Entering the Underworld:
A Formal Model of Prison Gang Recruitment*

Danilo Alves Mendes Freire

15th February 2015

Abstract

This paper offers a game-theoretical model to analyse both the incentives for a criminal to join a prison gang and how those groups are able to hire competent criminals under conditions of uncertainty and information asymmetry. The model suggests three findings. First, it stresses the role of informers in the gang's recruitment process. Informers allow the prison gang to keep low entry costs, so the gang can attract a larger pool of applicants and still be able to select competent candidates. Second, it indicates that there are cases in which joining a prison gang is not the best option for an inmate. When the detainee has enough skills to endure prison conditions by himself, the prisoner might be better off if he or she decides to "go it alone" and devote his or her ability exclusively to his own survival. Third, the model confirms the idea that the prison gang is not only a "school of crime", but perhaps most importantly, it is a highly effective screening device. Prison gangs increase the welfare of the inmates by providing an extremely valuable public good: reliable information. Policy implications and possible extensions of the current study are also discussed.

1 Introduction

Criminal games are played for very high stakes (Dixit, 2011, 2). In the underworld, rules are violently enforced (Liebling & Arnold, 2012; Von Lampe & Ole Johansen, 2004), mistakes are rarely tolerated, and commitments usually have prohibitive exit costs (Campana & Varese, 2013). Prisons take these problems to the extreme (Clemmer, 1940; Sykes & Messinger, 1960). Since prisoners share the same physical space for long periods of time, their interactions are maximised and the chances of fleeing and escaping retaliation are notably low (Skarbek, 2011). Moreover, prisoners have their schedules tightly controlled by correctional officers, what imposes clear limits on the detainees' ability to meet and negotiate (Skarbek, 2014, 7). Since the communication between inmates is limited, they have to rely upon costly signals, which are by definition expensive and risky, to convey a message (Gambetta, 2009).

It is also obvious that prisoners cannot ask the authorities to solve their issues, even if they wanted them to. First, the penal system has limited resources, so surveillance is usually imperfect. There is only a very small guarantee that the guards will be able to protect the

inmates as well as they should (Skarbek, 2014, 20). In the developing world, where prisons are overcrowded and officers are knowingly corrupt (Darke, 2013; Lemgruber, 2005; Silveira, 2007), this is indeed the normal state of affairs. Second, even if the penal authorities were honest and well-funded, officers may not have enough information on the illegal activities carried by prisoners, such as drug selling or requests for target killings (Kauffman, 1988), to protect inmates efficiently. Hence, they would be unable to mediate or prevent conflicts related to black markets. Additionally, being seen as an informer ("snitch") in the prison greatly increases the chances that an inmate will suffer severe retaliations (Åkerström, 1986), so even if the guards had this kind of information it would be unlikely that inmates would ask for their help. What would then be the best strategy for an inmate to adopt?

As the literature suggests, prison gangs are to a large extent a response to these problems (Camp & Camp, 1985; Fleisher & Decker, 2001). Gangs reduce transaction costs, enforce physical and property rights, and act as brokers in correction facilities (Buentello et al., 1991; DeLisi et al., 2004; Skarbek, 2011). Nevertheless, prison organisations also have their own interests, and they need to maximise their strength by recruiting "the most competetent" prisoners in the underworld. This is definitely not an easy task.

In the next pages I present a formal model to address these questions. My intention here is to devise an analytical framework that explains a criminal's decision to join a prison gang and how the organisation reacts to the potential new members' intentions.

2 The Model

The model consists of two players, a prisoner P and a prison gang leadership G[1]. Since the purpose of the present exercise is to evaluate the likelihood of P joining a gang, I shall assume that G is already present in a given prison. Therefore, G moves first, and sets a cost τ to be paid by the prisoner. This a very common gang practice: not only the need for financial contributions are specified in several gang constitutions (Leeson & Skarbek, 2010), but there is plenty of qualitative evidence suggesting that gang members regularly contribute with a weekly or monthly fee to support the organisation (e.g. Varese, 2001; Gambetta, 1996). Also, gang members are expect to engage in the group's risky plans, so other important non-monetary costs are also captured by τ. Here I assume that such cost is known by P, either because the information is described in the gang's statute, or because it has been disseminated by the inmates in a gang-controlled prison.

The prisoner can either comply (C) and bear the costs τ, or defect and be independent in the prison (D). G plays next, and after assessing if P meets its criteria, chooses whether to accept (A) or reject (R) the new member, and payoffs are realised.

Let S be a measure of the gang's capacity to ameliorate prison conditions. I assume that the true value of S is known to both the prisoner and the gang, what implies that the players have a perfect understanding of how G can help the P by providing both public and club goods.

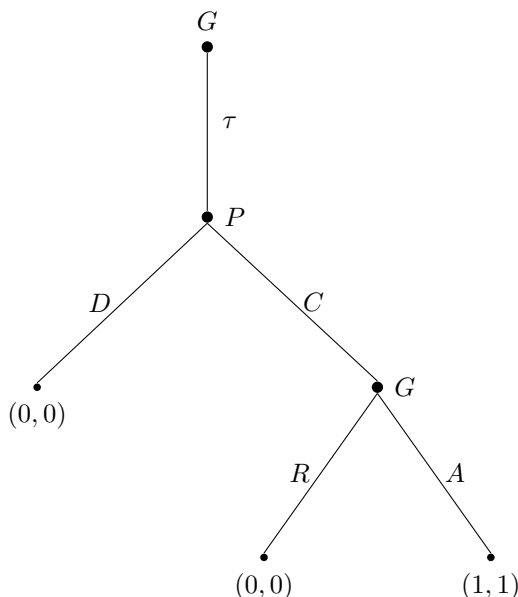[1]The notation employ here is similar to the one used in Lessing (2014).

Figure 1: Gang Prisoner Acceptance

The assumption is reasonable. The benefits given by the gangs, such as protection and financial assistance, are widely advertised in prisons and are precisely the reason why inmates would join a group (Fleisher & Decker, 2001; Trammell, 2012).

Let $s \sim U[0,1]$ be a series of skills that are useful for prison life. These can be physical strength or other abilities such as strategic thinking, technical expertise or a dense personal network. The prisoner knows his own value s_P, whereas G only knows that that s follows a uniform distribution. The utility of s_P increases monotonically, as more skills are always preferred to less.

The expected net utility can thus be represented as a simple difference between the benefit (b) of being in a prison gang and the benefit of remaining independent. As we have seen above, the costs for a prisoner to be in a gang are positive, since the prisoner has to pay a monthly monetary tax to the gang and follow the group's rules.

$$b_P(s,1) - b_P(s,0) \geq 0 \tag{1}$$

$$c_P(s,1) - c_P(s,0) \geq 0 \tag{2}$$

Nevertheless, there exists a value $\bar{s}$ such that

$$b_P(s,1) - c_P(s,1) \leq b_P(s,0) - c_P(s,0) \tag{3}$$

If $s_P \geq \bar{s}$. That is, above a given value of s it is interesting for a prisoner to "go it alone."[2] If P has enough skills to survive in prison by himself, P might ponder if it is really profitable for him to join any kind of group. For prisoners who are below this threshold, becoming a

---

[2]Obviously, the opposite is true whether $s_P < \bar{s}$.

member is a dominant position. The benefits are therefore $S_P = 1$ if the prisoner decides to comply with the gang's cost τ and the group accepts him as new member ($A_P = C$ and $A_G = A$). In contrast, $S_P = 0$ when $A_P = D$ or $A_G = R$, that is, when the prisoner defects or the gang refuses to have in its ranks.

As for the gang, there is a low threshold $\underline{s}$ under which any member acceptance implies more costs than benefits to the gang. Any prisoner below this value is likely to become a burden to G since he will not be competent enough to fulfil the tasks demanded by the group, and might lower the collective welfare by increasing other members' chances of being punished by the police. Therefore, there exists $\underline{s} \in [0, 1]$ such that

$$b_G(s, 1) - c_G(s, 1) \leq b_G(s, 0) - c_G(s, 0) \tag{4}$$

When $s \leq \underline{s}$, and reverse situation when $s \geq \underline{s}$. G will only choose a member that maximises its utility (a high value of s), and reject every prisoner it sees as unfit for the job due to his lack of competence in criminal activities.

2.1 Perfect Information Game

Assuming that the prisoner and the gang have complete information, there are three possible equilibria to this game. Whether $s_P \in [0, \underline{s}]$, P will comply and try join the group (C), while G will refuse him (R). If $s_P \in [\underline{s}, \overline{s}]$, the equilibrium is ($C, A$), in which P joins the group and G accepts him. Finally, if $s_P \in [\overline{s}, 1]$ the equilibrium is (D, A), and P will remain alone.

This situation would be valid if G could easily assess the qualities of every potential member. While prison gangs do their best to obtain credible information on the convicts, it is likely that a perfect information game would only be a good approximation to the facts in a very selected number of cases. It is true that prison records provide a reliable source of information on individuals, and gangs often have access to those data via inmates and sometimes even prison staff members (Gambetta, 2009). Not only this is a long, costly task, but may not be feasible in some circumstances. High security prisons makes this comprehensive evaluation process virtually impossible. In such environment, the gang cannot be completely sure of the inmate's past behaviour and present intentions, and since interactions between a potential gang member and their "godfathers in crime" are relatively few, criminals have difficulties to advertise their credentials to the gang. Recruitment with perfect information is therefore unlikely to happen.

In this regard, I now discuss the recruitment process of a prison gang under conditions of information asymmetry. This is a more realistic assumption. The gang's requirement that every prospective member should receive the approval of other current members before joining the gang is a proof that the gang needs to cross-check information before taking a decision. Thus, the *informants* play a crucial role in a prison gang's recruitment process, what is indeed expected since an inmate's commitment to the gang is usually for life. Also, informants have strong incentives to provide reliable information to the gangs, simply because they are held

accountable for future actions of their "godson". Bad selection choices can not only cost him prestige, but to the extreme lead to severe punishments as fines and expulsion from the group.

2.2 Imperfect Information Game

The next game has the same two players of the previous one, a prisoner P and a prison gang G. However, I add a third figure that does not take any action in the game *per se*, but is of extreme importance to the model: the *informer* (I). Here, I represents the current gang member who is in charge of assessing the criminal qualities of a potential candidate. While I has several reasons to provide credible information to the gang, he can also incur in errors and make wrong judgements about a prisoner's abilities. The information about the prisoner can thus be (T) or false (F)[3].

To reiterate, the game goes as follows.

1. Nature chooses the skills of the prisoner P, with $s \sim U[0,1]$.

2. The gang leader chooses t^*.

3. The prisoner decides if apply or not to the gang. If not, the game ends. If yes, the game follows to the next stage.

4. The gang leader then decides to accept or reject the join proposal. The game ends.

In the present model, I assume that the gang will decide whether it accepts the candidate or not based upon his additional productivity. Therefore, the gang will add a new member if and only if

$$\alpha\left(n+1\right) \geq \alpha\left(n\right) \tag{5}$$

Where α is the productivity function and n is the current number of members in the gang. We assume thereafter that exists a number of prisoners that optimises $\alpha(n)$, and we denote it as α^*.[4]

The game can be solved via backward induction. The prisoner applies when

$$U_P[Comply] \geq U_P[Defect] \tag{6}$$

The utility of joining is equal to

---

[3]By false I do not imply that wrong information given by I is intentionally manipulated. It can be merely the result of miscalculations, and the term here has no pejorative meaning.

[4]Notice further that the productivity can be expressed by a simple function that has limit 1 when $n \to \infty$ and limit 0 when $n \to 0$. We assume that this function has a maxima in n^* as well. An example could be $(an)^{\frac{b}{n}}$, for parameters a and b given.

$$\alpha(n+1)s - \tau \tag{7}$$

Where $\alpha(n+1)$ is the marginal productivity of the new member to the gang, s is the skill of the prisoner and τ is the fee (monetary and non-monetary) imposed by G on P. The prisoner's utility for not joining the gang is simply s, which is P's individual ability to survive in prison.

We can then expand the equations as

$$
\begin{aligned}
\alpha(n+1)s - \tau &\geq s \\
[\alpha(n+1) - 1]s &\geq \tau \\
s \geq \frac{\tau}{\alpha(n+1) - 1} &= s^*(\tau)
\end{aligned}
\tag{8}
$$

All prisoners with $s \geq s^*(\tau)$ will try to enter the gang. Notice that

$$\frac{ds^*(\tau)}{d\tau} = \frac{1}{\alpha(n+1) - 1} > 0 \tag{9}$$

Thus, the proportion of prisoners that intend to join the gang shrinks as the leadership rises the fee. In such scenario, G can choose qualified candidates only by increasing the entry barriers.

Let us now consider that $\tau = \underline{\tau} + t$, where $\underline{\tau}$ is an exogenous cost of joining the gang. To illustrate this situation, imagine that police forces might initiate a crackdown on the prison gang, or that there is a dispute between two different gangs in the same jail. In such scenario, t is the share of the cost that G is imposes in the selection process. The idea is to raise the level of its members, since more skills are needed. G will choose a value of τ that maximises its revenues and skills. This is given by

$$\alpha(n(s^*(\underline{\tau} + t))) \times (1 - s^*(\underline{\tau} + t)) \times E(s^*(\underline{\tau} + t)) \tag{10}$$

In which the first part of the equation is the effect caused by the productivity parameter, the second part is the proportion of prisoners that are accepted in the gang and the third the expected productivity of the gang as a whole. Furthermore, as the number of members is always decided by criteria of optimise the productivity $\alpha(n)$, we can without loss of generality change it by α^*. Thus, the gang optimises the ex-post utility by maximise the following equation

$$
\begin{aligned}
\max_{t \geq 0} \{\alpha^* \times (1 - s^*(\underline{\tau} + t)) \times E(s^*(\underline{\tau} + t))\} \\
s^*(t) = \frac{1}{3}
\end{aligned}
\tag{11}
$$

And solving for t we have that

$$t^* = \frac{1}{3}(\alpha^* - 1) - \underline{\tau} \tag{12}$$

And therefore, the duple t^* and $s^*(t^*)$ solves the game for the Gang and the Prisoner, respectively.

This equilibrium has interesting qualitative properties:

1) All prisoners with

$$s \geq s^* = \frac{\underline{\tau} + t^*}{\alpha(n+1) - 1}$$

Will join the prison gang G.

2) $t^* = 0$, when the exogenous shocks are high

3) $t^* > 0$ when there are many prisoners trying to join G, or then $\underline{\tau}$ is low. When a high number of prisoners are willing to join the prison gang, it can be implied that low-skilled inmates will also apply and get the spot.

4) n^* maximises the ex-post benefit for the gang. The reason is that, given that t^* had already been chosen, the only preoccupation for the gang is its productivity increase. As long as the new member P can rise G's productivity levels, the gang will accept him.

5) t^* maximises the ex-ante utility for the gang. The reasoning behind it is that the gang can control the admission of prisoners by selecting only inmates with high values of s^*.

Now, consider the situation when we add the informant. The informant acts as a probability of find out the true value of s of a given applicant. Let this probability be denoted by π, when the complement means that the information is wrong.

The time line for the second game goes as follows.

1. Nature chooses the skills of the prisoner P, with $s \sim U[0,1]$.

2. The gang leader chooses t^*.

3. The prisoner decides if apply or not to the gang. If not, the game ends. If yes, the game follows to the next stage.

4. With probability π the gang leader finds out the true type of the prisoner. With probability $1 - \pi$ the gang leader does not find it.

5. Finally, the gang leader then decides to accept or reject the join proposal.

Adding this new feature the the previous game empowers the gang leader by giving him a better screen process, where he can combine fees and a selection upon the revealing of the quality parameter by the informant.

Solving the game backwards we have that the gang accepts an offer, ex-post the choice of t whenever

$$\alpha(n+1)S + \alpha(n+1)\left[\pi s + (1-\pi)E[s*|t]\right] \geq \alpha(n)S \tag{13}$$

Where S is the sum of the abilities of all current gang members. Solving this equation we have that the gang accepts ex-post the new member when

$$s \geq \frac{S}{\pi}\left[\frac{\alpha(n) - \alpha(n+1)}{\alpha(n+1)}\right] - \frac{1-\pi}{\pi}\left[\frac{1 - s^*(t)^2}{2}\right] = s^{**} \tag{14}$$

This equation means that there will be a new limit, s^{**} that, upon the information of the productivity of the prisoner, the gang decides by imposing a higher threshold. This means that $s^{**} > s^*(t)$. The qualitative difference that this will generate is that the gang now can select more skilled prisoners without having to raise the fees t.

In the sequence, the prisoner ask to join the gang if and only if

$$U_P[Comply] \geq U_P[Defect]$$
$$\alpha(n+1)s - \tau \geq s \tag{15}$$
$$s \geq \frac{\tau + t}{\alpha(n+1) - 1} = s^*(t)$$

Which is the same as in the previous model. Finally, the ex-ante optimum for the gang will be to set the fees at zero. This because when there is informants inside the process we have that the ex-ante value that the gang maximises is equal to

$$\max_{t \geq 0}\{\alpha^* \times (\pi(1 - s^{**})\int_{s^{**}}^{1} \sigma d\sigma + (1-\pi)(s^{**} - s^*)\int_{s^*}^{s^{**}} \sigma d\sigma)\} \tag{16}$$

And optimising in t this expression leads us to

$$t = \left(-\frac{s^{**}}{3} - \tau\right)(\alpha^* - 1) \tag{17}$$

And as $t < 0$, $t^* = 0$ for all parameters, as long as $s^{**} > s^*$.

For the model above, we have the following qualitative properties:

1) Information minimises the use of entry barriers by the prison gang;

2) The quality of the information (π) provider by the informer I also matters. We see that π has a positive impact on s^{**}, as $\frac{ds^{**}}{d\pi} > 0$. In a nutshell, it is easy to observe that the more (and better) information available, the more precise the screening process is;

3) In the present model, the proportion of prisoners remains the same;

4) However, the difference we observe here is that the gang has now a much more qualified cadre. If G is capable of extracting good, reliable information, it has a significant impact on the abilities of the selected members.

2.3   Comparison between Models

The first and more important difference between models 1 and 2 is that in the latter the selection process is notably more sophisticated than in the earlier one. The use of information, here designed by π, helps the gang to choose more qualified members to its ranks.

When we reach the limit satisfaction, that is when $\pi \to 1$ (when the gang's information of the given prisoner is precise and sufficient), the revenue for the gang is clearly maximised. In a nutshell, for a prison gang information is a very necessary good, certainly crucial for it to recruit the correct prisoner[5].

In the first model, one may observe that $t^* \geq 0$, whereas in the second one $t^* = 0$. This is a result of information, which is enough to deter unskilled prisoners from entering the gang.

Interestingly, not only G has a benefit increase when information is added to the model. For P, adding information to the selection process eliminates the cost t^*, thus reducing the loss for the inmate.

Another interesting point is that such result holds for general α^* values as long as we consider that $\alpha(n)$ has those characteristics assumed in the text. If we assume that α reaches a point where it starts to generate benefits lower than 1, some prisoners will prefer to "go it alone" and not join the prison group. Obviously, the most skilled prisoners will be these ones that shall run alone and the others may free ride, as they are less skilled and can free ride in the group's productivity.

Finally, if we consider the possibility of a lower threshold $\underline{s}$ under which the gang prefers to refuse a prospective member, if $s^* < \underline{s}$, there is the possibility that the gang is indifferent for that choice. In the informer's equilibrium, there is a π probability that the inmates will get caught by a third party (the police forces or a rival gang in our example), and shall not join the gang. This measure avoids the fact that the gang leadership would become weaker.

---

[5]The true role of information may be distorted in these conclusions, but note that π also changes the second threshold s^{**}. This makes information the key to the last result.

3 Discussion

The model has three important findings. First, it shows that a prison gang uses the value of its initial costs as a first selection method, since only skillful prisoners will be able to comply with the gang's demands if the costs are too high. Although useful, this procedure generates costs for both the prisoner and the gang. For the prisoner, it is necessary to spend long periods of time to display all abilities required by the gang, what may lead to an increasing lack of interest for the criminal. The gang also has to bear costs, since the process of establishing and enforcing a high threshold demands considerable human and financial resources from the group. In order to reduce such gang costs, the group can make use of informers. With an informer, the gang can keep a lower entry cost, thus attracting a large pool of applicants while still being able to select competent candidates for the job.

Second, there are cases in which joining a prison gang is not the best option for an inmate. When the detainee has enough skills to endure prison conditions by himself, he might be better off if he decides to "go it alone" and devote his ability exclusively to his own survival. If a prison gang decides to lower its threshold to admit a higher number of members, a competent criminal will have its individual benefit reduced. Since the collective goods provided by the gang would be divided amongst several detainees, his net benefit will be lower than in the previous stage, where the selection process was stricter. However, the gang can eliminate this loss by raising the bar for admission, so that qualified criminals will only share the collective goods with those who can contribute to the group's welfare as much as they do.

Third, the models confirms the idea that the prison gang is not only a "school of crime", but perhaps most importantly, a highly effective screening device (Gambetta, 2009). If a given prisoner knows that the admission to a certain gang is difficult to obtain, as soon as he manages to join the gang he can identify peers who are at least as skillful (and trustworthy) as he is. Therefore, prison gang membership solves one of the most critical issues of the criminal underworld, that of *peer identification.* If a prison gang's admission process is deemed as satisfactory by the prisoners, if an inmate joins the gang he sends a signal that only true criminals could afford, thus dramatically reducing the chances that he might be a snitch. This also has the benefit of reducing violence between criminals, as they do not have to constantly "test" each others' integrity or to use violence as a communicative means to express one's toughness. Prison gangs thereby increase the welfare of the inmates by providing an extremely valuable public good: reliable information.

Finally, it should also be noted that this study can be expanded in several ways. The model shown here has to be tested with qualitative information from prison gangs in order to give it external validity. Moreover, due to serious lack of data on the topic, I was not able to incorporate the state dimension into the model. By including corrupt prison officers and weak public institutions to our formal analysis one could make the model even more realistic and derive more robust predictions from subsequent findings.

References

Åkerström, M. (1986). Outcasts in Prison: The Cases of Informers and Sex Offenders. *Deviant Behavior*, 7(1), 1–12.

Buentello, S., Fong, R. S., & Vogel, R. E. (1991). Prison Gang Development: A Theoretical Model. *The Prison Journal*, 71(2), 3–14.

Camp, G. M. & Camp, C. G. (1985). *Prison Gangs: Their Extent, Nature, and Impact on Prisons.* US Department of Justice, Office of Legal Policy, Federal Justice Research Program.

Campana, P. & Varese, F. (2013). Cooperation in Criminal Organizations: Kinship and Violence as Credible Commitments. *Rationality and Society*, 25(3), 263–289.

Clemmer, D. (1940). *The Prison Community.* Christopher Publishing House.

Darke, S. (2013). Inmate Governance in Brazilian Prisons. *The Howard Journal of Criminal Justice*, 52(3), 272–284.

DeLisi, M., Berg, M. T., & Hochstetler, A. (2004). Gang Members, Career Criminals and Prison Violence: Further Specification of the Importation Model of Inmate Behavior. *Criminal Justice Studies*, 17(4), 369–383.

Dixit, A. (2011). A Game-Theoretic Perspective on Diego Gambetta's Codes of the Underworld. *Global Crime*, 12(2), 134–145.

Fleisher, M. S. & Decker, S. H. (2001). An Overview of the Challenge of Prison Gangs. *Corrections Management Quarterly*, 5, 1–9.

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

Gambetta, D. (2009). *Codes of the Underworld: How Criminals Communicate.* Princeton University Press.

Kauffman, K. (1988). *Prison Officers and Their World.* Harvard University Press.

Leeson, P. T. & Skarbek, D. B. (2010). Criminal Constitutions. *Global Crime*, 11(3), 279–297.

Lemgruber, J. (2005). *The Brazilian Prison System: A Brief Diagnosis.*

Lessing, B. (2014). A Hole at the Center of the State: Prison Gangs and the Limits to Punitive Power. *Center of Democracy, Development, and the Rule of Law.*

Liebling, A. & Arnold, H. (2012). Social Relationships between Prisoners in a Maximum Security Prison: Violence, Faith, and the Declining Nature of Trust. *Journal of Criminal Justice.*

Silveira, V. J. (2007). A Realidade dos Presídios na Visão da Pastoral Carcerária. *estudos avançados*, 21(61), 209.

Skarbek, D. (2011). Governance and Prison Gangs. *American Political Science Review*, 105(4), 702–716.

Skarbek, D. (2014). *The Social Order of the Underworld: How Prison Gangs Govern the American Penal System*. Oxford University Press.

Sykes, G. M. & Messinger, S. L. (1960). The Inmate Social System. *Theoretical studies in social organization of the prison*, (pp. 5–19).

Trammell, R. (2012). *Enforcing the Convict Code: Violence and Prison Culture*. Lynne Rienner Publishers Boulder.

Varese, F. (2001). *The Russian Mafia: Private Protection in a New Market Economy*. Oxford University Press.

Von Lampe, K. & Ole Johansen, P. (2004). Organized Crime and Trust: On the Conceptualization and Empirical Relevance of Trust in the Context of Criminal Networks. *Global Crime*, 6(2), 159–184.