke cigarettes, attend church and drop out of school based on a two stage least squares analysis of the National Education Longitudinal Study. They claim that since students have no choice about their classmates, self-selection is less of a concern. Their results show that peer effects are positive and significant for all variables except for smoking, even when controlling for a wide variety of family characteristics. They suggest that moving a teenager from a school where none of his classmates use drugs, to one where half do, increases that probability of drug use by sixteen percentage points. A similar situation with regard to alcohol consumption increases the likelihood of drinking by roughly seventeen percentage points. The results were shown to be reasonably robust by including

[2]Instruments used for the peer effects were the metropolitan unemployment rate, median family income, the poverty rate, and the percentage of adults who completed college.

a wide variety of school effects (another source of correlation between regressors and errors).

McCarthy, Hagan, and Cohen 1998 also found significant peer effects for theft amongst street gangs in Toronto and Vancouver, although, after acknowledging the potential presence of endogeneity, they proceeded with an OLS analysis. They claimed to find little evidence of multicollinearity.

Whilst my work will no doubt suffer from a certain amount of endogeneity, I hope to minimise its effects by considering peer groups where the main attraction to members is not the costly activity they are currently engaged in. Returning to the drinking example, most students (I would argue) do not join sports clubs for drinking purposes, rather in order to keep fit, and possibly to compete for their university or college. Any self-selection is likely to be based upon sporting ability, guaranteed by the club's coaches. Although the social aspect may be appealing, clubs are more likely to attract individuals with a higher than average propensity for competition. This, if anything, is likely to ensure that members have a strong sense of loyalty towards their respective clubs, and are hence less likely to want to let the side down.

1.2.2 Theoretical Work

One argument, put forward by Sah 1991, is that peer pressure may be the manifestation of positive network externalities associated with particular forms of conformist behaviour. Consider, for example, a teenager deciding whether to start taking drugs. If he/she observes very few people taking drugs, he/she may be less inclined to begin experimenting. Given that the authorities use a fixed amount of resources on drug abuse recognition and punishment, the chance of being caught are quite high, since the resources are concentrated. Now suppose he or she observes lots of people taking drugs. In this situation, police resources will be spread thinly, and the probability of any one particular individual being caught is relatively low. One can foresee that, in the second case, taking drugs will seem far more appealing than in the first.

Sah models an individual's propensity for crime as a function of aggregate crime participation rates and the probability of punishment, both lagged over several periods. This reflects that the individual is unlikely to have perfect information about current levels, and must hence form beliefs about current participation and probability of punishment. He then shows then the propensity for crime is increasing in previous crime participation rates, and decreasing in the amount of resources spent on criminal apprehension. Sah then extends his analysis to consider a situation in which there are two distinct groups in the economy, both of which choose whether to engage in the criminal activity. Under general conditions, he then derives that there are positive spillovers between groups associated with criminal activity. Essentially, an increase in criminal participation rates for one group raises the rate of participation for the other, as the probability of being caught falls.

In Sah's model, apparent peer pressure could be viewed as the observed effects of these positive network externalities. Whilst this may be true of criminal behaviour, it is not necessarily the case for other costly activities. For example, the model fails to explain why, returning to sports clubs, agents would choose to engage in drinking games that, whilst immediately costly, cannot (by themselves) result in prison sentences.

Another model of peer pressure is Akerlof 1997. His paper presents two alternative quadratic utility functions. The first, given below, is designed to mimic the effects of status. Let x be the an individual's consumption of a particular good, and let $\overline{x}$ be the average individual consumption of that good in society.

Akerlof suggests that following utility function:

$$-d(\overline{x} - x) - ax^2 + bx + c \tag{1.1}$$

Clearly, agents gain additional utility from consuming above average quantities of the choice good. The second, more relevant model describes conformity. Akerlof suggests that agents care not only about their consumption in absolute terms, but also relative the societal average:

$$-d|x - \overline{x}| - ax^2 + bx + c \tag{1.2}$$

Given these preferences, agents have an added incentive to consume similar amounts to their peers. Akerlof shows that both these situations lead to the potential for superoptimal consumption compared with preferences that have no reference dependence, thus helping to explain why agents may engage in activities in groups that they would not necessarily undertake alone, such as drug abuse or binge drinking. One could interpret the model outlined in this paper as an attempt to find a foundation for the behaviour predicted by these utility functions.

Finally, sociologists have been concerned with peer effects for a considerable length of time. Many sociologists criticise game theory for the assumption that agents' only aim is to maximise their own personal utility. A classic example of such behaviour is the prisoners' dilemma. The socially optimal outcome is for each player to co-operate, thereby ensuring that neither player serves a long jail sentence. However, this is not a Nash equilibrium, as each has an incentive to deviate and defect. *Social dilemma theory* argues that, rather than being myopic, players are able to reason collectively. So called *collective rationality*, they argue, ensures that players not only aim to maximise their own utility, but also social welfare. Experimental economists have added weight to this idea, providing significant evidence of a fairness motive in simple games (see, for example, Hoffman, McCabe, and Smith 1996). With regard to peer pressure, this conflict between maximising individual and collective welfare may help to explain why agents choose to engage in actions that are personally costly, but make the group better off.

1.3 The Model

I model peer pressure in the context of a dynamic game with incomplete information and observed actions. There is a large (finite) number of players, $i = 1, ..., 2N$. Players are patient, i.e. do not discount future stages of the game. In the first stage, Nature allocates each player a type drawn from the set, $\Theta_i = \{$ *"Strong"*, *"Weak"*$\}$. *"Strong"* types have a higher alcohol tolerance. Each players' type is private information of the player, but Θ_i is common knowledge. Players have common priors over $\Theta = \times_{i=1}^{N} \Theta_i$, which assign probability p to the event that a particular player is *"Strong"* (and consequently $1 - p$ to them being *"Weak"*). Nature also allocates each player randomly to one of two opposing sports teams, both of size N. Given the common priors assumption, each team has a proportion p of *"Strong"*-types. A player's team automatically learns their type. However it remains unknown to their opponents.

Having been informed of their type and team, players then choose whether to continue to engage in the drinking game. If the player chooses to withdraw, the game ends and they pay a cost, $c(\theta_i, p)$, where $\theta_i \in \Theta_i$ is the player's type. This reflects the *peer pressure* their team inflicts for their refusal to continue. It is assumed that there is a small but increasing cost to the team from applying this pressure. Given this, the team will want to apply the minimum amount possible to ensure entry.

Why might a group apply such peer pressure? As previously mentioned, empirical research suggests that peer pressure is not ubiquitous. Why then would it be particularly pervasive amongst competitive peer groups? In situations where peer groups compete, there is a feeling that the pride of the group is at stake. Moreover, each member of the group cares about the group's reputation, and thus suffers disutility when it is damaged. Consequently, any player who does not engage in the drinking game inflicts a negative externality upon the other members of their team[3]. Avoidance of this externality may be sufficient incentive for members of the team to pressurise the player into entering the game. In particular, members of the team may then partially ostracise or ridicule the player. The fear of this occurrence is sufficient to ensure that the player participates in the game.

In section 1.5, I will return to the cost associated with peer pressure, in an attempt to discern the threshold amount of pressure that the team needs to apply in order to ensure continuation.

Suppose that a player decides to remain in the game. Nature randomly pairs them with another player from the opposing team of unknown type. Having been paired with an opponent, the player then has the opportunity to acquire a signal about their rival. For example, by offering to buy their opponent a drink, the player can glean important information about their attitude towards alcohol by the drink that they request. If they request a double scotch, it is likely that the player is *"Strong"*, as it would be too costly for a *"Weak"*-type to attempt to drink it. It is conceivable that this information will be sufficient, given the simple two types model, to accurately assess their type. Note, also, that the opponent observes that the signal has been acquired. This type of signalling is a strengthening of the concept of *hard information* (Nelson 1974).

[3] One could also argue that they would inflict a positive externality on members of the opposing peer group.

Hard information is something that cannot be faked, thus removing problems associated with truthful revelation. However, with hard information the signaller has the choice as to whether they reveal their signal. In this model, that is not the case. The purpose of acquiring the signal is that, by doing so, the player is better prepared to play the drinking game, where their payoffs will be type-dependent. However, their opponent is able to condition their drinking strategy upon whether a signal was sought.

Suppose that two players i and j are paired against each other. Their action in this stage is $s_i, s_j \in S \equiv \{Acquire, Not\ Acquire\}$.

Finally, players engage in the drinking game. They play a simple simultaneous moves game, choosing actions $a_i, a_j \in A \equiv \{C, D\}$. C could be described as *"Chicken Out"*, i.e. stay for the first drink, make apologies and leave. The action D would refer to the alternative, namely entering into a heavy drinking session. In each possible game, each type of player has a strictly dominant action. However, the dominant action is different, depending upon the type of their opponent. Specifically, I assume that any player, when faced with *"Strong"* opposition, would rather chicken out (take action C) than engage in a lengthy, and costly game. A *"Weak"* player would be soundly beaten. A *"Strong"* player would find it too costly (both financially, and in terms of the ensuing hangover). However, any player facing a *"Weak"* opponent will strictly prefer to engage them (take action D). *"Strong"* players will gain a victory for certain. Moreover, their pride would be hurt if they chickened out when facing a weaker opponent. *"Weak"* players will also want to play, as it provides an opportunity to engage in the game relatively safely.

Consider a *"Strong"* player in the drinking game. Suppose that Nature has paired him/her with a *"Strong"* opponent. The payoffs are described below:

		"Strong"	
		C	D
"Strong"	C	s_1^S, s_1^S	s_2^S, s_3^S
	D	s_3^S, s_2^S	s_4^S, s_4^S

I restrict the set of possible payoff so that they satisfy the following conditions.

Assumption 1 ((*"Strong"*, *"Strong"*) Payoffs) *The payoffs in the drinking game between two "Strong" players satisfy the following ordering:*

$$s_4^S < s_2^S < s_3^S < s_1^S < 0 \tag{1.3}$$

Thus taking action C is a strictly dominant strategy for a "Strong" player when faced with a "Strong" opponent in the perfect information subgame.

The two other possible player combinations are summarised below:

		"Weak"	
		C	D
"Strong"	C	s_1^W, w_1^S	s_2^W, w_3^S
	D	s_3^W, w_2^S	s_4^W, w_4^S

		"Weak"	
		C	D
"Weak"	C	w_1^W, w_1^W	w_2^W, w_3^W
	D	w_3^W, w_2^W	w_4^W, w_4^W

I similarly restrict the set of possible payoffs in the following ways

Assumption 2 ((*"Strong"*, *"Weak"*) Payoffs) *The payoffs in the final game between a "Strong" and a "Weak" player satisfy the following orderings:*

$$s_2^W < s_4^W < s_1^W < s_3^W < 0 \tag{1.4}$$

$$w_4^S < w_2^S < w_3^S < w_1^S < 0 \tag{1.5}$$

Thus a "Strong" player strictly prefers D to C when facing a "Weak" opponent, whereas a "Weak" player strictly prefers C to D when facing a "Strong" opponent

Assumption 3 ((*"Weak"*, *"Weak"*) Payoffs) *The payoffs in the final game between two "Weak" players satisfy the following ordering:*

$$w_1^W < w_3^W < w_2^W < w_4^W < 0 \tag{1.6}$$

Thus taking action D is a strictly dominant strategy for a "Weak" player when faced with a "Weak" opponent in the perfect information subgame.

Summarising, the model proceeds as follows:

1. Players are allocated to one of two teams. Both they and their team are informed of their type.

2. Nature pairs members of opposing teams at random.

3. Players have the opportunity to withdraw from the game:

   (a) If they withdraw, they suffer disutility associated with peer pressure and the game ends.

   (b) If they remain in the game, the game continues.

4. Players can then acquire an accurate signal about their opponent. Their opponent is informed if a signal is sought.

5. Players engage in a simultaneous drinking game, and payoffs are realised.

1.4 Equilibria

In what follows, I focus on pure strategy symmetric perfect Bayesian equilibria. In this setting, such an equilibrium will consist of:

1. A peer pressure cost for every type of player, $c(\theta, p) \geq 0$, conditional on the common priors of the teams.

2. A decision for each type regarding whether to withdraw from the game, conditional on the peer pressure costs. and the common priors.

3. A decision for each type regarding whether to acquire a signal, $s(\theta_i, c(\theta_i, p), p) \in S$, conditional on the peer pressure costs and the common priors.

4. An action for each type in the drinking game, $a(s_i, s_j, p) \in A$, conditional on any signals acquired or sought and the common priors.

For the moment, I assume that each team threatens sufficient peer pressure to ensure that no team members withdraw. The objective of the next section will be to discern the minimum amount of pressure required to ensure that this is the case. As such, the remainder of this section will focus on the final two components of any equilibrium, namely the decision to acquire signals and the subsequent action in the drinking game. Since all team members continue onto this stage, a player's initial beliefs about their opponent's type upon being paired with them are identical to their priors.

It is possible to distinguish between two types of equilibria, namely those in which signals are acquired by at least one type of player, and those in which neither type requests a signal. I consider each in turn, and derive necessary and sufficient conditions for their existence.

1.4.1 Acquired Signal Equilibria

For every possible priors, there exists an equilibrium in which all players acquire a signal about their opponent. These are summarised in the following:

Proposition 1 (All Acquire Signals) *Consider the drinking game model with sufficient peer pressure. For all $p \in [0,1]$ there exists an equilibrium in which both types of player acquire a signal. When facing a "Strong" opponent, they subsequently take action C. Otherwise they subsequently take action D. Out of equilibrium, should they fail to acquire a signal, then:*

1. *If $p \in \left[0, \min\left\{\frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W}, \frac{w_4^W - w_2^W}{w_2^S - w_4^S + w_4^W - w_2^W}\right\}\right]$, then both types will subsequently take action D;*
2. *If $p \in \left(\min\left\{\frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W}, \frac{w_4^W - w_2^W}{w_2^S - w_4^S + w_4^W - w_2^W}\right\}, \max\left\{\frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W}, \frac{w_4^W - w_2^W}{w_2^S - w_4^S + w_4^W - w_2^W}\right\}\right)$, then*
   - *(a) If $\frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W} < \frac{w_4^W - w_2^W}{w_2^S - w_4^S + w_4^W - w_2^W}$ then "Strong" players will take action C, whereas "Weak" players will take action D; or*
   - *(b) If $\frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W} > \frac{w_4^W - w_2^W}{w_2^S - w_4^S + w_4^W - w_2^W}$ then "Strong" players will take action D, whereas "Weak" players will take action C; or*

3. *If* $p \in \left[\max\left\{\frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W}, \frac{w_4^W - w_2^W}{w_2^S - w_4^S + w_4^W - w_2^W}\right\}, 1\right]$*, then both types will subsequently take action* C.

Proof. See Appendix 1.A. ■

By acquiring a signal, each player is able to act optimally against their opponent. Both their type and that of their opponent become common knowledge between them. They consequently play the unique Nash equilibrium of whichever payoff matrices they face. Out of equilibrium, a player who does not acquire a signal knows that their opponent knows their type, and will take a dominant action in the drinking game. Their best response thus depends upon both the payoffs they face, and their priors regarding the type of their opponent. If they believe it sufficiently likely that their opponent is a *"Weak"* player, they will optimally take action D. Conversely, if they believe that their opponent is *"Strong"* with sufficiently high probability then they will take action C.

One other symmetric equilibrium exists in which a signal is acquired, outlined below:

Proposition 2 (*"Strong"* Acquire Signals) *Consider the drinking game model with sufficient peer pressure. For all* $p \in [0, 1]$ *there exists an equilibrium in which only "Strong" players acquire a signal. In such an equilibrium:*

1. *"Strong" players will take action* C *if they face a "Strong" opponent, and* D *otherwise; and*

2. *"Weak" players will take action* C *if their opponent acquired a signal, and take action* D *otherwise.*

Proof. See Appendix 1.B. ■

Given the equilibrium strategies, a *"Weak"* player is always strictly better off when he does not acquire a signal. If his opponent is *"Strong"*, then there is no benefit associated with acquiring a signal. In equilibrium, he learns that his opponent is *"Strong"* by the very fact that they purchase a signal. Moreover, his opponent learns his type, so he cannot convince his opponent that he is in fact *"Strong"*. Conversely, if his opponent is *"Weak"*, purchasing a signal will make the opponent believe that he is *"Strong"*. Thus his opponent will take action C. However, by assumption 3, he prefers it when the opponent plays D. As such, a *"Weak"* player would never choose to acquire a signal.

If a *"Strong"* player were not to acquire a signal, her payoff would be unaffected when facing a *"Strong"* opponent. As it is common knowledge between them that her opponent acquired a signal, both players know their opponent's type with probability one. However, when facing a *"Weak"* opponent, she would lose out. By not acquiring a signal, she would lead them to believe that she is also *"Weak"*. This would result in her opponent taking action D, rather than C. This results in her receiving a lower payoff, as she always prefers that her opponent chicken out in the drinking game. Consequently, her expected payoff from not acquiring a signal is also strictly lower than that from acquiring a signal, supporting the equilibrium strategies.

This proves to be the only equilibrium in which types separate over the acquisition of a signal. In particular:

Proposition 3 (*"Weak"* Do Not Acquire Signals) *Consider the drinking game model with sufficient peer pressure. There exists no pure strategy symmetric perfect Bayesian equilibria in which "Weak" players acquire a signal and "Strong" players do not.*

Proof. See Appendix 1.C. ■

The lack of equilibria stems from the fact that *"Weak"* players wish to deceive *"Strong"* players. Since *"Strong"* players do not acquire signals, their beliefs about their opponent are formed by whether their opponent acquired a signal from them. Consequently, in an equilibrium in which *"Weak"* players acquire signals, a *"Strong"* player will assume that, if their opponent does not acquire a signal, their opponent is also *"Strong"*. As a result, they will take action C, rather than D, in the drinking game. Given the payoff structure when a *"Strong"* player meets a *"Weak"* player, the *"Weak"* player is strictly better off. Conversely, if their opponent is a *"Weak"* player, they acquire a signal in the proposed equilibrium. As such, they respond optimally to the fact that they are also a *"Weak"* player, and their payoff is unaffected by the choice not to acquire a signal. *"Weak"* players therefore receive a higher expected payoff through not acquiring a signal, contradicting the equilibrium strategies proposed.

1.4.2 No Signal Equilibria

There are, of course, several other equilibria in which no players acquire signals about their opponent. However, these equilibria are not supported for every possible set of priors. Nevertheless, they are important if the teams are going to apply sufficient peer pressure to ensure that their members always enter into drinking games. Since all of these equilibria involve pooling at the signal acquisition stage, team members must rely on their common priors when deciding upon which action to take. As such, these equilibria fall broadly into two categories. Firstly, those supported by p close to zero or one. In these cases, players are sufficiently confident about the type of their opponent to induce pooling. Secondly, there are those equilibria supported by p around one half. In these cases, there is sufficient uncertainty to warrant separation. Considering each in turn:

Proposition 4 (Pooling No Signal Equilibria) *Consider the drinking game model with sufficient peer pressure. The following equilibria exist, such that no player acquires a signal:*

1. *If* $p \in \left[\max\left\{\frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W}, \frac{w_3^W - w_1^W}{w_1^S - w_3^S + w_3^W - w_1^W}\right\}, 1\right]$*, all players take action* C*;*

and

2. *If $p = 0$, all player take action D.*

Proof. See Appendix 1.D. ■

Clearly, if p is sufficiently large, then players are confident that they are facing a *"Strong"* opponent. As such, they opt to take action C. This is supported in equilibrium by *"Strong"* players' out of equilibrium beliefs that anyone who purchases a signal must be *"Weak"*. They consequently then choose action D. Since both types of player prefer a *"Strong"* player to take action C, neither type has an incentive to deviate by acquiring a signal. If $p = 0$, players know almost surely that their opponent is a *"Weak"*. Consequently, they optimally choose action D. This is supported by a range of out of equilibrium beliefs. If players believe that an opponent who acquires a signal is *"Weak"* , then they will play D irrespective. As such, there is no incentive to acquire a signal. If, instead, they believe them to be *"Strong"*, they will choose action C. Since, having acquired a signal, a player's opponent is almost surely *"Weak"*, all types would prefer they play D, providing a strict incentive not to acquire a signal.

Finally, turning attention to separating equilibria:

Proposition 5 (Separating No Signal Equilibria) *Consider the drinking game model with sufficient peer pressure. If:*

$$p \in \left[\frac{s_4^W - s_2^W}{s_1^S - s_3^S + s_4^W - s_2^W}, \frac{w_4^W - w_2^W}{w_1^S - w_3^S + w_4^W - w_2^W}\right]$$

then there exists an equilibrium in which no players acquire signals and:

1. *"Strong" players will take action C; and*

2. *"Weak" players will take action D.*

Out of equilibrium beliefs are such that any player who faces an opponent who acquires a signal takes action D.

Proof. See Appendix 1.E. ■

This equilibrium exists only if $\frac{s_4^W - s_2^W}{s_1^S - s_3^S + s_4^W - s_2^W} < \frac{w_4^W - w_2^W}{w_1^S - w_3^S + w_4^W - w_2^W}$. *"Strong"* players are sufficiently convinced (p is high enough, given their payoffs) that their opponent is also *"Strong"* that they choose action C. Concurrently, p is low enough that, given their payoffs, *"Weak"* players still prefer to take action D. They are still sufficiently convinced that their opponent is also *"Weak"*.

Under a fairly innocuous assumption, this proves to be the unique separating no signal equilibrium:

Corollary 1 *Consider the drinking game model with sufficient peer pressure. If:*

$$w_4^W - w_1^W < w_1^S - w_2^S$$

then Proposition 5 fully characterises the set of pure strategy separating equilibria in which no players acquire a signal.

Proof. See Appendix 1.F. ■

The assumption in the corollary simply states that a *"Weak"* player has more to lose by making a mistake against a *"Strong"* opponent than can be gained against a *"Weak"* one. As such, when faced with an equilibrium in which *"Strong"* players are taking action D and other *"Weak"* players are taking action C, they always prefer to acquire information about their opponent.

1.5 Peer Pressure

Since entry into costly drinking games is often observed, it must be the case that each team applies sufficient peer pressure to ensure that the its members are weakly better off entering into a drinking contest than withdrawing. Whilst a simple approach would be to make the cost of withdrawing greater than the minimum payoff it is possible to receive in any game, this may prove costly for the team itself. Moreover, given the common priors, the peer pressure threatened can be conditioned on the realisation of p. This may ensure that a smaller amount of peer pressure needs to be applied, reducing the cost to the team.

Given the existence of multiple equilibria, it may not be possible for the team to distinguish precisely which equilibrium each pair is playing. As such, some individuals may need more peer pressure applied than others, as they are entering into a lower payoff equilibrium. If the team is unable to observe which pairs are going to play which equilibrium, or if the equilibrium is not determined for each team member prior to their being paired, it may be necessary for the team to consider the worst case scenario. To ensure that no individual will withdraw from the drinking game, it must therefore threaten a peer pressure cost at least as great as the worst payoff an individual of a given type can receive, given the distribution of types.

So, in order to calculate the amount of peer pressure required, it is necessary to compute the expected payoff to an individual of each type for every equilibrium. Clearly, all equilibria in which at least one type acquires a signal result in the same expected payoff. Prior to entering the drinking game, players know the type of their opponent almost surely. Consequently, they know which set of payoffs they face, and best respond accordingly. Given that each set of payoffs incorporates a dominant strategy, the expected payoff for *"Strong"* is:

$$ps_1^S + (1-p)\,s_3^W \tag{1.7}$$

Similarly for *"Weak"* players it is:

$$pw_2^S + (1-p)\, w_4^W \tag{1.8}$$

This reflects the fact that *"Strong"* players optimally play C against *"Strong"* opponents and D against *"Weak"* ones, whereas *"Weak"* players optimally choose D and C respectively.

If no players acquire signals, their payoff depends upon which equilibrium is played. Clearly, these equilibria do not always exist, and so the team may be confident about applying less peer pressure for some sets of priors. In the two pooling equilibria, the payoffs depend upon whether players are sufficiently convinced that their opponent is *"Strong"* or *"Weak"*. If $p = 0$, *"Strong"* players earn s_4^W whereas *"Weak"* players earn w_4^W. Conversely if $p \geq \max\left\{\frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W}, \frac{w_3^W - w_1^W}{w_1^S - w_3^S + w_3^W - w_1^W}\right\}$, *"Strong"* players earn $ps_1^S + (1-p)\, s_1^W$ whereas *"Weak"* players earn $pw_1^S + (1-p)\, w_1^W$. Finally, in a separating no signal equilibrium, *"Strong"* players earn $ps_1^S + (1-p)\, s_2^W$ whereas *"Weak"* players earn $pw_3^S + (1-p)\, w_4^W$.

In order to fully describe the peer pressure function, $c(\theta, p)$, it is necessary to discuss several cases. These are defined by the relationship between the sets of priors for which the various equilibria exist. In order to expedite this discussion, it is useful to define the following parameters:

Definition 1 *The boundary values for the sets of priors over which the various equilibria of the drinking game model exist are given by:*

- $q_1 \equiv \max\left\{\frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W}, \frac{w_3^W - w_1^W}{w_1^S - w_3^S + w_3^W - w_1^W}\right\}$*;*
- $q_2 \equiv \frac{s_4^W - s_2^W}{s_1^S - s_3^S + s_4^W - s_2^W}$*; and*
- $q_3 \equiv \frac{w_4^W - w_2^W}{w_1^S - w_3^S + w_4^W - w_2^W}$.

I will assume throughout that $q_2 < q_3$, so there are values of p for which the separating no signal equilibrium exists. Ambiguity remains, however, regarding the relationship between q_1, the minimum value of p for which the pooling no signal equilibrium exists, and q_2 and q_3. Considering each possibility in turn:

Case 1 *If $q_1 < q_2$ then the peer pressure functions for each type of player are given by:*

$$c\left(\text{"Strong"}, p\right) = \begin{cases} s_4^W & p = 0 \\ ps_1^S + (1-p)\, s_3^W & p \in (0, q_1) \\ ps_1^S + (1-p)\, s_1^W & p \in [q_1, q_2) \\ ps_1^S + (1-p)\, s_2^W & p \in [q_2, q_3] \\ ps_1^S + (1-p)\, s_1^W & p \in (q_3, 1] \end{cases}$$

$$c\left(\text{"Weak"}, p\right) = \begin{cases} pw_2^S + (1-p)\, w_4^W & p \in [0, q_1) \\ \min\left\{pw_2^S + (1-p)\, w_4^W, pw_1^S + (1-p)\, w_1^W\right\} & p \in [q_1, 1] \end{cases}$$

Case 2 *If $q_2 < q_1 < q_3$ then the peer pressure functions for each type of player are given by:*

$$c\left(\text{"Strong"}, p\right) = \begin{cases} s_4^W & p = 0 \\ ps_1^S + (1-p)\, s_3^W & p \in (0, q_2) \\ ps_1^S + (1-p)\, s_2^W & p \in [q_2, q_3] \\ ps_1^S + (1-p)\, s_1^W & p \in (q_3, 1] \end{cases}$$

$$c\left(\text{"Weak"}, p\right) = \begin{cases} pw_2^S + (1-p)\, w_4^W & p \in [0, q_1) \\ \min\left\{pw_2^S + (1-p)\, w_4^W, pw_1^S + (1-p)\, w_1^W\right\} & p \in [q_1, 1] \end{cases}$$

Case 3 *If $q_3 < q_1$ then the peer pressure functions for each type of player are* *given by:*

$$c\left(\text{"}Strong\text{"}, p\right) = \begin{cases} s_4^W & p = 0 \\ ps_1^S + (1-p)\, s_3^W & p \in (0, q_2) \\ ps_1^S + (1-p)\, s_2^W & p \in [q_2, q_3] \\ ps_1^S + (1-p)\, s_3^W & p \in (q_3, q_1) \\ ps_1^S + (1-p)\, s_1^W & p \in [q_1, 1] \end{cases}$$

$$c\left(\text{"}Weak\text{"}, p\right) = \begin{cases} pw_2^S + (1-p)\, w_4^W & p \in [0, q_1) \\ \min\left\{pw_2^S + (1-p)\, w_4^W, pw_1^S + (1-p)\, w_1^W\right\} & p \in [q_1, 1] \end{cases}$$

For *"Strong"* players, the best outcome would be to end up in an equilibrium in which at least one type of player acquires a signal. This enables them to select the dominant action against each type of opponent. In cases 1 and 2, this is the unique equilibrium when $p \in (0, \min\{q_1, q_2\})$. In case 3, it is the unique equilibrium for $p \in (0, q_2) \cup (q_3, q_1)$. As such, over these sets, the teams do not need to threaten large amounts of peer pressure to ensure that *"Strong"* players enter into the drinking game. When $p = 0$, however, they could find themselves in a pooling no signal equilibrium. They almost surely face a *"Weak"* opponent, who expects that they are also *"Weak"*. They consequently receive a lower payoff (s_4^W) than they would receive from an acquired signal equilibrium (s_3^W), and the peer pressure threatened must reflect this.

Finally, when $p \geq \min\{q_1, q_2\}$, other no signal equilibria become possible. In any of these equilibria, *"Strong"* players take action C. Whilst this is dominant when facing a *"Strong"* opponent (earning them s_1^S), it results in a strictly lower payoff when their opponent is *"Weak"*. As a result, more peer pressure is required. The exact amount depends upon how *"Weak"* players behave. If $p \geq q_1$, then equilibria exist in which *"Weak"* players also choose C. Although this results in a lower payoff than an acquired signal equilibrium, it still proves better than when

"Weak" opponents choose D. When $p \in [q_2, q_3]$, *"Strong"* players can receive $s_2^W < s_1^W$ instead when facing a *"Weak"* opponent.

For *"Weak"* players it is unclear what the best possible outcome is. For $p < \min\{q_1, q_2\}$ (cases 1 and 2) or $p \in [0, q_2) \cup (q_3, q_1)$ (case 3), the payoff players can expect to receive is unique, namely $pw_2^S + (1-p)\, w_4^W$. Whilst this is the best payoff they could hope for against a *"Weak"* opponent, they would be strictly better off if *"Strong"* opponents chose C rather than D. This is precisely the action they take in no signal equilibria, unless $p = 0$. If $p > q_1$, there exist equilibria in which all players take action C. This improves their payoff against *"Strong"* players, but reduces it against *"Weak"* opponents. Depending upon whether the reduction dominates the improvement, the amount of peer pressure the team needs to threaten may increase.

Note that *"Weak"* players receive a strictly higher payoff in a separating no signal equilibrium than in an acquired signal equilibrium. Whilst *"Weak"* opponents take the same action in both types of equilibria, *"Strong"* players are sufficiently convinced that their opponent will be *"Strong"* to take action C. Since this yields a higher payoff for a *"Weak"* player, separating no signal equilibria have no effect upon the amount of peer pressure each team must threaten its weaker members with to ensure that they enter the drinking game.

1.6 Conclusions

Peer pressure is the minimum cost that a team must threaten inflict upon a member to ensure they enter into a costly activity. Empirical work has shown that this is not as widespread as hitherto imagined. This results from the fact that individuals often have some degree of choice over who they associate with, leading to self-selection biases. In this paper, I have focused on a situation in which I believe peer pressure still plays a large role in conformist behaviour: those where peer groups compete. Under such conditions, there is likely to be pressure applied by teammates to ensure the individuals do not damage the pride of the peer group.

The amount of peer pressure that needs to be applied is heavily depend upon the magnitude of the payoffs associated with the costly activity, and on the prior distribution of players. In particular, as uncertainty regarding the type of an opponent increases, so too does the cost a team must threaten to inflict upon a player to ensure that they engage in a costly activity. This is especially interesting, as the agents in my model are risk neutral. By allowing individuals to communicate with their opponents, it is possible to reduce the size of the peer pressure, by enabling them to gain information about the type of their opponent. This facilitates players choosing a better strategy in the costly activity, so as to minimise their losses. This hypothesis could easily be tested in an experimental setting, or by attempting to derive a measure of the costs inflicted upon group members who do not engage in activities and comparing it to the payoffs of those who do.

Peer pressure results in inefficiency by ensuring that players enter into an activity that is personally costly. If no peer pressure is applied, players have no incentive to enter into the game, leading to an efficient outcome whereby all players receive a zero payoff. However, if peer pressure is widespread, the threat of peer pressure is sufficient to ensure that all players enter into the activity, providing negative payoffs. The social cost is reduced significantly if players are able to communicate with each other.

One future application of this model would be to coercive recruitment practices used by street gangs during gang wars (see, for example, Jankowski 1991). During these periods of violence, it is not uncommon for gangs to change their recruitment strategy. In peacetime, the emphasis is on getting individuals who are happy to join, and who will be loyal to the gang. However, when a war occurs, gangs massively increase their demand for labour in order to protect their territory. These recruits tend to be threatened or bullied into joining for the duration of the war. If one views fights within a gang war in a similar manner to a drinking game in this paper, then the amount of coercion required to recruit members is equivalent to the amount of peer pressure teams threaten their members with. A further extension would be to consider the optimal level of coercion a gang applies, given that the number of victories or losses it suffers within a war will be a function of both the number of members it has, and their ability to fight.

Acknowledgements

Thanks are owed to Leonardo Felli for numerous suggestions and help. Comments by participants at the LSE Microeconomic Theory Work In Progress seminar have proved useful.

1.A Proof of Proposition 1

Consider the proposed equilibrium. It is necessary to show that:

1. Both types acquire information; and
2. Should a player not acquire information, the action postulated in the equilibrium are best responses to the actions of their opponent, given their beliefs about their opponent's type.

In any equilibrium in which all players acquire a signal, players know the type of their opponent. As such, they can best respond in the perfect information game. If all players acquire signals, the Nash equilibrium of each perfect information subgame is reached. *Ex ante*, *"Strong"* players who acquire a signal receive:

$$ps_1^S + (1-p)\, s_3^W$$

For any p this is the highest they can receive. Thus, given the actions of the other players, *"Strong"* players will always wish to acquire signals.

"Weak" players, on the other hand expect to receive:

$$pw_2^S + (1-p)\, w_4^W$$

Since all players play D against a *"Weak"* player, this once again constitutes the highest payoff *"Weak"* players can expect to receive *ex ante*. Thus acquiring a signal is always a best response to the other players' postulated strategies.

It remains to prove that the out of equilibrium actions are also best responses. With probability one, a player's opponent will acquire a signal. Since each perfect information subgame has a dominant action, their opponent's action is unaffected by the player's lack of signal. *"Strong"* players who do not acquire a signal know that their opponent will play C. The expected payoff from taking action C is thus:

$$ps_1^S + (1-p)\, s_1^W$$

whereas the payoff from taking action D yields expected payoff:

$$ps_3^S + (1-p)\, s_3^W$$

A *"Strong"* player who has not acquired a signal will therefore taking action C if and only if:

$$\begin{aligned} ps_1^S + (1-p)\, s_1^W &\geq ps_3^S + (1-p)\, s_3^W \\ &\Leftrightarrow \;\; p \geq \frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W} \in (0,1) \end{aligned}$$

Similarly, *"Weak"* players who do not acquire a signal know that their opponent will play D. Their expected payoff from taking action C is:

$$pw_2^S + (1-p)\, w_2^W$$

whereas taking action D gives:

$$pw_4^S + (1-p)\, w_4^W$$

A *"Weak"* player who has not acquired a signal will therefore taking action C if and only if:

$$\begin{aligned} pw_2^S + (1-p)\, w_2^W &\geq pw_4^S + (1-p)\, w_4^W \\ &\Leftrightarrow \;\; p \geq \frac{w_4^W - w_2^W}{w_2^S - w_4^S + w_4^W - w_2^W} \in (0,1) \end{aligned}$$

So if $p < \min\left\{\frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W}, \frac{w_4^W - w_2^W}{w_2^S - w_4^S + w_4^W - w_2^W}\right\}$, both types take action D if they fail to acquire a signal. If $p > \max\left\{\frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W}, \frac{w_4^W - w_2^W}{w_2^S - w_4^S + w_4^W - w_2^W}\right\}$, both types take action C. Finally, if $p \in \left[\min\left\{\frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W}, \frac{w_4^W - w_2^W}{w_2^S - w_4^S + w_4^W - w_2^W}\right\}, \max\left\{\frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W}, \frac{w_4^W - w_2^W}{w_2^S - w_4^S + w_4^W - w_2^W}\right\}\right]$

and:

- $\frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W} < \frac{w_4^W - w_2^W}{w_2^S - w_4^S + w_4^W - w_2^W}$ then *"Strong"* players take action C, whereas *"Weak"* players take action D; or

- $\frac{w_4^W - w_2^W}{w_2^S - w_4^S + w_4^W - w_2^W} < \frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W}$ then *"Strong"* players take action D, whereas *"Weak"* players take action C.

This completes the proof. ■

1.B Proof of Proposition 2

In the proposed equilibrium, each *"Strong"* player learns the type of their opponent. Moreover, since only *"Strong"* players acquire signals, *"Weak"* players update their beliefs about their opponent based upon whether they received a signal, such that:

$$\mu^W \left(\text{Opponent's } \theta = \text{"Strong"} \,|\text{Opponent acquired signal}\right) = 1$$

Again, since each type has a dominant action in each perfect information subgame, the equilibrium payoffs are as before. For *"Strong"* players, it is:

$$ps_1^S + (1-p)\, s_3^W$$

and for *"Weak"* players it is:

$$pw_2^S + (1-p)\, w_4^W$$

If a *"Strong"* player does not acquire a signal, their payoff is not affected if their opponent is also *"Strong"*. However, if their opponent is a *"Weak"* player, they will be able to convince them that they are also *"Weak"*. Their opponent will then play D rather than C. Their expected payoff from this strategy is:

$$ps_1^S + (1-p)\, s_4^W$$

Clearly, since $s_3^W > s_4^W$ by Assumption 2, *"Strong"* players strictly prefer to acquire a signal in this equilibrium for any p.

If a *"Weak"* player acquires a signal, a *"Weak"* opponent will believe that they are *"Strong"*. They will thus choose action C. *"Strong"* opponents also acquire a signal, and consequently do not change their strategy. The expected payoff from acquiring a signal is thus:

$$pw_2^S + (1-p)\, w_3^W$$

By Assumption 3, $w_3^W < w_4^W$, so *"Weak"* players also strictly prefer playing their equilibrium strategy. This completes the proof. ■

1.C Proof of Proposition 3

Consider an equilibrium in which only *"Weak"* players acquire signals. I show that *"Weak"* players strictly prefer not to acquire signals. For a *"Strong"* player:

$$\mu^S\left(\text{Opponent's } \theta = \text{\textit{“Strong”}}\,|\text{Opponent acquired signal}\right) = 0$$

Consequently, when facing an opponent who has acquired a signal, they optimally take action D. Otherwise they take action C. The equilibrium payoff for a *"Weak"* player is:

$$pw_2^S + (1-p)\, w_4^W$$

Conversely, if they do not acquire a signal, they receive:

$$pw_3^S + (1-p)\, w_4^W$$

By Assumption 2, $w_3^S > w_2^S$, so *"Weak"* players prefer a unilateral deviation from their equilibrium strategy. No such equilibrium can exist. This completes the proof. ■

1.D Proof of Proposition 4

Suppose that no player acquires a signal. Then, for both types of player, their beliefs are such that:

$$\mu\left(\text{Opponent's } \theta = \text{“}\textit{Strong}\text{”} \,|\text{Opponent did not acquired signal}\right) = p$$

Suppose, firstly, that all players are taking action C. The equilibrium payoff for a *“Strong”* player is:

$$ps_1^S + (1-p)\, s_1^W$$

Conversely, were they to take action D,they would earn:

$$ps_3^S + (1-p)\, s_3^W$$

Following the equilibrium strategy yields a higher payoff if and only if:

$$\begin{aligned} ps_1^S + (1-p)\, s_1^W &\geq ps_3^S + (1-p)\, s_3^W \\ &\Leftrightarrow p \geq \frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W} \in (0,1) \end{aligned}$$

Similarly, a *“Weak”* player's expected equilibrium payoff is:

$$pw_1^S + (1-p)\, w_1^W$$

whereas a unilateral deviation to D would earn:

$$pw_3^S + (1-p)\, w_3^W$$

Following the equilibrium strategy yields a higher payoff if and only if:

$$p \geq \frac{w_3^W - w_1^W}{w_1^S - w_3^S + w_3^W - w_1^W} \in (0,1)$$

So, if $p \geq \max\left\{\frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W}, \frac{w_3^W - w_1^W}{w_1^S - w_3^S + w_3^W - w_1^W}\right\}$, all players take action C, as required, assuming no player acquires a signal. Acquiring a signal is an out of equilibrium action. Setting:

$$\begin{aligned}
\mu^S\left(\text{Opponent's } \theta = \text{“\textit{Strong}”} \,|\text{Opponent acquired signal}\right) &= 0 \\
\mu^W\left(\text{Opponent's } \theta = \text{“\textit{Strong}”} \,|\text{Opponent acquired signal}\right) &= 1
\end{aligned}$$

ensures that the equilibrium is always supported. *“Strong”* players take action D against an opponent who has acquired a signal, whereas *“Weak”* players take action C. In particular, by acquiring a signal, a *“Strong”* player earns:

$$ps_2^S + (1-p)\, s_3^W$$

Not acquiring a signal is therefore optimal if:

$$\begin{aligned}
ps_1^S + (1-p)\, s_1^W &\geq ps_2^S + (1-p)\, s_3^W \\
&\Leftrightarrow \; p \geq \frac{s_3^W - s_1^W}{s_1^S - s_2^S + s_3^W - s_1^W}
\end{aligned}$$

By Assumption 1, $s_2^S < s_3^S$, so $\frac{s_3^W - s_1^W}{s_1^S - s_2^S + s_3^W - s_1^W} < \frac{s_3^W - s_1^W}{s_1^S - s_3^S + s_3^W - s_1^W}$ and acquiring information is never optimal over the support of the equilibrium for *“Strong”* players. Similarly, the expected payoff from acquiring a signal for a *“Weak”* player is:

$$pw_2^S + (1-p)\, w_3^W$$

Not acquiring a signal is optimal if and only if:

$$p \geq \frac{w_3^W - w_1^W}{w_1^S - w_2^S + w_3^W - w_1^W}$$

By Assumption 2, $w_2^S < w_3^S$, so $\frac{w_3^W - w_1^W}{w_1^S - w_2^S + w_3^W - w_1^W} < \frac{w_3^W - w_1^W}{w_1^S - w_3^S + w_3^W - w_1^W}$ and acquiring information is never optimal for *"Weak"* players.

Now consider the equilibrium in which all players take action D. The equilibrium payoff for a *"Strong"* player is:

$$ps_4^S + (1-p)\, s_4^W$$

Conversely, were they to take action C,they would earn:

$$ps_2^S + (1-p)\, s_2^W$$

Following the equilibrium strategy yields a higher payoff if and only if:

$$\begin{aligned} ps_4^S + (1-p)\, s_4^W &\geq ps_2^S + (1-p)\, s_2^W \\ &\Leftrightarrow \quad p \leq \frac{s_4^W - s_2^W}{s_2^S - s_4^S + s_4^W - s_2^W} \in (0,1) \end{aligned}$$

Similarly, a *"Weak"* player's expected equilibrium payoff is:

$$pw_4^S + (1-p)\, w_4^W$$

whereas a unilateral deviation to C would earn:

$$pw_2^S + (1-p)\, w_2^W$$

Following the equilibrium strategy yields a higher payoff if and only if:

$$p \leq \frac{w_4^W - w_2^W}{w_2^S - w_4^S + w_4^W - w_2^W} \in (0, 1)$$

In this equilibrium, suppose that:

$$\begin{aligned} \mu^S \left(\text{Opponent's } \theta = \text{“}\textit{Strong}\text{”} \,|\text{Opponent acquired signal}\right) &= 1 \\ \mu^W \left(\text{Opponent's } \theta = \text{“}\textit{Strong}\text{”} \,|\text{Opponent acquired signal}\right) &= 0 \end{aligned}$$

“Strong” players take action C against an opponent who has acquired a signal, whereas *“Weak”* players take action D. By acquiring a signal, a *“Strong”* player earns:

$$ps_1^S + (1 - p)\, s_4^W$$

Not acquiring a signal is optimal if and only if:

$$\begin{aligned} ps_4^S + (1 - p)\, s_4^W &\geq ps_1^S + (1 - p)\, s_4^W \\ &\Leftrightarrow p \leq 0 \end{aligned}$$

“Weak” players expect to earn:

$$pw_1^S + (1 - p)\, w_4^W$$

Again, acquiring information is not optimal if and only if $p \leq 0$. So, if $p = 0$, an equilibrium in which no player acquires a signal and all players take action D is possible. This completes the proof. ■

1.E Proof of Proposition 5

As with the proof of Proposition 1.D, in equilibrium:

$$\mu\left(\text{Opponent's } \theta = \text{\textit{“Strong”}} \,|\text{Opponent did not acquired signal}\right) = p$$

Suppose that no player acquires information, and consider the payoff to *“Strong”* players. By taking action C, they expect to earn:

$$ps_1^S + (1-p)\,s_2^W$$

whereas taking action D yields:

$$ps_3^S + (1-p)\,s_4^W$$

A *“Strong”* player prefers to take action C if and only if:

$$\begin{aligned} ps_1^S + (1-p)\,s_2^W &\geq ps_3^S + (1-p)\,s_2^W \\ p &\geq \frac{s_4^W - s_2^W}{s_1^S - s_3^S + s_4^W - s_2^W} \in (0,1) \end{aligned}$$

By taking action C, a *“Weak”* player expects:

$$pw_1^S + (1-p)\,w_2^W$$

whereas D yields:

$$pw_3^S + (1-p)\,w_4^W$$

“Weak” players prefer D to C if and only if:

$$p \leq \frac{w_4^W - w_2^W}{w_1^S - w_3^S + w_4^W - w_2^W} \in (0,1)$$

So if $p \in \left[\frac{s_4^W - s_2^W}{s_1^S - s_3^S + s_4^W - s_2^W}, \frac{w_4^W - w_2^W}{w_1^S - w_3^S + w_4^W - w_2^W}\right]$ and no players acquire information, the equilibrium is supported. As acquiring a signal is an out of equilibrium action, define:

$$\begin{aligned} \mu^S \left(\text{Opponent's } \theta = \text{“\textit{Strong}”} \,|\text{Opponent acquired signal}\right) &= 0 \\ \mu^W \left(\text{Opponent's } \theta = \text{“\textit{Strong}”} \,|\text{Opponent acquired signal}\right) &= 0 \end{aligned}$$

All players take action D against an opponent who acquires a signal. For a *"Strong"* player, acquiring a signal thus yields an expected payoff of:

$$ps_2^S + (1-p)\, s_4^W$$

They therefore prefer not to acquire a signal if and only if:

$$p \geq \frac{s_4^W - s_2^W}{s_1^S - s_2^S + s_4^W - s_2^W}$$

Once again, by Assumption 1, this is always true over the specified support of the equilibrium. For a *"Weak"* player, the expected payoff from acquiring a signal is:

$$pw_2^S + (1-p)\, w_4^W$$

By Assumption 2, this is always strictly smaller than the equilibrium payoff, since $p \leq \frac{w_4^W - w_2^W}{w_1^S - w_3^S + w_4^W - w_2^W} < 1$. Thus no player wishes to acquire a signal. This completes the proof. ■

1.F Proof of Corollary 1

Firstly note that, since the support of the equilibrium in Proposition 5 is bounded by each type's incentive compatibility constraints, no other support can exist for an equilibrium in which *"Strong"* players do not acquire a signal and choose action C, whereas *"Weak"* players do not acquire a signal and take action D.

Consider the decision of a *"Weak"* player in an equilibrium in which *"Strong"* players do not acquire a signal and choose action D, whereas *"Weak"* players do not acquire a signal and take action C. In equilibrium, they receive an expected payoff of:

$$pw_2^S + (1-p)\,w_1^W$$

The out of equilibrium beliefs about a player who acquires a signal can lead to one of four possible pure strategy profiles:

1. All players choose action C;
2. *"Strong"* players choose action C, whereas *"Weak"* players choose action D;
3. *"Strong"* players choose action D, whereas *"Weak"* players choose action C; or
4. All players choose action D.

Considering each in turn:

1.F.1 All players choose action C

By acquiring a signal, a *"Weak"* player chooses C against a *"Strong"* opponent, and D against a *"Weak"* opponent. Their expected payoff is:

$$pw_1^S + (1-p)\,w_3^W$$

They will forego a signal if and only if:

$$\begin{aligned} pw_2^S + (1-p)\, w_1^W &\geq pw_1^S + (1-p)\, w_3^W \\ &\Leftrightarrow p\left(w_2^S - w_1^S + w_3^W - w_1^W\right) \geq w_3^W - w_1^W \end{aligned}$$

By the assumption in the corollary, the left hand side is negative. Conversely, by Assumption 3, the right hand side is positive. This yields a contradiction.

1.F.2 *"Strong"* players choose action C, whereas *"Weak"* players choose action D

By acquiring a signal, a *"Weak"* player chooses C against a *"Strong"* opponent, and D against a *"Weak"* opponent. Their expected payoff is:

$$pw_1^S + (1-p)\, w_4^W$$

They will forego a signal if and only if:

$$\begin{aligned} pw_2^S + (1-p)\, w_1^W &\geq pw_1^S + (1-p)\, w_4^W \\ &\Leftrightarrow p\left(w_2^S - w_1^S + w_4^W - w_1^W\right) \geq w_4^W - w_1^W \end{aligned}$$

By the assumption in the corollary, the left hand side is negative. Conversely, by Assumption 3, the right hand side is positive. Once again, this yields a contradiction.

1.F.3 *"Strong"* players choose action D, whereas *"Weak"* players choose action C

By acquiring a signal, a *"Weak"* player chooses C against a *"Strong"* opponent, and D against a *"Weak"* opponent. Their expected payoff is:

$$pw_2^S + (1-p)\, w_3^W$$

They will forego a signal if and only if:

$$\begin{aligned} pw_2^S + (1-p)\, w_1^W &\geq pw_2^S + (1-p)\, w_3^W \\ &\Leftrightarrow w_1^W \geq w_3^W \end{aligned}$$

By Assumption 3, this is a contradiction.

1.F.4 All players choose action D

By acquiring a signal, a *"Weak"* player chooses C against a *"Strong"* opponent, and D against a *"Weak"* opponent. Their expected payoff is:

$$pw_2^S + (1-p)\, w_4^W$$

They will forego a signal if and only if:

$$\begin{aligned} pw_2^S + (1-p)\, w_1^W &\geq pw_2^S + (1-p)\, w_4^W \\ &\Leftrightarrow w_1^W \geq w_4^W \end{aligned}$$

By Assumption 3, this is a contradiction.

Since no other pure strategy is possible, the separating no signal equilibrium in Proposition 5 must be unique. This completes the proof. ■

Chapter 2

Initiation and Protection Rackets

2.1 Introduction

Thus far, I have discussed groups whose members engage in activities closely related to crime. In this chapter, and the one to follow, I take a more direct approach by shifting the focus to organised crime. As noted towards the end of Chapter 1, criminal organisations may engage in coercive recruitment during times of need. This is tantamount to the peer pressure applied by sports teams (although much more severe). However, this tends to be an option of last resort. Those coercively recruited tend to have little allegiance to the organisation, undermining its internal cohesion (Jankowski 1991). Instead, criminal groups prefer to hire in much the same way as regular firms – they offer a wage in exchange for effort. It is this approach, and its implications for policy, that I explore in the remainder of the thesis.

Organised crime has long been an area of concern for policy makers and academics alike. It is the source of numerous negative externalities. These include victimisation, the fear of crime, lost human capital and issues surrounding drug use. Moreover, the cost of administering justice (incorporating prevention, detection and punishment) is increasingly a matter for debate. In the United States, it is estimated that there are around 731,000 gang members (Egley and Howell 2011) inflicting a cost on society of around $465 billion per annum (based upon estimates by Cohen and Piquero 2009). In many northern European cities, the profile of organised crime is growing (Bennett and Holloway 2004, Klein, Weerman, and Thornberry 2006). For example, Metropolitan Police Service (2012) suggests that there are around 4,800 gang members in London. These individuals are responsible for 22% of serious violence, 50% of shootings and 14% of rape in the city. In southern Europe, the problem obviously has a very long history.

Of the various types of criminal organisation, the protection racket has received the most attention from economists (see, for example, Gambetta 1996, Skaperdas and Syropoulos 1997, Garoupa 2000, Dixit 2007). These groups use violent intimidation to reduce the competition faced by their customers. Since a comparative advantage for violence is necessary for the existence of any criminal group (they have no legal protection), many such organisations start out as protection rackets, before moving into other illegal markets.

This paper presents a model of a protection racket, hereafter known as the *Mafia*[1]. Youths live in the Mafia's territory. These youths are heterogeneous in two dimensions, reflecting their intrinsic suitability for work in the primary labour market and criminal sector respectively. Initially, youths are given the opportunity to engage in juvenile crime. By doing so, they acquire criminal skills. Juvenile crime is costly, but the cost is declining in a youth's criminal ability. They then choose which sector to work in. If they opt for the primary labour market, they receive a wage related to their formal ability. If instead they become a criminal, they join the Mafia (becoming a *mafioso*).

The Mafia operates in the market for protection. Its customers (firms or individuals) pay a fee to *avoid* violent intimidation by rivals. As such, mafiosi are unable to signal the level of criminal skills they possess in equilibrium. Instead,

[1] Gambetta 1996 states that protection is the defining characteristic of mafia groups.

the amount the Mafia's customers are willing to pay depends upon the size of the racket and its *reputation for violence* - the amount of criminal skill customers believe each mafioso has accrued. This will be the same for all mafiosi, and equal to the expected amount of criminal skill across the organisation. As such, each member of the racket generates identical revenue.

When youths join the Mafia, they are forced to undertake an initiation. This initiation is costly, but the cost is declining in the criminal skill a recruit possesses. It serves two purposes. Firstly, it allows the racket to accurately assess the recruit's level of criminal skill, and hence determines the recruit's standing within the organisation's hierarchy. Secondly, the difficulty of initiation (although not the cost each recruit suffers) acts as a signal to customers regarding both the number of recruits the organisation is accepting and the distribution of criminal skills amongst those recruits. They use this signal to update the Mafia's reputation and hence their willingness to pay for protection.

Mafiosi are paid a wage that is proportional both to the revenue each member generates and their standing within the organisation, subject to the Mafia's budget constraint (Chang, Lu, and Chen 2005). Mafias are often (despite their portrayal in the media) relatively weak collectives of individuals (Jankowski 1991, Klein, Weerman, and Thornberry 2006). As such, there is some evidence to suggest that such organisations take a fairly utilitarian approach to recruitment. Specifically, they choose recruits to maximise the payoff to current members. Since the number of current members is fixed, this is equivalent to maximising the revenue that each member is able to generate. By doing so, it is able to offer a high wage to each member. It optimally adjusts initiation difficulty to trade off reputational gains with size. If, instead it were to maximise total revenue, it may be able to increase this further by recruiting more members. However, these additional recruits would need to be paid, potentially reducing the bounty available to its original members. Jankowski 1991 provides some evidence of precisely this approach. When interviewing a New York gang leader about their recruitment strategy, they said:

> "Man, we don't let all the dudes who want to be let in in. We can't do that, or I can't, 'cause right now we're sitting good. We got a good bank account and the whole gang is getting dividends. But if we let in a whole lot of other dudes, everybody will have to take a cut unless we come up with some more money, but that don't happen real fast. So you know the brothers ain't going to dig a cut, and if it happens, then they going to be on me and the rest of the leadership's ass and that ain't good for us." [Jankowski 1991]

It is not, however, guaranteed that the Mafia is able to do this. When initiation becomes more difficult, recruits with relatively low levels of criminal skill decide to join the primary labour market instead. This improves the Mafia's reputation, but also reduces the remaining members' relative standing within the racket. In turn, this causes more members to leave. The reputation improves once more, and those who still remain suffer a further decline in their standing. This process could continue until the gang losses (almost) all of its members.

To ensure that increases in initiation difficulty leads to a finite increase in reputation, and associated decline in size, it is sufficient for the hazard rate of the distribution of criminal abilities (conditional on formal ability) to be bounded. Each improvement in reputation will lead to fewer and fewer members leaving. As such, each incremental reputational gain will decline to zero. This suggests that the key player policies proposed by several papers on network economics (for example Ballester, Calvó-Armengol, and Zenou 2004 and 2010, Schwartz and Rouselle 2009, Liu, Patacchini, Zenou, and Lee 2011) may be more effective than previously believed. By targeting those with relatively high criminal ability, these policies alter the distribution of abilities within the organisation, potentially making it more difficult for the Mafia to control both reputation and size through initiation. Conversely, targeting foot soldiers may prove detrimental.

2.2 Stylised Facts About Protection Rackets

2.2.1 Reputation in Protection Markets

The importance of a reputation for violence in determining a protection racket's ability to generate revenue is widely accepted by criminologists. With reference to the Sicilian Mafia, Gambetta 1996 goes so far as to claim that a reputation for violence is its most important commodity. A strong reputation means that customers of the Mafia are less likely to face intimidation by rivals. Rivals' expected cost of intimidation is higher, as the Mafia is more likely to inflict violent retribution. As such, the Mafia can charge a higher price to its customers, and will need to engage in less violence (Reuter 1995, which is generally considered to be costly). Katz 1988 makes similar observations with regard to American street gangs. He suggests that many individuals are attracted to gangs to become part of a, "Street Elite". By associating with a gang, youths not only feel safer, but are also able to engage in low level crime without fear of recrimination. Levitt and Venkatesh 2000 confirm the value of association. In a unique analysis of a drug gang's finances, they show that many youths actually pay a fee to the gang in order to become associated with it. In this sense, whilst they call themselves gang members (see, for example, Poutvaara and Priks 2009 and 2011) they are in fact protection customers.

To illustrate the importance of reputation, consider a simple example in Figure 2.1. The protection racket can either be tough (with probability t) or weak. The racket's customer, C, faces a threat from a rival (a loss of $-K$). If the pro-

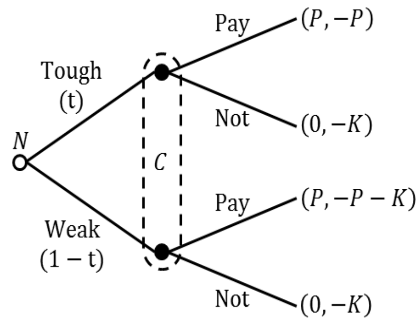

Figure 2.1: A simple model of protection

tection racket is tough, it can protect its customer from the rival. If it is weak, it cannot. The customer does not know how tough the racket is. In this simple framework, the customer decides whether to employ the racket or not, at price P. Their maximum price they are willing to pay is tK. Clearly, if t is larger (that is to say, the racket has a better reputation) they are able to charge a higher fee. A similar argument can be made for an extortion racket (e.g. Bueno De Mesquita and Hafer 2008), whereby the customer pays to avoid violence that only a tough racket could inflict. In both of these examples, the customer pays to avoid being exposed to violence. In a more dynamic setting, violence is therefore not an equilibrium outcome. The racket's reputation is sufficient to ensure that the customer is protected, or that they pay extortion fees when required.

There are several other avenues through which a reputation for violence benefits a Mafia. Rackets with strong reputations are less likely to face competition within their territory (Silverman 2004). As such, they have a stronger monopoly, and their members are able to extract larger rents (à la Milgrom and Roberts 1982). Once again, violence is not an equilibrium outcome, as a strong reputation is sufficient to deter entry.

The racket cannot operate in isolation. Inevitably, its members and customers will acquire evidence of the organisation's illegal activities. In order to prevent them from sharing this information with law enforcers, a credible threat of violence is required. Indeed Baccara and Bar-Isaac 2008 suggest that the threat of violence enables the racket to develop a relatively efficient hierarchical structure. This allows it to transmit information within the organisation with relative ease, safe in the knowledge that its members will not inform the authorities.

2.2.2 Initiation as a Determinant of Reputation

It is clear from the above discussion that asymmetric information is rife within the criminal sector. As violence is rare in equilibrium, there are few opportunities for mafiosi to signal their skills to other members, or for the Mafia itself to signal to its customers. One widely used approach to circumventing this problem is the use of initiation rituals. I define initiation as any activity that recruits to the racket undergo prior to becoming fully-fledged members. Initiation tends to have three transient features:

1. It enables the protection racket to learn how skilled its recruits are;
2. it is costly to the recruit, but this cost is declining in their criminal skill; and
3. the procedure, if not the result, is publicised within the Mafia's territory.

The primary reason for initiating recruits is to discover their aptitude for violence (Skaperdas 2001). Criminal gangs tend to organise into a hierarchical structure. An individual member's standing within that structure is often directly related to how violent they are. Thus initiation provides an opportunity for the Mafia to discover where the recruit fits into this hierarchy. Those who perform well during initiation may be promoted relatively quickly. Moreover, they are likely to be allowed to skim additional earnings from the revenue they raise on behalf of the organisation. The payoff they receive thus depends upon their criminal skill relative to those of other mafiosi.

There is also widespread evidence to suggest that initiation is costly. Decker 1996, for example, discussing several initiation rituals used by a drugs gang in the United States. Some recruits were forced to fend off a continued attack by several members of the gang. They were accepted based upon their performance. Jankowski 1991 reports similar rites in the gangs he studied. Decker also reported that other gang members were required to walk through enemy neighbourhoods wearing gang colours. They were accepted if they returned alive. In both of these examples, there is a clear physical cost associated with initiation. However, given that the gangs wished to determine aptitude, it is necessarily the case that tougher recruits find it easier. By observing the disutility a recruit suffers, the Mafia can thus learn their criminal skill. Furthermore, a difficult initiation increases the marginal benefit of acquiring criminal skills, leading to higher skill levels across all recruits to the organisation.

Other criminal organisations do not, at first glance, appear to have costly initiation. For example, Paoli 2003 suggests that initiation amongst the Italian Mafia families is largely ritualistic[2]. The initiate smears a few drops of blood on a picture of a saint. This picture is then set alight, and whilst the picture burns in their hands, the initiate recites an oath of allegiance. However, in order to be initiated, the recruit must be vouched for by a, “Man of honour,” who has been responsible for assessing their reliability prior to the ceremony. As such, the ceremony serves more as a graduation. In a sense, the recruit’s initiation took place whilst the man of honour was assessing them. In particular, it is common that a recruit must first, “Make their bones,” before being allowed to undergo initiation. This tends be mean being involved in a Mafia murder, or other violent activity, upon which the men of honour can judge the recruits readiness for initiation (see, for example, Raab 2006). This element of the process is clearly

[2]Morgan 1960 and Iwai 1986 discuss similarly ritualistic initiations within triad gangs in Hong Kong and the Yakuza in Japan respectively.

closer to those of the drug gangs.

Finally, initiation acts as a signal to the protection racket's customers regarding the quality of its members. Whilst they are unlikely to be able to observe precisely who is being initiated, or how they performed, they will be aware of how difficult the initiation was. They may observe members injuries resulting from a beating, or be made aware of which organisation carried out a particular murder. This aspect of initiation, to my knowledge so far unexplored, enables customers to update their beliefs about the gang's reputation. Whilst they cannot observe the precise aptitude for violence of each individual member, by observing initiation difficulty they learn something about the distribution of skills. Only relatively capable youths find it profitable to join the racket when the initiation is difficult. As such, it will be small, but with a tough reputation. Conversely, if the initiation is easy, customers will know that a lot of youths join the gang, and that some of them will consequently be reasonably inept. Since these beliefs will dictate customers' willingness to pay for the racket's protection, it is essential that it provides a clear advert. Given the relative scarcity of opportunities to engage in violence, it is an important avenue that the protection racket exploits. I therefore assume that the structure of the contract the Mafia offers is fixed, so that initiation difficulty contains all the information required for customers to get an accurate picture of the Mafia's attributes.

2.2.3 Acquisition of Criminal Skills

How do youths become tough? Numerous works by Athens (summarised in Rhodes 2001) suggest that many individuals convicted of violent crime went through a common developmental process, dubbed violentisation. One important aspect of this process is violent coaching. After initial bad experiences with violence, Athens proposes that youths are coached to believe that violence is a correct and appropriate response to minor provocation. Youths not only become desensitised to the violence they inflict, it becomes an automatic response to confrontation or perceived insult. Jankowski 1991 discusses a similar process called defiant individualism. He suggests that, as a result of scarce (physical and emotional) resources, youths in poor neighbourhoods learn to become both highly competitive and self-sufficient very early in their lives. Over time, this leads them to engage in violence, as it makes them more likely to acquire larger shares of available resources.

A common feature of violent coaching is engagement in juvenile crime (often in a delinquent gang). Esbensen and Lynskey 2001 interviewed fourteen year-old gang members in the United States. Of the respondents, twenty-five percent claimed to have shot at someone. Whilst much of this may be bravado, it nevertheless indicates an acceptance of violence and an understanding of the importance of signalling toughness. Similar (although less extreme) statistics were reported by Esbensen and Weerman 2005 and Salagaev, Shashkin, Sherbakova, and Touriyanskiy 2005 regarding Dutch and Russian juvenile gang members respectively.

Violent coaching can take two forms; learning by doing and learning by observing. Ballester, Calvó-Armengol, and Zenou 2010 interpret this in a network setting. Youths learn by engaging in crime or by observing those to whom they are connected. As expected, those with a higher degree of centrality within the network acquire more skill. Bayer, Pintoff, and Pozen 2004 provide empirical evidence to back up this assertion with regard to network formation in juvenile detention centres. This heterogeneity will play an important part in the analysis to come.

The remainder of this paper proceeds as follows. In the next section, I present a model of organised crime in which a protection racket uses initiation to trade off reputational gains against fewer recruits. In section 2.4 I outline the equilibria of the model. Section 2.5 considers intuitive conditions under which the Mafia can regulate its size and reputation through initiation. Section 2.6 discusses some implications of the results, and section 2.7 concludes. Major proofs are left to the appendices.

2.3 The Model

The model focuses on the territory of a neighbourhood protection racket known as the *Mafia*. The Mafia recruits members from a measure N of youths who live in the neighbourhood. These youths are heterogeneous in two respects; their *formal* and *criminal abilities*. Formal ability can be high or low, $\theta_i \in \{H, L\}$, and dictates the range of opportunities a youth has in the primary labour market. The probability that a particular youth has high formal ability is given exogenously by h. Criminal ability can take a range of values, $\sigma_i \in [0, 1]$. Youths with higher criminal ability find it less costly to acquire the skills required for a successful criminal career. If youths have a lot of friends who are also engaging in juvenile crime, they are able to learn from their friends mistakes. Conversely, if they are relatively isolated they must engage in relatively costly trial and error. The extent to which these two avenues are open to youths creates a large amount of heterogeneity in criminal ability. σ_i measures a youth's centrality within a delinquent network, and is distributed according to some $G(\sigma)$. G is continuously differentiable, with density function $g(\sigma)$, and has full support on $[0, 1]$.

I allow for the possibility of correlation between formal and criminal abilities. Some skills, such as basic numeracy, are equally important in both the primary labour market and the market for protection. Others, such as a short temper, may prove helpful when offering protection, but would be detrimental to a successful career in the legitimate economy. As such, I denote by $G_\theta(\sigma)$ the distribution of criminal abilities conditional on youths' formal ability. Whilst one might broadly expect positive correlation, the results herein are equally valid if the relationship were negative (or, indeed, if abilities are fully independent).

The focus of much of the economics of crime literature has been on the decision to become involved in crime. Following this approach, I consider youths' choices. Initially, they have the opportunity to engage in juvenile crime. Juvenile crime is costly, but enables youths to acquire criminal skills. Thereafter, they choose between a career in the *primary labour market* or in crime. If they opt for a criminal career, I assume that they join the Mafia.

Criminal skills can be thought of as an aptitude for inflicting violence. A youth with criminal ability σ_i can acquire criminal skill c_i by engaging in juvenile crime at a cost $J\left(\frac{c_i}{\sigma_i}\right)$. J is increasing, convex and continuously differentiable, so the cost of acquiring criminal skill is increasing in the amount acquired, but decreasing in criminal ability (reflecting the fact that higher ability youths can learn from observing others). For simplicity, I assume that $J(0) = 0$, so there are no fixed costs associated with criminal skills, and that $J'(0) = 0$ so acquiring some criminal skill is almost always affordable

Should youths then opt to work in the primary labour market, they receive a wage that is commensurate with their formal ability. A high ability youth earns w_H whereas a low ability youth earns $w_L < w_H$. If, on the other hand, they choose to become a criminal, the Mafia makes use of their criminal skills. I denote the set of recruits to the Mafia by R, so $i \in R$ if and only if youth i joins the Mafia.

Before describing the payoff to a youth from joining the Mafia, it is expedient to discuss its operations in more detail. Since the Mafia is a protection racket, it relies on two features to generate revenue: its members, and its reputation for violence. Protection customers pay the Mafia to avoid encounters with violence.

The amount they are willing to pay thus depends upon their beliefs about the skills possessed by their protector. However, since mafiosi are unable to demonstrate their violent tendencies to their customers (violence is not an equilibrium outcome), the expected level of skill will be the same across all members. I call this expected level of criminal skill, $\bar{c}_t$, the Mafia's *reputation*. By extension, each member must generate the same revenue for the Mafia:

$$r\left(M_t, \bar{c}_t\right) \tag{2.1}$$

where M_t denotes the Mafia's *size* - the measure of its membership - reflecting possible network externalities (Sah 1991, Cook, Ludwig, Venkatesh, and Braga 2007) or congestion effects. I assume that individual revenue has finite, positive and diminishing marginal returns in both size and reputation. Moreover, a larger Mafia may be more able to take advantage of reputational gains, so $r_{M\bar{c}}\left(M_t, \bar{c}_t\right) > 0$. Since the Mafia is already established, I endow it with initial size M_0 and reputation $\bar{c}_0$. Having recruited new members from the neighbourhood, its size and reputation become:

$$M_1 = (1-p) M_0 + |R| \tag{2.2}$$

$$\bar{c}_1 = (1-p)\frac{M_0}{M_1}\bar{c}_0 + \frac{|R|}{M_1}E\left[c_i \left| i \in R\right.\right] \tag{2.3}$$

where p is the probability that current mafiosi are arrested, and $|R| \in [0, N]$ is the measure of youths whom the Mafia recruits.

The Mafia's objective is to maximise $r\left(M_t, \bar{c}_t\right)$. Many criminal organisations are relatively loose collectives. As such, in order to maintain cohesion, it attempts to maximise the payoff it can offer its members, subject to the wage bill not exceeding its revenue. By increasing size, it will be able to generate more revenue. However, as it will also have to pay more members, it may be the case that the payoff each member receive actually falls on average. Focusing on individual revenue negates this.

Returning to youth payoffs, recruits to the Mafia undergo an initiation. Initiation serves several purposes. Firstly, by watching how tough the recruit finds the process, it allows the Mafia to observe the criminal skills of its recruits. Furthermore, initiation is costly. It often involves the initiate engaging in a violent activity during which they are at a disadvantage. By adjusting the difficulty of the initiation, the Mafia can provide an incentive to acquire criminal skills (conditional on a youth still wishing to join). Finally, whilst the performance of recruits may be difficult for protection customers to observe, the difficulty of initiation is often commonly known. Hence it acts as a signal to protection customers regarding the Mafia's reputation. In terms of the model, the Mafia publicly announces its initiation difficulty, $I \geq 0$. Youth $i \in R$ suffer disutility $\frac{I}{c_i}$ from undergoing initiation, reflecting the fact that more difficult initiation is more costly, but more skilled recruits suffer less. Note that $I = 0$ is equivalent to no initiation, as all recruits find initiation equally straightforward. In this case, the Mafia is unable to differentiate between different skill levels, and must treat all recruits homogeneously. The Mafia chooses initiation difficulty to maximise the revenue each member is able to generate.

Having undergone initiation, mafiosi are paid according to:

$$\left(u + (1-u)\frac{\tilde{c}_i}{\bar{c}_1}\right) r\left(M_1, \bar{c}_1\right) \tag{2.4}$$

where $u \in [0,1]$ is fixed and $\tilde{c}_i$ is the level of criminal skill the gang believes youth i to possess. Irrespective of the value of u, (2.4) satisfies budget balance so the organisation's wage bill is equal to its total revenue. The form of contract offered by the Mafia nests several examples by Chang, Lu, and Chen 2005. If $u = 1$, all Mafia members receive identical wages. Chang et al dub this a *uniform sharing* *scheme.* Such a scheme attracts a large number of youths to the Mafia, but causes them to acquire relatively low levels of criminal skill[3]. If $u = 0$ (and *ability-adherent sharing scheme*) mafiosi receive wages commensurate with the level of skills they have acquired. Whilst this solves the low skill levels, the budget constraint creates an externality. By acquiring more skills, a recruit improves both the Mafia's reputation and their own standing within the organisation's hierarchy. Since individual revenue suffers from diminishing marginal returns, this reduces the payoff to other mafiosi. In turn, ability adherence could lead to a fall in the Mafia's size.

In summary, the payoff from joining the Mafia is:

$$-\frac{I}{c_i} + \left(u + (1-u)\frac{\tilde{c}_i}{\bar{c}_1}\right) r\left(M_1, \bar{c}_1\right) \tag{2.5}$$

A youth i's career decision is denoted $j_i \in \{0, 1\}$, taking unitary value if and only if they decide to join the Mafia.

The timing is as follows. Firstly, the Mafia (whose initial size and reputation are common knowledge) publicly announce an initiation difficulty. Youths privately learn their formal and criminal abilities, and decide how much criminal skill to acquire. They then decide upon a career. Protection customers update their beliefs about the Mafia's reputation, and all mafiosi (including recruits) generate revenue for the organisation. Finally, the Mafia shares out its earnings according to (2.4).

2.4 Equilibrium

This model supports several perfect Bayesian equilibria, consisting of:

1. An initiation difficulty, $I^* \geq 0$, that maximises the revenue each mafioso is

[3] As their model does not incorporate initiation, it leads to pure free-riding.

able to generate, and Mafia beliefs, $\widetilde{c}_i$, about each recruit's criminal skill, given the Mafia's initial size and reputation.

2. Criminal skill, c_i^*, and career decisions, j_i^*, for every youth, given the Mafia's initiation difficulty, initial size and reputation.

3. An updated reputation, $\bar{c}_1$, based upon customers' beliefs about the distribution of criminal skill within the Mafia, given the Mafia's initiation difficulty, initial size and reputation.

I focus on symmetric equilibria. In order to discuss youths' decisions, it is important to distinguish between the case where the Mafia employs an initiation and the case where it does not. In the former situation, the organisation learns precisely the level of criminal skill each youth has acquired. In the latter, it does not.

2.4.1 Youth Decisions with Initiation

Let us begin by assuming that $I > 0$. In this setting $\widetilde{c}_i = c_i$, as the Mafia know I and observe $-\frac{I}{c_i}$. A youth with ability profile (θ_i, σ_i) faces the following utility maximisation problem:

$$\max_{c_i \geq 0, j_i \in \{0,1\}} \left\{ (1 - j_i) \, w_{\theta_i} + j_i \left[-\frac{I}{c_i} + \left(u + (1-u) \, \frac{c_i}{\bar{c}_1} \right) r \left(M_1, \bar{c}_1 \right) \right] - J \left(\frac{c_i}{\sigma_i} \right) \right\} \tag{2.6}$$

Suppose that $j_i = 0$, so youth i seeks employment in the primary labour market. Criminal skills are of no use in their career. However, there is still a cost associated with acquiring them. Consequently, youth i chooses to avoid juvenile crime, selecting $c_i = 0$. Conversely, if $j_i = 1$, the youth gains a great deal from acquiring criminal skill. Not only do they find the initiation less challenging, but they also receive greater remuneration upon successfully joining the Mafia. They choose $c_i^* \equiv c^*(\sigma_i, I)$ to satisfy:

$$\frac{I}{c_i^*} + (1-u)\frac{c_i^*}{\bar{c}_1} r(M_1, \bar{c}_1) \equiv \frac{c_i^*}{\sigma_i} J'\left(\frac{c_i^*}{\sigma_i}\right) \tag{2.7}$$

The marginal benefit to acquiring criminal skills is increasing in both the difficulty of initiation and the size of the Mafia. A larger Mafia increases the revenue that each youth generates, increasing the size of the bounty to be shared amongst the mafiosi. On the other hand, a tougher reputation actually reduces the marginal benefit of criminal skill acquisition. Acquiring additional criminal skill does relatively little to improve a recruit's standing within the Mafia's hierarchy, as competition is stiffer.

Clearly, any increase in recruits' criminal skills will lead to an increase in the Mafia's reputation. This creates an externality, which will prove important in determining whether initiation is an effective method for controlling youths' incentives. By acquiring more criminal skill, a recruit effectively claims a larger share of the Mafia's bounty, at the expense of their fellow members. This reduces youths' incentive to join the Mafia in the first place.

Turning attention to career decisions, a youth i will become a recruit to the Mafia if and only if they satisfy the following participation constraint:

$$-\frac{I}{c_i^*} + \left(u + (1-u)\frac{c_i^*}{\bar{c}_1}\right) r(M_1, \bar{c}_1) - J\left(\frac{c_i^*}{\sigma_i}\right) \geq w_{\theta_i} \tag{2.8}$$

The left hand side is strictly increasing in the criminal ability of the youth. Consider two youths i and j, with $\sigma_i < \sigma_j$. If both youths were to acquire the same level of criminal skill, they would suffer the same disutility from initiation and receive the same earnings upon joining the Mafia. However, acquiring criminal skill is easier for youth j, so their overall payoff from Mafia membership must be strictly higher.

On the other hand, the payoff from joining the primary labour market is independent of criminal ability. Hence we can say that, if a particular youth joins the Mafia, all youths with the same formal ability and higher criminal ability also join the Mafia. In other words:

Lemma 1 *For any initial Mafia size and reputation, any initiation difficulty, and for each $\theta \in \{H, L\}$, there exists $\sigma_\theta \in (0, 1]$ such that:*

$$i \in R \iff \theta_i = \theta \text{ and } \sigma_i \geq \sigma_\theta$$

For each $\theta_i \in \{H, L\}$ I call the youth with ability profile $(\theta_i, \sigma_{\theta_i})$ the *marginal youth*. Since the Mafia initiate its members, $\sigma_\theta > 0$. A youth with no criminal ability finds it prohibitively costly to acquire criminal skill. Without criminal skills, they find initiation unbearable, and will never join. Assuming that a positive measure of youths with both formal abilities join the Mafia, the identity of the marginal youth is given by the participation constraint (2.8) binding. They are exactly indifferent between joining the primary labour market and being recruited to the Mafia. Clearly, any decrease in the payoff from being recruited (caused by, say, an increase in initiation difficulty) will cause the abilities of the marginal youths to increase.

Finally, we can turn attention to the resulting Mafia size and reputation:

$$M_1 = (1-p) M_0 + Nh (1 - G_H(\sigma_H)) + N (1-h) (1 - G_L(\sigma_L)) \tag{2.9}$$

and:

$$\bar{c}_1 = (1-p) \frac{M_0}{M_1} \bar{c}_0 + \frac{N}{M_0} \left(h \int_{\sigma=\sigma_H}^{1} c^*(\sigma, I)\, dG_H(\sigma) + (1-h) \int_{\sigma=\sigma_L}^{1} c^*(\sigma, I)\, dG_L(\sigma) \right) \tag{2.10}$$

Given that I is common knowledge, the Mafia's customers are able to infer both the abilities of the marginal youths, and the amount of criminal skill acquired by any youth of a given criminal ability. However, without knowledge of the ability profile of any particular youth, they cannot determine the criminal skill of the particular mafioso who asks them to pay their protection money.

These strategies admit up to three partial equilibria, displayed by Figure 2.2. Heuristically, suppose that both high and low formal ability youths become recruits. The Mafia is large. Moreover, since recruits with high formal ability will also have relatively high criminal ability, the Mafia's reputation is strong. Each mafioso generates a lot of revenue, and the payoff to all members is relatively high. This makes joining the Mafia worthwhile for high formal ability youths, supporting the equilibrium ($0 < \sigma_L^{H\&L} < \sigma_H^{H\&L} < 1$).

Secondly, suppose that only low formal ability youths become recruits. The Mafia is smaller than before, and there is a relatively high proportion of lower criminal ability recruits. As a result, each mafioso generates less revenue, and the payoff from joining the Mafia is lower than in the previous case. Whilst it is still worthwhile joining the Mafia for those with a low outside option ($\sigma_L^L < 1$), those with high formal ability find it more appealing to work in the primary labour market ($\sigma_H^L = 1$).

Finally, suppose neither low nor high formal ability youths choose to join the Mafia. The organisation's size falls dramatically. Reputation is unchanged at $\bar{c}_0$. The payoff from joining the Mafia may be even lower than before. As such, both high and low formal ability youths find it more worthwhile to join the primary labour market ($\sigma_L^N = \sigma_H^N = 1$).

2.4.2 Youth Decisions without Initiation

If $I = 0$, the Mafia's beliefs about a youth's criminal skill are independent of the youth's ability, $\widetilde{c}_i = \widetilde{c}$. A youth with ability profile (θ_i, σ_i) faces the following

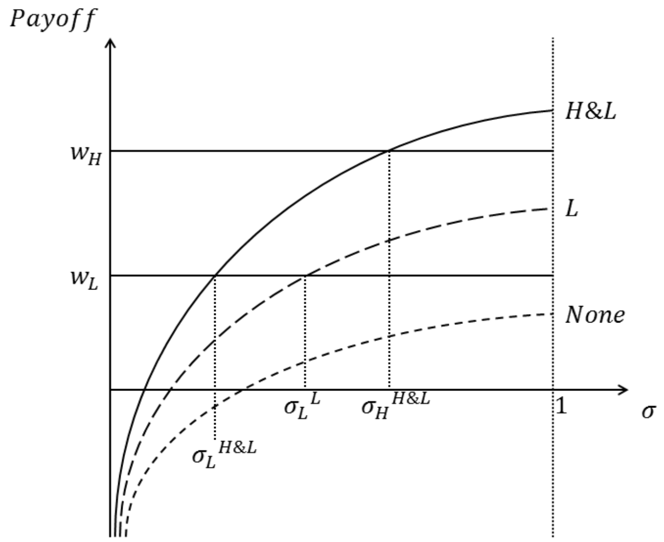

Figure 2.2: Equilibria with high and low recruits, only low recruits and no recruits.

utility maximisation problem:

$$\max_{c_i \geq 0, j_i \in \{0,1\}} \left\{ (1 - j_i)\, w_{\theta_i} + j_i \left[\left(u + (1-u)\, \frac{\widetilde{c}}{\bar{c}_1} \right) r\,(M_1, \bar{c}_1) \right] - J \left(\frac{c_i}{\sigma_i} \right) \right\} \quad (2.11)$$

As before, if the youth joins the primary labour market, criminal skills are of no use to them. As criminal skills are costly, youths do not acquire any. Suppose they decide to join the Mafia instead. Without an initiation ritual, the Mafia is unable to observe their skills. This creates a strict incentive to free-ride on others' criminal skills. Once again, they choose not to acquire any, $c_i^* = 0$. Knowing this, the Mafia sets $\widetilde{c} = 0$, and customers set:

$$\bar{c}_1 = (1 - p)\, \frac{M_0}{M_1} \bar{c}_0 \quad (2.12)$$

A youth with ability profile (θ_i, σ_i) will therefore join the Mafia if and only if:

$$ur\,(M_1, \bar{c}_1) \geq w_{\theta_i} \quad (2.13)$$

Since both payoffs are independent of the criminal ability of the youth, we have that:

Lemma 2 *For any initial Mafia size and reputation, and for each $\theta \in \{H, L\}$, if $I = 0$ then there exists $\sigma_\theta \in \{0, 1\}$ such that:*

$$i \in R \iff \theta_i = \theta \text{ and } \sigma_i \geq \sigma_\theta$$

In this case, either all youths of a particular formal ability have a strict incentive to join the Mafia, or none do. Combined with the youths' criminal skill decisions, we have:

$$M_1 = (1-p) M_0 + Nh (1 - G_H (\sigma_H)) + N (1-h) (1 - G_L (\sigma_L)) \tag{2.14}$$

Once again, this admits up to three partial equilibria. If

$$ur \left((1-p) M_0 + N, (1-p) \frac{M_0}{M_1} \bar{c}_0 \right) \geq w_H \tag{2.15}$$

then there exists an equilibrium in which all youths join the Mafia ($\sigma_H^{H\&L} = \sigma_L^{H\&L} = 0$). The organisation is large, buts its reputation is relatively diluted. Recruits make up a large proportion of members, and none of them have acquired criminal skills. Nevertheless, even high formal ability youths receive a strictly higher payoff from joining a large Mafia.

If:

$$w_L \leq ur \left((1-p) M_0 + N (1-h), (1-p) \frac{M_0}{M_1} \bar{c}_0 \right) \leq w_H \tag{2.16}$$

then there exists an equilibrium in which only youths will low formal ability join the Mafia ($\sigma_H^L = 1$, $\sigma_L^L = 0$). A smaller Mafia is less able to generate revenue through protection. However, its reputation is less diluted, as there are fewer mafiosi with no criminal skills. Even so, the payoff to each member is lower, and only low formal ability youths find it profitable to join. Finally, if:

$$ur\left(\left(1-p\right)M_0, \bar{c}_0\right) \leq w_L \tag{2.17}$$

then there exists an equilibrium in which nobody joins the Mafia ($\sigma_H^N = \sigma_L^N = 1$). The Mafia is too small to generate enough revenue to be able to attract new members, despite having the strongest reputation of any of the equilibria.

2.4.3 Mafia Decisions

The Mafia chooses initiation difficulty to maximise the revenue each member is able to generate:

$$I^* \in \arg\max_{I \geq 0} \left\{r\left(M_1, \bar{c}_1\right)\right\} \tag{2.18}$$

The choice over whether to initiate recruits is not trivial. It knows that, in any equilibrium, if it does not initiate then recruits will not acquire criminal skills. However, it will attract the maximum number of recruits available to it in that equilibrium. Consequently, if the Mafia is small and has a good reputation, then it may find it profitable not to initiate. In particular, the revenue each member generates is discontinuous around $I = 0$. Without initiation, each individual generates:

$$r\left(\widetilde{M}_1, \left(1-p\right)\frac{M_0}{\widetilde{M}_1}\bar{c}_0\right) \tag{2.19}$$

where $\widetilde{M}_1$ depends upon the equilibrium. Conversely, as $I \to 0^+$, the Mafia can still distinguish between youths. As such, recruits will acquire:

$$\underline{c}\left(\sigma_i\right) \equiv \sigma_i \left(J'\right)^{-1}\left(\sigma_i\left(1-u\right)\frac{r\left(M_1, \bar{c}_1\right)}{\bar{c}_1}\right) > 0 \tag{2.20}$$

Moreover, for each $\theta_i \in \{H, L\}$ there exists $\underline{\sigma}_{\theta_i} > 0$ such that:

$$\left(u + (1-u)\frac{\underline{c}\left(\underline{\sigma}_{\theta_i}\right)}{\bar{c}_1}\right) r\left(M_1, \bar{c}_1\right) - J\left(\frac{\underline{c}\left(\underline{\sigma}_{\theta_i}\right)}{\underline{\sigma}_{\theta_i}}\right) = w_{\theta_i} \tag{2.21}$$

and youths will only join if $\theta_i = \theta$ and $\sigma_i \geq \underline{\sigma}_\theta$. The Mafia will be strictly smaller, but with a better reputation. Assuming that the Mafia does choose to initiate its recruits, in each equilibrium it will choose an initiation difficulty satisfying:

$$\frac{\frac{\partial r}{\partial M_1}}{\frac{\partial r}{\partial \bar{c}_1}} = -\frac{\frac{\partial M_1}{\partial I}}{\frac{\partial \bar{c}_1}{\partial I}} \tag{2.22}$$

This choice of difficulty optimally balances increases in revenue through reputational gains against lost revenue resulting from fewer mafiosi. Letting the resulting reputation and Mafia size be $\bar{c}_1^*$ and M_1^* respectively, the Mafia will choose to initiate its recruits if and only if:

$$r\left(\widetilde{M}_1, (1-p)\frac{M_0}{\widetilde{M}_1}\bar{c}_0\right) \leq r\left(M_1^*, \bar{c}_1^*\right) \tag{2.23}$$

namely when the reputational gains from initiation are sufficient to offset the loss of members. Since revenue suffers from diminishing marginal returns to both size and reputation, the Mafia will only forego initiation if it already has a strong reputation and relatively few members.

2.5 Initiation as a Signal

If the Mafia is to utilise initiation, it must be able to control its size and reputation by varying its difficulty. Increasing difficulty has two direct effects. Firstly, it reduces the payoff to each youth from becoming mafiosi. The disutility they suffer from undergoing initiation is greater. Those towards the bottom of the criminal ability distribution suffer the most, as they acquire relatively few criminal skills.

As such, these individuals switch away from a criminal career towards the primary labour market. The abilities of the marginal youths increase, and the Mafia's size shrinks.

Secondly, the marginal benefit of acquiring criminal skills increases. Since recruits suffer greater disutility when they undertake initiation, they choose to offset this by acquiring more criminal skill. They opt to suffer more by engaging in more intense juvenile crime in order to be in better shape for initiation. In turn, this improves the Mafia's reputation. Since all recruits acquire more criminal skill, increased initiation difficulty has relatively little direct effect upon their relative standing within the Mafia's hierarchy.

There are, however, additional endogenous effects. When the Mafia's size falls, it is those with relatively low levels of criminal skill who leave. This further increases the average criminal skill in the Mafia, and its reputation improves again. Moreover, with fewer members, each mafioso's standing within the hierarchy falls. In turn, this reduces their payoff from joining the Mafia, since the organisation operates with a certain amount of ability adherence. More youths opt for the primary labour market.

Initiation is therefore effective, insofar as if it causes an increase in reputation, it also causes a decline in recruits of both formal criminal abilities and vice-versa:

Proposition 6 *Consider the effect of increasing initiation difficulty in any equilibrium of the initiation model. Then:*

$$\frac{\partial \bar{c}_1}{\partial I} > 0 \iff \frac{\partial \sigma_H}{\partial I} > 0 \text{ and } \frac{\partial \sigma_L}{\partial I} > 0 \tag{2.24}$$

Proof. See Appendix 2.A. ■

As noted above, improvements in reputation cause an endogenous increase in the abilities of the marginal youths. The converse is simply a result of the optimisation behaviour of the Mafia. If the abilities of the marginal youths are increasing in initiation difficulty, the Mafia's size must be declining. For initiation to be worthwhile, it must be the case that there are reputational gains associated with initiation. Otherwise, the Mafia could strictly improve the payoff to its members by reducing initiation difficulty.

The endogenous effects could be the source of instability within the Mafia, precluding its use of initiation. When initiation difficulty increases, youths at the bottom of the hierarchy suffer most, opting instead for a career in the primary labour market. This, combined with an initial increase in the criminal skills acquired by remaining mafiosi, improves the Mafia's reputation.

The improvement in reputation reduces the ability-adherent payoff to the remaining mafiosi. As the Mafia has a stronger reputation, each mafioso sees their relative standing with the hierarchy decline. There is simply more intense competition over the organisation's income. Those who now find themselves towards the bottom of the hierarchy find that they are receiving a smaller income than previously. They therefore opt to join the primary labour market. There is a further decline in low skilled mafiosi, and the reputation improves further. Concurrently, the marginal benefit to acquiring criminal skills declines. When faced with stiffer competition, a relatively large increase in criminal skill is required to secure a modest improvement in a mafioso's standing. This partially offsets the endogenous reputational gain, but is not sufficient to cause an endogenous reduction.

The endogenous reputational gain causes further declines in Mafia size which, in turn, leads to further reputational gains. Under certain circumstances, this process could continue until the Mafia has a very strong reputation but (almost) no members. In this sense, the criminal organisation is relatively unstable. It cannot use initiation difficulty to control its reputation and size because the endogenous effects dominate. It is therefore unable to take advantage of this important signalling opportunity.

The following condition ensures that the use of initiation is *stable*, in the sense that it leads to a finite increase in reputation, and associated decline in Mafia size.

Proposition 7 *Consider any equilibrium of the initiation model. There exists* $\Lambda > 0$ *such that if, for each* $\theta \in \{H, L\}$*, the hazard rate of the conditional distribution of criminal abilities is sufficiently bounded:*

$$\frac{g_{\theta}(\sigma)}{1 - G_{\theta}(\sigma)} \leq \Lambda \tag{2.25}$$

then the Mafia is stable.

Proof. See Appendix 2.B. ■

A bounded hazard rate suggests that the conditional distribution of criminal abilities is severely skewed. The vast majority of youths of any formal ability have relatively low criminal abilities. Moreover, the density of the distribution is rapidly declining. When the Mafia increases the difficulty of its initiation, the organisation will see a relatively large initial decline in size, and associated improvement in reputation. A relatively large proportion of recruits have abilities equal to, or very close to, that of the marginal youths. Consequently, a small increase in the abilities of the marginal youths will result in a relatively large measure of youths opting for a career in the primary labour market. Combined with the increase in the criminal skills acquired by those who remain committed to becoming mafiosi, the reputational gains are substantial.

However, as the endogenous effects begin to operate, each incremental increase in the abilities of the marginal youths results in a smaller and smaller measure of youths switching to careers in the primary labour market. As such, whilst the Mafia still lose recruits towards the bottom of its skills distribution, each resulting incremental gain in reputation is smaller. Combined with the endogenous decline in criminal skills resulting for increased competition over the organisation's bounty, the Mafia's reputation converges to a new, finite, level. Similarly, the organisation's size declines. The Mafia reaches point where its reputation and size have converged, and initiation has been successful in signalling both size and the distribution of criminal skills.

The implications of Proposition 7 can be shown diagrammatically in the equilibrium with initiation in which only youths with low formal ability join the Mafia. First, consider Figure 2.3. Suppose that there is a small increase in initiation difficulty. The expected cost of joining the Mafia increases, affecting the ability of the marginal youth, as shown by point A. Similarly, since youths acquire more criminal skills, the immediate impact on the Mafia's reputation is given by B. With the increase in reputation, each mafioso's standing in the organisation falls, reducing their ability adherent payoff. The ability of the marginal youth increases endogenously, given by C. Since those with relatively low levels of criminal skill no longer join the Mafia, its reputation further improves, shown by D. This leads to a further loss of members (point E), and the process continues. If Proposition 7 holds, as is the case in the diagram, the size of each endogenous increase in the responsiveness of the Mafia's reputation and ability of the marginal youth declines, until eventually they converge to $\left(\frac{\partial \bar{c}_1}{\partial I}\right)^*$ and $\left(\frac{\partial \sigma_L}{\partial I}\right)^*$ respectively.

Now consider a situation in which Proposition 7 does not hold. Initiation may not be stable, as shown in Figure 2.4. As before, the exogenous impact of an increase in initiation on the ability of the marginal youth and the Mafia's reputation are given by A and B respectively. As before, we can trace the endogenous effect of an improvement in the Mafia's reputation upon the ability of the marginal youth as a movement from B to C. Subsequent endogenous effects take us to D, then to E etc. However, in this case, the distribution of crimi-

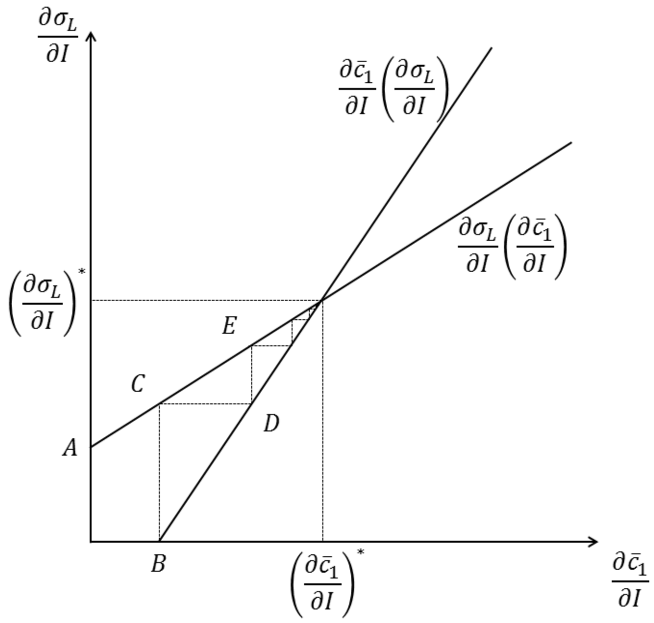

Figure 2.3: A situation in which Proposition 7 holds.

nal abilities is such that the endogenous effects actually increase in magnitude. Rather than converging to finite impacts, $\frac{\partial \bar{c}_1}{\partial I}$ and $\frac{\partial \sigma_L}{\partial I}$ increase towards infinity. Any small increase in the initiation difficulty the Mafia prescribe to its recruits could potentially leads to a complete collapse in membership.

2.6 Discussion

The results from the previous section suggest that protection rackets are likely to thrive in neighbourhoods with extremely skewed distributions of criminal abilities. In such areas, initiation difficulty act as a sufficient statistic for the distribution of criminal skills within the racket. By publicising changes in the difficulty of initiation that the Mafia's recruits undertake, customers can easily update their beliefs about the value of the protection on offer and consequently adjust their willingness to pay for such services.

There is evidence to suggest that protection rackets do indeed operate in such areas. Recruits with higher criminal abilities acquire more criminal skill,

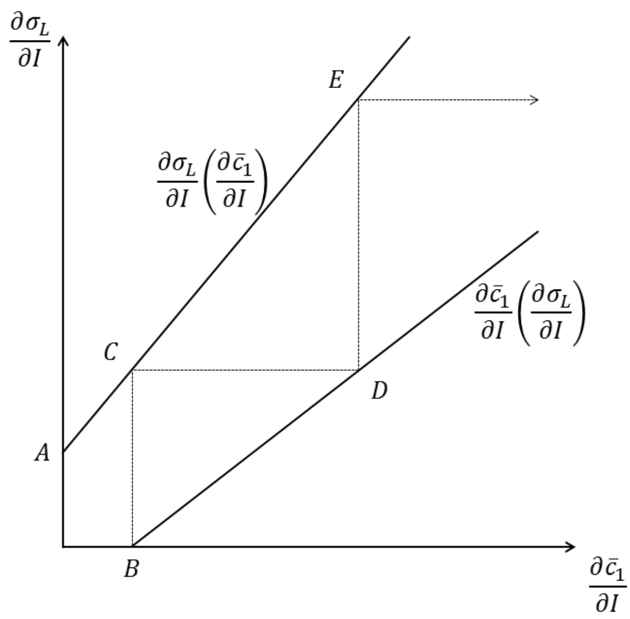

Figure 2.4: A situation in which Proposition 7 does not hold.

and those with more criminal skill have a higher rank in the organisation. As such, a bounded hazard rate would suggest that the vast majority of organised criminals would receive a low wage. Moreover, relatively few members would have higher ability, resulting in dramatically higher wages. This is precisely the distribution of wages found by Levitt and Venkatesh 2000. The foot soldiers they studied in a Chicago gang earned a flat rate of around $2,400 per annum. This suggests that they all have fairly similar, low criminal ability (certainly compared to a gang leader, who could earn in excess of $200,000 per annum).

This result lends further support to the use of key player policies (Ballester, Calvó-Armengol, and Zenou 2004 and 2010). These policies seek to identify those individuals who play an integral part within a criminal network. They could be individuals who interact with a lot of other members, or they could act as couriers between otherwise independent groups (Liu, Patacchini, Zenou, and Lee 2011). By apprehending them, law enforcers make criminal interactions more difficult, thus disrupting the wider organisation.

In the model presented here, these individuals are towards the top of the criminal ability distribution. By removing them, law enforcers change the distribution of criminal ability (and hence criminal skill) within the Mafia in two ways (Ballester, Calvó-Armengol, and Zenou 2004). Firstly, there is the direct effect of removing the individuals from the network, reducing the domain of the resulting distribution of criminal abilities. In particular, suppose that all those with criminal ability greater that $\widetilde{\sigma}$ are removed. The resulting hazard rate of the conditional distribution of criminal abilities is:

$$\frac{g_{\theta}(\sigma)}{G_{\theta}(\widetilde{\sigma}) - G_{\theta}(\sigma)} \tag{2.26}$$

Since $G(\widetilde{\sigma}) < 1$, clearly this is larger than before. As such, it is less likely that the Mafia will be able to effectively signal their size and propensity for violence through initiation. However, removing all those at the top of the distribution has a countervailing effect. It leads to an immediate reduction in the Mafia's reputation. In effect, this increases the required upper bound on the hazard rate, Λ, as the standing of all members in the resulting organisation is improved.

The second avenue through which key player policies operate involves reducing the abilities of all remaining youths. By removing some of the most connected youths in the juvenile network, law enforcers also remove other youths' ability to interact with, and learn from, these individuals. As a result, each remaining youth has to rely more heavily upon learning by doing. This increases the proportion of mafiosi with relatively low ability. Once again, this makes it less likely that the Mafia will be able to use initiation to control its reputation. If it were to attempt to increase initiation difficulty, a larger proportion of potential recruits would switch to employment in the primary labour market. As a result, the Mafia's reputation improves more than it would have done previously. This triggers an endogenous increase in the abilities of the marginal youths. Since every ability has a higher weighting than previously, the reputational gain causes a larger proportional fall in membership. In turn, this causes a greater endogenous reputational gain. Once again, the hazard rates are likely to be increased. However, unlike the direct removal of those at the top end of the abilities distribution, this second avenue has no incremental effect on the upper bound.

Targeting foot soldiers, or those on the periphery of the Mafia will have the opposite effect. Apprehending them reduces the proportion of the group who have relatively low criminal ability. As a result, when initiation difficulty increases, the Mafia's reputation converges to a higher level more rapidly. Each endogenous increase in the abilities of the marginal youths results in a relatively few recruits opting for a career in the primary labour market. The subsequent reputational gain is therefore relatively modest.

2.7 Conclusions

Criminal organisations inflict a large cost upon society. Many organisations start out as protection rackets, where a reputation for violence is necessary and sufficient to operate. They become a *Mafia.* With relatively few opportunities to signal an aptitude for violence, many Mafias utilise initiation rituals. Whilst all recruits find these costly, those who have engaged in juvenile crime have acquired sufficient skills to make undertaking them worthwhile. A difficult initiation thus attracts relatively few, highly capable, recruits. Customers update their beliefs about the skills mafiosi possess, and are consequently willing to pay more for protection services.

When the Mafia adjusts its initiation difficulty, recruits acquire more criminal skills. Concurrently, recruits with relatively low levels of criminal skills opt to join the primary labour market instead. Both effects lead to improvements in the Mafia's reputation. Since the average skill level within the organisation is higher, each individual's relative skill level (and hence their standing within the organisation's hierarchy) declines. This leads more recruits to leave, and results in further endogenous improvements in reputation. If the hazard rate of the distribution of underlying criminal abilities is sufficiently bounded, these endogenous declines in the size of the organisation (and the associated reputational gains) are small enough to enable the Mafia to use initiation difficulty to optimally control the reputation versus size trade off.

This result has implications for policy. In particular, it provides further weight to the notion that law enforcers should target key players within criminal organisations. By doing so, the distribution of criminal abilities is altered, placing greater weight upon those with relatively low abilities. The hazard rate may increase as a result, increasing the endogenous effects to the point that changes in initiation difficulty result in a terminal decline in the size of the organisation. As such, this approach not only reduces the Mafia's ability to interact, but may also deprive it of one of the most common approaches to advertising their reputation for violence.

Conversely, those policies which focus on the organisation's foot soldiers may partially backfire. In contrast to key player policies, this approach adjusts the distribution to place greater weight on those with relatively high ability. This reduces the endogenous effects, causing the organisation's reputation to converge more rapidly. This makes it easier for the racket to vary its size and reputation through changing initiation difficulty.

Acknowledgements

Thanks are owed to Leonardo Felli and Alwyn Young for numerous suggestions and help. Comments by participants at various LSE Microeconomic Theory Work In Progress seminars have also proved useful. Time at the Study Center in Gersenzee is also acknowledged. All errors remain my own.

2.A Proof of Proposition 6

The proof that follows is for an equilibrium in which both high and low formal ability youths join the mafia. The proof for the case in which only low formal ability youths become mafiosi is similar, and therefore omitted. We have two statements to prove. Considering each in turn:

2.A.1 $\frac{\partial \bar{c}_1}{\partial I} > 0 \Rightarrow \frac{\partial \sigma_H}{\partial I} > 0$ and $\frac{\partial \sigma_L}{\partial I} > 0$

For each given formal ability, the marginal youth is defined by:

$$-\frac{I}{c_\theta^*} + \left(u + (1-u)\frac{c_\theta^*}{\bar{c}_1}\right) r\left(M_1, \bar{c}_1\right) - J\left(\frac{c_\theta^*}{\sigma_\theta}\right) = w_\theta$$

where $c_\theta^* \equiv c^*\left(\sigma_\theta, I\right)$. Taking first order conditions with respect to I and applying the envelope theorem:

$$\frac{\partial \sigma_\theta}{\partial I} = \frac{\sigma_\theta^2}{c_\theta^* J'\left(\frac{c_\theta^*}{\sigma_\theta}\right)}\left[\frac{1}{c_\theta^*} + (1-u)\frac{c_\theta^*}{\bar{c}_1}\frac{r\left(M_1, \bar{c}_1\right)}{\bar{c}_1}\frac{\partial \bar{c}_1}{\partial I} - \left(u + (1-u)\frac{c_\theta^*}{\bar{c}_1}\right)\left(r_M\frac{\partial M_1}{\partial I} + r_{\bar{c}}\frac{\partial \bar{c}_1}{\partial I}\right)\right]$$

Since the gang seeks to maximise the revenue each member is able to generate $r_M\frac{\partial M_1}{\partial I} + r_{\bar{c}}\frac{\partial \bar{c}_1}{\partial I} \equiv 0$. So, if $\frac{\partial \bar{c}_1}{\partial I} > 0$ then $\frac{\partial \sigma_\theta}{\partial I} > 0$, as required.

2.A.2 $\frac{\partial \sigma_H}{\partial I} > 0$ and $\frac{\partial \sigma_L}{\partial I} > 0 \Rightarrow \frac{\partial \bar{c}_1}{\partial I} > 0$

We have that:

$$r_M\frac{\partial M_1}{\partial I} = -r_{\bar{c}}\frac{\partial \bar{c}_1}{\partial I}$$

with $r_M > 0$ and $r_{\bar{c}} > 0$. So $\frac{\partial \bar{c}_1}{\partial I} > 0$ if and only if $\frac{\partial M_1}{\partial I} < 0$. Now:

$$\frac{\partial M_1}{\partial I} = -Nhg_H\left(\sigma_H\right)\frac{\partial \sigma_H}{\partial I} - N(1-h)\,g_L\left(\sigma_L\right)\frac{\partial \sigma_L}{\partial I}$$

Since $g_\theta(\sigma) > 0$ for all $\theta \in \{H, L\}$ and $\sigma \in [0,1]$, if $\frac{\partial \sigma_H}{\partial I} > 0$ and $\frac{\partial \sigma_L}{\partial I} > 0$ then

$\frac{\partial M_1}{\partial I} < 0$ as required.

This completes the proof. ■

2.B Proof of Proposition 7

Once again, the proof that follows is for an equilibrium in which both high and low formal ability youths join the mafia. The proof for the case in which only low formal ability youths become mafiosi is similar, and therefore omitted.

We have that:

$$\begin{aligned}\frac{\partial \bar{c}_1}{\partial I} &= \frac{N}{M_1}\left[h\int_{\sigma=\sigma_H}^{1}\frac{\partial c^*}{\partial I}dG_H(\sigma) + (1-h)\int_{\sigma=\sigma_L}^{1}\frac{\partial c^*}{\partial I}dG_L(\sigma)\right] \\ &\quad + \frac{N}{M_1}\left[hg_H(\sigma_H)(\bar{c}_1 - c_H^*)\frac{\partial \sigma_H}{\partial I} + (1-h)g_L(\sigma_L)(\bar{c}_1 - c_L^*)\frac{\partial \sigma_L}{\partial I}\right]\end{aligned}$$

Substituting for $\frac{\partial c^*}{\partial I}$, $\frac{\partial \sigma_H}{\partial I}$ and $\frac{\partial \sigma_L}{\partial I}$ yields:

$$\begin{aligned}\frac{\partial \bar{c}_1}{\partial I} & \left[\frac{M_1}{N} + (1-u)\frac{r(M_1, \bar{c}_1)}{\bar{c}_1}\left\{\begin{array}{c} h\left(\begin{array}{c}\int_{\sigma=\sigma_H}^{1}\frac{c^{*3}}{2I+\frac{c^{*3}}{\sigma^2}J''\left(\frac{c^*}{\sigma}\right)}dG_H(\sigma) \\ -\frac{\sigma_H^2 g_H(\sigma_H)}{J'\left(\frac{c_H^*}{\sigma_H}\right)}(\bar{c}_1 - c_H^*)\end{array}\right) \\ +(1-h)\left(\begin{array}{c}\int_{\sigma=\sigma_L}^{1}\frac{c^{*3}}{2I+\frac{c^{*3}}{\sigma^2}J''\left(\frac{c^*}{\sigma}\right)}dG_L(\sigma) \\ -\frac{\sigma_L^2 g_L(\sigma_L)}{J'\left(\frac{c_L^*}{\sigma_L}\right)}(\bar{c}_1 - c_L^*)\end{array}\right)\end{array}\right\}\right]\Bigg] \\ &= h\left[\int_{\sigma=\sigma_H}^{1}\frac{c^{*3}}{2I+\frac{c^{*3}}{\sigma^2}J''\left(\frac{c^*}{\sigma}\right)}dG_H(\sigma) + \frac{\sigma_H^2 g_H(\sigma_H)}{J'\left(\frac{c_H^*}{\sigma_H}\right)}(\bar{c}_1 - c_H^*)\right] \\ &\quad + (1-h)\left[\int_{\sigma=\sigma_L}^{1}\frac{c^{*3}}{2I+\frac{c^{*3}}{\sigma^2}J''\left(\frac{c^*}{\sigma}\right)}dG_L(\sigma) + \frac{\sigma_L^2 g_L(\sigma_L)}{J'\left(\frac{c_L^*}{\sigma_L}\right)}(\bar{c}_1 - c_L^*)\right]\end{aligned}$$

A sufficient condition for $\frac{\partial \bar{c}_1}{\partial I} > 0$ is that, for each $\theta \in \{H, L\}$, we have:

$$\int_{\sigma=\sigma_\theta}^{1}\frac{c^{*3}}{2I+\frac{c^{*3}}{\sigma^2}J''\left(\frac{c^*}{\sigma}\right)}dG_\theta(\sigma) > \frac{\sigma_\theta^2 g_\theta(\sigma_\theta)}{J'\left(\frac{c_\theta^*}{\sigma_\theta}\right)}|\bar{c}_1 - c_\theta^*|$$

In order to show that the boundedness of the hazard function is sufficient for this to be true, it is necessary to prove some intermediate results.

Lemma 3 *For all $I > 0$, the optimal criminal skills of all recruits, i, is bounded:*

$$0 < \underline{c} \leq c_i^* \leq \widehat{c}$$

Proof. For a youth with criminal ability σ_i, the marginal benefit of acquiring criminal skill is bounded below:

$$\frac{I}{c_i^{*2}} + (1-u)\frac{r(M_1, \bar{c}_1)}{\bar{c}_1} > (1-u)\frac{r(M_1, \bar{c}_1)}{\bar{c}_1}$$

Given the marginal cost of criminal skill, we have that:

$$c^*(\sigma_i, I) \geq \sigma_i (J')^{-1}\left(\sigma_i (1-u)\frac{r(M_1, \bar{c}_1)}{\bar{c}_1}\right) \equiv \underline{c}$$

Now, individual revenue is also bounded above: $r(M_1, \bar{c}_1) \leq b_0 \in \mathbb{R}_+$. Also, for each M_1, $\frac{r(M_1, \bar{c}_1)}{\bar{c}_1}$ is bounded above by assumption. Moreover, since $M_1 \in [(1-p)M_0, (1-p)M_0 + N]$, there exists $b_1 \in \mathbb{R}_+$ such that $\frac{r(M_1, \bar{c}_1)}{\bar{c}_1} \leq b_1$. The payment the gang offers is thus bounded above by a straight line:

$$\left(u + (1-u)\frac{c_i^*}{\bar{c}_1}\right) r(M_1, \bar{c}_1) \leq ub_0 + (1-u)b_1 c_i^*$$

The total cost of joining the gang, on the other hand is bounded below by an increasing, convex function. Since $\sigma_i \in [0, 1]$:

$$J\left(\frac{c_i^*}{\sigma_i}\right) + \frac{I}{c_i^*} \geq J(c_i^*)$$

So, there exists $\hat{c} \in \mathbb{R}_+$ satisfying:

$$ub_0 + (1-u)\, b_1 \hat{c} = J(\hat{c})$$

such that, for all $c_i^* \geq \hat{c}$, the payoff to a youth from joining the gang is negative. Thus, no youth would choose would optimally choose $c_i^* \geq \hat{c}$. This completes the proof. ■

An implication of the lemma is that the reputation of the gang is also bounded above:

$$\bar{c}_1 \leq \bar{c}_0 + \hat{c}$$

Similarly, there is an upper bound on the initiation difficulty the gang will set. Since the gang pays at most $ub_0 + (1-u)\, b_1 c_i^*$, and the minimum disutility caused by initiation is $-\frac{I}{\hat{c}}$, nobody will choose to join the gang in any equilibrium if:

$$I \geq \hat{I} \equiv \hat{c}\left(ub_0 + (1-u)\, b_1 \hat{c}\right)$$

thus setting an initiation difficulty above $\hat{I}$ is never worthwhile for the gang.

The second intermediate result is as follows:

Lemma 4 *For all $I > 0$, the identity of the marginal youth with lower formal ability, σ_L, is bounded above zero:*

$$\sigma_L \geq \underline{\sigma} > 0$$

Proof. For sufficiently low criminal ability, the payoff from joining the gang is negative for all $c_i^* \geq 0$. Consequently, for sufficiently small $\sigma_i \leq \underline{\sigma}$ the youth would never join the gang. ■

Given these results, a sufficient condition for $\frac{\partial \bar{c}}{\partial I} > 0$ is:

$$\int_{t=\sigma}^{1} \frac{\underline{c}^3}{2\hat{I} + \frac{\hat{c}^3}{\underline{\sigma}^2} J''\left(\frac{\hat{c}}{\underline{\sigma}}\right)} dG_\theta(t) \quad > \quad \frac{g_\theta(\sigma)}{J'(\underline{c})}\left(\bar{c}_0 + \hat{c} - \underline{c}\right)$$

$$\Longleftrightarrow \quad \frac{g_\theta(\sigma)}{1 - G_\theta(\sigma)} < \frac{\underline{c}^3 J'(\underline{c})}{\left(2\hat{I} + \frac{\hat{c}^3}{\underline{\sigma}^2} J''\left(\frac{\hat{c}}{\underline{\sigma}}\right)\right)\left(\bar{c}_0 + \hat{c} - \underline{c}\right)}$$

Defining $\Lambda \equiv \frac{\underline{c}^3 J'(\underline{c})}{\left(2\hat{I} + \frac{\hat{c}^3}{\underline{\sigma}^2} J''\left(\frac{\hat{c}}{\underline{\sigma}}\right)\right)\left(\bar{c}_0 + \hat{c} - \underline{c}\right)}$ completes the proof. ■

Chapter 3

Recruitment to Organised Crime

3.1 Introduction

Over recent years, a myriad of policies have been suggested to tackle youth involvement in crime. The reason for this is simple: crime inflicts a cost upon society. Recent estimates (Cohen and Piquero 2009) place the average present value to society of saving one high-risk eighteen year-old from a life of crime between \$2.6 million and \$5.3 million. This includes between £675,000 and \$1 million in lost productivity. Over the last decade, estimates of these costs have increased substantially. In an earlier study employing a similar approach (Cohen 1998), the headline cost was between \$1.7 million and \$2.3 million, with lost productivity accounting for around £155,000. Admittedly, a large proportion of the increase is the result of improvements in measurement techniques. Nevertheless, crime is much more costly than hitherto imagined. The individual involved suffers from foregone education, likely drug use, and potential punishment. Wider society is forced to invest in security, pay for public prosecution and incarceration, and suffer from victimisation and the fear of crime. The recent intensification of research into these policies is therefore unsurprising.

All of the policies put forward are based upon the same argument: increasing the opportunity cost of crime reduces its amount and intensity. In turn, this reduces the social cost of crime. Broadly speaking, this appears sound. However, in one special case - that of organised crime - it may fail. Criminal organisations are proactive. They are more able to able to adapt to changes in their environment. They can respond to new policies, counteracting their effects by adjusting their inputs and the wages they offer. Under certain conditions, discussed herein, the social cost of crime may even increase.

How large is the problem? Wage data from a Chicago gang (Levitt and Venkatesh 2000) suggest that the cost of youths joining criminal organisations may be towards the top end of Cohen and Piquero's estimates. They report that the gang studied valued the lives of its members at somewhere between $8,000 and $500,000. Even the upper estimate is only 10% of the average value of a human life in the literature, suggesting an average loss to society of around $4.5 million. A recent survey (Egley and Howell 2011) estimated that there are 28,100 gangs active in the US, employing some 731,000 individuals. Combined with Cohen and Piquero's estimates, the annual cost of youth involvement in organised crime may be as high as $465 billion in the US, or 3% of US GDP (Bureau of Economic Analysis 2011)[1].

This paper develops a simple framework in which to study a criminal organisation's likely reaction to the implementation of policy. A profit-maximising gang has two available inputs, labour and violence, which it uses generate revenue through a range of activities, from prostitution and drug dealing, to people trafficking and protection. Heterogeneous youths grow up in the gang's neighbourhood. During their early years, they have the opportunity to acquire criminal skills. They then decide where to seek employment. If they join the formal economy, they are paid a flat wage. If instead they opt for a criminal career,

[1] Cohen and Piquero use a constant 2% discount rate. Author's estimates are based upon youths being active from age 18 to 26 and use 2009 US nominal GDP.

they join the neighbourhood gang. By joining the gang, they are required to inflict violence. The exact amount of violence, and the form of compensation provided, depends upon the gang's ability to discriminate between youths with different abilities. At its most basic, the gang will offer a flat wage, and require all youths to inflict the same amount of violence. A discriminating gang, on the other hand, may offer a range of different wages, each associated with a different level of violence.

When a policy is introduced that affects the youths' incentives, the gang's reaction depends upon how variation of inputs affect profits. If violence and size are complementary in the profit function, any policy that aims to reduce the incentive to join the gang also reduces the amount of violence the gang employs, unequivocally lowering the social loss from crime. If, on the other hand, violence and size are profit substitutes, a variety of outcomes may arise. Falling size may cause the gang to substitute towards violence (in a similar manner to Poutvaara and Priks 2009 and 2011) or vice versa. Perversely, this could increase the social cost of crime. Policies are therefore most effective when they not only reduce gang membership, but also hamper the gang's ability to intensify violence. For example, prevention of juvenile crime may prove particularly effective. It not only increases the opportunity cost of crime, but also reduces youths' ability to learn criminal skills. Conversely, improving labour market conditions, conditional on a youth still choosing to become a criminal, have no impact upon their incentives to learn criminal skills. Whilst membership is reduced, such policies unambiguously increases the marginal profitability of violence, causing the gang to intensify its activities.

3.1.1 Criminal Organisations as Firms

Two approaches are often employed when considering the activities of organised crime, reflecting two different literatures. Those considering the origins of criminal organisations (for example Gambetta 1996, Skaperdas and Syropoulos 1997 or Dixit 2007) think of them as pseudo-states, filling the void left by weak law enforcement. This literature views a local monopoly over violence as the defining characteristics of organised crime. Conversely, it is through the lens of a profit-maximising firm that established criminal organisations are most successfully analysed (for example Garoupa 2000, Chang, Lu, and Chen 2005 or Kugler, Verdier, and Zenou 2005). This paper adopts the second approach, whilst incorporating elements of the first.

Some tailoring of the standard theory of the firm apparatus is necessary. Criminal organisations are not regular firms; their property rights are not protected by statute; nor are their activities constrained by it. Moreover, their activities are under constant (violent) threat from law enforcement agencies, as well as competitors. As such, their factors of production are slightly different. The economic and criminological literatures point to two inputs being common to all flavours of criminal organisation: number of members and violence.

As with all firms, labour is critical for criminal organisations. In addition to its traditional role in production (Chang, Lu, and Chen 2005), there are several economies of scale that a larger organisation is able to take advantage of. Transactions within illegal markets are fraught with risk. Trading partners are prone to cheating one another. There is a constant threat from undercover police officers. These risks lead to a reduction in trade (Cook, Ludwig, Venkatesh, and Braga 2007) giving rise to the usual inefficiencies (from the criminal organisation's perspective, at least). Large criminal organisations can internalise a lot of these transactions. By vouching for its members, and inflicting severe punishments on those who renege, a criminal organisation enables individuals to trade with one another in safety. Membership acts as a guarantee. They also make it harder for police to infiltrate them by being more self-sufficient.

Sah 1991 notes that larger criminal organisations can also stretch police resources. As the number of members increases, the probability than any one individual will be arrested may diminish. Consequently, larger organisations suffer proportionally less from police disruption.

Violence is the other key component of criminal enterprises. At first pass, the reasons seem obvious. Firms operating a protection racket must be willing and able to use violence against those who disrupt their clients' businesses (Gambetta 1996, Dixit 2007). As protection often evolves into extortion, violence may also be required to ensure that clients continue to pay their fees (Garoupa 2000). Violence (more frequently, the threat of violence) can thus be seen as a direct input into the criminal organisation's production function.

Violence is equally important as a mechanism for reducing disruption to the organisation's other operations. As it is impossible for the organisation to operate in isolation, numerous stakeholders will, over time, gain information that could implicate it in various illegal activities. Violence is used to ensure that the cost of informing the police is prohibitively high.

Of these stakeholders, the ones with the greatest potential to harm the organisation are its members. Baccara and Bar-Isaac 2008 consider the problem of information diffusion within a network design framework. Where the threat of violence is credible, and sufficient to prevent members implicating the organisation, a hierarchical structure proves to be optimal. Information is passed throughout the organisation efficiently. Organisations, such as terrorist groups, who cannot rely on their members not to share information, are forced into a much less efficient cell structure. The use of extreme violence against members-turned-informants is well documented, particularly with regard to the Italian Mafia's code of secrecy; the *omertà* (see, for example, Gambetta 1996, Paoli 2003, Raab 2006). Similarly, violence may be employed to prevent residents of the organisation's territory (where its activities are profuse) from interfering, or cooperating with law enforcers.

Finally, criminal organisations use violence to protect their local monopolies from competitors. Silverman 2004 provides a relevant economic analysis, admittedly in a broader economics of crime context. He shows that there is an incentive for individuals to develop a reputation for violence, even if they are not inherently violent. By doing so, only those with a genuine predilection for violence will stand up to them. Similarly, criminal organisations may develop a reputation for violence to deter rivals.

Employing these inputs, criminal organisations' output is varied. They tend to be active in a variety of markets, from drugs and prostitution to protection and smuggling. Irrespective of which market(s) the organisation is operating in, size and violence are revenue complements. Members of more violent groups suffer less disruption, enjoy stronger monopolies, and may even be able to extort higher revenues. In other words, the aggregate marginal revenue product of labour is increasing in violence, and vice versa.

Criminal organisations' costs, on the other hand, are primarily wages paid to their members. Whilst the organisation's leaders dictate the extent of violence employed, it is the foot soldiers that face the cost of implementing it during the commission of their crimes. Levitt and Venkatesh 2000 show that, over a four-year period, members of drug-selling gang in Chicago had a 25% chance of dying (versus a 0.4% chance nationwide in the same demographic). Over the same period, they also suffered an average of two non-fatal injuries (ranging from gunshots and knife wounds to beatings). In order to attract and retain members, wages must incorporate compensation for the violence they are forced to inflict. This provides an incentive for criminal organisations to substitute between size and violence. By increasing the amount of violence it requires its members to employ, the organisation's wage bill grows substantially, as each member must be compensated. In this sense, the marginal cost of size is also increasing in violence, and vice versa.

Whether violence and size are complement or substitute inputs in the organisation's profit function depends upon which effect dominates. When the gang's size increase, the marginal revenue it derives from violence also increases. For example, with more members, the gang is able to extort money more successfully from its neighbourhood. This provides an incentive to increase violence. However, any increase in violence requires that the gang compensate its members. With a larger gang, more members need to be compensated. As such, an increase in size also increases the marginal cost of violence. Whether violence's profitability increases when size increases clearly depends upon whether marginal revenue or marginal cost increases more. This will play a key role in determining the effectiveness of policy.

3.1.2 Opportunities in a Criminal Neighbourhood

Youths growing up within a criminal organisation's territory face several difficulties when seeking work in the primary labour market. Such neighbourhoods tend to develop reputations for criminal activity. If this leads employers to engage in discrimination, and reduce their expectations regarding prospective employees' productivity from that neighbourhood, finding a job could be more challenging (Verdier and Zenou 2004). In turn, poor prospects provide less incentive to acquire human capital, creating a self-fulfilling prophecy (Lundberg and Startz 1983). Austen-Smith and Fryer 2005 suggest that this may be compounded by peer pressure. Attempting to gain a good education can be interpreted as a negative signal by a youth's peer group. This provides a further disincentive to acquire the skills necessary to seek employment. Combined, these two effects can lead to situations where the opportunities open to youths from such a neighbourhood are severely limited in the primary labour market.

In sharp contrast, opportunities abound in the informal economy. Success in this industry requires the acquisition of a different set of skills. Youths must develop an acceptance of, and willingness to use, violence. Ballester, Calvó-Armengol, and Zenou 2010 explore the acquisition of criminal skill in a network. As no schools offer certificates in intimidation (to this author's knowledge), they suggest that youths may acquire such skills through a mixture of trial and error (juvenile crime), and observing others' mistakes. This second channel - learning from others' mistakes - can cause of a great deal of heterogeneity in youths' intrinsic ability to acquire criminal skills. In their analysis, Ballester *et al* suggest that youths who feature centrally in a juvenile network will find skill acquisition far easier. They can observe others more easily, suffering less from trial and error. Once they have joined the criminal organisation, they will find inflicting violence far less costly as a result.

There is a broad range of criminological evidence in support of a learning process by which individuals become accustomed to using violence. Various works by Athens (summarised in Rhodes 2001) identify a common process of 'violentization' undergone by a large sample of prisoners incarcerated for violent crime. During this process, individuals are first desensitised to violence, before learning (through positive reinforcement) that it is an appropriate response to minor provocations. Juvenile gangs play an important role in this system. Esbensen and Lynskey 2001 interviewed fourteen year-olds in the US who claimed to be members of a juvenile gang. Of those, 25% claimed to have shot at someone.

FBI statistics back up this claim. 80% of gang murders in 2009 were attributed to juvenile gangs (Federal Bureau of Investigation 2010). This is not merely a US phenomenon, however. Esbensen and Weerman 2005 conducted a similar study in The Netherlands. They found that members of a juvenile gang were four times more likely to be involved in violence than those not involved in a gang. Similarly, Salagaev *et al* 2005 found that Russian gang youths were seven times more likely to use violence to appropriate money or other goods. Whilst a proportion of these results may be attributed to bravado, they nevertheless indicate an acceptance of violence as a means to an end.

3.1.3 Tackling Organised Crime

Policies designed to tackle crime, and organised crime in particular, are motivated by the fact that crime is costly to society. The scale of the total loss is ultimately related to the organisation's two transient features - intensity of violence and membership size. In particular, criminal organisations inflict three main externalities. As with all crime, by far the largest (Cohen and Piquero 2009) is the fear of crime and victimisation costs that crime generates. With a larger gang, one would expect that more individuals are going to be victims of crime. Moreover, the loss each individual suffers will be increasing in the amount of violence inflicted during the commission of each crime.

Secondly, society must expend resources protecting itself from the gang. This includes the cost of preventative measures, from the policies discussed herein, to the measures taken by individuals to protect themselves. It also includes the resources expended investigating and prosecuting criminals, as well as the cost of enforcing punishment. Once more, we would expect these costs to be increasing in both size and violence. With more members, crime is more prevalent. As such, the chance of becoming a victim of crime is higher. More crimes will need investigation, and will lead to more prosecutions. If violence is higher, a victim suffers a greater loss. As such, the return to investing in protection is higher and more investment will occur.

Finally, criminal organisations can be the cause of economic discrimination à la Lundberg and Startz 1983. It is possible that even youths who gain positive surplus by joining the gang would do better in the primary labour market if the gang were not present. Employers in the primary labour market may believe that youths from the gang's neighbourhood are less productive than those from elsewhere. As such, they will offer lower wages to anyone applying from that neighbourhood. This reduces the incentive to acquire primary labour market skills, and creates a self-fulfilling prophecy. Without the gang, no such signal would be generated and youths would be offered higher wages. In this sense, the gang is a source of economic discrimination[2]. Again, we would expect the size of the discrimination to be increasing in both size and violence. If the gang is larger, fewer youths will acquire primary labour market skills. As such, the average amount of skill acquired will be lower. This will lead to lower wage offers. Similarly, if the gang is more violent, employers are likely to have a more critical opinion of those from the neighbourhood. Again, lower wage offers will result.

Since the inception of the economics of crime, economists have been suggesting policies to reduce criminal activity. In Becker's seminal paper of 1968, he proposed that a relatively cheap way to reduce crime was to simply increase the fine incurred when caught. Since then, an abundance of policies have been put forward, each aiming to manipulate the incentives of would-be criminals. This paper considers four broad categories:

1. Increasing the severity of punishment;

[2]Indeed, the gang has an incentive to maximise the discrimination against youths from its neighbourhood. Discrimination will enable it to offer lower wages for the same amount of violence.

2. increasing arrest and conviction rates;

3. primary labour market policies; and

4. prevention of juvenile crime.

The idea of increasing the severity of punishment dates back to Becker. By increasing the size of fine, the expected payoff to committing crime is reduced, given a constant arrest rate. This reduces the incentive to become involved in crime, relative to staying honest. Related to this is increasing arrest and conviction rates. At first pass, the effects should be equivalent. Certainly, Burdett, Lagos, and Wright 2004 find that, in addition to reducing the incentive to commit crime, both policies increase average wages and reduce inequality. However, when we turn our attention to organised crime, this equivalence begins to break down. Increasing arrest rates reduces the number of members available for the organisation to utilise. This, in turn, will impact upon the wages they are willing to offer, and even their optimal levels of violence. Whilst increasing the length of prison terms may have similar effects, other increases in severity may not.

Primary labour market policies cover an extremely broad range of suggestions, all aiming to increase the wage paid in the formal economy. A by-product of this is a fall in crime. Two cases stand out, however, for their direct targeting of high-risk youths. The now famous Perry Preschool Programme (see Parks 2000), focused on poor black preschool children with low IQs in the 1960s. Participants attended preschool classes for 2.5 hours per day, five days per week. The preschool teachers also engaged with parents, visiting them for a further 1.5 hours each week. Participants were tracked over forty years, creating a reasonably comprehensive data set on their educational, employment and criminal outcomes. Recent analysis, whilst downgrading previous measures of success, still suggest that the project yielded an internal rate of return of 7%-10% (Heckman, Moon, Pinto, Savelyev, and Yavitz 2010). The second example relates to a range of case studies by the Education Innovation Laboratory at Harvard University. The Harlem Children's Zone in New York combines an intensive education programme with access to community services (Dobbie and Fryer 2011). EdLabs are also involved in numerous projects across the US aimed at high school students (Fryer 2010). For example, the Paper Project in Chicago targets ninth and tenth grade students. The organisers pay the students for passing their classes. They can earn up to $2,000 per year, with 50% payable upon graduation from high school.

Prevention of juvenile crime increases the cost of acquiring criminal skills, and is at the heart of the arguments put forward by Ballester, Calvó-Armengol, and Zenou 2010. By disrupting juvenile networks, the authorities are able to increase the cost of acquiring criminal skills. Youths are forced to learn in isolation, and are unable to learn from other mistakes. They are thus less likely to join criminal organisations, as they will find inflicting violence to be too costly.

Each policy works by reducing a criminal organisation's demand for members and/or violence. As mentioned above, if size and violence are complementary in the profit function, a reduction in one creates an incentive to reduce the other. Conversely, if they are substitutes, it creates an incentive to increase the other. Under certain conditions, this countervailing effect may actually cause net increases in the demand for size or violence, increasing the loss suffered by society at the hands of organised crime. It is this mechanism that the rest of the paper is devoted to investigating.

The remainder of the paper proceeds as follows. In the next section, I introduce a model of recruitment to organised crime, and define two different types of gang. In section 3.3 I discuss the behaviour of the first of these types: one which can only offer a single wage and violence contract. Having done this, section

3.4 discusses the likely impact of policy under various conditions, and highlights where policies may have unintended consequences. Section 3.5 extends the analysis, discussing a situation in which the gang can offer a range of jobs to youths. Section 3.6 considers the effects of policy in this more complicated setting. Finally, section 3.7 concludes.

3.2 A Model of Recruitment to Organised Crime

I present a model of organised crime recruitment. The economic environment, hereafter referred to as the *neighbourhood*, consists of two sectors: the *primary labour market*[3] and *the gang*[4]. A mass N of heterogeneous youths grow up in the neighbourhood. After investing in appropriate skills, they decide where to seek employment. Whilst the primary labour market is passive, the gang acts as a monopsonist employer of criminals in the neighbourhood. Gang leaders adjust their approach to recruitment in order to maximise the gang's profits.

The gang offers a *contract schedule* comprising of a series of wage and violence intensity pairs $(g(s), V(s))$. Recruits to the gang are able to choose any contract, s, they wish from the available menu. I assume that contracts are binding on both sides, so by choosing a contract the recruit commits to inflicting a given level of violence in exchange for the associated wage.

The revenue each gang member, i, generates, be it from drug sales or extortion, prostitution or people trafficking, depends primarily upon two characteristics: the size of the gang (total membership, denoted by $M \geq 0$) and the intensity of violence they inflict, $V(s_i)$. Each individual's revenue stream is given by the function[5], $r(M, V(s_i))$. I assume that, at the gang level, $Mr(M, V)$ is

[3]This terminology follows Huang, Laing, and Wang 2004, who develop a similar model of predation.

[4]Whilst the terminology used refers to a street gang, the model presented is equally relevant to alternative forms of organised crime.

[5]The black box nature of revenue (as opposed to production) is purely for notational ease.

subject to positive but diminishing marginal returns and constant returns to scale. The constant returns to scale assumption is not necessary, but does make for a simpler characterisation of the equilibrium and policy effects. Moreover, there is some degree of complementarity between violence and size, insofar as $r(M,0) = r(0,V) = 0$ for all $M, V \geq 0$.

As each gang member receives a wage, $g(s_i)$, the profit they generate for the gang is given by:

$$\pi(M, s_i) \equiv r(M, V(s_i)) - g(s_i) \tag{3.1}$$

The gang chooses its contract schedule to maximise total profits:

$$\Pi\left(M, (g(s), V(s))_{s\geq 0}\right) \equiv ME\left[\pi(M, s_i)\,|i \text{ joins}\right] \tag{3.2}$$

The contract schedule becomes common knowledge, and is announced prior to any decisions by youths.

Youths vary in their intrinsic criminal ability, denoted by σ and distributed exponentially with parameter $\lambda > 0$[6]. Youths simultaneously make two decisions. Firstly, they choose how much criminal skill to acquire. They then decide which sector to work in. Acquiring criminal skill is a costly process. However, those with a higher criminal ability find it easier than those with a lower ability. Denoting the amount of criminal skill acquired by youth i by c_i, the cost of acquiring criminal skill is given by $kJ\left(\frac{c_i}{\sigma_i}\right)$. $J(\cdot)$ is a strictly increasing and convex function. $k > 0$ reflects the fact that policy can influence how easy it is to acquire criminal skills.

For simplicity, the primary labour market pays an exogenously given flat wage rate, $w \geq 0$. One can consider this wage to be net of any cost of education, as

One can think about it as an indirect revenue function: the one resulting from the optimal allocation of inputs across the wide range of activities the gang engages in. Kugler *et al* 2005 consider a more structured approach, decomposing revenue into the number of crimes committed, and the bounty collected from each crime.

[6] Note that this automatically ensures that the hazard rate on the distribution of criminal abilities is bounded, as suggested in the previous chapter.

well as incorporating the probability of unemployment. Should the youth join the gang instead, they choose a contract, s_i, from the available menu. However, as discussed in the introduction, being involved in the gang is a dangerous affair. There is a possibility of arrest and conviction. Following Becker 1968, arrest occurs with probability p, resulting in a fine of size f[7] and wages being confiscated[8]. Moreover, gangs are violent enterprises. Whilst the gang leaders choose the level of violence the gang is known for, it is the members who must bear the cost of inflicting that violence. It is at this point that acquiring criminal skill pays off. By investing effort in learning to be a criminal, youths become desensitised to violence. So, whilst all gang members suffer disutility from having to engage in violence, those who have acquired large amounts of criminal skill suffer less. In particular, each youth suffer disutility $-\frac{V(s_i)}{c_i}$. I assume that arrests are always made after a crime has been committed. Since youths inflict violence during their crimes, they therefore suffer this disutility irrespective of whether they are subsequently arrested. The payoff from joining the gang is therefore:

$$G\left(s_i;\sigma_i\right) \equiv (1-p)\,g\left(s_i\right) - pf - \frac{V\left(s_i\right)}{c_i} - kJ\left(\frac{c_i}{\sigma_i}\right) \tag{3.3}$$

In the remainder of the paper, I will distinguish between two extreme types of gang. A gang is *simple* if it can only offer a single contract, (g, V), to all members. In this sense, it is a single-price monopsonist. All recruits receive the same wage, and inflict identical levels of violence. At the other end of the spectrum, a gang is *separating* if it can offer a full range of contracts, $\left(g\left(s\right), V\left(s\right)\right)_{s\geq 0}$, and, in particular, if these contracts satisfy the youths' incentive compatibility constraints. As

[7] Whilst f is a constant in this model, the results that follow would apply equally to situations in which different crimes receive different punishments. In that case, an increase in f would be equivalent to a situation in which all punishments became more severe, but the gradient of the punishment schedule remained unchanged.

[8] This is a simplification. In reality, there is some evidence to suggest that criminal organisations pay members' families whilst they are incarcerated. However, as they are unable to take advantage of other membership benefits, their gang wage does go down.

a results, recruits fully separate according to their abilities; equivalent to second degree price discrimination[9]. Whilst I am agnostic regarding which type of gang better represents real criminal organisations, this enables me to demonstrate the robustness of my policy results.

3.3 The Simple Gang

The model with a simple gang yields a pooling perfect Bayesian equilibrium. The gang announces its choices of V and g to maximise profits. Youths then acquire criminal skills and choose a career, conditional on the announced V and g, as well as their criminal ability, σ. As per usual, the equilibrium is found by backwards induction.

3.3.1 Youth Decisions

Taking the announced level of violence and gang wage as given, a youth with criminal ability σ_i faces the following utility maximisation problem:

$$\max_{j\in\{0,1\},c\geq 0}\left\{(1-j)\,w + j\left[(1-p)\,g - pf - \frac{V}{c}\right] - kJ\left(\frac{c}{\sigma_i}\right)\right\} \tag{3.4}$$

where $j \in \{0,1\}$ takes value one when the youth chooses to join the gang and zero otherwise.

Consider first the choice of criminal skill, conditional upon career choice. If the youth chooses to join the primary labour market, criminal skill is of no use to them. They consequently do not incur the cost of acquiring it, and select $c_i^* = 0$.

[9] One can easily reinterpret this model in the terms described in Chapter 2. Since initiation difficulty is sufficient to describe how the gang's reputation changes, the revenue it generates could be redefined as a function of gang size and initiation difficulty. Thus interpreting violence in this model as initiation difficulty is relatively straightforward. Note, however, that the gang's objective is different - profit maximisation.

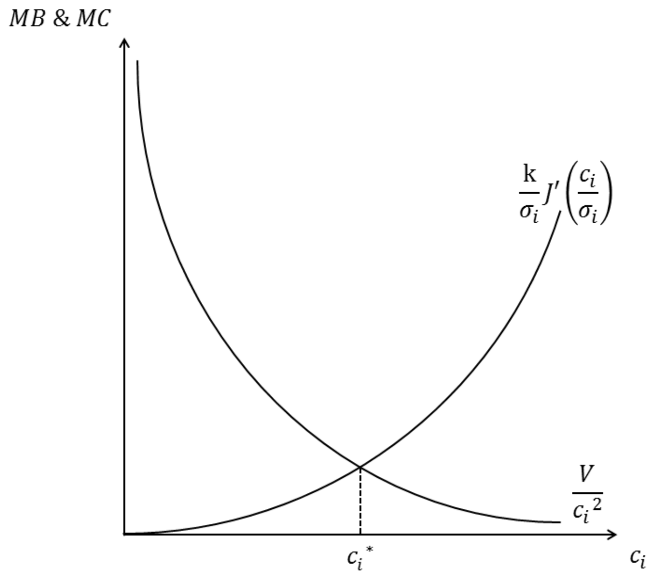

Figure 3.1: A σ-ability youth's choice of criminal skill.

Conversely, if they decide to join the gang, they choose $c^* (\sigma_i, V)$ satisfying:

$$\frac{V}{c^* (\sigma_i, V)^2} \equiv \frac{k}{\sigma_i} J' \left(\frac{c^* (\sigma_i, V)}{\sigma_i} \right) \tag{3.5}$$

Equation (3.5) is represented in Figure 3.1. The resulting $c^* (\sigma_i, V)$ is strictly positive, and increasing in both the level of violence employed by the gang and the criminal ability of the youth. More violence increases the marginal benefit of acquiring criminal skills, whereas increasing criminal ability reduces the marginal cost.

Given youths' choice of criminal skill, it is straightforward to show that the payoff to joining the gang is strictly increasing in criminal ability. Consider a youth with ability $\sigma' > 0$. Suppose that they join the gang, and acquire the optimal amount of criminal skill, $c^* (\sigma', V)$. Now consider a youth with ability $\sigma'' > \sigma'$. If this youth joins the gang and acquires the same amount, $c^* (\sigma', V)$, of criminal skill, then they will enjoy the same wage and suffer the same disutility from violence. However, since they have higher criminal ability, the cost of acquiring $c^* (\sigma', V)$ is lower. They can therefore guarantee themselves

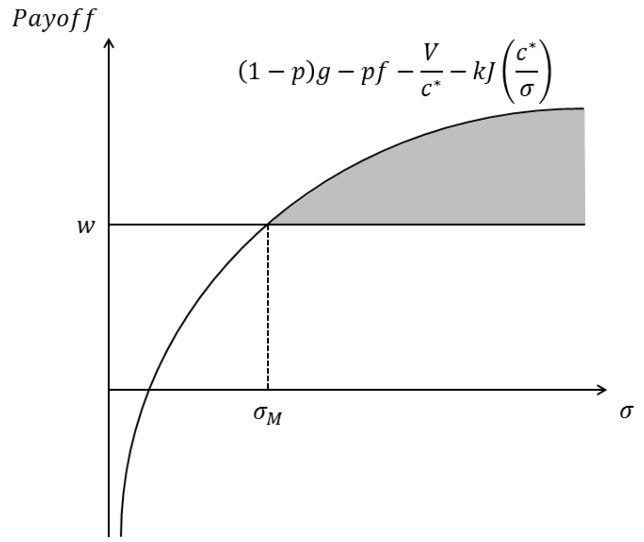

Figure 3.2: Youths' payoffs from gang membership versus the primary labour market.

a strictly higher payoff than the youth with criminal ability σ', and will do even better by acquiring more criminal skill. Conversely, the payoff from joining the primary labour market, w, is independent of a youth's criminal ability. We can therefore conclude that there exists some σ_M such that a youth will join the gang if and only if $\sigma_i \geq \sigma_M(V, g)$. σ_M is defined by:

$$(1-p)\,g - pf - \frac{V}{c^*(\sigma_M, V)} - kJ\left(\frac{c^*(\sigma_M, V)}{\sigma_M}\right) \equiv w \tag{3.6}$$

We call the youth with ability σ_M the *marginal youth*. Since all youths with ability above σ_M join the gang, its size will be given by $M = N(1-p)\,e^{-\lambda\sigma_M}$. A proportion $e^{-\lambda\sigma_M}$ of the mass N youths join the gang. However, a proportion p are arrested and convicted, making them unproductive (what Levitt 1996 calls the *incapacitation effect*).

Youths' career choices are represented by Figure 3.2. An increase in the wage offered by the primary labour market increases the opportunity cost of joining the gang, raising the ability of the marginal youth. Similarly, increases in the conviction rate, severity of punishment, the degree of violence employed by the gang, or the cost of acquiring criminal skill reduces the payoff from joining the gang. Once again, this raises the ability of the marginal youth and, by extension, lowers the gang's size. The shaded area represents the surplus accruing to gang members. Since the gang is a single price monopsonist in this setting, all gang members receive a positive surplus through membership. Moreover, since the cost of violence decreases with criminal ability, higher ability youths receive a larger surplus than those with lower ability. Note that this does not imply that organised crime generates a social surplus. It simply suggests that those who join the gang are better off doing so in equilibrium. As noted in the introduction, gang activity have a tendency to suppress formal wages and reduce the incentive to invest in formal human capital. Thus it may be the case that, were the gang not present, youths could guarantee themselves an even higher payoff in the primary labour market.

3.3.2 Gang Leader Decisions

During the exposition of the model, the gang leadership's profit maximisation problem was described as a decision regarding the degree of violence it expected gang members to engage in, V, and a wage rate it offered members, g. As a result of these decisions, some membership size, M, was induced. It turns out that a computationally easier way to view the gang leadership's problem is to think about them choosing the degree of violence, and then compensating gang members sufficiently to induce a chosen gang size. The extent of the compensation is derived as follows. In order to acquire gang size of precisely M, it is necessary that the marginal youth have criminal ability:

$$\sigma_M = \frac{\ln N + \ln(1-p) - \ln M}{\lambda} \tag{3.7}$$

This youth must therefore be indifferent between the gang and the primary labour market. In order to ensure this with degree of violence V, the gang must offer a wage:

$$g(M,V) \equiv \frac{w+pf}{1-p} + \frac{V}{c_M^*(1-p)} + \frac{k}{1-p} J\left(\frac{\lambda c_M^*}{\ln N + \ln(1-p) - \ln M}\right) \tag{3.8}$$

where $c_M^*(M,V)$ is the criminal skill of the marginal youth, defined implicitly by:

$$\frac{V}{c_M^*(M,V)^2} \equiv \frac{\lambda k}{\ln N + \ln(1-p) - \ln M} J'\left(\frac{\lambda c_M^*(M,V)}{\ln N + \ln(1-p) - \ln M}\right) \tag{3.9}$$

The gang leadership's profit maximisation problem thus becomes:

$$(M^*, V^*) = \underset{M \in [0,N], V \geq 0}{\arg\max} \{Mr(M,V) - Mg(M,V)\} \tag{3.10}$$

Before continuing to outline the solution to (3.10), it is expedient to discuss an issue alluded to in the introduction. When violence increases, the marginal revenue product of size increases. Gang members face less disruption, a stronger monopoly, and may even be able to extort higher prices. Concurrently, however, the marginal cost of labour also increases. Each member is being forced to engage in more violence, increasing the loss they suffer as a result. The gang must offer additional compensation according to (3.8), in order to prevent those with relatively low criminal ability from opting to join the primary labour market instead. These two effects counteract one another, and consequently, the net effect on the marginal profit generate by size is unclear. Determining which effect dominates is not only helpful when describing the equilibrium, but proves to have important implications for the impact of policy in this environment. It is straightforward to show that:

$$\Pi_{MV} = \frac{1}{Mc_M^* (1-p)} \left[\eta - \frac{1}{\lambda \sigma_M} \frac{1+\varepsilon_M}{2+\varepsilon_M} \right] \tag{3.11}$$

where $\eta = \frac{Mr_{MV}(M,V)}{r_V(M,V)} > 0$ is the cross elasticity of the individual marginal revenue product of violence with respect to labour, and $\varepsilon_M = \frac{\frac{c_M^*}{\sigma_M} J''\left(\frac{c_M^*}{\sigma_M}\right)}{J'\left(\frac{c_M^*}{\sigma_M}\right)}$ is the elasticity of the marginal cost of acquiring criminal skills with respect to $\frac{c_M^*}{\sigma_M}$. Clearly, η represents the relative increase in marginal revenue product of violence. The second term in parenthesis in (3.11) reflects the relative increase in the wage the gang offers. As the degree of violence increases, the amount of compensation required increases. However, this effect is tempered by the fact that youths invest in more criminal capital.

These terms depend upon the functional forms of $r(\cdot,\cdot)$ and $J(\cdot)$ respectively. It is therefore helpful to make one of two assumptions:

Assumption 4 (Complements) *The marginal revenue product of violence is sufficiently elastic with respect to size to ensure that* Π_{MV} *is always positive.*

Assumption 5 (Substitutes) *The marginal revenue product of violence is sufficiently inelastic with respect to size to ensure that* Π_{MV} *is always negative.*

These assumptions are illustrated in Figure 3.3. When η is large relative to M, the marginal cost of compensating youths for increasing amounts of violence is relatively small compared to the increase in marginal revenue. As such, Assumption 4 holds. In particular, if $\eta > \frac{1}{\lambda\sigma_M}$, then the revenue effect always dominates the cost effect, irrespective of the functional form of $J(\cdot)$. Conversely, if η is small relative to M, the opposite is true. If $\eta < \frac{1}{2\lambda\sigma_M}$, the relative increase in marginal revenue is insufficient to compensate for the relative increase in marginal cost, irrespective of the functional form of $J(\cdot)$. For η in between these

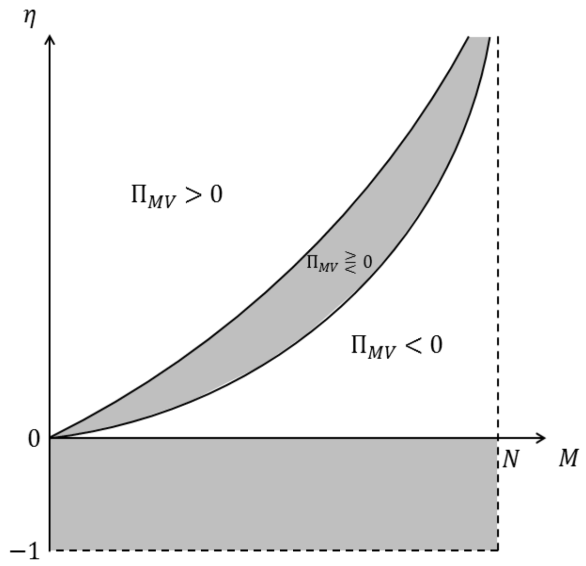

Figure 3.3: The range of values over which Assumption 4 and Assumption 5 hold.

two values, the situation is less clear, and the convexity of the criminal skill cost function becomes important.

We are now in a position to outline the gang leaders' choices.

Proposition 8 (Profit Maximisation) *Suppose that $\eta > 0$, and that either Assumption 4 or Assumption 5 holds. Then the gang leadership's profit maximisation problem given by (3.10) has a unique solution, with $V^* > 0$ and $0 < M^* < N$.*

Proof. See Appendix 3.A. ■

Requiring that $\eta > 0$ serves two purposes and is sufficient to ensure a maximum exists under both Assumptions 4 and 5. The marginal revenue product of violence declines as violence increases. However, as the gang requires greater feats of violence from its members, each youth optimally invests more heavily in acquiring criminal skills. As a result, each marginal increase in violence, dV, has a smaller impact upon the cost youths bear from inflicting violence, $\frac{dV}{c^*}$, since c^* is larger. Consequently, the gang must increase its compensation for inflicting violence by smaller amounts as violence increases: the marginal cost of violence is also decreasing. In order for a maximum to exist, we require that marginal revenue decline faster than the marginal cost. A sufficient condition for ensuring this is that $\eta > -\frac{1}{2}$.

Under Assumption 4, $\eta > -\frac{1}{2}$ guarantees that a unique maximum exists. Unfortunately, if violence and size are substitutes, the incentive to substitute may be strong enough to move the gang towards one of the extremes (high violence, tiny membership or vice versa). To counteract this, we require that a decline in violence reduces the marginal revenue product of size sufficiently as to admit interior maximum. This, in turn, requires the slightly stronger condition that $\eta > 0$.

Assumptions 4 and 5 are not necessary, but are sufficient to ensure uniqueness of the equilibrium. To see this, consider the restricted factor demand functions, $\widetilde{V}(M)$ and $\widetilde{M}(V)$. These are derived directly from the first order conditions (below), and give the gang's optimal choice of violence and membership size respectively, holding the other input constant:

$$r_V\left(M, \widetilde{V}\right) - \frac{1}{c_M^*\,(1-p)} \equiv 0 \quad (3.12)$$

$$r_M\left(\widetilde{M}, V\right) + Mr\left(\widetilde{M}, V\right) - g\left(V, \widetilde{M}\right) - \frac{V}{\lambda\sigma_{\widetilde{M}}c_{\widetilde{M}}^*\,(1-p)} \equiv 0 \quad (3.13)$$

By varying size in equation 3.12, it is possible to trace the gang's optimal choice of violence in (M, V)-space. Similarly, by varying violence in equation 3.13, one acquires the gang's optimal size. Maintaining the assumption that $\eta > 0$, Figure 3.4 displays the restricted demand functions for complements and substitutes. In both cases, the gang leaders' equilibrium choices are described by the intersection of the two curves, where their choice of violence is optimal given their size, and their choice of membership size is optimal given their level of violence. It is clear from Figure 3.4, however, that the comparative statics are

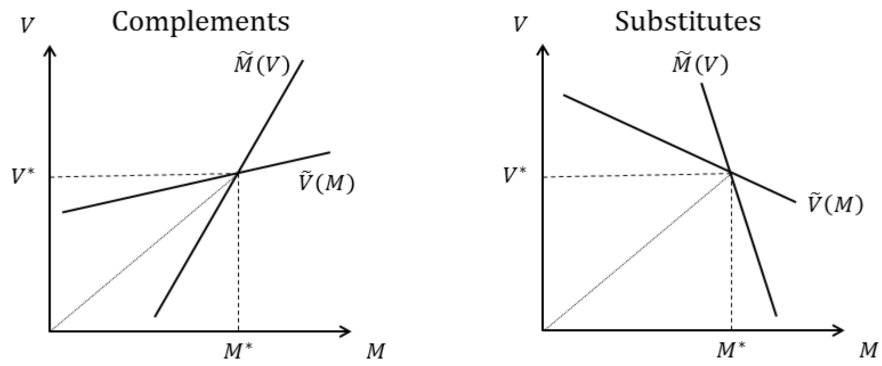

Figure 3.4: Restricted factor demands with complements and substitutes.

different. When membership size and violence are complementary, both restricted demand functions slope upwards[10]. In this case, an exogenous increase in, say, the gang's restricted demand for violence ($\widetilde{V}(M)$ shifts upwards) makes increasing size more profitable. Consequently, both increase concurrently. In contrast, when size and violence are substitutes, both curves slope downwards. An exogenous increase in the gang's restricted demand for violence makes size less profitable. In this case, the gang optimally reduces size as violence increases.

3.4 Policy with a Simple Gang

Using the framework developed in the previous section, we will now discuss the effect on the social loss from organised crime of the four broad policy areas outlined in the introduction. Under certain conditions to be made clear below, we will find that policies can have unintended consequences. In particular, anti-crime policies could, perversely, increase the amount of violence the gang employs

[10] The slopes of the restricted demand curves are derived by a simple application of the Implicit Function Theorem. $\widetilde{V}(M)$ is defined by $\Pi_V\left(M, \widetilde{V}\right) \equiv 0$. This yields:

$$\widetilde{V}_M = -\frac{\Pi_{MV}}{\Pi_{VV}}$$

Since $\eta > -\frac{1}{2}$, $\Pi_{VV} < 0$. Under Assumption 4, $\Pi_{MV} > 0$, so $\widetilde{V}_M > 0$. Under Assumption 5, $\Pi_{MV} < 0$, so $\widetilde{V}_M < 0$. An equivalent argument holds for $\widetilde{M}_V$, noting that the slope of the curve in Figure 3.4 is $\frac{1}{\widetilde{M}_V}$.

(but never its membership size).

Each of the policies in the introduction is associated with a parameter in the model. Specifically:

1. Increasing the severity of punishment increases f.
2. Primary labour market policies increase w.
3. Increasing the arrest or conviction rate increases p.
4. Prevention of juvenile crime disrupts the ability of youths to learn criminal skills, increasing k.

Note that the aim of this section is not to discuss *optimal* policy. That would require modelling of the technologies involved in manipulating these parameters, the associated cost functions, and a more detailed discussion of the exact nature of the social loss function. Rather, the aim is more modest: to highlight conditions under which policies designed to combat organised crime may in fact worsen one of its features. Each of these policies are under active discussion in both academic and political circles. I hope to help inform these debates by comparing each in the context of my model. I will first derive results for a generic policy, ϕ, before turning attention to the specific policies above.

3.4.1 Results for a Generic Policy

We can view the impact of a generic policy by considering its effects on the restricted demand functions. When a policy is implemented, it can change youth's incentives in two ways. Firstly, it may reduce the net benefit they gain from joining the gang. In order to retain members, this may necessitate the gang raising the wage they offer, or reducing the cost of violence they inflict upon their membership. Higher wages increase the marginal cost of size for the gang, as each new youth must be paid more, as shown in (3.13). Faced with the increased marginal cost, and no equivalent increase in marginal revenue, a profit-maximising gang will reduce its restricted demand for members, $\widetilde{M}$ as marginal profit derived from size, Π_M, becomes negative.

Secondly, a policy may affect how youths respond to changes in violence or gang wages. Some policies increase youths' sensitivity to violence. Once again, this will force the gang to increase its wage at every gang size, to retain the services of the marginal youth. Moreover, any increases in the level of violence the gang enforces will now require a larger amount of compensation. As such, the marginal cost of violence will also increase. As there is no equivalent increase in marginal revenue, the marginal profit derived from violence, Π_V, will also become negative. In this case, the gang will optimally reduce its restricted demand for violence, $\widetilde{V}$.

This intuition, combined with the first panel of Figure 3.4, leads us very quickly to our first result:

Proposition 9 (Policy with Complements) *Suppose that $\eta > 0$ and that Assumption 4 holds. Then any policy which reduces either Π_M or Π_V and does not increase the other, reduces both the amount of violence the gang employs, and the number of members that the gang chooses to recruit.*

This result is illustrated in Figure 3.5. If Π_M declines, then the restricted demand for size shifts inwards. Similarly, if Π_V declines, then the restricted demand for violence shifts inwards. Since both curves are upward sloping, any inward shift leads, unambiguously, to a fall in both size and violence. Now, any fall in, say, size reduces the marginal revenue product of violence. However, since fewer gang members need to be compensated for changes in violence, the marginal cost of violence also falls. If violence and size are complements, the fall in marginal revenue exceeds the fall in marginal cost, and the gang reduces

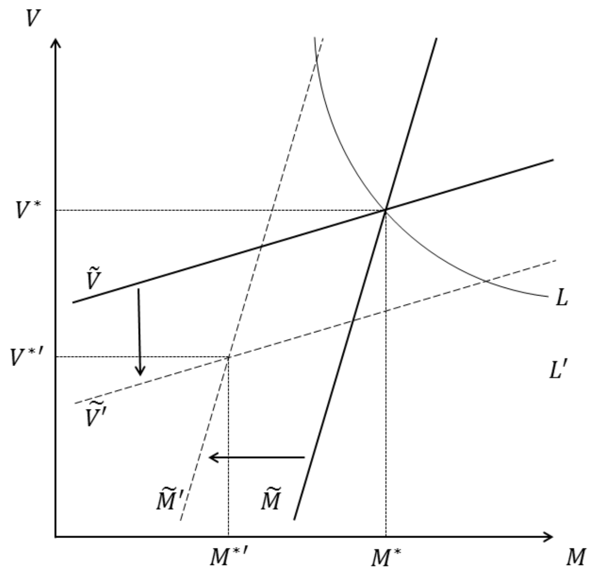

Figure 3.5: Policy effects with complements.

its optimal level of violence. In turn, this causes a further reduction in size. This cycle reinforces the decline in both size and violence, leading to the result. We can therefore conclude that, if membership size and violence are sufficiently strong revenue complements, any policy will be effective in reducing the loss society suffers at the hands of the gang. In Figure 3.5, L is the locus of all size and violence combinations giving the same social loss as (M^{*}, V^{*}); it is an iso-loss curve. As each policy leads to a reduction in both inputs, the resulting profit-maximising combination lies below L, indicating a smaller loss.

Unfortunately, the case with substitutes is not so clear cut, as shown in Figure 3.6. In contrast to complements, reductions in size cause the gang to substitute towards violence, and vice versa. As with complements, a fall in, say, size reduces both the marginal revenue product and the marginal cost of violence. Now, however, the decline in marginal cost exceeds the decline in marginal revenue, causing an endogenous increase in the restricted demand for violence. Similarly, a decline in violence makes size more profitable, causing an endogenous increase in the restricted demand for size. Consequently, if one of these effects were

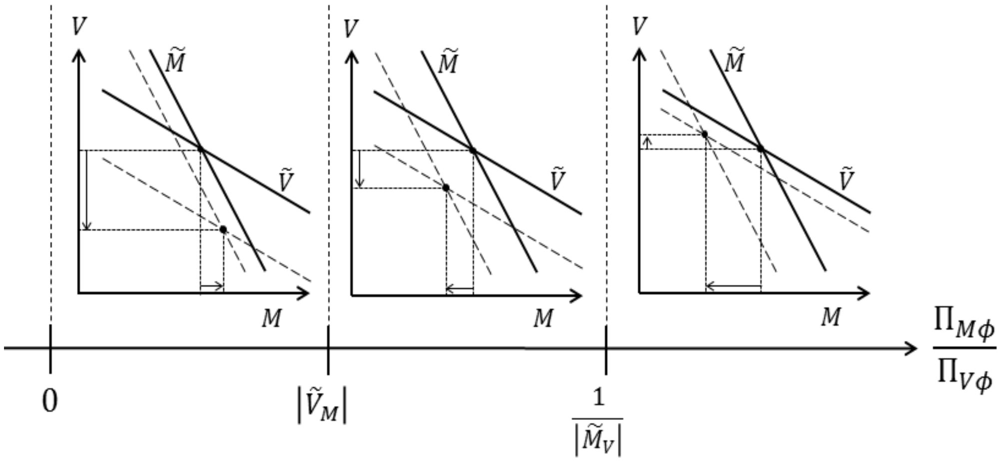

Figure 3.6: Policy effects with substitutes.

to offset the initial impact of the policy, we could have a situation in which either membership size or violence increases (but, fortunately, not both). As a result, it may even be possible that a policy designed to reduce the social loss from organised crime may actually increase it. The term which determines these effects is $\frac{\Pi_{M\phi}}{\Pi_{V\phi}}$, as the following proposition makes clear:

Proposition 10 (Policy with Substitutes) *Suppose that $\eta > 0$ and that Assumption 5 holds. Consider any policy which reduces either Π_M or Π_V and does not increase the other:*

1. *If $\frac{\Pi_{M\phi}}{\Pi_{V\phi}} < \left|\widetilde{V}_M\right|$ then the policy reduces violence, but increases size.*
2. *If $\frac{\Pi_{M\phi}}{\Pi_{V\phi}} \in \left[\left|\widetilde{V}_M\right|, \frac{1}{\left|\widetilde{M}_V\right|}\right]$ then the policy reduces both size and violence.*
3. *If $\frac{\Pi_{M\phi}}{\Pi_{V\phi}} > \frac{1}{\left|\widetilde{M}_V\right|}$ then the policy reduces size, but increases violence.*

Each case is outlined in Figure 3.6. The fact that $\eta > 0$ ensures that all three scenarios are feasible, i.e. that $\left|\widetilde{V}_M\right| < \frac{1}{\left|\widetilde{M}_V\right|}$. The immediate effect of the policy is to (weakly) reduce the restricted demand for each input. Which case occurs depends, crucially, upon the size of each shift. In case one, the restricted demand for violence shifts down by more than the restricted demand for size. $\widetilde{V}$ is defined by $\Pi_V = 0$. Fixing size, we have that the vertical shift, $\widetilde{V}_\phi$, is given by:

$$\begin{aligned} \Pi_{VV}\widetilde{V}_\phi + \Pi_{V\phi} &= 0 \\ \iff \widetilde{V}_\phi &= -\frac{\Pi_{V\phi}}{\Pi_{VV}} \end{aligned} \tag{3.14}$$

Similarly, $\widetilde{M}$ is defined by $\Pi_M = 0$. Fixing size again, and considering the vertical shift, we have that:

$$\begin{aligned} \Pi_{VM}V_\phi + \Pi_{M\phi} &= 0 \\ \iff V_\phi &= -\frac{\Pi_{M\phi}}{\Pi_{MV}} \end{aligned} \tag{3.15}$$

So if the restricted demand for violence shifts downwards by more than the restricted demand for size:

$$\begin{aligned} \frac{\Pi_{V\phi}}{\Pi_{VV}} &> \frac{\Pi_{M\phi}}{\Pi_{MV}} \\ \iff \left|\widetilde{V}_M\right| &> \frac{\Pi_{M\phi}}{\Pi_{V\phi}} \end{aligned} \tag{3.16}$$

Intuitively, the marginal profit accruing to violence declines by much more than the marginal profit accruing to size. As such, the restricted demand for violence decreases dramatically. This fall in violence decreases both the marginal revenue and marginal cost of size. However, since the two inputs are substitutes, marginal cost reduces more, offsetting the initial fall in marginal profit. In case one, the fall in violence is so large that the marginal profit accruing to size actually becomes positive, creating an incentive for the gang to become larger.

Identical arguments can be made for the remaining two cases. In case two, the horizontal shift of the restricted demand for violence must exceed the horizontal shift in the restricted demand for size, and vice versa for the vertical shift. Neither input suffers a particularly large fall in marginal profit, and the substitution effects are not sufficient to counteract the initial declines in demand. In case three, the horizontal shift of the restricted demand for size must exceed the equivalent shift for violence. This time, the marginal profit accruing to size falls dramatically, creating a strong incentive for the gang to substitute away from size, towards violence.

The cross-elasticity, η, plays a large role in determining which case arises, as it is the key parameter in describing how strongly the gang chooses to substitute between violence and size. In particular, when η is large, $\left|\widetilde{M_V}\right|$ and $\left|\widetilde{V_M}\right|$ both decrease, as the two inputs are strong revenue complements, leading to large declines in their respective marginal revenue products. As such, case two becomes more prominent. The size of the internal $\left[\left|\widetilde{V_M}\right|, \frac{1}{\left|\widetilde{M_V}\right|}\right]$ grows. Conversely, if η is small, there is a high degree of substitutability between size and violence, and it becomes increasingly likely that one of the two extreme cases occurs.

3.4.2 Results for Specific Policies

The direct effect of each of class of policy is to reduce the marginal profit of either violence, size or both inputs. As such, under Assumption 4, Proposition 9 holds, and both violence and size are unequivocally diminished. The loss society suffers as a result of the gang will always decline. For each policy, we will therefore focus on what happens under Assumption 5. Under this assumption, it is possible that a fall in demand for one input will cause an increase in demand for the other. The social loss the gang inflicts may therefore increase.

Severity of Punishment (f)

When violence and size are substitutes, the effects of an increase in the severity of punishment are unambiguous. The situation is illustrated in Figure 3.7. When severity increases, the marginal cost of size increases (from (3.13)). Recruits

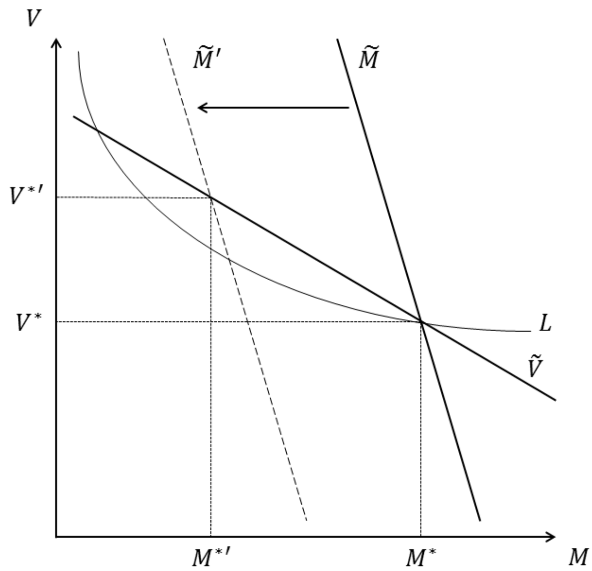

Figure 3.7: Effect of an increase in the severity of punishment (f) or the market wage (w)

require a greater degree of compensation for the possibility of being punished, increasing the wage the gang offers for any given gang size. As a result, the gang's restricted demand for members falls ($\Pi_{Mf} < 0$). However, if they still decide to join the gang, greater severity of punishment has no impact upon youths' willingness to acquire criminal skills (given by (3.5)). In particular, for each given gang size, the criminal skills acquired by the marginal youth remain unchanged. Consequently, there is no exogenous change in the gang's marginal cost of violence. By (3.12), the gang's restricted demand for violence remains unchanged ($\Pi_{Vf} = 0$). We are firmly in case three.

As the gang reduces its size, the marginal cost of violence also declines. Not only are there fewer members to compensate for changes in violence, but those that are left are less sensitive to those changes. The marginal revenue product of violence is also reduced, as membership size and violence are revenue complements. Under Assumption 5, the fall in marginal cost exceeds the fall in marginal revenue, and the gang chooses to increase the amount of violence it inflicts. Increasing the severity of punishment causes the gang to unambiguously substitute away from membership size, towards a greater degree of violence:

Lemma 5 *Suppose that $\eta > 0$ and that Assumption 5 holds. Then any increase in the severity of punishment will result in fewer gang members, but more violence.*

If society suffers sufficiently from increases in the intensity of violence that the gang chooses to inflict (the marginal rate of substitution of the iso-loss curve, L, is sufficiently low), then this policy could lead to an increase in the social loss the gang creates.

Primary Labour Market Policies (w)

The effect of an increase in the market wage is similar to that of an increase in the severity of punishment. The scenario is, once again, shown in Figure 3.7. The opportunity cost of joining the gang is increased. The gang must offer higher wages at every gang size, increasing the marginal cost of size. Consequently, by (3.13), the marginal cost of size exceeds the marginal revenue it generates, and the restricted demand for membership size declines ($\Pi_{Mw} < 0$). Upon deciding to join the gang, youths' incentives to acquire criminal skill are unaffected by the increase in w. As such, for each given size, the marginal cost to the gang of increasing violence is unchanged ($\Pi_{Vw} = 0$). As a result, by (3.12), the restricted demand for violence once again remains the same.

Gang size falls. Whilst the marginal revenue product of violence declines, smaller gangs require less compensation for violent behaviour. There are fewer members to pay, and those who remain are relatively insensitive to violence. The marginal cost of violence declines by more than its marginal revenue, and the gang increases the amount of violence it inflicts:

Lemma 6 *Suppose that $\eta > 0$ and that Assumption 5 holds. Then any improvement in the primary labour market will result in fewer gang members, but more* *violence.*

Again, if society is particularly sensitive to changes in the intensity of violence, relative to changes in size, labour market policies could actually increase the social cost of the gang.

Arrest and Conviction Rate (p)

The arrest and conviction rate has the most complex effect of any of the policies areas I consider. The situation is shown in Figure 3.8. An increase in the probability of conviction reduces the restricted demand for membership size for three reasons. For a given level of violence, an increase in p increases the wage the gang must offer to maintain its membership. Not only does it become more likely that gang members will be punished (increasing pf), but the probability that they will be deprived of their wages also rises. Youths discount for this in (3.8), and consequently require even more pay in order to join. These two outcomes constitute Levitt's *deterrence effect* (1996). Moreover, maintaining the same size of gang involves recruiting lower ability members, as more members are locked away (and are thus unproductive). Since lower ability individuals are more sensitive to violence, a third increase in the wage the gang offers is required. This is Levitt's incapacitation effect. Combined, these three wage increases raise the marginal cost of size in (3.13). As marginal revenue is thus far unaffected, the restricted demand for size declines ($\Pi_{Mp} < 0$).

In contrast to the previous two policies, increasing the probability of conviction also reduces the restricted demand for violence. For given gang size, the ability of the marginal youth is lower. Lower ability youths find it more costly to invest in criminal skills. The marginal youth is thus relatively sensitive to changes in violence (c_M^* falls in (3.9)). As a result, any increase in violence requires a greater increase in the wage the gang offers, in order to retain their

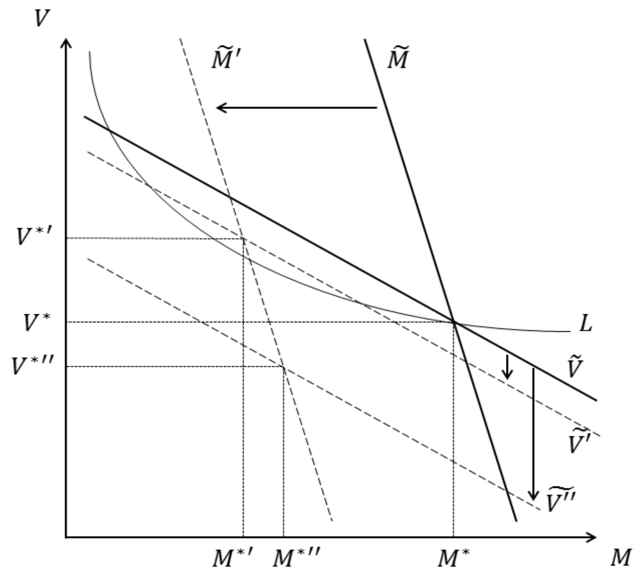

Figure 3.8: Effect of an increase in the arrest and conviction rate (p) or in the prevention of juvenile crime (k).

membership. In addition, youths incur the cost of violence irrespective of whether they are arrested or not. Since there is a greater chance that they will not receive their wages, they require proportionally more compensation should the amount of violence they inflict increase. Both these effects increase the marginal cost of violence in (3.12), reducing the restricted demand for violence ($\Pi_{Vp} < 0$).

In sum, both the restricted demand curves shift inward. Consequently, depending upon the size of the shifts, any of the cases outlined in Proposition 10 appear feasible. This turns out not to be the case, as we have the following result:

Lemma 7 *Suppose that $\eta > 0$ and that Assumption 5 holds. Then any improvement in the arrest and conviction rate may result in:*

1. *Fewer gang members, but more violence; or*

2. *fewer gang members and less violence.*

It can never be the case that more gang members, but less violence results.

The result hinges upon η. If η is large, then membership size and violence are relatively weak profit substitutes. As such, when the restricted demand for violence and size decline, the incentive to substitute is insufficient to cause the gang to increase their demand for either input. Conversely, if η is low, there is a strong incentive to substitute and the gang could potentially increase demand for either input. In particular, there is a threshold value of η, call it $\eta_V^p > -1$, such that if $\eta < \eta_V^p$ then the gang has a strong enough incentive to increase demand for violence. Similarly, there exists $\eta_M^p > -1$ such that if $\eta < \eta_M^p$ then the gang increases membership size. It is possible to show that $\eta_M^p < 0$. As we require that $\eta > 0$, we can rule out the gang's increasing its membership as a result of increased arrests and convictions. The increase in the marginal cost of membership size always dominates the increase in marginal revenue, even taking into account changes in violence.

The same cannot be said for η_V^p. It is therefore possible that the level of violence the gang inflicts actually increases. However, if the gang size is either very large or very small, violence will certainly decline. If it is small, then the ability of the marginal youth is high. When the arrest and conviction rate is increased, the marginal cost of size remains relatively unchanged. Wages are very close to $\frac{w+pf}{1-p}$, as the marginal youth does not require a lot of compensation for the violence they inflict. As a result, since the marginal revenue product of size is large for a small gang, gang size also remains relatively stable. This provides a relatively weak incentive to substitute. The increase in the marginal revenue product of violence is insufficient to counteract the exogenous fall in the restricted demand for violence.

When gang size is very large, the ability of the marginal youth is low. They therefore require a large amount of compensation for the violence the gang requires them to inflict. Any fall in the restricted demand for size will thus decrease the marginal cost of size substantially. Once again, the result is that the gang's size does not decline very much. The incentive to substitute is dominated by the decline in the restricted demand for violence.

Relative to the previous two policies, increasing the arrest and conviction rate is relatively successful at reducing the loss society suffers. Whilst it may be the case that violence increases, this only happens when size and violence are extremely strong profit substitutes ($(M^{*\prime}, V^{*\prime})$ in Figure 3.8, in this case lying above L). Otherwise, even under Assumption 5, both size and violence decline, leading to an unambiguous fall in the social loss ($(M^{*\prime\prime}, V^{*\prime\prime})$ in Figure 3.8).

Prevention of Juvenile Crime (k)

The effects of an increase in k are shown in Figure 3.8. As with the conviction rate, increasing the effort to prevent juvenile crime impacts upon both the restricted demand for size, and the restricted demand for violence. For each degree of violence, youths find it more difficult to acquire criminal skills (from (3.5)). They suffer more from the violence the gang requires them to inflict. Youths therefore require a larger wage to retain their membership. Moreover, the marginal youth is disproportionately affected by the policy. As the youth with the lowest ability, they are more sensitive to changes to the cost of acquiring criminal skills (see (3.9)). Any attempt by the gang to increase its membership therefore require a larger increase in the wage than before the introduction of the policy. Both the higher wage and the larger wage increase required to recruit more members increase the marginal cost of size for the gang (in (3.13)), reducing its restricted demand ($\Pi_{Mk} < 0$).

Since all youths acquire fewer criminal skills, the marginal cost of violence increases as well. In particular, for each gang size, increasing violence whilst retaining the marginal youth is more expensive. With fewer criminal skills, they are more sensitive to violence, and consequently require a greater increase in their wage to compensate them for the additional violence the gang wishes them to inflict. As the marginal revenue product of violence is unaffected by the policy, the increased marginal cost induces the gang to reduce its restricted demand for violence as well ($\Pi_{Vk} < 0$).

Once more, it appears that all three cases in Proposition 10 are feasible. Both restricted demand curves have shifted inwards. Again, this turns out to be incorrect:

Lemma 8 *Suppose that $\eta > 0$ and that Assumption 5 holds. Then any improvement in the prevention of juvenile crime may result in:*

1. *Fewer gang members, but more violence; or*

2. *fewer gang members and less violence.*

It can never be the case that more gang members, but less violence results.

The value of η is again critical in determining which outcome occurs. As with changes in the arrest and conviction rate, if η is large, then membership size and violence are weak profit substitutes. The endogenous increase in profitability of each input resulting from the decline in the other is never sufficient to offset the exogenous decline resulting from the policy change. However, there exists $\eta_V^k > 0$ such that, if $\eta < \eta_V^k$ then the endogenous increase in the profitability of violence resulting from the decline in size dominates the initial decline. In this case, the gang will increase its demand for violence. Similarly, there is $\eta_M^k > -1$ such that, if $\eta < \eta_M^k$ then demand for membership size increases.

As before, it is possible to rule out increases in membership size, since $\eta_M^k < 0$. The increase in the marginal revenue product of size resulting from an exogenous decline in violence is never large enough to offset the increase in the marginal cost caused by the increase in juvenile crime prevention. We can also rule out increases in violence for very large or very small gangs. In both cases, the improvement in the prevention of juvenile crime will result in relatively small changes in gang size, creating a very weak incentive to substitute towards violence.

Prevention of juvenile crime also proves to be relatively effective at reducing the loss society suffers. Once again, size and violence need to be extremely weak revenue complements to cause an increase in gang violence. It is quite possible that both will decline as a result of the policy, reducing the social cost of the gang.

3.5 The Separating Gang

The analysis performed in the previous few sections was done under the assumption that the gang was unable to discriminate between individuals. It offered one wage rate, and one level of violence. I now relax that assumption, and instead ask how the various policies perform when the gang is free to implement any wage scheme and associated violence schedule. I consider symmetric separating perfect Bayesian equilibrium. In this setting, such an equilibrium consists of a profit-maximising contract schedule, $(g(s), V(s))_{s\geq 0}$, implemented by the gang, followed by career, criminal skill and contract decisions for each youth. I focus on direct mechanisms. As before, I proceed by backwards induction.

3.5.1 Youth Decisions

In the model with a separating gang, a youth with criminal ability σ_i faces the following utility maximisation problem:

$$\max_{j\in\{0,1\},c\geq 0,s\geq 0}\left\{(1-j)\,w+j\left[(1-p)\,g(s)-pf-\frac{V(s)}{c}\right]-kJ\left(\frac{c}{\sigma_i}\right)\right\} \tag{3.17}$$

The youth's choice of criminal skill and career are identical to before. If they join the gang, they will acquire criminal skills $c^*(\sigma_i, s_i^*)$ satisfying:

$$\frac{V(s_i^*; \boldsymbol{\phi})}{c^*(\sigma_i, s_i^*)^2} \equiv \frac{k}{\sigma_i} J'\left(\frac{c^*(\sigma_i, s_i^*)}{\sigma_i}\right) \tag{3.18}$$

otherwise, they will not invest anything. Moreover, they will join the gang if:

$$(1-p)\, g(s_i^*) - pf - \frac{V(s_i^*)}{c^*(\sigma_i, s_i^*)} - kJ\left(\frac{c^*(\sigma_i, s_i^*)}{\sigma_i}\right) \geq w \tag{3.19}$$

If a youth of ability σ_i joins the gang, it is straightforward to show that all youths with ability $\sigma > \sigma_i$ also join the gang. Suppose that the youth with ability σ_i chooses contract s_i^*. If they join, it must be the case that $G^*(s_i^*; \sigma_i) \geqslant w$. Now, any youth choosing contract s_i^* is offered the same $(g(s_i^*), V(s_i^*))$. Moreover, as discussed in Section 3.3.1, for given g and V, a youth's payoff is increasing in σ. So any youth with ability greater than σ_i can guarantee themselves a payoff greater than w by joining the gang and choosing contract s_i^*. The payoff from joining the gang must once again be strictly increasing in ability. Consequently, there exists a marginal youth, who has the lowest ability of any gang member, $\sigma_M \geq 0$.

Finally, each youth who decides to join the gang chooses a contract to maximise the payoff they receive from their membership:

$$(1-p)\frac{\partial g}{\partial s}(s^*(\sigma_i)) \equiv \frac{1}{c^*(\sigma_i, s_i^*)}\frac{\partial V}{\partial s}(s^*(\sigma_i)) \tag{3.20}$$

A youth will therefore truthfully reveal their ability if and only if σ_i satisfies the above equation, i.e. $s^*(\sigma_i) = \sigma_i$.

3.5.2 Gang Leader Decisions

The gang leaders' decisions are significantly more complicated. They must now choose a profit-maximising contract schedule, subject to its being implementable. The form of an implementable contract schedule is given by the following:

Proposition 11 (Implementable Contracts) *A contract schedule,* $\{(g(s), V(s))\}_{s\geq 0}$, *is implementable if and only if it is of the form:*

$$g(s) = \frac{w + pf}{1-p} + \frac{V(s)}{c^*(1-p)} + \frac{k}{1-p} J\left(\frac{c^*}{s}\right) + \frac{1}{1-p}\int_{t=\sigma_M}^{s} \frac{V(t)}{tc^*} dt \qquad (3.21)$$

Proof. See Appendix 3.B. ■

This relationship between wages and violence bears a tremendous similarity to that of the marginal youth in a simple gang, given by (3.8). The first three terms simply state that each youth must be compensated for the expected costs incurred by joining the gang. The difference arises in the final term of (3.21), and is most succinctly explained by considering a youth's net payoff from joining the gang:

$$w + \int_{t=\sigma_M}^{\sigma} \frac{V(t)}{tc^*} dt \qquad (3.22)$$

In a perfect information setting, the gang would choose the level of violence they wished each youth to engage in, and then simply pay them enough to make them indifferent between joining the gang or joining the primary labour market. Youths would earn w irrespective of their career choice. However, the gang must elicit each youth's ability. Since a higher ability youth can gain positive surplus by pretending to be of a lower ability, the gang must pay an informational rent to ensure that they are made at least as well off by revealing their true ability.

An example of an implementable contract schedule is shown in Figure 3.9, displaying indifference curves for two youths. Higher ability youths are less sensitive to violence. As such, less compensation is required when violence is increased in

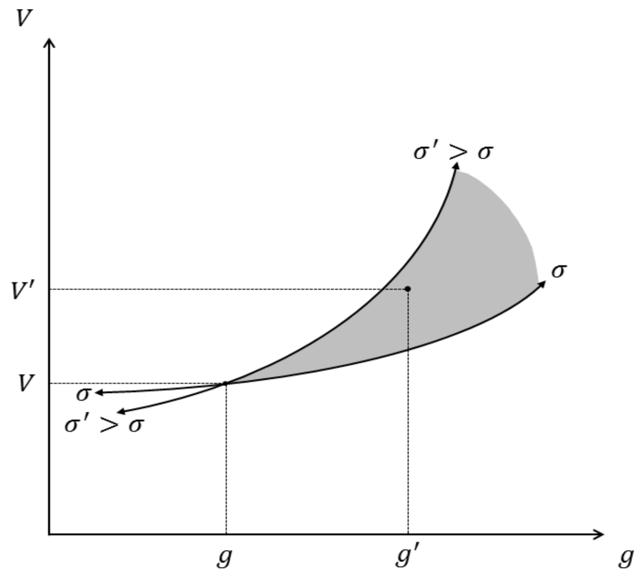

Figure 3.9: An example of an implementable contract

order to maintain indifference: their indifference curves are steeper. The contract schedule is designed so that each youth weakly prefers the contract designed for their ability to all others. For example, the youth with ability $\sigma' > \sigma$ prefers (g', V') to (g, V).

An important feature of an implementable contract is made clear by Figure 3.9: wages and violence must both be increasing in ability. To see this, note that for a youth with ability σ' to prefer a bundle (g', V') to (g, V) it must lie to the south-east of the σ'-indifference curve passing through (g, V). Similarly, for a youth with ability σ to prefer (g, V) over (g', V'), (g', V') must lie to the north-west of the σ-indifference curve passing through (g, V). Only contracts in the shaded region satisfy both properties, so g and V must both be increasing in ability.

Restricting attention to implementable contracts, we now turn our attention to the gang's profit maximisation decision. Since the gang's choice of violence uniquely determines the wage it must pay to its members, it is sufficient once again to think of the gang as maximising profits with respect to $\{V(s)\}_{s\geq 0}$ and

M. Its optimal choice of $g(s)$ will then be given by (3.21). In other words, the gang solves:

$$\begin{aligned}\left(\left(V^*(s)\right)_{s\geq 0}, M^*\right) &= \underset{(V(s))_{s\geq 0}, M\geq 0}{\arg\max} \left\{N(1-p)\int_{s=\sigma_M}^{\infty} \pi(s,M)\,\lambda e^{-\lambda s} ds\right\} \quad (3.23)\\ \text{subject to} &: (3.21)\end{aligned}$$

The solution is described in two stages. Firstly, for each ability and each gang size, I describe the optimal choice of violence. This provides a restricted violence schedule, dependent upon the gang's membership size $\left(\widetilde{V}(s,M)\right)_{s\geq 0}$. Then, incorporating this restricted violence schedule into (3.23), the gang chooses membership size to maximise its profits. Without further ado:

Proposition 12 (Restricted Violence Schedule) *Suppose that, for each σ, $\eta(\sigma) > -\frac{1}{2}$ and $1 + \frac{\frac{c^*}{\sigma} C'''\left(\frac{c^*}{\sigma}\right)}{C''\left(\frac{c^*}{\sigma}\right)} > \varepsilon\left(\frac{c^*}{\sigma}\right)$. Then, for each gang size, M, there exists a unique violence schedule, $\left(\widetilde{V}(s,M)\right)_{s\geq 0}$ that maximises profits.*

Proof. See Appendix 3.C. ■

For each gang size and each $\sigma \geq \sigma_M$, the gang selects $\widetilde{V}(\sigma)$ to satisfy:

$$r_V\left(M, \widetilde{V}(\sigma)\right) - \frac{1}{c^*(1-p)} - \frac{1}{c^*(1-p)}\frac{1+\varepsilon}{\lambda\sigma(2+\varepsilon)} \equiv 0 \qquad (3.24)$$

where $c^* = c^*(\sigma,\sigma)$ and $\varepsilon = \frac{\frac{c^*}{\sigma} J''\left(\frac{c^*}{\sigma}\right)}{J'\left(\frac{c^*}{\sigma}\right)}$, as before. This expression is very similar to (3.12). The marginal benefit to the gang of increasing a member's violence comes in the form of additional revenue they will be able to generate. The marginal cost comprises two elements. Firstly, it is necessary to compensate the individual for the disutility they suffer from inflicting more violence. The second element of the marginal cost represents the need to increase the informational rent paid to members. Increasing violence for a youth with ability σ increases the informational rent to all members with ability greater than σ. The rent paid to those with lower ability is unaffected (see (3.21)). As increasing size involves attracting lower ability members to the gang, the increase in informational rent is unaffected by the size of the gang.

Substituting the restricted violence schedule into (3.23), we are now in a position to calculate the gang's optimal size.

Proposition 13 (Profit Maximisation with Separation) *Suppose that Proposition 12 is satisfied, and that* $r_{MM} > 0$, $r_{MMV} < 0$, $r_{MV} r_{MVV} < 0$ *and* $\frac{\partial^3 \pi}{\partial V^2 \partial \sigma} < 0$. *Then there exists a unique gang size,* $0 < M^* < N$, *that maximises profits.*

Proof. See Appendix 3.D. ■

Given a violence schedule, the gang choose a membership size to satisfy:

$$r\left(\widetilde{M}, V\left(\sigma_{\widetilde{M}}\right)\right) - g\left(\sigma_{\widetilde{M}}\right) + N(1-p)\int_{s=\sigma_{\widetilde{M}}}^{\infty} r_M\left(\widetilde{M}, V(s)\right)\lambda e^{-\lambda s} ds - \frac{V\left(\sigma_{\widetilde{M}}\right)}{\lambda \sigma_M c_M^* (1-p)} \equiv 0 \tag{3.25}$$

where $c_M^* = c^*(\sigma_M, \sigma_M)$.When the gang increases its size, its new members generate revenue equal to $r\left(\widetilde{M}, V\left(\sigma_{\widetilde{M}}\right)\right)$. The marginal cost has several components. First, each new member must be paid. Their wage is given by (3.21), but does not need to incorporate any informational rent since a youth would never choose to over perform during initiation. Secondly, since the aggregate revenue function has diminishing marginal returns, it must be the case that an increase in size reduces the revenue each inframarginal member is able to generate. Finally, the gang must increase the wages paid to all youths who were already planning to join the gang. Otherwise, there would exist a higher ability youth who would choose to acquire the low wage-low violence contract, $s = \sigma_M$.

3.6 Policy with a Separating Gang

In section 3.4, the extent to which violence and membership size were revenue complements was critical in determining the effects of policy. When a policy reduced demand for, say, size, the marginal revenue product and marginal cost of violence fell. With fewer members, the gang was less able to convert violence into higher revenue. However, the gang needed to compensate fewer members for the violence they were required to inflict. When size and violence were strong revenue complements, the fall in the marginal revenue product of violence exceeded the fall in marginal cost. The demand for violence fell. Conversely, if they were weak revenue complements, the opposite was true.

This intuition proves to hold when the gang is capable of separating out recruits. When gang size increases, there are two opposing effects on the marginal revenue product of violence. Firstly, as there are more members inflicting violence, aggregate revenue from violence increases. The increased membership will also impact upon the personal marginal revenue products of violence of those already in the gang. Gang members may be able to take advantage of network externalities and returns to scale, increasing their personal marginal revenue products of violence. Conversely, congestion effects may reduce the personal marginal revenue product, as more individuals are attempting to extract rents from the neighbourhood. So long as the first effect dominates the second at the gang level, we are consistent with the model of the previous section, as the marginal revenue product of violence will rise on aggregate. However, the degree of revenue complementarity will be strongly affected by whether the personal marginal revenue product of violence increases or decreases. This leads us to make one of two assumptions:

Assumption 6 (Separating Complements) *An individual gang member's mar-*

ginal revenue product of violence is strictly increasing in the gang's size:

$$r_{MV}(M, V) > 0$$

Assumption 7 (Separating Substitutes) *An individual gang member's marginal revenue product of violence is strictly decreasing in the gang's size:*

$$r_{MV}(M, V) < 0$$

Clearly, if Assumption 6 holds, then size and violence are very strong revenue complements. When the gang's membership increases, additional members increase the amount of violence the gang is able to bring to bear, and enable existing members to take advantage of network externalities. This makes violence more profitable for each member. If Assumption 7 holds, on the other hand, size and violence are weak revenue complements. Whilst the additional members enable the gang to inflict more violence, the neighbourhood becomes satiated and existing gang members see their personal returns to violence fall.

3.6.1 Results for a Generic Policy

The impact of a generic policy is similar to that discussed in section 3.4. Each policy still acts to increase the gang's marginal cost of membership size and (in certain cases) violence. For a given violence schedule, policies reduce the surplus youths receive from joining the gang. As such, any increase in membership necessitates paying higher wages than before. This is compounded by the need to maintain truthful revelation. If the gang offers higher wages to new (low ability) members, and does not ask them to inflict more violence, then there will exist a higher ability youth who will strictly prefer to accept the contract offered to the new members. The gang must therefore increase the wages of all its members (informational rent payments may increase). Given this increase in marginal cost, the gang will optimally choose to reduce its size.

Policy may also affect youths' response to changes in violence. In contrast to the previous model, however, each youth inflicts different levels of violence, so the gang adjusts its violence schedule on a youth-by-youth basis. Assuming that a policy affects every youth in a similar way, each youth will require an increase in their wages to compensate them for increases in the level of violence they are required to inflict. For given gang size, the marginal revenue product of violence is unaffected. So, once again, the gang will optimally reduce the amounts of violence it requires its members to inflict.

The overall effect also depends upon the strength of revenue complementarity between membership size and violence. If they are strong complements, we have the following result, equivalent to Proposition 9 in section 3.4:

Proposition 14 (Policy with Separation and Complements) *Suppose the conditions given in Proposition 13 and Assumption 6 hold. Then any policy which reduces either π_V or Π_M and does not increase the other reduces both the amount of violence each member of the gang inflicts and the number of members the gang chooses to recruit.*

Consider a policy that reduces the gang's size. Under Assumption 6, gang members are no longer able to take advantage of economies that were previously available to them. As a result, the marginal revenue product of violence declines for each gang member. The marginal cost of violence is unaffected by changes in size. Consequently, $\pi_V < 0$ for every gang member, and the gang chooses to reduce the amount of violence it requires its members to inflict.

Now consider a policy that reduces the amount of violence the gang's members inflict for each M. By the Envelope Theorem, the only effect on the marginal profitability of size manifests itself in a decline of the personal marginal revenue products of size of the inframarginal recruits. Consequently, $\Pi_M < 0$ and the gang chooses to reduce its size.

These two effects reinforce declines in both gang size and the amount of violence each member inflicts, giving rise to the result in Proposition 14. If Assumption 6 is satisfied, just as before, any policy will be effective at reducing the loss society suffers at the hands of the gang.

If increases in membership cause congestion, violence and membership are relatively weak revenue complements. As with the simple gang environment, this makes the policy effects much more difficult to predict. A policy that reduces the gang's size increases the marginal revenue product of violence for each gang member. With fewer members, each individual is able to extract greater rents from the neighbourhood by employing violence. It could be the case that this increase in marginal revenue exceeds any increase in marginal cost resulting from the direct impact of the policy. If so, the gang will increase the amount of violence each member inflicts.

Similarly, any policy which reduces violence increases the marginal revenue product of size. With the neighbourhood less satiated with violence, there are greater opportunities for new members to generate revenue. Once again, if this increase in marginal revenue exceeds the increase in marginal cost resulting from the policy, the gang will opt to increase its membership.

This intuition is formalised in the following proposition, equivalent to Proposition 10 in section 3.4:

Proposition 15 (Policy with Separation and Substitutes) *Suppose the conditions given in Proposition 13 and Assumption 7 hold. Consider any policy which reduces either π_V or Π_M and does not increase the other:*

1. *If $\Pi_{M\phi} < M \int_{s=\sigma_M}^{\infty} \pi_{V\phi} |V_M| \lambda e^{-\lambda(s-\sigma_M)} ds$ then the policy reduces the amount of violence each member of the gang inflicts, but increases size.*

2. *If* $\Pi_{M\phi} > M \int_{s=\sigma_M}^{\infty} \pi_{V\phi} |V_M| \lambda e^{-\lambda(s-\sigma_M)} ds$ *then the policy reduces the size of the gang, and:*

   (a) *Individual gang members for which* $\frac{\pi_{V\phi}}{r_{MV}} > |M_\phi|$ *inflict less violence.*

   (b) *Individual gang members for which* $\frac{\pi_{V\phi}}{r_{MV}} < |M_\phi|$ *inflict more violence.*

The conditions given in Proposition 15 are equivalent to those seen before. Size is defined by $\Pi_M = 0$. If, after a policy was enforced, size were to remain unchanged, we would have:

$$\Pi_{M\phi} + M \int_{s=\sigma_M}^{\infty} \pi_{V\phi} V_M \lambda e^{-\lambda(s-\sigma_M)} ds = 0 \tag{3.26}$$

The direct decline in the marginal profitability of size resulting from the policy would be exactly offset by the increase in marginal profitability resulting from a fall in violence. The condition in case one states that there is an overall increase in the marginal profitability of size, resulting in a net increase in members. In case two, the opposite is true. Note that, if every gang member inflicts the same amount of violence, (3.26) becomes $\Pi_{M\phi} - M\pi_{V\phi} |V_M| = 0 \iff |V_M| = \frac{\Pi_{M\phi}}{M\pi_{V\phi}} = \frac{\Pi_{M\phi}}{\Pi_{V\phi}}$, exactly as before.

As different youths engage in different levels of violence, the effect of policy on violence is significantly more complex than in Proposition 10. In particular, there is no such thing as $\widetilde{M_V}$. Instead, consider the effect on the marginal profitability of violence of policy for a given level of violence:

$$\pi_{V\phi} + r_{MV} M_\phi \tag{3.27}$$

When a policy is implemented, its direct effect may incorporate an immediate decline in the profitability of violence. However, if size also declines, the neighbourhood becomes less congested. This enables each youth to generate more revenue through violence. The marginal cost of each youth increasing violence is unaffected by size, as changes in violence only affect the informational rents of those youths with higher ability. If the decline in gang size is sufficiently large, the increase in the marginal revenue product of violence may dominate the fall in profitability caused by the policy. In this case, $\pi_{V\phi} + r_{MV}M_{\phi} > 0$ and the gang optimally increases the amount of violence the youth is required to inflict. The conditions given in Proposition 15 are thus precisely those that dictate whether the marginal profitability of violence increases or decline.

3.6.2 Results for Specific Policies

When the gang can discriminate between its members, and Assumption 6 holds, the effects of any of the policies I consider are unambiguous. Since all policies' direct effects include reducing the marginal profit generated by membership size, size declines. This causes a reduction in the marginal revenue product of violence for all members, in turn endogenously reducing violence. Proposition 14 holds. Under Assumption 7, the results are less clear, and are outlined below.

Severity of Punishment (f)

Increasing the severity of punishment only affects the restricted demand for size. It increases the expected cost of joining the gang. However, once a youth has decided to join, it leaves their incentive to acquire criminal skills (and hence their sensitivity to violence) unaltered. In order to maintain the size of the its membership, the gang must increase the wages it pays to all of its members for a given violence schedule (see (3.21)). If it were to attempt to attract new members, it would need to offer them a higher wage as well. The marginal cost of size has increased, as shown in (3.25). Since the violence schedule is as yet unaffected, the marginal revenue product of size remains constant, leading to a fall in the marginal profitability of size, $\Pi_M < 0$. The gang chooses to reduce its membership.

The decline in size increases the personal marginal revenue product of violence. There is less congestion in the neighbourhood than there otherwise would be, enabling each recruit to generate greater revenues through violence. As the marginal cost of violence is accrued by each individual, it is unaffected by the falling gang size. The marginal profitability of violence therefore increases, $\pi_V > 0$, and the gang increases the amount of violence each youth engages in:

Lemma 9 *Suppose the conditions given in Proposition 13 and Assumption 7 hold. Then any increase in the severity of punishment will result in fewer gang members, but each member will increase the amount of violence they inflict.*

As in section 3.4, if society is relatively sensitive to increases in gang violence, introducing more severe punishment across the board could lead to a increase in the damage the gang inflicts on society. Whilst membership does decline, reducing the loss, each remaining member may become increasingly violent. This can more than offsets these gains.

Primary Labour Market Policies (w)

The effect of improvement in the primary labour market is, once again, identical to an increase in the severity of punishment. As the opportunity cost of joining the gang increases, the gang must offer the marginal youth a higher wage, given the violence schedule. This increases the marginal cost of size in (3.25), whilst having no effect upon its marginal revenue product. The marginal profitability of size declines, and the gang optimally reduces its size.

Under Assumption 7 this reduces congestion in the neighbourhood, increasing the marginal revenue product of violence for each member above its marginal cost.

The gang therefore increases the amount of violence it requires each member to inflict:

Lemma 10 *Suppose the conditions given in Proposition 13 and Assumption 7 hold. Then any improvement in the primary labour market will result in fewer gang members, but each member will increase the amount of violence they inflict.*

Unsurprisingly, even under this more complicated wage setting, improvements in the primary labour market can also yield greater social losses from the gang.

Arrest and Conviction Rate (p)

When there is an increase in the arrest and conviction rate, the marginal cost of size increases. In a similar manner to the previous two policies, the expected cost of joining the gang has increased, as it is more likely that an individual will be punished (pf is higher in (3.21)). Furthermore, if youths are caught, their wages are confiscated. When deciding upon whether to join the gang, they discount their wages for this possibility, and consequently require higher wages to encourage them into a criminal career. With no equivalent increase in the marginal revenue product, the marginal profitability of size declines in (3.25), and the gang chooses to reduce its size.

The marginal cost of violence also increases for every youth. Gang members are only arrested after committing crime. However, it is during the commission of crime that they inflict violence. Irrespective of whether they are caught, they therefore incur the cost of violence. The compensation they receive for doing so, on the other hand, are conditional on their evading arrest. Consequently, when the arrest rate increases, and gang members discount their wage further, any increase in violence requires a more substantial increases in pay. With no change in the marginal revenue product of violence (holding size constant), this causes the gang to reduce the levels of violence it requires its members to inflict.

Combined, these two effects give rise to the following result, equivalent to Lemma 7:

Lemma 11 *Suppose the conditions given in Proposition 13 and Assumption 7 hold. Then any improvement in the arrest and conviction rate may result in:*

1. *More gang members and less violence;*

2. *Fewer gang members and less violence; or*

3. *Fewer gang members and more violence.*

Under Assumption 7, an exogenous decrease in violence for every member increases the marginal revenue product of size. Similarly, a reduction in size increases the marginal revenue product of violence for each member of the gang. It is therefore possible that the reduction in violence may cause a sufficient increase in the marginal revenue product of size to offset the increase in marginal cost. This would create an incentive for the gang to recruit more members. In this case, the marginal revenue product of violence would decline further, leading to an additional reduction in violence, reinforcing the growth in gang size.

Conversely, the fall in gang size may cause a sufficiently large increase in the marginal revenue product of violence as to offset the increase in marginal cost. As such, the gang would increase the amount of violence it required its member to inflict (compensating them accordingly). In turn, this would further reduce the marginal profitability of size, leading to a further decline in membership.

Prevention of Juvenile Crime (k)

Increases in the prevention of juvenile crime reduces the restricted demand for size. The cost of acquiring criminal skills increases. For a given violence schedule, every prospective member of the gang invest less, and consequently suffers a greater disutility from the violence they are forced to inflict. The marginal youth is particularly affected. As the youth with the lowest intrinsic ability, they are more sensitive to changes in the cost of acquiring criminal skills, and consequently reduce their skills dramatically. The cost of retaining their membership increases. They require higher wages. This leads to increases in the informational rent paid to all inframarginal members. The marginal cost of size increases. As, given the violence schedule, there is no equivalent increase in the marginal revenue product of size, the marginal profit associated with size declines ($\Pi_M < 0$). The gang optimally reduces the number of members it recruits.

Concurrently, the policy also reduces the restricted demand for violence. As each youth incurs a higher cost of acquiring criminal skills, they reduce their investment. In turn, this causes them suffer a greater disutility from the violence they are required to inflict. Moreover, as they have lower levels of criminal skill, they also become more sensitive to changes in violence. Each youth therefore requires a greater level of compensation for any changes in the level of violence they are required to inflict. The marginal cost of violence increases for every gang member. For a given gang size, there is no change in the marginal revenue product of violence. The marginal profitability of violence declines for all members ($\pi_V < 0$). The gang reduces every element of its violence schedule.

We have the following result:

Lemma 12 *Suppose the conditions given in Proposition 13 and Assumption 7 hold. Then any improvement in the prevention of juvenile crime may result in:*

1. *More gang members and less violence;*
2. *Fewer gang members and less violence; or*
3. *Fewer gang members and more violence.*

As both the restricted demand for size and the restricted violence schedule decline, the overall effect is uncertain. The direct fall in demand for size increases the marginal revenue product of violence. This may be sufficient to offset the increase in the marginal cost of violence caused by greater difficulties in acquiring criminal skills. As a result, the gang may choose to raise the amount of violence it requires of its members. This would lead to a further decline in the marginal revenue product of size, reducing size and reinforcing the increase in violence.

The opposite may also be true. The reduction in the restricted violence schedule increases the marginal revenue product of size. If this exceeds the increase in marginal cost caused by having to compensate youths more for acquiring criminal skills, the gang will increase its size. In turn, this would reduce the marginal revenue product of violence for each individual member, again reinforcing the increase in size.

Finally, it could be the case that neither endogenous increase in marginal revenue product outweighs the direct increase in marginal costs caused by improvements in the prevention of juvenile crime. Both size and violence would therefore decline.

3.7 Conclusions

Over recent years, numerous policies have been put forward to combat the social loss associated with crime. These policies aim to decrease individuals' incentive to engage in crime and, in doing so, reduce the amount of crime that occurs. However, when applied to neighbourhoods where organised crime is prevalent, this argument breaks down. When a policy is implemented, criminal organisations may adjust its inputs, substituting towards increasing the intensity of violence. This may increase the loss society suffers at the hands of organised crime.

This paper has shown the effects of several popular policies in such an environment. As criminal organisations tend to operate within a well-defined geographical territory, they act as a monopsonist employer for all criminals within that territory. Irrespective of whether the organisation operated a single wage or more complicated recruitment strategy, results were shown to be robust.

The effects of policy depend upon the degree of complementarity between inputs in the criminal organisation's revenue function. If they are strong complements, policies that one input reduce the marginal profitability of the other. Both size and violence decline. Conversely, if they are weak revenue complements, the organisation may choose to substitute between size and violence, possibly undoing the some of the effects of the policy. In this case the loss society suffers may increase.

When there is an incentive to substitute, policies which simply increase the opportunity cost of joining a criminal organisation, such as improved labour market wages or more severe punishment, fair badly. As they do not affect youths' incentive to acquire criminal skill, they actually reduce the marginal cost of violence. Those who chose to remain in the organisation after the policy is implemented are highly skilled. They do not require as much compensation when violence is intensified. As such, criminal organisations will always choose to increase violence, at the expense of membership.

Other policies prove more effective. Prevention of juvenile crime and improved arrest or conviction rates may cause an intensification of violence, but only in relatively extreme circumstances. Otherwise, these policies diminish both the organisation's size and violence. Preventing juvenile crime not only increases the opportunity cost of joining a criminal organisation, but also reduces the incentive to acquire criminal skill. By doing so, it increases not only the marginal cost of acquiring members, but also the marginal cost of violence. Improving arrest rates have a similar effect. As youths may be prevented from receiving their wages, they require more compensation *ex ante* for the violence they inflict. As such, the marginal cost of violence once again increases. If the degree of substitutability between size and violence is particularly large, then the criminal organisation may still choose to substitute away from size towards violence. Otherwise, it will reduce both its size and the violence it inflicts.

In summary, anti-crime policies are most effective against organised crime when they not only reduce the incentive of youths to join the organisation, but also hamper its ability to increase violence.

Acknowledgements

I owe thanks to Alwyn Young and Leonardo Felli for their invaluable help. Comments at various stages by Philipp Kircher, Ashwini Natraj, Francesco Nava, Michele Piccione, Andrea Prat, Ronny Razin and Jan Zapal have proved useful, as have those of participants at the LSE Microeconomic Theory Work In Progress seminars, the 7th Royal Economic Society Ph.D. Presentation Meeting, and seminars at Stockholm University and Cardiff Business School. Time spent at the Study Center in Gersenzee is also acknowledged. All errors remain my own.

3.A Proof of Proposition 8

Firstly, note that $M = N$ is never profit maximising, as it involves the gang paying infinitely large wages. Also, if $V^* = 0$ or $M^* = 0$, equilibrium profit for the gang, Π^*, is non-positive. So, to prove that the gang will operate with positive V and M, it will be necessary to show that positive profits will result.

Now, the first order conditions for profit maximisation are:

$$\begin{aligned}
\Pi_V &= Mr_V\left(M^*,V^*\right) - \frac{M^*}{c_M^*\left(1-p\right)} \equiv 0 \\
\Pi_M &= r\left(M^*,V^*\right) + Mr_M\left(M^*,V^*\right) - g\left(M^*,V^*\right) - \frac{V^*}{\lambda\sigma_M c_M^*\left(1-p\right)} \equiv 0
\end{aligned}$$

Given that the revenue function has constant returns to scale, $r_V + Mr_{MV} = -Vr_{VV} = -\frac{M}{V}\left(2r_M + Mr_{MM}\right)$. Substituting appropriately, this yields second-order conditions:

$$\begin{aligned}
\Pi_{MM} &= -\frac{V^*}{M^*}\frac{1}{M^*c_M^*\left(1-p\right)}\left[\eta + 1 + \frac{1}{\lambda\sigma_M}\left(1 + \frac{1}{\lambda\sigma_M} + \frac{1}{\lambda\sigma_M}\frac{1+\varepsilon_M}{2+\varepsilon_M}\right)\right] \\
\Pi_{MV} &= \frac{1}{M^*c_M^*\left(1-p\right)}\left[\eta - \frac{1}{\lambda\sigma_M}\frac{1+\varepsilon_M}{2+\varepsilon_M}\right] \\
\Pi_{VV} &= -\frac{M^*}{V^*}\frac{1}{M^*c_M^*\left(1-p\right)}\left[\eta + 1 - \frac{1}{2+\varepsilon_M}\right]
\end{aligned}$$

Necessary and sufficient conditions for a maximum are that $\Pi_{MM} < 0$, $\Pi_{VV} < 0$, and $\Pi_{VV}\Pi_{MM} - \Pi_{MV}^2 > 0$. The first of these conditions is satisfied unambiguously upon inspection, since $r_V + Mr_{MV} > 0$. The second condition is satisfied if and only if $\eta > \frac{1}{2+\varepsilon_M} - 1$. A sufficient condition is that $\eta > -\frac{1}{2}$. Thirdly, we require that $\Pi_{VV}\Pi_{MM} > \Pi_{MV}^2$. With relatively little work, it can be shown that this is satisfied if and only if:

$$\eta > \frac{1 + \frac{1}{\lambda\sigma_M}\left(1 + \frac{\varepsilon}{2+\varepsilon}\right) + \frac{1}{\lambda^2\sigma_M^2}}{2 - \frac{1}{2+\varepsilon} + \frac{1}{\lambda\sigma_M}\left(2 + \frac{\varepsilon}{2+\varepsilon}\right) + \frac{1}{\lambda^2\sigma_M^2}\left(1 + \frac{\varepsilon}{2+\varepsilon}\right)} - 1$$

A sufficient condition is that $\eta > 0$. So any point where both first order conditions are satisfied constitutes a local maximum. Rearranging these conditions yields:

$$\left|\frac{\Pi_{MV}}{\Pi_{VV}}\right| < \left|\frac{\Pi_{MM}}{\Pi_{MV}}\right|$$

Note that these inequalities do not simply hold at a point of profit maximisation - they hold everywhere. Therefore, assuming that the sign of Π_{MV} never changes, any profit maximising point will be unique.

Finally, it remains to show that the profit derived by the gang in any such equilibrium is positive. From the first-order conditions, we have that $V^* M^* r_V(M^*, V^*) = \frac{M^* V^*}{c_M^*(1-p)}$ and $M^*(r(M^*, V^*) + M^* r_M(M^*, V^*) - g(M^*, V^*)) = \frac{M^* V^*}{c_M^*(1-p)(\ln N + \ln(1-p) - \ln M^*)}$. Noting that the gang level revenue function is homogeneous of degree one, it is clear that:

$$\Pi^* = \frac{M^* V^* (1 + \ln N + \ln(1-p) - \ln M^*)}{c_M^*(1-p)(\ln N + \ln(1-p) - \ln M^*)} > 0$$

This completes the proof. ■

3.B Proof of Proposition 11

A contract is implementable if it is incentive compatible and individually rational. Considering first the issue of incentive compatibility, a youth has a strict incentive to truthfully reveal their type if:

$$\sigma = \arg\max_{s \geq 0} \left\{ (1-p) g(s) - pf - \frac{V(s)}{c^*} - kJ\left(\frac{c^*}{\sigma}\right) \right\}$$

where c^* is a function of both σ and V. Taking first-order conditions, this is equivalent to:

$$(1-p) \frac{\partial g}{\partial s}(\sigma) \equiv \frac{1}{c^*} \frac{\partial V}{\partial s}(\sigma)$$

Integrating both sides over the range $[\sigma_M, \sigma]$ yields:

$$g(\sigma) = g(\sigma_M) + \frac{V(\sigma)}{c^*(1-p)} - \frac{V(\sigma_M)}{c_M^*(1-p)} + \frac{1}{1-p} \int_{t=\sigma_M}^{\sigma} \frac{V(t)}{c^{*2}} \frac{\partial c^*}{\partial t} dt$$

Now, for σ_M to be the marginal youth, it must be the case that $G(\sigma_M, \sigma_M) = w$. Otherwise, if $G(\sigma_M, \sigma_M) > w$, a lower ability youth will be able to gain a larger payoff by joining the gang and sending signal $s_i = \sigma_M$, contradicting the fact that σ_M is the marginal youth. On the other hand, if $G(\sigma_M, \sigma_M) < w$, then the marginal youth would strictly prefer to join the primary labour market, again providing a contradiction. So:

$$g(\sigma_M) = \frac{w + pf}{1-p} + \frac{V(\sigma_M)}{c_M^*(1-p)} + \frac{k}{1-p} J\left(\frac{c_M^*}{\sigma_M}\right)$$

Also, we have that:

$$\begin{aligned} \frac{\partial}{\partial \sigma}\left(kJ\left(\frac{c^*}{\sigma}\right)\right) &= \frac{k}{\sigma} J'\left(\frac{c^*}{\sigma}\right)\frac{\partial c^*}{\partial \sigma} - \frac{kc^*}{\sigma^2} J'\left(\frac{c^*}{\sigma}\right) \\ &= \frac{V(\sigma)}{c^{*2}}\frac{\partial c^*}{\partial \sigma} - \frac{kc^*}{\sigma^2} J'\left(\frac{c^*}{\sigma}\right) \end{aligned}$$

So, by (3.18):

$$\begin{aligned} \int_{t=\sigma_M}^{\sigma} \frac{V(t)}{c^{*2}}\frac{\partial c^*}{\partial t} dt &= k\int_{t=\sigma_M}^{\sigma} \frac{\partial}{\partial t}\left(J\left(\frac{c^*}{t}\right)\right) dt + \int_{t=\sigma_M}^{\sigma} \frac{kc^*}{t^2} J'\left(\frac{c^*}{t}\right) dt \\ &= kJ\left(\frac{c^*}{\sigma}\right) - kJ\left(\frac{c_M^*}{\sigma_M}\right) + \int_{t=\sigma_M}^{\sigma} \frac{V(t)}{tc^*} dt \end{aligned}$$

Substituting, we have that:

$$g(\sigma) = \frac{w+pf}{1-p} + \frac{V(\sigma)}{c^*(1-p)} + \frac{k}{1-p} J\left(\frac{c^*}{\sigma}\right) + \frac{1}{1-p}\int_{t=\sigma_M}^{\sigma} \frac{V(t)}{tc^*} dt$$

Finally, we must show that this is individually rational. The implementable payoff from joining the gang is:

$$w + \int_{t=\sigma_M}^{\sigma} \frac{V(t)}{tc^*} dt$$

For any youth with $\sigma > \sigma_M$, the payoff from joining the gang strictly exceeds the wage they would earn in the primary labour market. For the marginal youth, the two are equal. For any youth with ability less than the marginal youth, they strictly prefer the primary labour market. This completes the proof. ■

3.C Proof of Proposition 12

Before evaluating the profit maximisation problem, consider the expected cost of informational rent for the gang:

$$I = \frac{1}{1-p}\int_{s=\sigma_M}^{\infty}\int_{t=\sigma_M}^{s}\frac{V(t)}{tc^*}dt\lambda e^{-\lambda(s-\sigma_M)}ds$$

Performing a standard integration by parts yields:

$$I = \frac{1}{1-p}\int_{s=\sigma_M}^{\infty}\frac{V(s)}{\lambda sc^*}\lambda e^{-\lambda(s-\sigma_M)}ds$$

so the gang leadership's objective function becomes:

$$N(1-p)\int_{s=\sigma_M}^{\infty}\left[r(M,V(s)) - \frac{w+pf}{1-p} - \frac{V(s)}{c^*(1-p)}\left(1+\frac{1}{\lambda s}\right) - \frac{k}{1-p}J\left(\frac{c^*}{s}\right)\right]\lambda e^{-\lambda s}ds$$

Now, given gang size, for each $\sigma \geq \sigma_M$, the restricted demand for V must maximise $\pi(M,V)$. It must therefore satisfy:

$$\pi_V = r_V\left(\widetilde{V},M\right) - \frac{1}{c^*(1-p)}\left[1+\frac{1}{\lambda\sigma}\frac{1+\varepsilon}{2+\varepsilon}\right] \equiv 0$$

The associated second-order condition is:

$$\begin{aligned}
\pi_{VV} &= r_{VV}\left(M,\widetilde{V}\right)+\frac{1}{\widetilde{V}c^{*}\left(1-p\right)\left(2+\varepsilon\right)}\left[1+\frac{1}{\lambda\sigma}\frac{1+\varepsilon}{2+\varepsilon}\right] \\
&\quad -\frac{\varepsilon}{\widetilde{V}c^{*}\left(1-p\right)\left(2+\varepsilon\right)^{3}}\frac{1}{\lambda\sigma}\left(1+\frac{\frac{c^{*}}{\sigma}J'''\left(\frac{c^{*}}{\sigma}\right)}{J''\left(\frac{c^{*}}{\sigma}\right)}-\varepsilon\right) \\
&= r_{VV}\left(M,\widetilde{V}\right)+\frac{r_{V}\left(\widetilde{V},M\right)}{\widetilde{V}\left(2+\varepsilon\right)}-\frac{\varepsilon}{\widetilde{V}c^{*}\left(1-p\right)\left(2+\varepsilon\right)^{3}}\frac{1}{\lambda\sigma}\left(1+\frac{\frac{c^{*}}{\sigma}J'''\left(\frac{c^{*}}{\sigma}\right)}{J''\left(\frac{c^{*}}{\sigma}\right)}-\varepsilon\right)
\end{aligned}$$

By assumption, the final term is positive. So the second-order condition is unambiguously negative if:

$$\begin{aligned}
r_{VV}\left(M,\widetilde{V}\right)+\frac{r_{V}\left(M,\widetilde{V}\right)}{\widetilde{V}\left(2+\varepsilon\right)} &< 0 \\
\iff \quad -\frac{\widetilde{V}r_{VV}\left(M,\widetilde{V}\right)}{r_{V}\left(M,\widetilde{V}\right)} &> \frac{1}{2+\varepsilon}
\end{aligned}$$

Defining $\eta\left(\sigma\right)\equiv\frac{Mr_{MV}}{r_{V}}=-1-\frac{Vr_{VV}}{r_{V}}$, as before, a sufficient condition is $\eta\left(\sigma\right)>-\frac{1}{2}$. This completes the proof. ■

3.D Proof of Proposition 13

Making liberal use of the envelope theorem, the gang's optimal size, M^{*}, satisfies the following first-order condition:

$$\begin{aligned}
\Pi_{M}=\hat{r}-\frac{w+pf}{1-p}-\frac{\hat{V}}{\hat{c}\left(1-p\right)}\left(1+\frac{1}{\lambda\sigma_{M}}\right)-\frac{k}{1-p}J\left(\frac{\hat{c}}{\sigma_{M^{*}}}\right) \\
+N\left(1-p\right)\int_{s=\sigma_{M^{*}}}^{\infty}r_{M}\left(M^{*},\widetilde{V}\left(s,M^{*}\right)\right)\lambda e^{-\lambda s}ds \equiv 0
\end{aligned}$$

where $\hat{r}\equiv r\left(M^{*},\widetilde{V}\left(\sigma_{M^{*}},M^{*}\right)\right)$ and $\hat{V}\equiv\widetilde{V}\left(\sigma_{M^{*}},M^{*}\right)$. The associated second-order condition is:

$$\begin{aligned}
\Pi_{MM} &= 2\hat{r}_M - \frac{\hat{V}}{\lambda\sigma_{M^*}M^*\hat{c}\,(1-p)}\left[1+\frac{1}{\lambda\sigma_{M^*}}\frac{3+2\varepsilon_{M^*}}{2+\varepsilon_{M^*}}\right] \\
&\quad + N\,(1-p)\int_{s=\sigma_{M^*}}^{\infty} r_{MV}\left(M^*,\widetilde{V}\,(s,M^*)\right)\frac{\partial\widetilde{V}}{\partial M}\lambda e^{-\lambda s}ds \\
&\quad + N\,(1-p)\int_{s=\sigma_{M^*}}^{\infty} r_{MM}\left(M^*,\widetilde{V}\,(s,M^*)\right)\lambda e^{-\lambda s}ds \\
&= 2\hat{r}_M - \frac{\hat{V}}{\lambda\sigma_{M^*}M^*\hat{c}\,(1-p)}\left[1+\frac{1}{\lambda\sigma_{M^*}}\frac{3+2\varepsilon_{M^*}}{2+\varepsilon_{M^*}}\right] \\
&\quad + N\,(1-p)\int_{s=\sigma_{M^*}}^{\infty}\left[r_{MM}-\frac{r_{MV}^2}{\pi_{VV}}\right]\lambda e^{-\lambda s}ds
\end{aligned}$$

Consider the final term:

$$\begin{aligned}
I &= N\,(1-p)\int_{s=\sigma_{M^*}}^{\infty}\left[r_{MM}-\frac{r_{MV}^2}{\pi_{VV}}\right]\lambda e^{-\lambda s}ds \\
&= M\hat{r}_{MM} - M\frac{\hat{r}_{MV}^2}{\hat{\pi}_{VV}} + N\,(1-p)\int_{s=\sigma_{M^*}}^{\infty} r_{MMV}\frac{\partial V}{\partial s}e^{-\lambda s}ds \\
&\quad -2N\,(1-p)\int_{s=\sigma_{M^*}}^{\infty}\frac{r_{MV}r_{MVV}}{\pi_{VV}}\frac{\partial V}{\partial s}e^{-\lambda s}ds + N\,(1-p)\int_{s=\sigma_{M^*}}^{\infty}\frac{r_{MV}^2}{\pi_{VV}^2}\frac{\partial^3\pi}{\partial V^2\partial s}e^{-\lambda s}ds
\end{aligned}$$

where $\frac{\hat{r}_{MV}^2}{\hat{\pi}_{VV}} = \frac{\partial\widetilde{V}}{\partial M}\,(M^*)$.By assumption:

$$\begin{aligned}
N\,(1-p)\int_{s=\sigma_{M^*}}^{\infty} r_{MMV}\frac{\partial V}{\partial s}e^{-\lambda s}ds - 2N\,(1-p)\int_{s=\sigma_{M^*}}^{\infty}\frac{r_{MV}r_{MVV}}{\pi_{VV}}\frac{\partial V}{\partial s}e^{-\lambda s}ds & \\
+N\,(1-p)\int_{s=\sigma_{M^*}}^{\infty}\frac{r_{MV}^2}{\pi_{VV}^2}\frac{\partial^3\pi}{\partial V^2\partial s}e^{-\lambda s}ds &< 0.
\end{aligned}$$

So:

$$\Pi_{MM} < 2\hat{r}_M - \frac{\hat{V}}{\lambda\sigma_{M^*}M^*\hat{c}\,(1-p)}\left[1+\frac{1}{\lambda\sigma_{M^*}}\frac{3+2\hat{\varepsilon}}{2+\hat{\varepsilon}}\right] + M^*\hat{r}_{MM} - M^*\frac{\hat{r}_{MV}^2}{\hat{\pi}_{VV}}$$

The right hand side is negative if and only if:

$$\begin{aligned}
2\hat{r}_M\hat{\pi}_{VV} - \frac{\hat{V}}{\lambda\sigma_{M^*}M^*}\hat{r}_V\hat{\pi}_{VV} - \frac{\hat{V}}{\lambda^2\sigma_{M^*}^2 M^*\hat{c}\,(1-p)}\hat{\pi}_{VV} + M^*\hat{r}_{MM}\hat{\pi}_{VV} - M^*\hat{r}_{MV}^2 &> 0 \\
\iff \hat{r}_M\hat{r}_{VV} + \hat{r}_M\gamma + \frac{1}{\lambda\sigma_{M^*}}\hat{r}_M\left(\hat{r}_{VV}+\gamma\right) - \frac{\hat{V}}{\lambda^2\sigma_{M^*}^2 M^*\hat{c}\,(1-p)}\left(\hat{r}_{VV}+\gamma\right) & \\
+M^*\hat{r}_{MM}\left(\gamma - \frac{\hat{r}_V}{\hat{V}}\right) - \frac{\hat{r}_M\hat{r}_V}{\hat{V}} &> 0
\end{aligned}$$

where $\gamma \equiv \frac{\hat{r}_V}{\hat{V}}\frac{1}{2+\varepsilon} - \frac{\varepsilon}{\widetilde{V}c^*(1-p)(2+\varepsilon)^3}\frac{1}{\lambda\sigma}\left(1 + \frac{\frac{c^*}{\sigma}J'''\left(\frac{c^*}{\sigma}\right)}{J''\left(\frac{c^*}{\sigma}\right)} - \varepsilon\right)$. Substituting for γ, and given the properties of $r(\cdot,\cdot)$, it is possible to show that this is unambiguously positive. Thus the second-order condition is invariably negative.

Now, in order to prove that this is profit-maximising, we need to show that equilibrium profits are positive. Otherwise, the gang will simply shut down. Note that, by the first-order condition, $\pi(\sigma_M, M) > 0$, since $r_M < 0$. Moreover, the profit each youth generates for the gang is increasing in ability for any given size:

$$\pi_\sigma(\sigma, M) = \frac{kc^*}{\sigma^2(1-p)}J'\left(\frac{c^*}{\sigma}\right) + \frac{\widetilde{V}_i}{c^*(1-p)}\frac{3+2\varepsilon_i}{\lambda\sigma(2+\varepsilon_i)} > 0$$

So the gang must generate positive profits in equilibrium. This completes the proof. ■

Bibliography

AKERLOF, G. A. (1997): "Social Distance and Social Decisions," *Econometrica*, 65(5), 1005–1028.

ALLAIS, M. (1953): "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine," *Econometrica*, 21(4), pp. 503–546.

AUSTEN-SMITH, D., AND R. G. FRYER, JR. (2005): "An Economic Analysis of "Acting White"," *The Quarterly Journal of Economics*, 120(2), pp. 551–583.

BACCARA, M., AND H. BAR-ISAAC (2008): "How to Organize Crime," *The Review of Economic Studies*, 75(4), pp. 1039–1067.

BALLESTER, C., A. CALVÓ-ARMENGOL, AND Y. ZENOU (2004): "Who's Who in Crime Networks: Wanted - The Key Player," CEPR Discussion Papers 4421, C.E.P.R. Discussion Papers.

——— (2010): "Delinquent Networks," *Journal of the European Economic Association*, 8(1), 34–61.

BAYER, P. J., R. PINTOFF, AND D. E. POZEN (2004): "Building Criminal Capital Behind Bars: Social Learning in Juvenile Corrections," Yale School of Management Working Papers ysm418, Yale School of Management.

Becker, G. S. (1968): "Crime and Punishment: An Economic Approach," *The Journal of Political Economy*, 76(2), pp. 169–217.

Bennett, T., and K. Holloway (2004): "Gang Membership, Drugs and Crime in the UK," *British Journal of Criminology*, 44(3), 305–323.

Bueno De Mesquita, E., and C. Hafer (2008): "Public Protection Or Private Extortion?," *Economics and Politics*, 20(1), 1–32.

Burdett, K., R. Lagos, and R. Wright (2004): "An On-the-Job Search Model of Crime, Inequality, and Unemployment," *International Economic Review*, 45(3), pp. 681–706.

Bureau of Economic Analysis (2011): "National Economic Accounts," Discussion paper, United States Department of Commerce.

Cabinet Office (2003): "Alcohol misuse: How much does it cost?," Discussion paper, Strategy Unit.

Chang, J.-J., H.-C. Lu, and M. Chen (2005): "Organized Crime or Individual Crime? Endogenous Size of a Criminal Organization and the Optimal Law Enforcement," *Economic Inquiry*, 43(3), 661–675.

Cohen, M., and A. Piquero (2009): "New Evidence on the Monetary Value of Saving a High Risk Youth," *Journal of Quantitative Criminology*, 25, 25–49.

Cohen, M. A. (1998): "The Monetary Value of Saving a High-Risk Youth," *Journal of Quantitative Criminology*, 14, 5–33.

Cook, P. J., J. Ludwig, S. Venkatesh, and A. A. Braga (2007): "Underground Gun Markets," *The Economic Journal*, 117(524), pp. F588–F618.

Decker, S. H. (1996): "Collective and Normative Features of Gang Violence," *Justice Quarterly*, 13(2), 243–264.

Department of Health (2008): "The cost of alcohol harm to the NHS in England: An update to the Cabinet Office (2003) study," Discussion Paper 10277, Health Improvement Analytical Team.

Dixit, A. (2007): *Lawlessness and Economics: Alternative Modes of Governance*, The Gorman lectures in economics. Princeton University Press.

Dobbie, W., and R. G. Fryer, Jr. (2011): "Are High-Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children's Zone," *American Economic Journal: Applied Economics*, 3(3), 158–87.

Egley, Jr., A., and J. C. Howell (2011): "Highlights of the 2009 National Youth Gang Survey," *OJJDP Juvenile Justice Fact Sheet, U.S. Department of Justice*.

Esbensen, F.-A., and D. P. Lynskey (2001): *The Eurogang Paradox: Street Gangs and Youth Groups in the U.S. and Europe*chap. Youth Gang Members in a School Survey, pp. 93–114. Kluwer Academic Publishers.

Esbensen, F.-A., and F. M. Weerman (2005): "Youth Gangs and Troublesome Youth Groups in the United States and the Netherlands," *European Journal of Criminology*, 2(1), 5–37.

Evans, W. N., W. E. Oates, and R. M. Schwab (1992): "Measuring Peer Group Effects: A Study of Teenage Behavior," *Journal of Political Economy*, 100(5), 966–91.

Faden, V. B., and M. L. Baskin (2002): "An Evaluation of College Online Alcohol-Policy Information," *Journal of American College Health*, 51(3), 101–107.

Federal Bureau of Investigation (2010): "Crime in the United States 2009," Discussion paper, U.S. Department of Justice.

FRYER, JR., R. G. (2010): "Financial Incentives and Student Achievement: Evidence from Randomized Trials," Working Paper 15898, National Bureau of Economic Research.

GAMBETTA, D. (1996): *The Sicilian Mafia: The Business of Private Protection.* Harvard University Press.

GAROUPA, N. (2000): "The Economics of Organized Crime and Optimal Law Enforcement," *Economic Inquiry*, 38(2), 278–88.

GAVIRIA, A., AND S. RAPHAEL (2001): "School-Based Peer Effects And Juvenile Behavior," *The Review of Economics and Statistics*, 83(2), 257–268.

HECKMAN, J. J., S. H. MOON, R. PINTO, P. A. SAVELYEV, AND A. YAVITZ (2010): "The rate of return to the HighScope Perry Preschool Program," *Journal of Public Economics*, 94(1-2), 114 – 128.

HOFFMAN, E., K. MCCABE, AND V. L. SMITH (1996): "Social Distance and Other-Regarding Behavior in Dictator Games," *American Economic Review*, 86(3), 653–60.

HOME OFFICE (2005): "Findings from the 2003 Offending, Crime and Justice Survey: Alcohol-related Crime and Disorder," Findings 261, Research, Development and Statistics Directorate.

——— (2006): "Underage Drinking: Findings from the 2004 Offending, Crime and Justice Survey," Findings 277, Research, Development and Statistics Directorate.

HUANG, C.-C., D. LAING, AND P. WANG (2004): "Crime and Poverty: A Search-Theoretic Approach," *International Economic Review*, 45(3), pp. 909–938.

Iwai, H. (1986): *Organized Crime: A Global Perspective*chap. Organized Crime in Japan, pp. 208–233. Rowman & Littlefield.

Jankowski, M. (1991): *Islands in the street: gangs and American urban society.* University of California Press.

Kahneman, D., and A. Tversky (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(2), pp. 263–292.

Katz, J. (1988): *Seductions of Crime: Moral and Sensual Attractions in Doing Evil.* Basic Books.

Keynes, J. (1936): *The General Theory of Employment, Interest and Money.* Macmillan and Co., London.

Klein, M. W., F. M. Weerman, and T. P. Thornberry (2006): "Street Gang Violence in Europe," *European Journal of Criminology*, 3(4), 413–437.

Kugler, M., T. Verdier, and Y. Zenou (2005): "Organized Crime, Corruption and Punishment," *Journal of Public Economics*, 89(9-10), 1639 – 1663.

Levitt, S. D. (1996): "The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation," *The Quarterly Journal of Economics*, 111(2), pp. 319–351.

Levitt, S. D., and S. A. Venkatesh (2000): "An Economic Analysis of a Drug-Selling Gang's Finances," *The Quarterly Journal of Economics*, 115(3), pp. 755–789.

Liu, X., E. Patacchini, Y. Zenou, and L.-F. Lee (2011): "Criminal Networks: Who is the Key Player?," Research Papers in Economics 2011:7, Stockholm University, Department of Economics.

LUNDBERG, S. J., AND R. STARTZ (1983): "Private Discrimination and Social Intervention in Competitive Labor Market," *The American Economic Review*, 73(3), pp. 340–347.

MATTHEWS, K., J. SHEPHERD, AND V. SIVARAJASINGHAM (2006): "Violence-related injury and the price of beer in England and Wales," *Applied Economics*, 38(6), 661–670.

MCCARTHY, B., J. HAGAN, AND L. E. COHEN (1998): "Uncertainty, Cooperation, and Crime: Understanding the Decision to Co-offend," *Social Forces*, 77(1), pp. 155–184.

METROPOLITAN POLICE SERVICE (2012): "Met launches major crackdown on gang crime," Press release, MPS Intelligence Systems, London.

MILGROM, P., AND J. ROBERTS (1982): "Predation, reputation, and entry deterrence," *Journal of Economic Theory*, 27(2), 280–312.

MORGAN, W. (1960): *Triad societies in Hong Kong*. Govt. Press.

NELSON, P. (1974): "Advertising as Information," *Journal of Political Economy*, 82(4), 729–54.

PAOLI, L. (2003): *Mafia Brotherhoods: Organized Crime, Italian Style*, Studies in crime and public policy. Oxford University Press.

PARKS, G. (2000): "The High/Scope Perry Preschool Project," *OJJDP Juvenile Justice Bulletin, U.S. Department of Justice.*

POUTVAARA, P., AND M. PRIKS (2009): "Hooliganism and Police Tactics," *Journal of Public Economic Theory*, 11(3), 441–453.

——— (2011): "Unemployment and gang crime: can prosperity backfire?," *Economics of Governance*, 12(3), 259–273.

RAAB, S. (2006): *Five Families: The Rise, Decline, and Resurgence of America's Most Powerful Mafia Empires*. St. Martin's Press.

REUTER, P. (1995): "The Decline of the American Mafia," *The Public Interest*, (120), 89–99.

RHODES, R. (2001): *Why They Kill: The Discoveries of a Maverick Criminologist*. Vintage Books.

SAH, R. K. (1991): "Social Osmosis and Patterns of Crime," *The Journal of Political Economy*, 99(6), pp. 1272–1295.

SALAGAEV, A., A. SHASHKIN, I. SHERBAKOVA, AND E. TOURIYANSKIY (2005): *European Street Gangs and Troublesome Youth Groups*chap. Contemporary Russian Gangs: History, Membership, and Crime Involvement, pp. 169–192, Violence prevention and policy series. AltaMira Press.

SCHWARTZ, D., AND T. ROUSELLE (2009): "Using Social Network Analysis to Target Criminal Networks," *Trends in Organized Crime*, 12, 188–207, 10.1007/s12117-008-9046-9.

SILVERMAN, D. (2004): "Street Crime and Street Culture," *International Economic Review*, 45(3), pp. 761–786.

SKAPERDAS, S. (2001): "The political economy of organized crime: providing protection when the state does not," *Economics of Governance*, 2(3), 173–202.

SKAPERDAS, S., AND C. SYROPOULOS (1997): *The Economics of Organised Crime*chap. Gangs as Primitive States, pp. 61–82. Cambridge University Press.

VERDIER, T., AND Y. ZENOU (2004): "Racial Beliefs, Location, and the Causes of Crime," *International Economic Review*, 45(3), pp. 731–760.