Optimal Accomplice-Witnesses Regulation under Asymmetric Information*

Giovanni Immordino[†] Salvatore Piccolo[‡]

September 27, 2011

Abstract

We study the problem of a Legislator designing immunity for cooperating accomplices. The objective of the paper is to highlight a positive externality between expected returns from crime and the information rents that must be granted by the Legislator to those whistleblowers that own insider information. We identify the accomplices' incentives to release distorted information and characterize the second-best policy limiting this behavior. The central finding is that this externality leads to a policy that purposefully allows cooperating accomplices not to fully reveal their information. Moreover, our results are consistent with the legislative provisions: accomplices must fulfill minimal information requirements to be admitted into the program (rationing), a bonus must be awarded to those accomplices that provide 'good-quality' information and rewards to a self-reporting boss can increase welfare.

Keywords: Accomplice-witnesses, Adverse Selection, Criminal Organizations, Leniency.

*We are grateful to Luigi Balletta, Alberto Bennardo, Tracy Lewis, Bentley MacLoad, David Martimort, Marco Pagano, Maarten Pieter Schinkel and Klaus Schmidt for useful comments. Seminar participants in Munich, Palermo, Naples and Queen Mary College (London) are also acknowledged for useful suggestions.

[†]University of Salerno and CSEF.

[‡]University of Naples Federico II and CSEF.

1 Introduction

Successful prosecution of criminal organizations' heads often requires to draw upon the uncorroborated evidence provided by cooperating accomplices. One of the main features of organized crime is that the most culpable and dangerous individuals rarely do the 'dirty job' — see, e.g., Jeffries and Gleeson (1995). Even if these people are ultimately responsible for the crimes committed by their 'soldiers', it is often hard to prove them guilty because they mainly deal through intermediaries and push their own participation up to behind-the-scenes control and guidance. This feature has lead many countries to introduce innovative legal rules facilitating the use of insider information in criminal proceedings. Among those laws, leniency programs, which are designed to encourage former mobsters to plea guilty and testify against their partners in exchange of charging or sentencing concessions, are perhaps the most relevant.

The use of accomplices as witnesses in criminal proceedings is a long-standing practice. Nevertheless, ever since these programs have been introduced, they have met with political scepticism as well as with outrage among relatives of victims of mafia-related crimes. Is it really necessary to promise judicial leniency to mobsters that have planned, ordered and even executed numerous murders and other criminal activities? What is the economic logic behind these programs? Why these benefits appear to be more generous than what one would normally expect?

Allowing for uncorroborated testimonies is usually viewed as a crucial advantage in the prosecution of organized crime. Insider information can provide a richly detailed context to a case — e.g., that a criminal organization met at a particular location and that the witness was in a position to know about the types of criminal acts at issue — that can help making the public proceeding against a defendant compelling. However, accomplices are willing to testify only if they perceive that there will be some legal benefits flowing to them from the arrangement, and this form of 'horse-trading' exacerbates the greater is the risk of intimidation and reprisal by their former partners. As explained by Schur (1988), crime instigators' most effective tool against prosecution is murdering whistleblowers. Intimidation in criminal proceedings has, in fact, both pervasive and perverse effects: the risk of biased and untruthful testimonies is, for instance, potentially staggering.[1] Why is it then reasonable to promise judicial leniency to flipping criminals that will probably not fully perform their task?

The purpose of this paper is twofold. First, we wish to understand the forces that lead informants to misbehave by releasing biased information. Second, we characterize the instruments that the Legislator can use to limit such behavior. Our main conclusion is that the risk of witness intimidation leads to distorted second-best policies that purposefully allow whistleblowers to hide part of their private information, but that at the same time require strict eligibility criteria for them to be admitted into the program.

To make this point, we study the optimal design of a leniency program in a simple game involving a

[1] There is much controversy concerning accomplice witnesses both on the efficiency and fairness grounds. In Germany, for instance, as underlined by Huber (2001), arguments against the use of accomplice witnesses are based on : "The principle of equal treatment and principles of proportionality and legality...Additionally, there have been doubts expressed about the level of truthfulness in the testimony of accomplice witnesses." Other countries, like those of Anglo-Saxon tradition, mainly underline the necessary role played by cooperating accomplices in criminal justice, especially when a state of emergency is justified because of organized crime.

hierarchical criminal organization and a Legislator. The criminal organization is formed by two mobsters: a principal (boss) and an agent (fellow), each with specific skills. The boss plans the crime and delegates its execution to the fellow. After the crime has been committed, some evidence about the boss and his involvement in the crime materializes, this evidence is observed only by the criminals but neither by prosecutors nor by the jury ruling the trial. The crime triggers an investigation and, at this stage, the agent can opt to whistle. The prize for cooperation entails an amnesty announced by the Legislator at the outset of the game. Moreover, the Legislator can also enforce restrictions on the selection process regulating the program's admission policy — i.e., only accomplices whose information satisfies minimal standards are eligible for the program.

We first show the beneficial value of leniency in this baseline set-up and characterize the second-best policy. Then we extend the analysis to a setting where the Legislator is also allowed to award the benefit of an amnesty to a self-confessing boss. The key driving force is that, as long as the relationship between the Legislator and the informants is plagued by asymmetric information, there is a positive externality between the need to grant rents to whistleblowers in order to elicit truthful information and the crime monetary return flowing to the boss. The point is that more precise and reliable testimonies imply a higher conviction probability for the boss, whose intimidation and reprisal ability weakens when convicted and jailed. Hence, in those states of nature where the evidence that can be gathered against the boss is quite reliable irrespective of whether the agent testifies or not, under-reporting pays off. Essentially, pretending to be in a situation where the risk of reprisal is substantially high allows the accomplice to ask for better amnesties than what would be necessary. In equilibrium, this possibility generates an ex-post information rent for the agent that stifles the reservation wage he needs to be offered in order to accept the illegal deal and, in turn, spurs the boss' net gain from the crime.

We show that this externality is the main source of a marked difference between the optimal policy under complete and incomplete information. In the former case the accomplice cannot distort the testimony because his information is the same as that available to the prosecutor. The optimal policy is then chosen so as to make an accomplice indifferent between talking and facing the trial, and it entails no entry restrictions to the program. By contrast, under asymmetric information the optimal policy awards better deals to those who provide more productive information. Moreover, in order to minimize the rents that privately informed accomplices can grab, the Legislator is forced to restrict the access to the program (*rationing*), to require distorted testimonies (*partial disclosure*) and to extend the benefit of an amnesty (sometimes) also to the boss (*self-reporting*).

Interestingly, under asymmetric information, the optimal policy entails an *excessively generous amnesty* relative to the complete information benchmark: in order to elicit truthful information a bonus needs to be awarded to those accomplices that reveal 'good quality' information. Similarly, although it is always convenient to draw upon accomplices' testimonies to fight organized crime, rationing the access to the program by way of an *information floor*, below which agents are sent to trial, is necessary insofar as it has a negative effect on the boss's expected profit and thus reduces the crime rate. Essentially, reducing the set of contingencies in which an accomplice can access the program stifles the uncertainty faced by the

Legislator when announcing the policy, whereby making mimicking less profitable. Once again, this effect is welfare enhancing because lower information rents for the agent shift onto higher break-even wages and thus imply higher costs for the boss.

Finally, we also explain why allowing the boss to *self-report* can be socially beneficial. This is true for reasons which are completely different from the previous literature, arguing that self-reporting saves enforcement resources and reduces risk because self-confessing individuals bear certain rather than uncertain sanctions — see, e.g., Kaplow and Shavell (1994). In our model, granting an amnesty to a self-reporting boss has an effect on the agent's information rent similar to that explained for the information floor: allowing the boss to plea guilty and report information that enables to convict with certainty the whole organization in exchange of legal benefits, reduces the subset of contingencies where the agent talks and thus stifles his information rent.

All these results are consistent with a number of legal provisions characterizing accomplice-witnesses regulations across the world, and show that the benefits of those programs in terms of reduced crime may justify, at least from an efficiency point of view, the risk of biased testimonies and the recognition of pronounced legal benefits to cooperating accomplices.

The rest of the paper is organized as follows. Section 2 links our contribution to the earlier literature. Section 3 sets up the baseline model and develops first the benchmark where the agent's information can be verified by the prosecutors at no cost and then the second-best policy. Section 4 extends the baseline model to the case where the benefit of an amnesty is also awarded to a self-reporting boss. Section 5 concludes. All proofs are in the Appendix.

2 Related literature

Our analysis is related to the literature on antitrust law enforcement studying the effects of leniency programs on cartel formation in oligopolistic markets. The first paper addressing the effects of leniency programs on cartels is Motta and Polo (2003). They analyze the impact of reduced fines for cartel members that cooperate with the antitrust authority and show that it can be efficient to reduce fines even when the authority has already started an investigation, but has not yet obtained evidence of misbehavior. Among many other differences, this paper takes leniency rules as exogenous, while the characterization of the first and second-best leniency policy is key to our analysis.[2] As Chen and Rey (2007), which study the optimal design of leniency programs in a framework where the verifiability of the insider information is not a concern, we also take a mechanism design approach to leniency, but in a more realistic context plagued by asymmetric information. Always in an antitrust setting, Aubert et al. (2006), analyze a model where leniency programs could have a positive social value insofar as they create a conflict of interests between members of different organizations. They also discuss informally the idea that leniency programs could be desirable insofar as they generate conflicts between the members of the same organizations (e.g., firms). Our paper innovates upon both these contributions. To the best of our knowledge, ours is the

[2] See also Rey (2003) and Spagnolo (2003), who take into account the role of rewards to former criminals by studying their determinants and social value.

first paper that introduces adverse selection into the agency relationship between the Legislator setting up the leniency program and the applicants to this program.

More recently, Acconcia et al. (2009) have developed a model of hierarchical criminal organizations where the Legislator grants legal benefits to low-rank criminals who decide to cooperate with the justice. By using data collected for Italy, they argue that the Italian accomplice-witness program introduced in 1991 did affect in a significative manner organized crime in those Italian regions where the mafias have been historically more pervasive. More specifically, they identify the positive effect of the policy on prosecution and argue that it also strengthened deterrence. The current paper is motivated by the evidence collected in Acconcia et al. (2009) and it also extends their theoretical analysis of the determinants of the optimal amnesty rate in three main directions. First, in contrast to them, we consider an adverse selection setting where the accomplice's information is non-verifiable. Second, we enlarge the Legislator's set of instruments to include, besides the amnesty rate, an information floor below which an agent's testimony is not accepted. Finally, we also consider the possibility of awarding an amnesty also to a self-reporting boss and show that this is sometimes necessary to fight organized crime in an efficient way. In this respect, our work also relates to the 'self-reporting' literature. In Kaplow and Shavell (1994), for instance, self-reporting saves enforcement resources because individuals who report their harmful acts need not be detected, and it reduces risk because individuals who report their behavior bear certain rather than uncertain sanctions. In our model, instead, the welfare enhancing effect of self-reporting stems from the hierarchical nature of criminal organizations, and it becomes more relevant the more severe is the adverse selection problem between the Legislator and the cooperating accomplices.

Another strand of literature that shares common features with our paper is that on organized crime. Traditionally, this literature has stressed welfare comparisons between monopoly and competitive supply of bads — see, e.g., Buchanan (1973) and Backhaus (1979) — while more recently Jennings (1984), Polo (1995), Konrad and Skaperdas (1994, 1997) and Garoupa (2000) started to model criminal organizations as vertical structures where the principal has the necessity to discipline its members.[3] But these models have overlooked the role of accomplice-witnesses programs as a tool to generate conflict within criminal organizations and their optimal design, which are instead key to our analysis.

Clearly, to formalize the idea that testimonies are not verifiable and so accomplices might gain by mimicking, our paper draws on the adverse selection literature — see, e.g., Laffont and Martimort, 2002. Our model contributes to this literature by showing that games between a Legislator and hierarchical criminal organizations have a peculiar trade-off: policies that grant higher information rents to the agents generate a positive externality on the boss who can offer lower wages to hire these agents. Interestingly, this trade-off naturally leads to the result that creating less conflict inside the organization by limiting the access to the accomplice witnesses program improves welfare. This result give rise both to a novel source of rationing and (as said before) to the optimality of awarding an amnesty also to the boss who decides to self-report.

[3] See also Fiorentini and Peltzman (1995), Kugler, Verdier and Zenou (2005), Mansour et al. (2006) and Baccara and Bar-Isaac (2008).

3 The Baseline Model

In this section we set-up the simplest possible model capturing the main features of accomplice-witnesses programs and, at the same time, accounting in a natural way for asymmetric information.

The criminal organization: Consider a game where a benevolent Legislator and two members of a criminal organization, the principal (boss) and his agent (fellow), interact sequentially. The Legislator, having forbidden welfare reducing criminal acts, designs an accomplice-witnesses program. Each member of the criminal organization owns specific skills: the boss plans the crime and delegates its execution to the fellow who materially commits the illegal act.

The crime yields a monetary return π which is stochastic and distributes over the compact support $\Pi \equiv [0, \overline{\pi}]$ according to the cumulative distribution function (cdf) $G(\pi)$. The boss hires the fellow after having observed the realization of the crime return π, he has full bargaining power and proposes a wage w. If accepted, the wage is paid after the crime is committed. If the agent refuses the offer, the game ends and both parties get a reservation utility $\underline{u}$. Committing the crime triggers an investigation with probability α. We normalize $\underline{u} = 0$ and $\alpha = 1$ with no loss of generality.

Information: After the crime has been committed and the wage w has been paid, some evidence about the boss and his involvement into the crime materializes. As we shall explain below, this evidence can be used against the boss and therefore affects the outcome of the investigative and judicial process. It is modeled as the realization of a random variable $\tilde{\theta}$ which distributes over the compact support $\Theta \equiv [0, \overline{\theta}]$ according to the twice continuously differentiable and atomless cdf $F(\theta)$, with density $f(\theta)$. As a convention, we assume that larger values of θ reflect better and more reliable evidence — i.e., higher values of θ mean that more evidence against the boss can be gathered by the judicial authority and brought at trial.

Judicial system and legal regimes: Two independent parties contribute to determine the (stochastic) outcome of the judicial system: the jury and the prosecutor. The jury is a body of citizens and public officials summoned by law and sworn to hear and deliver a verdict upon a case presented in court. The prosecutor is the public official that institutes and conducts the legal proceedings against criminals. Both these parties are unaware of the realization of the random variable $\tilde{\theta}$. There are two possible legal regimes, one with leniency and one without leniency:

No leniency: Under this regime the agent must face the trial: with probability p he is convicted and bears a sanction S_a. The principal is convicted and bears the sanction S_p with probability $q(\theta)$. Otherwise, both mobsters are acquitted with probability $1-p$ and $1-q(\theta)$, respectively. The probability of convicting the boss $q(\theta)$ under no leniency is increasing in θ. The link between θ and $q(.)$ can be interpreted as the outcome of the prosecutor's investigative activity (e.g., shadowing the agent or tapping his phone) that maps the available evidence θ onto the jury's final decision, which we assume is stochastic and reflects the jury's objective function. For simplicity, we will not explicitly model neither the prosecutor's investigative activity nor the jury's objective function.

Leniency: When a leniency program is in place, the agent can decide to whistle on the basis of the observed realization of $\tilde{\theta}$. If so, he enjoys an amnesty $\phi(s)$ in exchange of a testimony s. While better evidence θ helps the investigative activity of the prosecutor, which in turn influences the jury's final decision as explained above, the accomplice's testimony s delivers a public signal to the jury providing additional uncorroborated evidence. This signal, together with the state of nature θ, determines the probability of convicting the boss $Q(s, \theta)$. The probability $Q(.)$ reflects the stochastic outcome of the unmodeled jury's decision given the total evidence brought at trial. For consistency, we assume that $Q(0, \theta) = q(\theta)$ for all θ.

Drawing upon the cooperating accomplices' testimonies is usually viewed as a crucial advantage in the prosecution of organized crime. Testimonies of former mobsters, which are typically bargained with prosecutors before the trial[4], can provide a richly detailed context to a case that can help make the public accusation against a defendant compelling. Clearly, juries are cautioned to use care in evaluating the accomplices' testimony. Because juries may be sceptical of such testimony, prosecutors rarely rely solely on this evidence (Fyfe and Sheptycki, 2006). This is why throughout we shall treat θ and s as complements in determining the boss' conviction probability $Q(.)$ when a leniency program is in place.

Direct revelation mechanism: There is no loss of generality in invoking the Revelation Principle in this framework. So, we restrict attention to direct revelation mechanisms such that, when launching a leniency program, the Legislator commits to a policy $\mathcal{L} = \left\{\phi(\hat{\theta}), s(\hat{\theta})\right\}_{\hat{\theta}\in\Theta}$ specifying an amnesty $\phi(.)$, with $\phi : \Theta \to [0, 1]$, and a testimony $s(.)$, with $s : \Theta \to \Theta$, both contingent on the agent's report $\hat{\theta}$, which, as explained below, is interpreted as a private signal sent by the whistleblower to the prosecutor. Essentially, the prize for the agent's cooperation entails a reduction $\phi(\hat{\theta})$ of the original sanction S_a, and requires a testimony $s(\hat{\theta})$ for every report $\hat{\theta}$.[5]

Finally, to capture the idea that, in practice, accomplice-witnesses programs often impose restrictions on the minimal information that is required to whistleblowers for being admitted in the program, we will also allow the Legislator to set an information floor $\underline{\theta}$: below this threshold a testimony is not accepted. Clearly, if the optimal policy entails $\underline{\theta} > \overline{\theta}$ the program is shut down.

Intimidation risk and retaliation: Criminal organizations seek to punish whistleblowers and, when they succeed, the agent bears a loss $R > 0$, normalized to 1 without loss of generality. We assume that this

[4] As explained in Cassidy (2004), based on the information provided by the cooperating accomplice, the prosecutor is the public official who is actually in charge of proposing and motivating to the jury leninecy for the whistleblower.

[5] In practice, prosecutors and accomplices privately agree on the latters' testimonies and on the legal benefits to be awarded as a prize for this cooperation — see, e.g., Cassidy (2004). For instance, during their long lasting cooperation, judge Falcone and Tommaso Buscetta, met several times before the Palermo Maxi Trial. During these meetings Buscetta revealed the existence and workings of the Sicilian Mafia Commission. This confidential information, coupled with his testimony at trial, enabled Falcone to argue at the trial that Cosa Nostra was a unified hierarchical structure ruled by a Commission and that its leaders — who normally would not dirty their hands with criminal acts — could be held responsible for criminal activities that were committed to benefit the organization. However, it often happens that the confidential information an accomplice reveals to the prosecutor is not fully disclosed at trial. Some of Buscetta's confessions were never disclosed and used at subsequent trials. For instance, as long as Falcone was alive, Buscetta never talked about the links between politicians and the Mafia, even though he subsequently admitted that he knew of such ties. As it will become clear, we capture this scenario by allowing $s(\theta)$ to fall short of θ in a truthful equilibrium.

loss materializes only in the state of nature where the boss is acquitted, which occurs with probability $1 - Q(s, \theta)$ in the state of nature θ given the testimony s. This is with no loss of insights under the hypothesis that the retaliation power of the boss weakens once he is convicted and jailed.

The historical evidence offers ample support to the idea that retaliation is an important source of deterrence for whistleblowers. Many accomplices in Italy have been murdered after their testimony in a mafia trial. The first member of the Sicilian mafia that publicly acknowledged its existence, Leonardo Vitale, was murdered after his testimony.[6] His cooperation lead to a number of alleged mobsters being indicted, but all were acquitted. In the end, only Vitale was imprisoned and killed by his former partners.[7] Buscetta received a different treatment. After his collaboration with judges Giovanni Falcone and Paolo Borsellino, the fugitive members of the corleonesi family murdered two of his sons and many among his closest relatives and friends. Similarly, in 1995, the fourteen years son of the accomplice Santino Di Matteo was brutally murdered after his father's cooperation by members of the corleonesi family that, in spite of the many trials opened against them, were still free of charges or fugitive at that time.

Timing: The timing of the game is as follows:

t=0 The Legislator decides whether to launch a leniency program and accordingly commits to a policy $\wp = (\mathcal{L}, \underline{\theta})$ that, as explained before, entails a mechanism $\mathcal{L} = \left\{\phi(\hat{\theta}), s(\hat{\theta})\right\}_{\hat{\theta} \in \Theta}$ and the floor $\underline{\theta}$.

t=1 Uncertainty about π resolves and the boss decides whether to commit the crime. If the crime is not executed the game ends. Otherwise, once the illegal act is committed, the boss pays the wage w to the accomplice and the game proceeds to the next stage.

t=2 The realization of the random variable $\tilde{\theta}$ materializes and only the criminals learn it.[8]

t=3 The investigation opens. If the leniency program is in place, the agent decides whether to whistle. If so, he sends a private message $\hat{\theta}$ to the prosecutor who then grants him an amnesty $\phi(\hat{\theta})$ in exchange of a testimony $s(\hat{\theta})$.

t=4 The trial uncertainty resolves and sanctions (including the retaliation loss) are imposed.

Actions and equilibrium concept: An action profile for the principal involves a wage offer w. An action profile for the agent involves a participation rule, which depends on the difference between the wage and his expected sanction, and a confession and reporting decision $(y, \hat{\theta})$, where $y = 1$ if the agent

[6] He walked into a Palermo police station on the evening of March 29, 1973, and declared that he was a member of the Mafia and confessed to various acts of extortion, arson and two homicides. In front of dumbfounded police officers he explained how a Mafia family is organised and revealed the existence of the Mafia Commission, long before the pentito Tommaso Buscetta exposed Mafia secrets to judges who were prepared to listen.

[7] According to judge Giovanni Falcone the Mafia understood the importance of Vitale's revelations much better than the Italian justice system at the time and killed him when the time was most opportune — see, e.g., Falcone (1991).

[8] This hypothesis captures the idea that cooperation between mobsters generates information. Only after interacting with the boss and executing its orders, the agent is able to learn some relevant information that, once released to prosecutors, can harm the crime instigator at the judicial stage.

cooperates and $y = 0$ otherwise. The Legislator simply announces a policy $\wp$. We shall look for the Subgame Perfect Nash Equilibrium (SPNE) of this game.

Technical assumptions: The analysis will be conducted under the following assumptions.

A1 The probability function $Q(.)$ is increasing and concave in θ, it is single peaked with respect to s and satisfies $Q_s(s,\theta) = 0$ for $s = \theta$ and $Q(s,\theta) > Q(0,\theta) = q(\theta)$ for all θ. Moreover, it also features increasing differences in θ and s — i.e., $Q_{s\theta}(.) > 0$ for all (θ, s).[9]

Assuming that $Q(.)$ is single peaked and has a unique maximum at $s = \theta$ simply implies that the closer is the informant's testimony to the true state of nature — i.e., the more precise is this testimony — the stronger is its effect on the boss' prosecution risk. Essentially, neither under-reporting nor making up information can improve the probability of convicting the boss.

Complementarity — i.e., $Q_{s\theta}(.) > 0$ — instead reflects the idea that the marginal impact of a better testimony on the probability of convicting the boss is stronger in states of nature where the available evidence is better — i.e., given $s' > s$ the difference $Q(s',\theta) - Q(s,\theta)$ is increasing in θ:

$$\frac{\partial\left(Q(s',\theta) - Q(s,\theta)\right)}{\partial\theta} = \int_s^{s'} Q_{s\theta}(x,\theta)\,dx > 0 \qquad \forall\theta\in\Theta.$$

This hypothesis means that more precise testimonies are relatively of little help when the evidence corroborating and supporting the informant's assertions is poor.[10]

A2 The inverse hazard rate associated to $F(.)$ is monotone and decreasing — i.e.,

$$\frac{\partial}{\partial\theta}\left(\frac{1-F(\theta)}{f(\theta)}\right) \leq 0 \qquad \forall\theta\in\Theta. \tag{A2}$$

This assumption is standard in the screening literature and, as we will show, ensures that the agent's incentive problem is well behaved. To focus on separating equilibria it will be convenient to impose the following additional technical assumptions:

A3 $Q_{s\theta\theta}(.) \leq 0$ and $Q_{ss\theta}(.) \geq 0$ for all s and θ.[11]

[9] For instance, one can easily check that the quadratic specification $Q(s,\theta) = \sigma(\theta - \frac{1}{2}s)s + \gamma\sqrt{\theta}$, with $\sigma > 0$ and $\gamma > 0$, satisfies these assumptions.

[10] Once again, the case of Leonardo Vitale is emblematic to make this point. Already in 1973 Vitale started to cooperate with the justice by indicating the names of many mafiosi and the roots of their main traffics, he also explained how a Mafia family is organized and revealed the existence of the Mafia Commission. In spite of this testimony, the evidence surrounding the trial that was opened on the basis of the mere Vitale's assertions was so tiny that all defendants were in the end acquitted. As pointed out by judge Falcone, this testimony turned out to be very important for the subsequent fights against organized crime in Italy, that is, when more evidence, gathered by Falcone and his investigative group, supported the testimonies of Vitale and the subsequent whistleblowers.

[11] One can immediately check that the class of quadratic functions — e.g., $Q(s,\theta) = \sigma(\theta - \frac{1}{2}s)s + \gamma\sqrt{\theta}$ — satisfies these requirements.

As standard in the screening literature we shall restrict attention to the class of continuous and almost everywhere differentiable mechanisms. As a *tie-breaking* condition we assume that whenever indifferent between joining the program and facing the trial, the agent will whistle. All players are risk neutral. Moreover, following the literature, all sanctions will be interpreted as the monetary equivalent of the imprisonment terms, fines, damages, and so forth, to which the criminals expose themselves.

Social goal: We finally introduce the Legislator's objective function. Given the policy $\wp$, let $C(\wp)$ and $w(\wp)$ denote the boss' expected sanction and the agent's break-even wage, respectively. Then committing the crime yields a non negative expected utility to the boss if and only if the return π exceeds the (total) expected costs — i.e.,

$$\pi \geq C(\wp) + w(\wp) \equiv \underline{\pi}(\wp).$$

Hence, the Legislator's optimal policy $\wp$ will be chosen so as to minimize the (expected) crime rate

$$W(\wp) = 1 - \Pr(\pi \leq \underline{\pi}(\wp)), \tag{1}$$

subject to the relevant incentive and participation constraints.

3.1 First-best Policy

In this section we develop the benchmark where the realization of the random variable $\tilde{\theta}$ is common knowledge to all players. The analysis of this case is fairly simple and intuitive. The Legislator commits to a policy such that each eligible type $\theta \geq \underline{\theta}$ is indifferent between talking or facing the trial. A simple backward induction logic allows to solve the game. Let

$$u(\theta) = -(1 - \phi(\theta)) S_a - (1 - Q(s(\theta), \theta)),$$

be the utility of a type-θ agent who enters the program: he provides a testimonies $s(\theta)$, enjoys an amnesty $\phi(\theta)$ and bears the retaliation loss $R = 1$ with probability $1 - Q(s(\theta), \theta)$ — i.e., in the event that the boss is acquitted. Moreover, let $u_0 = -pS_a$ be the expected utility that the agent obtains when facing the trial: a sanction S_a is imposed with probability p in this case. In each state θ where the agent can apply to the program — i.e., $\theta \geq \underline{\theta}$ — the Legislator chooses the amnesty rate $\phi(\theta)$ and a testimony $s(\theta)$, so as to equalize $u(\theta)$ and u_0 — i.e.,

$$(1 - \phi(\theta)) S_a + (1 - Q(s(\theta), \theta)) = pS_a. \tag{2}$$

Hence, for any policy $\wp$ such that (2), the boss commits the crime if and only if the revenue π exceeds his expected costs — i.e.,

$$\pi \geq C(\wp) + w(\wp) \equiv \underline{\pi}(\wp), \tag{3}$$

where the boss' expected sanction $C(\wp)$ is:

$$C(\wp) = \int_0^{\underline{\theta}} q(\theta) S_p f(\theta) d\theta + \int_{\underline{\theta}}^{\overline{\theta}} Q(s(\theta), \theta) S_p f(\theta) d\theta, \tag{4}$$

and the agent's break even wage $w(\wp)$ solves the following participation constraint:

$$w(\wp) + \int_0^{\underline{\theta}} u_0 f(\theta) d\theta + \int_{\underline{\theta}}^{\overline{\theta}} u(\theta) f(\theta) d\theta = 0 \Longrightarrow w(\wp) = pS_a \qquad \forall \underline{\theta}. \tag{5}$$

It then follows that:

$$\underline{\pi}(\wp) \equiv pS_a + \int_0^{\underline{\theta}} q(\theta) S_p f(\theta) d\theta + \int_{\underline{\theta}}^{\overline{\theta}} Q(s(\theta), \theta) S_p f(\theta) d\theta,$$

so that the Legislator's program writes as:

$$\max_{\wp} \Pr(\pi \leq \underline{\pi}(\wp)), \tag{6}$$

$$\text{subject to} \qquad u(\theta) = u_0.$$

which simply amounts to solve:

$$\max_{\underline{\theta},\ s(\theta)} \underline{\pi}(\wp) = \max_{\underline{\theta},\ s(\theta)} \left\{ \int_0^{\underline{\theta}} q(\theta) S_p f(\theta) d\theta + \int_{\underline{\theta}}^{\overline{\theta}} Q(s(\theta), \theta) S_p f(\theta) d\theta \right\}.$$

The solution of this program immediately leads to the following intuitive result.

Proposition 1 ***(First-best policy)*** *Assume* ***A1****, then the first-best policy $\wp^{fb}$ features the following properties:*

- ***(No rationing)*** *The agent is always admitted into the program — i.e., $\underline{\theta}^{fb} = 0$.*
- ***(Full disclosure)*** *In every state of nature the whistleblower fully discloses his information — i.e., $s^{fb}(\theta) = \theta$.*
- ***(Zero-rent amnesty)*** *The amnesty rate $\phi^{fb}(\theta)$ is chosen so as to make the agent indifferent between talking and facing the trial in every state θ — i.e.,*

$$\phi^{fb}(\theta) = 1 - p + (1 - Q(\theta, \theta)) \frac{1}{S_a}. \tag{7}$$

This amnesty is decreasing in θ — i.e.,

$$\dot{\phi}^{fb}(\theta) = -Q_\theta(\theta, \theta) \frac{1}{S_a} < 0.$$

Under complete information there is no reason to distort the agent's testimony: he fully reveals the state of nature, so to provide the most productive testimony. The 'no rationing' result has also a simple economic interpretation. As the state of nature can be verified at no cost by the prosecutor (and the jury), and the agent's information is always productive — i.e., $Q(\theta, \theta) > q(\theta)$ for all $\theta > 0$ — it will be optimal to always welcome a cooperating accomplice into the program. Finally, to explain why the first-best amnesty rate is decreasing in θ, note that an accomplice with better information faces a lower risk of retaliation because his testimony induces a conviction for the boss with a higher probability (since $Q(\theta, \theta)$ is increasing in θ).

Next section studies how these conclusions change once asymmetric information on the state of nature θ is taken into account.

3.2 Second-best Policy

We now turn to analyze the case where neither the prosecutors nor the jury can observe the realized state of nature θ. In this scenario the game is one of imperfect information and the privately informed agent can gain from mimicking when allowed to join the program. More precisely, depending on the shape of the mechanism $\mathcal{L} = \left\{\phi(\hat{\theta}), s(\hat{\theta})\right\}_{\hat{\theta}\in\Theta}$ announced by the Legislator, in state θ the whistleblower might gain from providing an untruthful report $\hat{\theta}$ in order to enjoy a lighter sanction. These mimicking opportunities force the Legislator to distort the optimal policy for rent extraction reasons: in addition to the participation constraint already considered above, now a set of incentive constraints must be satisfied in order for the policy to elicit truthful information revelation.

To characterize the incentive feasible allocations, let

$$u(\hat{\theta}, \theta) = -(1 - \phi(\hat{\theta}))S_a - (1 - Q(s(\hat{\theta}), \theta)),$$

be the agent's utility in state θ given the report $\hat{\theta}$. An incentive feasible allocation must induce truthful information revelation by those agents that are admitted into the program and, if the floor $\underline{\theta}$ exceeds the lower-bound 0, it must also be such that rationed accomplices do not find it profitable to lie in order to join the program.

The allocation $\{\phi(.), s(.)\}_{\theta\in\Theta}$ must then satisfy the following first- and second- order local conditions for truth-telling:

$$u_{\hat{\theta}}(\hat{\theta}, \theta)\Big|_{\hat{\theta}=\theta} = 0 \qquad \Longrightarrow \qquad \dot{\phi}(\theta)S_a + Q_s(s(\theta), \theta)\,\dot{s}(\theta) = 0 \qquad \forall\theta \geq \underline{\theta}, \tag{8}$$

$$u_{\hat{\theta}\theta}(\hat{\theta}, \theta)\Big|_{\hat{\theta}=\theta} \geq 0 \qquad \Longrightarrow \qquad \dot{s}(\theta)\,Q_{s\theta}(\theta, \theta) \geq 0 \qquad \forall\theta \geq \underline{\theta}, \tag{9}$$

plus the participation constraint:

$$u(\theta) \geq u_0 \qquad \forall\theta \geq \underline{\theta}. \tag{10}$$

These conditions ensure that (locally) the cooperating accomplice has no incentive to manipulate his information and that he prefers to join the program rather than facing the trial.[12] As standard, an envelope argument allows to rewrite the first-order incentive compatibility constraint as:

$$\dot{u}(\theta) = Q_\theta(s(\theta), \theta) \qquad \forall \theta \geq \underline{\theta}. \tag{11}$$

Hence, under **A1** the information rent is increasing — i.e., $\dot{u}(\theta) > 0$. Agents with better information have an incentive to mimic downward because the risk of retaliation is higher in worst states of nature — i.e., the probability $1 - Q(s(\theta), \theta)$ is decreasing in θ — and thus they request a more generous amnesty in exchange of a testimony. This induces the agent to under-report in order to enjoy lighter (expected) sanctions than it would be necessary from the Legislator's point of view.

Integrating equation (11) we have:

$$u(\theta) = u(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} Q_\theta(s(x), x)\, dx \qquad \forall \theta \geq \underline{\theta}. \tag{12}$$

An important point to note is that the rent increases with the testimony $s(x)$. A more precise testimony amplyfies the incentive to mimic downward for the same reasons of an higher θ. Moreover, a tighter information floor — i.e., a larger $\underline{\theta}$ — also stifles the information rent: when the access to the program is rationed there are fewer mimicking possibilities.

Finally, for all types that are excluded from the program the following rationing constraint must hold:

$$u_0 \geq \max_{\hat{\theta} \geq \underline{\theta}} u(\hat{\theta}, \theta) \qquad \forall \theta < \underline{\theta}, \tag{13}$$

that is, rationed types must prefer facing the trial instead of mimicking those who can access the program.

We can then turn to the boss' and Legislator's optimization problems. As before, the boss will commit the crime if and only if its return exceeds the expected costs — i.e., the sum of the expected sanction $C(\wp)$ and the agent's break-even wage $w(\wp)$. The lower-bound $\underline{\pi}(\wp)$ above which the crime will be committed can again be derived going through the same analysis leading to equations (3)-(5). This procedure directly yields:

$$\underline{\pi}(\wp) \equiv \int_0^{\underline{\theta}} (q(\theta) S_p - u_0) f(\theta)\, d\theta + \int_{\underline{\theta}}^{\overline{\theta}} (Q(s(\theta), \theta) S_p - u(\theta)) f(\theta)\, d\theta,$$

hence, the Legislator's program can be rewritten as:

$$\max_{\underline{\theta},\ s(.),\ u(.)} \Pr(\pi \leq \underline{\pi}(\wp)),$$

$$\text{subject to } \ (9), (10), (12), (13).$$

Neglecting the second-order local incentive constraint (9) and the rationing constraint (13) (which will be both verified ex-post), inserting (12) into the maximand and integrating by parts, one gets the following

[12] We shall verify in the Appendix that, when these conditions hold, mimicking is unprofitable also globally.

relaxed program $\mathcal{P}$:

$$\max_{\underline{\theta},\ s(.)} \left\{ \int_0^{\underline{\theta}} (q(\theta) S_p - u_0) f(\theta) \, d\theta + \int_{\underline{\theta}}^{\overline{\theta}} \left(Q(s(\theta), \theta) S_p - u(\underline{\theta}) - Q_\theta(s(\theta), \theta) \frac{1 - F(\theta)}{f(\theta)} \right) f(\theta) \, d\theta \right\},$$

$$\text{subject to } u(\underline{\theta}) \geq u_0.$$

In the analysis that follows we show that the solution of $\mathcal{P}$ yields an optimal policy. Before turning to the result, though, it is worthwhile describing the key forces at play. What are the trade-offs at stake when the inability to verify the agent's insider information creates scope for manipulation? How does the second-best policy change with respect to the first-best?

The key difference between the case of asymmetric information and the complete information benchmark rests on a very simple fact. In order to elicit truthful information revelation the Legislator needs to give up an information rent that generates a positive externality on the boss' expected profit: excessive rents granted by the Legislator ex post, translate onto lower wages that the boss has to pay to the agent ex ante, thus making the crime more profitable everything else being kept constant. By the same token, limiting the subset of types eligible for the program — i.e., a tighter floor $\underline{\theta}$ — also stifles the boss' crime return. This restriction, however, comes at a cost: excluding agents from the program generates a positive externalities on the boss's expected utility as long as the information of these excluded types is very productive — i.e., if $Q(s(\theta), \theta) - q(\theta)$ is not negligible.

The next proposition shows that, taken together, these effects force the Legislator to: (i) distort the testimony required to whistleblowers, and (ii) ration the access to the program in the attempt to minimize the information rent and the positive externality on the boss's expected profit.

Proposition 2 ***(Second-best policy)*** *Assume **A1-A3**, then the second-best policy $\wp^{sb}$ features the following properties:*

- ***(Rationing)*** *There exists a lower bound $\underline{\theta}^{sb} > 0$ such that all types θ above $\underline{\theta}^{sb}$ are admitted into the program and prefer to talk in equilibrium, all types below $\underline{\theta}^{sb}$ prefer to opt-out and face the trial. The bound $\underline{\theta}^{sb}$ is determined by the following condition:*

$$(Q(s^{sb}(\underline{\theta}^{sb}), \underline{\theta}^{sb}) - q(\underline{\theta}^{sb}))S_p - Q_\theta(s^{sb}(\underline{\theta}^{sb}), \underline{\theta}^{sb}) \frac{1 - F(\underline{\theta}^{sb})}{f(\underline{\theta}^{sb})} = 0, \tag{14}$$

- ***(Partial disclosure)*** *All agents admitted into the program provide a downward distorted testimony — i.e., $s^{sb}(\theta) \leq \theta$ with equality only at $\theta = \overline{\theta}$ and $s^{sb}(\theta)$ being the solution of*

$$Q_s(s^{sb}(\theta), \theta) S_p - \frac{1 - F(\theta)}{f(\theta)} Q_{s\theta}(s^{sb}(\theta), \theta) = 0, \tag{15}$$

with $\dot{s}^{sb}(\theta) > 0$.

- ***(Excessive amnesty)*** *The second-best amnesty* $\phi^{sb}(\theta)$ *is larger than the first-best one in every state* θ *— i.e.,*

$$\phi^{sb}(\theta) = \underbrace{1 - p + (1 - Q(s^{sb}(\theta), \theta)\frac{1}{S_a}}_{\textit{Zero Rent Amnesty}} + \underbrace{\frac{1}{S_a}\int_{\underline{\theta}^{sb}}^{\theta} Q_\theta(s^{sb}(x), x)dx}_{\textit{Bonus}} > \phi^{fb}(\theta), \tag{16}$$

with $\dot{\phi}^{sb}(\theta) < 0$.

The second-best policy trades off the social costs and benefits of a leninecy program. The information floor $\underline{\theta}^{sb}$ is determined so as to account for the rent-effect that asymmetric information adds to the entry process into the program. A smaller support of types admitted into the program — i.e., a higher $\underline{\theta}^{sb}$ — stifles the agent's mimicking possibilities, whereby reducing his ex post information rent. This rent-reduction effect due to rationing translates onto the boss' expected utility: lower ex post rents for the agent mean higher expected wages and thus higher costs for the boss. On the other side, however, a smaller support of types also stifles the boss' risk of prosecution wherebe reducing his expected profits. On the balance, the second-best policy calls for stricter eligibility criteria relative to the first-best one. Interestingly, by creating less conflict between the boss and the agent, stricter eligibility criteria increase the wage that the former has to pay to the latter, whereby stifling the equilibrium crime rate. The same type of intuition also explains why the second-best policy does not feature full disclosure: to limit mimicking opportunities, and the implied information rents, the Legislator is forced to require downward distorted testimonies.

Finally, note that the amnesty rate will be set so as to satisfy the local incentive compatibility constraint (8) and to make the marginal type $\underline{\theta}^{sb}$ indifferent between talking or facing the trial. This leads to a second-best amnesty that, besides the zero utility level characterizing the complete information benchmark, also grants a bonus increasing with the quality of information provided by the agent — i.e., increasing in θ. Overall, however, the optimal amnesty is still decreasing in θ because a cooperating accomplice with worse information faces a higher likelihood of retaliation and need to get compensated for such extra risk.

4 When the Boss can Whistle

This section provides an extension of the baseline model that analyzes the case where the benefit of an amnesty is also awarded to a self-reporting boss. Up to now, we have only considered a policy that grants an amnesty to the agent. What would happen if this benefit is also awarded to the boss? Would this policy be optimal? If so, under what conditions?

The historical evidence supports the idea that sometimes even leaders of criminal organizations decide to cooperate with the justice by cheating their relatives, former allies and 'employees'.[13] The collaboration

[13] For instance, in 1996, Giovanni Brusca, one of the most powerful leaders of the corleonesi family and self-confessed multiple murderer (e.g., he was convicted for the bombings that killed judges Chinnici and Falcone), started his cooperation

of these leaders, often charged with several life sentences, usually meets with political scepticism and outrage among relatives of victims of Mafia-related crimes. Is it really necessary to promise judicial leniency to mobsters that have planned and ordered numerous murders and other criminal activities?

In this section we modify the baseline model of hierarchical criminal organizations to encompass this possibility. The objective is to show that, in the context of our model, dealing directly with a self-confessing boss is sometimes necessary to efficiently fight organized crime.

Suppose that the Legislator grants an amnesty, say Φ, to the boss as a reward for his cooperation (self-reporting). The structure of the game is similar to that analyzed before with the following modifications:

- In period **t=0** the Legislator commits to a policy $\wp_\Phi = (\mathcal{L}, \underline{\theta}, \Phi)$ that, as explained before, includes the direct mechanism $\mathcal{L} = \left\{\phi(\hat{\theta}), s(\hat{\theta})\right\}_{\hat{\theta}\in\Theta}$ with the floor $\underline{\theta}$, and in addition also awards a discount Φ to the self-reporting boss.
- In period **t=2**, after the random variable $\tilde{\theta}$ has realized, the boss can self-report. If so, he enjoys a discount Φ on the sanction S_p, while the agent is convicted with probability 1.
- In period **t=3** the agent can whistle and enter the program if and only if the boss has not already self-reported at **t=2**.

Hence, when the boss self-reports, the agent cannot avoid a conviction. This seems reasonable to the extent that the information provided by a self-confessing boss is more reliable than the agent's imperfect testimony. The hypothesis that boss self-reports before the agent is also with no loss of insights. Indeed, since prosecutors will typically value more a confession of a 'big-fish', which leads to a conviction with probability 1 of the whole organization, than the imperfect evidence provided by a simple 'soldier', letting the boss and the agent self-report at the same time would not change our qualitative results but only complicate the equilibrium description.[14]

We shall retain the same assumptions made in the baseline model — i.e., **A1-A3**. Moreover, to capture the idea that the boss may be more reluctant to talk than the agent, we denote by $\delta > 0$ the additional cost that he bears from cheating the organization. This parameter reflects those psychological costs incurred by a mobster that gives up a powerful 'command position', or it could simply capture the reluctance of an individual grown up with the values and codes of the mafia to renege his criminal culture.

Throughout we will make the following hypothesis:

A4 The cost δ is large enough relative to the boss' expected sanction in the absence of leniency — i.e.,

$$\delta > q(\overline{\theta})S_p.$$

with the justice by releasing relevant information that turned out to be crucial for the capture and conviction of many of his former partners, among which the powerful 'Godfather' Bernardo Provenzano whose hiding lasted for over forty years.

[14] To make things simple we will therefore assume that the self-reporting game is sequential and the boss has a first-mover advantage.

This assumption avoids that the boss might find it convenient to whistle when the agent is not allowed to talk. This seems perfectly in line with the historical evidence.[15]

Before characterizing formally the optimal policy it is worthwhile discussing how the equilibrium of the game looks like in this new setting. To make the problem interesting, we will focus on equilibria where, to improve efficiency, the Legislator grants a positive amnesty to the boss (as otherwise the result of Proposition 2 would emerge). Then, two extra features characterize the on- and off-equilibrium strategies under self-reporting. First, since the boss decides to self-report upon observing the realization of the state of nature θ, the equilibrium description must specify the subset of states of nature where this occurs. Second, the 'off-equilibrium' actions and policy matter to sustain equilibria with self-reporting. To understand this point, suppose that the boss is expected to talk in state θ, and assume that he deviates by not talking. Then what happens in the continuation game following such an unexpected action? What would the agent do in this unexpected contingency?

Under **A4** this deviation is unprofitable as long as the agent talks in state θ and the policy $\wp_\Phi$ is such that:

$$Q\left(s\left(\theta\right),\theta\right)S_p \geq \left(1-\Phi\right)S_p+\delta, \tag{17}$$

implying that the boss' expected sanction in the deviation exceeds the utility from self-reporting.[16]

We shall look for a cut-off equilibrium where: in some states neither the agent nor the boss talk, in some other states only the agent talks, and in the remaining states the boss self-reports. In order to describe such an equilibrium two relevant thresholds need to be characterized: (i) $\underline{\theta}>0$ below which no agent is admitted into the program (precisely as before), and (ii) $\underline{\theta}\left(\Phi\right)$ above which the boss self-reports.

Given the policy $\wp_\Phi=(\mathcal{L},\underline{\theta},\Phi)$ and a state $\theta>\underline{\theta}$, the boss decides to self-report if and only if inequality (17) holds. Consider any incentive compatible policy specifying a disclosure rule $s\left(\theta\right)$, such that $s\left(\theta\right)\leq\theta$ and $\dot{s}\left(\theta\right)\geq 0$. Denote by $\underline{\theta}\left(\Phi\right)$ the solution to (17) taken as an equality with respect to θ, then in all states above this cutoff — i.e., $\theta\geq\underline{\theta}\left(\Phi\right)$ — the boss will find it convenient to self-report. Of course, if δ is large enough this will never happen — i.e., the solution $\underline{\theta}\left(\Phi\right)$ lies outside the support Θ. Hence, the next lemma follows:

Lemma 3 *Suppose that $\delta>Q(\overline{\theta},\overline{\theta})S_p$, then the boss will never talk in equilibrium and the optimal policy is the one characterized in Proposition 2.*

The intuition for this result is straightforward. If the boss never finds it convenient to self-report, because the cost δ is very large, the optimal policy is the one characterized in the baseline model: only the agent talks in equilibrium. This is, for instance, the case of organizations such as the 'Ndrangheta, where the leadership is often inherited on a 'blood relationship' basis. But, it seems not to be true

[15] In Italy, for instance, the earliest whistleblowers were simple soldiers, even Buscetta (the first important 'pentito') never reached the status of leader within the organization, it was only a few years after the introduction of the accomplice-witnesses program that the first important bosses started their cooperation — e.g., Giovanni Brusca and Giuseppe Di Cristina in Sicily, Carmine Alfieri and Domenico Bidognetti in Campania and Francesco Fonti in Calabria (Falcone, 1991).

[16] Also in this case we assume the tie-breaking condition that the boss prefers to talk whenever indifferent between cooperating and facing the uncertainty of the trial.

for the Mafia and the Camorra: for these organizations the command positions are taken over typically after bloody interior fights. To make the problem interesting, from now on we will impose the following assumption:

A5 The cost of whistling δ is such that the boss talks at least in some states:

$$Q(\overline{\theta},\overline{\theta})S_p > \delta. \tag{18}$$

Intuitively, **A5** implies that for Φ sufficiently close to 1 the boss is going to talk in state $\overline{\theta}$ and in its neighborhood.[17] Finally, to guarantee uniqueness of the optimal policy we also make the next additional technical requirement:

A6 The function $Q(\theta,\theta)$ is strictly concave in θ and satisfies the Inada condition $Q_\theta(0,0) = +\infty$.[18]

Assuming that, when the agent tells the truth, the probability of convicting the boss $Q(\theta,\theta)$ exhibits 'decreasing marginal returns' of information seems a mild and reasonable assumption.

We now begin the analysis with the following preliminary result.

Lemma 4 *Under **A4-A5** for any disclosure rule $s(\theta)$, such that (i) $0 \leq s(\theta) \leq \theta$, (ii) $s(0) = 0$ and $s(\overline{\theta}) = \overline{\theta}$, and (iii) $\dot{s}(\theta) \geq 0$, there exits an upper-bound $\overline{\Phi} < 1$ and a lower-bound $\underline{\Phi} < \overline{\Phi}$ such that:*

- *for all $\Phi < \underline{\Phi}$ the boss never talks;*
- *for all $\Phi > \overline{\Phi}$ the boss always talks;*
- *for every $\Phi \in (\underline{\Phi},\overline{\Phi})$ there exists a cut-off $\underline{\theta}(\Phi) > 0$ such that the boss talks for all $\theta \geq \underline{\theta}(\Phi)$, while he does not talk for $\theta < \underline{\theta}(\Phi)$.*

The intuition for this result is straightforward. Under **A5** a too generous amnesty — i.e., Φ larger than $\overline{\Phi}$ — incentivizes the boss to plea guilty and cheat his fellow, while a too restrictive policy — i.e., Φ smaller than $\underline{\Phi}$ — discourages self-reporting. Hence, for intermediate values of Φ there is a subset of realizations of $\tilde{\theta}$ where the boss does not talk and a ragion where he talks.

As a preliminary benchmark let us briefly illustrate the first-best policy when the boss is allowed to self-report. Under **A5** the analysis of the first-best policy changes with respect to the one illustrated in Section 4.1:

Proposition 5 ***(First-best with self-reporting)*** *Assume **A1-A6**, then the first-best policy $\wp^{fb}_{\Phi}$ features the same properties as in Proposition 1 — i.e., no rationing and full disclosure. Moreover, under self-reporting the following is also true:*

[17] Of course, this is only a sufficient condition and is made only for simplicity. More generally, for any given δ, there exists a Φ sufficiently large such that the equivalent of (18) holds.

[18] The Inada condition above allows to safely focus on corner solutions, and can be easily relaxed by $Q_\theta(0,0) = K$ with K finite but large enough.

- ***(Self-reporting)*** *The boss cooperates for all* $\theta \geq \underline{\theta}(\Phi^{fb})$*, with* $0 < \underline{\theta}(\Phi^{fb}) < \overline{\theta}$*, where the threshold* $\underline{\theta}(\Phi^{fb})$ *and the optimal amnesty* $\Phi^{fb} \in (0,1)$ *solve:*

$$Q(\underline{\theta}(\Phi^{fb}), \underline{\theta}(\Phi^{fb}))S_p = (1-\Phi^{fb})S_p + \delta, \tag{19}$$

$$\underbrace{\frac{(1-p)S_a}{Q_\theta(\underline{\theta}(\Phi^{fb}), \underline{\theta}(\Phi^{fb}))}}_{\textbf{\textit{Domino effect}}} = \underbrace{\frac{1-F(\underline{\theta}(\Phi^{fb}))}{f(\underline{\theta}(\Phi^{fb}))}S_p}_{\textbf{\textit{Crime Enhancing effect}}} \quad . \tag{20}$$

Hence, in equilibrium, the agent cooperates only in the states of nature $\theta \leq \underline{\theta}(\Phi^{fb})$*.*

Hence, allowing the boss to self-report is socially beneficial even under complete information. However, the reason is completely different from the traditional result in Kaplow and Shavell (1994), where self-reporting saves enforcement resources because individuals who report their harmful acts need not be detected, and it reduces risk because individuals who report their behavior bear certain rather than uncertain sanctions. The interpretation is as follows: as long as the boss is willing to self-report, the Legislator will rely on a policy that awards him an amnesty Φ^{fb} that tades off the costs and benefits of self-reporting on the ex ante crime rate. First, when the boss talks, the agent is convicted with certainty as opposed to the case where the trial opens and the agent ends up being convicted only with probability p. This reflects a vertical externality, which we call *domino effect*, that spurs the agent's conviction risk ex ante and translates onto a higher reservation wage that, in turn, reduces the crime rate.[19] Second, there is a cost associated with the recognition of an amnesty to the boss. This cost stems from the simple fact that in all states larger than $\underline{\theta}(\Phi^{fb})$ the boss enjoys a reduced sanction. Of course, this weakens deterrence and enhances the ex ante crime return whereby reducing welfare: a *crime enhancing* effect.

We can now turn to study the case of asymmetric information where the choice of the boss' amnesty will be also affected by the rent that the agent obtains in equilibrium. As a first step we define the incentive feasible set. It is easy to verify that the participation, rationing and (local) incentive compatibility constraints are as before. We assume up-front, and verify ex post, that $\overline{\theta} > \underline{\theta}(\Phi) > \underline{\theta}$; hence, the boss' expected sanction can be written as:

$$C(\wp_\Phi) \equiv \int_0^{\underline{\theta}} q(\theta) S_p f(\theta)\, d\theta + \int_{\underline{\theta}}^{\underline{\theta}(\Phi)} Q(s(\theta), \theta) S_p f(\theta)\, d\theta + \int_{\underline{\theta}(\Phi)}^{\overline{\theta}} ((1-\Phi) S_p + \delta) f(\theta)\, d\theta.$$

The agent's break-even wage is instead defined by the participation constraint:

$$w(\wp_\Phi) + \int_0^{\underline{\theta}} u_0 f(\theta)\, d\theta + \int_{\underline{\theta}}^{\underline{\theta}(\Phi)} u(\theta) f(\theta)\, d\theta + \int_{\underline{\theta}(\Phi)}^{\overline{\theta}} S_a f(\theta)\, d\theta = 0,$$

[19] This domino effect echoes Baccara and Bar-Isaac (2005).

Hence, following the same procedure as before, we have:

$$\underline{\pi}\left(\wp_{\Phi}\right) \equiv \int_{0}^{\underline{\theta}} (q\left(\theta\right) S_p - u_0) f\left(\theta\right) d\theta + \int_{\underline{\theta}}^{\underline{\theta}(\Phi)} \left(Q\left(s\left(\theta\right),\theta\right) S_p - u\left(\theta\right)\right) f\left(\theta\right) d\theta +$$
$$+ \int_{\underline{\theta}(\Phi)}^{\overline{\theta}} \left(S_a + (1-\Phi) S_p + \delta\right) f\left(\theta\right) d\theta.$$

The Legislator's program is then:

$$\max_{\underline{\theta},\ s(.),\ u(.),\ \Phi} \Pr\left(\pi \leq \underline{\pi}\left(\wp_{\Phi}\right)\right),$$

$$\text{subject to } \ (9),\ (10),\ (12),\ (13),\ \phi\left(\theta\right) \in [0,1] \text{ and } s(\theta) \leq \theta.$$

Neglecting the second-order local incentive constraint (9) and the rationing constraint (13) (which will be both verified ex-post), inserting (12) into the maximand and integrating by parts, one gets the following relaxed program $\mathcal{P}_{\Phi}$:

$$\max_{\underline{\theta},\ s(.),\ \Phi} \left\{ \int_{0}^{\underline{\theta}} (q\left(\theta\right) S_p - u_0) f\left(\theta\right) d\theta + \int_{\underline{\theta}(\Phi)}^{\overline{\theta}} \left(S_a + (1-\Phi) S_p + \delta\right) f\left(\theta\right) d\theta + \right.$$
$$\left. + \int_{\underline{\theta}}^{\underline{\theta}(\Phi)} \left(Q\left(s\left(\theta\right),\theta\right) S_p - u\left(\underline{\theta}\right) - Q_{\theta}\left(s\left(\theta\right),\theta\right) \frac{F(\underline{\theta}\left(\Phi\right)) - F(\theta)}{f(\theta)} \right) f\left(\theta\right) d\theta \right\}.$$

$$\text{subject to } u\left(\underline{\theta}\right) \geq u_0,\ \Phi \in [0,1] \text{ and } s\left(\theta\right) \leq \theta.$$

Therefore, for an equilibrium where the boss talks in some states while the agent talks in some other states to exist, the policy must solve $\mathcal{P}_{\Phi}$. The main result of this section then follows.

Proposition 6 ***(Second-best with self-reporting)*** *Assume* ***A1-A6****, then the optimal policy features the following properties:*

- ***(Rationing)*** *There exists a lower bound $\underline{\theta}_{\Phi}^{sb} > 0$ such that all types θ above $\underline{\theta}_{\Phi}^{sb}$ are admitted into the program and in equilibrium prefer to talk, all types below $\underline{\theta}_{\Phi}^{sb}$ prefer to opt-out. The cut-off $\underline{\theta}_{\Phi}^{sb}$ is determined by the following condition:*

$$(Q(s_{\Phi}^{sb}(\underline{\theta}_{\Phi}^{sb}),\underline{\theta}_{\Phi}^{sb}) - q(\underline{\theta}_{\Phi}^{sb}))S_p - Q_{\theta}(s_{\Phi}^{sb}(\underline{\theta}_{\Phi}^{sb}),\underline{\theta}_{\Phi}^{sb}) \frac{F(\underline{\theta}(\Phi^{sb})) - F(\underline{\theta}_{\Phi}^{sb})}{f(\underline{\theta}_{\Phi}^{sb})} = 0, \tag{21}$$

- ***(Partial disclosure)*** *For all states $\theta \in [\underline{\theta}_{\Phi}^{sb}, \underline{\theta}(\Phi^{sb}))$ the agent talks in equilibrium and must provide a testimony $s_{\Phi}^{sb}\left(\theta\right)$ solving the following equation:*

$$Q_s(s_{\Phi}^{sb}\left(\theta\right),\theta)S_p - Q_{s\theta}(s_{\Phi}^{sb}\left(\theta\right),\theta) \frac{F(\underline{\theta}(\Phi^{sb})) - F\left(\theta\right)}{f\left(\theta\right)} = 0, \tag{22}$$

with $s_{\Phi}^{sb}(\theta) \leq \theta$ with equality only at $\theta = \underline{\theta}(\Phi^{sb})$ and $\dot{s}_{\Phi}^{sb}(\theta) > 0$.

- ***(Excessive self-reporting)*** *In equilibrium the boss cooperates if $\theta \geq \underline{\theta}(\Phi^{sb})$, with $\underline{\theta}_{\Phi}^{sb} < \underline{\theta}(\Phi^{sb}) < \overline{\theta}$. The threshold $\underline{\theta}(\Phi^{sb})$ and the optimal amnesty Φ^{sb} solve:*

$$Q(\underline{\theta}(\Phi^{sb}), \underline{\theta}(\Phi^{sb}))S_p = (1 - \Phi^{sb})S_p + \delta, \tag{23}$$

$$\underbrace{\frac{(1-p)S_a}{Q_\theta(\underline{\theta}(\Phi^{sb}), \underline{\theta}(\Phi^{sb}))}}_{\textit{Domino effect}} + \underbrace{\frac{\int_{\underline{\theta}_{\Phi}^{sb}}^{\underline{\theta}(\Phi^{sb})} Q_\theta(s_{\Phi}^{sb}(\theta), \theta)d\theta}{Q_\theta(\underline{\theta}(\Phi^{sb}), \underline{\theta}(\Phi^{sb}))}}_{\textit{Rent Saving effect}} = \underbrace{\frac{1 - F(\underline{\theta}(\Phi^{sb}))}{f(\underline{\theta}(\Phi^{sb}))}S_p}_{\textit{Crime Enhancing effect}} \ . \tag{24}$$

Moreover, under asymmetric information the boss self-reports more often than in the complete information benchmark — i.e., $\underline{\theta}(\Phi^{sb}) < \underline{\theta}(\Phi^{fb})$ and $\Phi^{sb} > \Phi^{fb}$.

- ***(Excessive amnesty)*** *The amnesty $\phi_{\Phi}^{sb}(\theta)$ is larger than the first-best level — i.e.,*

$$\phi_{\Phi}^{sb}(\theta) = \underbrace{1 - p + (1 - Q(s_{\Phi}^{sb}(\theta), \theta))\frac{1}{S_a}}_{\textit{Zero Rent Amnesty}} + \underbrace{\frac{1}{S_a}\int_{\underline{\theta}_{\Phi}^{sb}}^{\theta} Q_\theta(s_{\Phi}^{sb}(x), x)dx}_{\textit{Bonus}} > \phi_{\Phi}^{fb}(\theta), \tag{25}$$

with $\dot{\phi}_{\Phi}^{sb}(\theta) < 0$.

- ***(Off-equilibrium allocation)*** *In all states $\theta > \underline{\theta}\left(\Phi^{sb}\right)$ the policy is such that an accomplice gets the truthful disclosure rule $\tilde{s}(\theta) = \theta$ and a constant amnesty $\tilde{\phi} = \phi_{\Phi}^{sb}\left(\underline{\theta}(\Phi^{sb})\right)$ if the boss has not self reported.*

There is one novel force shaping the second best amnesty Φ^{sb} in addition to the *domino* and the *crime enhancing* effects emphasized in Proposition 5. Essentially, granting an amnesty to the boss has a beneficial *rent saving effect* that goes through the incentive constraints. The logic is similar to that explaining the rationing result. An higher amnesty for the boss increases the states of nature where the boss decides to self-report, whereby decreasing the measure of agents admitted into the program — recall that an agent can whistle if and only if the boss has not already done so. This stifles the agent's mimicking possibilities, whereby reducing his ex post information rent. This rent-reduction effect reinforces the domino effect whereby leading the boss to self-report more often than in the complete information case.

The second-best policy under self-reporting also differs from the policy characterized in Proposition 2. Since asymmetric information introduces mimicking opportunities that lead the Legislator to require biased testimonies, and these distortions are positively linked with the measure of agents that talk in equilibrium, granting an amnesty to the self-reporting boss allows to mitigate the rent-efficiency trade-off:

Corollary 7 *Assume **A1-A7**, then $s_{\Phi}^{sb}(\theta) > s^{sb}(\theta)$ for all $\theta \in [\underline{\theta}_{\Phi}^{sb}, \underline{\theta}(\Phi^{sb})]$ — i.e., in equilibrium, cooperating accomplices provide better testimonies when the boss is allowed to self-report relative to the case* *where he is not. Moreover, $\underline{\theta}_{\Phi}^{sb} < \underline{\theta}^{sb}$ — i.e., under self-reporting the program becomes less requiring in terms of the minimal amount of information needed to access to it. The effect of self-reporting on the amnesty is ambiguous.*

When the boss can self-report there is less need to distort the agent's testimony simply because in equilibrium there will be a lower fraction of agents that whistle — as reflected by the modified distortion in equation (22). Hence, the Legislator needs to waste less rents to elicit truthful information revelation, and this allows more precise testimonies. The same reasoning implies that under self-reporting there is also less need for rationing.

The reason why the effect of self-reporting on the amnesty is ambiguous is due to the presence of two countervailing effects. On the one hand, the bonus under self-protection increases because of both better testimonies (higher s) and less rationing (lower $\underline{\theta}$). On the other hand, better testimonies also reduce the retaliation risk faced by an agent entering the program, which in turn decreases the need for compensating this risk.

5 Concluding remarks

This paper has studied the problem of a policy maker designing immunity for accomplice-witnesses. We have focused on a pyramidal criminal organization to capture in the easiest possible way the basic trade-offs emerging when the efficacy of an accomplice-witnesses program is undermined by informants' private information, a point never made before in the earlier literature. In so doing, we have identified the forces that may lead those agents to release distorted information, so as to characterize the second best policy preventing untruthful information revelation.

Our results are consistent with the legislative provisions showing that accomplices must fulfill minimal information requirements to be admitted into the program, that a bonus should be granted to those who provide more productive information and that allowing also the boss to whistle is socially beneficial.

Some simplifying hypothesis have been made through the analysis. The model could be extended to encompass horizontal criminal organizations as well as instances where there are spillovers between the information disclosed by different accomplices. For simplicity, we have decided to treat these interesting topics in some future research.

6 Appendix

Proof of Proposition 1: Differentiating the objective function of the first-best program with respect to $s(.)$ and $\underline{\theta}$ respectively we have:

$$\frac{\partial \underline{\pi}(\wp)}{\partial s} = Q_s(s(\theta), \theta),$$

$$\frac{\partial \underline{\pi}(\wp)}{\partial \underline{\theta}} = (Q(\underline{\theta}^{fb}, \underline{\theta}^{fb}) - q(\underline{\theta}^{fb}))S_p.$$

Under **A1** these equations immediately imply $s^{fb}(\theta) = \theta$ and $\underline{\theta}^{fb} = 0$ — i.e., full disclosure and no rationing. Moreover, the Legislator will induce agents to apply to the program by granting the reservation amnesty $\phi(.)$ that satisfies the participation constraint as equality — i.e.,

$$(1 - \phi^{fb}(\theta))S_a + (1 - Q(\theta, \theta)) = pS_a,$$

which immediately implies that $\dot{\phi}^{fb}(\theta) < 0$ because $Q(\theta, \theta)$ is increasing in θ. ■

Incentive feasible allocations: The characterization of the incentive compatibility constraints is standard, see for instance Laffont and Martimort (2002, Ch., 3). Equation (8) is straightfoward, while (9) comes from the usual total differentiation technique which implies $u_{\hat{\theta}\hat{\theta}}(\theta, \theta) \leq 0$ and thus $u_{\hat{\theta}\theta}(\theta, \theta) \geq 0$. Finally, the expression for $\dot{u}(\theta)$ follows immediately from (8) together with an application of the Envelope Theorem; **A1** also implies that this rent is positive. ■

Proof of Proposition 2: First, since the objective of $\mathcal{P}$ is decreasing in $u(\underline{\theta})$ at the optimum one necessarily has $u(\underline{\theta}) = u_0$. The Legislator's (relaxed) optimization program then becomes:

$$\max_{\underline{\theta},\ s(.)} \left\{ \int_0^{\underline{\theta}} (q(\theta) S_p - u_0) f(\theta) d\theta + \int_{\underline{\theta}}^{\overline{\theta}} \left(Q(s(\theta), \theta) S_p - u_0 - Q_\theta(s(\theta), \theta) \frac{1 - F(\theta)}{f(\theta)} \right) f(\theta) d\theta \right\}.$$

Optimizing pointwisely with respect to $s(\theta)$ one gets immediately the first-order condition (15), which directly implies $s^{sb}(\theta) \leq \theta$ for all θ with equality only at $\overline{\theta}$ by **A1**. Moreover, optimizing with respect to $\underline{\theta}$ one has the first-order condition (14). Given the pair $(s^{sb}(\theta), \underline{\theta}^{sb})$, the second-best amnesty has to satisfy two requirements: (i) it has to ensure that the agent's incentive compatibility constraint is met — i.e., it must satisfy (8) evaluated at $s^{sb}(\theta)$; (ii) and its lower bound $\phi^{sb}\left(\underline{\theta}^{sb}\right)$ must be such that the cut-off type $\underline{\theta}^{sb}$ has to be indifferent between entering the program and facing the trial (as reflected by the participation constraint $u(\underline{\theta}^{sb}) = u_0$). From (12) one then has:

$$-(1 - \phi^{sb}(\theta))S_a - (1 - Q(s^{sb}(\theta), \theta)) = u_0 + \int_{\underline{\theta}^{sb}}^{\theta} Q_\theta(s^{sb}(x), x) dx,$$

hence, using $u_0 = -pS_a$ one gets:

$$\phi^{sb}(\theta) = 1 - p + (1 - Q(s^{sb}(\theta), \theta))\frac{1}{S_a} + \frac{1}{S_a}\int_{\underline{\theta}^{sb}}^{\theta} Q_\theta(s^{sb}(x), x)dx.$$

which immediately implies that in state $\underline{\theta}^{sb}$ the agent is indifferent between entering the program and facing the trial. Showing that $\dot{\phi}^{sb}(\theta)$ follows immediately from (8).

We now prove that the first-order necessary conditions (14) and (15) are also sufficient for an optimum by showing that the objective of the Legislator's relaxed program $\mathcal{P}$ is strictly concave under **A1**-**A3**. To begin with, observe that for any given $\underline{\theta}$ the objective of $\mathcal{P}$, let us denote it $W(.)$ with a little abuse of notation, is strictly concave in $s(.)$ — i.e.,

$$\frac{\partial^2 W(.)}{\partial s(.)^2} = Q_{ss}(s^{sb}(\theta), \theta)S_p - Q_{ss\theta}(s^{sb}(\theta), \theta)\frac{1 - F(\theta)}{f(\theta)} < 0.$$

Now, differentiating twice with respect to $\underline{\theta}$ the objective $W(.)$ evaluated at the disclosure rule $s^{sb}(\theta)$ that solves (15) one has:

$$\frac{\partial^2 W(.)}{\partial \underline{\theta}^2} = \left\{-Q_\theta(.) - \dot{s}^{sb}(.)\left(Q_s(.)S_p - Q_{\theta s}(.)\frac{1 - F(\theta)}{f(\theta)}\right) + \right.$$
$$\left. + \left(Q_{\theta\theta}(.)\frac{1 - F(\theta)}{f(\theta)} + Q_\theta(.)\frac{\partial}{\partial\theta}\left(\frac{1 - F(\theta)}{f(\theta)}\right)\right)\frac{1}{S_p}\right\}_{\theta = \underline{\theta}^{sb}},$$

which by (15) implies:

$$\frac{\partial^2 W(.)}{\partial \underline{\theta}^2} = \left\{-Q_\theta(.) + \left(Q_{\theta\theta}(.)\frac{1 - F(\theta)}{f(\theta)} + Q_\theta(.)\frac{\partial}{\partial\theta}\left(\frac{1 - F(\theta)}{f(\theta)}\right)\right)\frac{1}{S_p}\right\}_{\theta = \underline{\theta}^{sb}}.$$

A2 then implies that $\frac{\partial^2 W(.)}{\partial \underline{\theta}^2} < 0$ since $Q_\theta(.) > 0$ and $Q_{\theta\theta}(.) \leq 0$. Moreover, in order to show that $\underline{\theta}^{sb}$ is in the interior of Θ, observe that, under **A1,** the left-hand side of (14) is positive, increasing and null at $\theta = 0$. In order to establish the optimality of setting a floor $\underline{\theta}^{sb} \in (0, \overline{\theta})$ it is then enough to verify that the right-hand side of (14) — i.e.,

$$\Omega(\underline{\theta}^{sb}) \equiv Q_\theta(s^{sb}(\underline{\theta}^{sb}), \underline{\theta}^{sb})\frac{1 - F(\underline{\theta}^{sb})}{f(\underline{\theta}^{sb})S_p},$$

satisfies the following conditions: (i) $\Omega(0) > 0$, (ii) $\Omega(\overline{\theta}) < Q(\overline{\theta}, \overline{\theta}) - q(\overline{\theta})$, and (iii) $\Omega(\theta)$ being continuous. Showing that $\Omega(0) > 0$ is immediate since $\lim_{\theta\to 0}\frac{1-F(\theta)}{f(\theta)} = 1/f(0) > 0$. Moreover, showing that $\Omega(\overline{\theta}) < Q(\overline{\theta}, \overline{\theta}) - q(\overline{\theta})$ is simple since $\lim_{\theta\to\overline{\theta}}\frac{1-F(\theta)}{f(\theta)} = 0$ implies $s^{sb}(\overline{\theta}) = \overline{\theta}$ from (15). The fact that $\Omega(\theta)$ must be continuous follows from the hypothesis that the functions $Q(.)$, $q(.)$ and $F(.)$ are twice continuously differentiable.

Finally, we need to verify that the policy characterized by (14), (15) and (16) satisfies (9) as well as the global incentive compatibility constraint and the rationing constraint (13). First, showing that (9) is met under **A1-A3** is simple. Indeed, since $Q_{s\theta}(.) > 0$ by **A1** we only need to show that $\dot{s}^{sb}(\theta) \geq 0$. This is straightforward, using (15) the Implicit Function Theorem implies:

$$\dot{s}^{sb}(\theta) = \frac{Q_{s\theta}(.)\left(S_p - \frac{\partial}{\partial\theta}\left(\frac{1-F(\theta)}{f(\theta)}\right)\right) - Q_{s\theta\theta}(.)\frac{1-F(\theta)}{f(\theta)}}{-Q_{ss}(.)S_p + Q_{ss\theta}(.)\frac{1-F(\theta)}{f(\theta)}},$$

which is positive under **A2** and **A3** —- i.e., $\dot{s}^{sb}(\theta) > 0$.

Second, in order to show that the global incentive compatibility constraint holds we need to verify that in no state $\theta \geq \underline{\theta}^{sb}$ the agent wants to lie by announcing a state $\theta' \geq \underline{\theta}^{sb}$ with $\theta \neq \theta'$ — i.e.,

$$u(\theta,\theta) - u(\theta',\theta) \geq 0 \quad \forall \ (\theta,\theta') \in [\underline{\theta}^{sb},\overline{\theta}]^2,$$

which by definition of $u(\theta',\theta)$ implies

$$(1-\phi^{sb}(\theta'))S_a + (1-Q(s^{sb}(\theta'),\theta)) \geq (1-\phi^{sb}(\theta))S_a + (1-Q(s^{sb}(\theta),\theta)). \tag{26}$$

Assume $\theta' > \theta$, with $s^{sb}(\theta') \leq \theta$, with no loss of generality, then (26) yields:

$$-\int_{\theta}^{\theta'}\left\{\dot{\phi}^{sb}(x)S_a + \dot{s}^{sb}(x)\,Q_s(s^{sb}(x),\theta)\right\}dx \geq 0,$$

using (8) and substituting for $\dot{\phi}^{sb}(x)S_a = -\dot{s}^{sb}(x)\,Q_s(s^{sb}(x),x)$ for $x \geq \theta$ we have:

$$0 \leq -\int_{\theta}^{\theta'}\left\{\dot{s}^{sb}(x)\,Q_s(s^{sb}(x),\theta) - \dot{s}^{sb}(x)\,Q_s(s^{sb}(x),x)\right\}dx =$$

$$= -\int_{\theta}^{\theta'}\left\{\dot{s}^{sb}(x)\int_{x}^{\theta}Q_{s\theta}(s^{sb}(x),y)dy\right\}dx. \tag{27}$$

which immediately implies the result since $\dot{s}^{sb}(x) \geq 0$, $x \geq \theta$ and $Q_{s\theta}(.) \geq 0$. Next, observe that, by definition, an agent in state θ cannot claim to be in state θ' as long as $s^{sb}(\theta') > \theta$.

Showing that no type $\theta \geq \underline{\theta}^{sb}$ can profit by mimicking a type $\theta' < \underline{\theta}^{sb}$ is obvious given the fact that $\dot{u}(\theta) > 0$. Finally, we show that the rationing constraint is satisfied — i.e., in no state $\theta < \underline{\theta}^{sb}$ the agent can profitably lie by announcing a state $\theta' \geq \underline{\theta}^{sb}$. We need to verify that:

$$u_0 = -pS_a \geq u(\theta',\theta) \quad \forall \quad \theta < \underline{\theta}^{sb} \ \text{ and } \ \theta' \geq \underline{\theta}^{sb}. \tag{28}$$

First, observe that by definition of the marginal type $\underline{\theta}^{sb}$ equation (26) can be rewritten as:

$$u(\underline{\theta}^{sb}) \geq u(\theta',\theta), \tag{29}$$

moreover, since $Q_\theta(.) > 0$ and $\underline{\theta}^{sb} > \theta$ it must be $u(\theta', \underline{\theta}^{sb}) > u(\theta', \theta)$. Inequality (29) must then hold as long as the following is true:

$$u(\underline{\theta}^{sb}) \geq u(\theta', \underline{\theta}^{sb}) \quad \forall \;\; \theta' > \underline{\theta}^{sb},$$

which is true precisely by the same argument used to show that the global incentive compatibility constraint holds for all types $(\theta, \theta') \in [\underline{\theta}^{sb}, \theta)^2$. This concludes the proof. ■

Proof of Lemma 3: The proof of this result is straightforward. Even when the maximal amnesty $\Phi = 1$ is granted to the boss, he prefers not to talk because the maximal (expected) sanction from the trial $q(\bar{\theta}, \bar{\theta})S_p$ falls short of his self-reporting cost δ. ■

Proof of Lemma 4: Take any continuously differentiable disclosure rule $s(\theta)$ such that: (i) $s(\theta) \leq \theta$, (ii) $s(\overline{\theta}) = \overline{\theta}$ and $s(0) = 0$, and (iii) $\dot{s}(\theta) \geq 0$. Then, given $s(\theta)$, let $\underline{\theta}(\Phi)$ be the solution in θ of

$$Q(s(\theta), \theta) S_p = (1 - \Phi) S_p + \delta. \tag{30}$$

By assumption $\underline{\theta}(\Phi)$ exists for some $\Phi > 0$ because $Q(s(\theta), \theta)$ is increasing in θ. Moreover, by assumption, we also know that, if it exists, $\underline{\theta}(\Phi)$ is decreasing in Φ — i.e.,

$$\frac{\partial \underline{\theta}(\Phi)}{\partial \Phi} = -\frac{1}{Q_s(s(\theta), \theta)\, \dot{s}(\theta) + Q_\theta(s(\theta), \theta)} < 0.$$

We then need to show that there exists a $\Phi \in (0, 1)$ such that $\underline{\theta}(\Phi) \in (0, \overline{\theta})$. This simply follows from the fact that **A5** implies that at $\Phi = 1$ one has $\underline{\theta}(\Phi) < \overline{\theta}$ — i.e., the boss self-reports for θ large enough. Moreover, **A4** implies that at $\Phi = 1$ equation (30) cannot hold for θ close to 0 and therefore the boss does not self-report in these states. Hence, $\underline{\theta}(\Phi) > 0$. By continuity, this implies that there must exist a non-empty open set $(\underline{\Phi}, \overline{\Phi})$ such that $\underline{\theta}(\Phi) \in (0, \overline{\theta})$ for all $\underline{\Phi} < \Phi < \overline{\Phi}$. ■

Proof of Proposition 5: To begin with, note that in an equilibrium where the boss self-reports — i.e., $\underline{\theta}(\Phi) < \overline{\theta}$ — and the agent talks but is rationed in some states — i.e., $0 < \underline{\theta} < \underline{\theta}(\Phi)$ — the optimal policy must solve the following (relaxed) program:

$$\max_{\underline{\theta},\ s(.),\ \Phi} \left\{ \int_0^{\underline{\theta}} (q(\theta) S_p - u_0) f(\theta)\, d\theta + \int_{\underline{\theta}(\Phi)}^{\overline{\theta}} (S_a + (1-\Phi) S_p + \delta) f(\theta)\, d\theta + \right.$$
$$\left. + \int_{\underline{\theta}}^{\underline{\theta}(\Phi)} (Q(s(\theta), \theta) S_p + p S_a) f(\theta)\, d\theta \right\}.$$

$$\text{subject to } \Phi \in [0, 1] \text{ and } s(\theta) \leq \theta.$$

Differentiating the above objective function with respect to $s(.)$ and $\underline{\theta}$ it is immediate to show that, in an interior solution, the first-best policy with self-reporting features full disclosure and no rationing — i.e., $s_\Phi^{fb}(\theta) = \theta$ for all θ and $\underline{\theta}_\Phi^{fb} = 0$. Differentiating with respect to Φ one obtains the first-order condition:

$$-(S_a + (1-\Phi^{fb})S_p + \delta)\frac{\partial \underline{\theta}(\Phi^{fb})}{\partial \Phi} - \frac{\int_{\underline{\theta}(\Phi^{fb})}^{\overline{\theta}} S_p f(\theta)\, d\theta}{f(\underline{\theta}(\Phi^{fb}))} + \frac{\partial \underline{\theta}(\Phi^{fb})}{\partial \Phi}(Q(\underline{\theta}(\Phi^{fb}), \underline{\theta}(\Phi^{fb}))S_p + pS_a) = 0.$$

Using the fact that if $\underline{\theta}(\Phi^{fb}) \in \text{int}\Theta$ then (19) must hold by definition — i.e.,

$$Q(\underline{\theta}(\Phi^{fb}), \underline{\theta}(\Phi^{fb}))S_p = (1-\Phi^{fb})S_p + \delta.$$

Using **A1** and the implicit Function Function Theorem, one gets

$$\frac{\partial \underline{\theta}(\Phi^{fb})}{\partial \Phi} = -\frac{1}{Q_\theta(\underline{\theta}(\Phi^{fb}), \underline{\theta}(\Phi^{fb}))} < 0.$$

Then, the first-order condition above boils down to (20).

Now, in order to show that $\underline{\theta}(\Phi^{fb}) < \overline{\theta}$ note that as long as Φ^{fb} is such that $\underline{\theta}(\Phi^{fb}) = \overline{\theta}$ equation (20) implies

$$\frac{(1-p)\,S_a}{Q_\theta(\overline{\theta}, \overline{\theta})} > 0,$$

hence $\underline{\theta}\left(\Phi^{fb}\right) < \overline{\theta}$. Moreover, to show that $\underline{\theta}\left(\Phi^{fb}\right) > 0$ note that for $\underline{\theta}\left(\Phi^{fb}\right) = 0$ equation (20) together with **A6** implies immediately:

$$\frac{(1-p)\,S_a}{Q_\theta\,(0,0)} - \frac{S_p}{f\,(0)} < 0.$$

Uniqueness of the optimal policy follows from the simple fact that under **A6** the left hand side of (20) is decreasing in Φ and $\frac{\partial \underline{\theta}(\Phi^{fb})}{\partial \Phi} < 0$, while **A2** together with $\frac{\partial \underline{\theta}(\Phi^{fb})}{\partial \Phi} < 0$ imply that the right hand side of (20) is decreasing in Φ. Finally, Lemma 4 together with $\underline{\theta}\left(\Phi^{fb}\right) \in \text{int}\Theta$ directly imply $\Phi^{fb} \in (0,1)$. ■

Proof of Proposition 6: The derivation of the objective function of $\mathcal{P}_\Phi$ is straightforward. Differentiating this objective with respect to $\underline{\theta}$ and $s\,(.)$ yields immediately the first-order conditions (21) and (22). Using the same techniques developed in the proof of Proposition 2 it follows immediately that: (i) $\underline{\theta}_\Phi^{sb} \in \text{int}\Theta$, (ii) $s_\Phi^{sb}\,(\theta) \leq \theta$ with equality only at $\theta = \underline{\theta}(\Phi^{sb})$ and $\dot{s}_\Phi^{sb}\,(\theta) \geq 0$, and (iii) that (25) yields the agent's state contingent amnesty.

Next, we must show that (24) identifies the optimal amnesty for the self-reporting boss, and that $\underline{\theta}_\Phi^{sb} < \underline{\theta}(\Phi^{sb}) < \overline{\theta}$. First, differentiating the objective of $\mathcal{P}_\Phi$ with respect to Φ and using the fact that $s_\Phi^{sb}(\underline{\theta}\left(\Phi^{sb}\right)) = \underline{\theta}\left(\Phi^{sb}\right)$ one gets:

$$-(S_a + (1-\Phi^{sb})S_p) + \delta\frac{\partial \underline{\theta}(\Phi^{sb})}{\partial \Phi} - \frac{\int_{\underline{\theta}(\Phi^{sb})}^{\overline{\theta}} S_p f(\theta)\, d\theta}{f(\underline{\theta}(\Phi^{sb}))} +$$
$$+ (Q(\underline{\theta}(\Phi^{sb}), \underline{\theta}(\Phi^{fb}))S_p + pS_a)\frac{\partial \underline{\theta}(\Phi^{sb})}{\partial \Phi} - \frac{\partial \underline{\theta}(\Phi^{sb})}{\partial \Phi}\int_{\underline{\theta}_\Phi^{sb}}^{\underline{\theta}(\Phi^{sb})} Q_\theta(s_\Phi^{sb}(\theta), \theta)d\theta = 0.$$

Using the fact that when $\underline{\theta}(\Phi^{sb}) \in$int$\Theta$ condition (23) must hold by definition — i.e.,

$$Q(\underline{\theta}(\Phi^{sb}), \underline{\theta}(\Phi^{sb}))S_p = (1 - \Phi^{sb})S_p + \delta,$$

the first-order condition above boils down to:

$$(1-p)\, S_a \frac{\partial \underline{\theta}(\Phi^{sb})}{\partial \Phi} + \frac{\int_{\underline{\theta}(\Phi^{sb})}^{\overline{\theta}} S_p f(\theta)\, d\theta}{f(\underline{\theta}(\Phi^{sb}))} + \frac{\partial \underline{\theta}(\Phi^{sb})}{\partial \Phi} \int_{\underline{\theta}_{\Phi}^{sb}}^{\underline{\theta}(\Phi^{sb})} Q_\theta \left(s_{\Phi}^{sb}(\theta), \theta \right) d\theta = 0. \tag{31}$$

Differentiating the indifference condition (23) and using the fact that by (22) one must have $s_{\Phi}^{sb}(\underline{\theta}\left(\Phi^{sb}\right)) = \underline{\theta}(\Phi^{sb})$, and therefore $Q_s(\underline{\theta}(\Phi^{sb}), \underline{\theta}(\Phi^{sb})) = 0$ by **A1**, we have:

$$\frac{\partial \underline{\theta}(\Phi^{sb})}{\partial \Phi} = -\frac{1}{Q_\theta(\underline{\theta}(\Phi^{sb}), \underline{\theta}(\Phi^{sb}))} < 0. \tag{32}$$

Implying that the first-order condition (31) rewrites as:

$$\frac{(1-p)\, S_a + \int_{\underline{\theta}_{\Phi}^{sb}}^{\underline{\theta}(\Phi^{sb})} Q_\theta(s_{\Phi}^{sb}(\theta), \theta) d\theta}{Q_\theta(\underline{\theta}(\Phi^{sb}), \underline{\theta}(\Phi^{sb}))} = \frac{1 - F\left(\underline{\theta}(\Phi^{sb})\right)}{f(\underline{\theta}(\Phi^{sb}))} S_p, \tag{33}$$

which is (24).

Now, to show that $\underline{\theta}(\Phi^{sb}) < \overline{\theta}$ note that for $\underline{\theta}(\Phi^{sb}) = \overline{\theta}$ the first-order condition (33) yields:

$$\frac{(1-p)\, S_a + \int_{\underline{\theta}_{\Phi}^{sb}}^{\overline{\theta}} Q_\theta(s_{\Phi}^{sb}(\theta), \theta) d\theta}{Q_\theta(\overline{\theta}, \overline{\theta})} > 0,$$

hence $\underline{\theta}(\Phi^{sb}) < \overline{\theta}$ and so $\Phi^{sb} < 1$ by Lemma 5. We need to show that $\underline{\theta}(\Phi^{sb}) > \underline{\theta}_{\Phi}^{sb} > 0$. First, suppose that $\underline{\theta}\left(\Phi^{sb}\right) = \underline{\theta}_{\Phi}^{sb}$, then (33) rewrites as

$$\frac{(1-p)\, S_a}{Q_\theta(\underline{\theta}_{\Phi}^{sb}, \underline{\theta}_{\Phi}^{sb})} = \frac{1 - F(\underline{\theta}_{\Phi}^{sb})}{f(\underline{\theta}_{\Phi}^{sb})} S_p. \tag{34}$$

Substituting $\underline{\theta}(\Phi^{sb}) = \underline{\theta}_{\Phi}^{sb}$ into the first-order condition (14) one has $\underline{\theta}_{\Phi}^{sb} = 0$. Then **A6** implies that (34) cannot hold, and therefore $\underline{\theta}\left(\Phi^{sb}\right) \neq \underline{\theta}_{\Phi}^{sb}$. Second, showing that $\underline{\theta}(\Phi^{sb}) < \underline{\theta}_{\Phi}^{sb}$ follows from a simple revealed preference argument. Suppose that there exists an equilibrium where only the boss talks, and denote by $\hat{\Phi}^{sb}$ the optimal amnesty such that $0 < \underline{\theta}(\hat{\Phi}^{sb}) < \hat{\underline{\theta}}_{\Phi}^{sb}$. Then in such equilibrium there will be a subset of states of nature where the trial takes place — i.e., for all $\theta \in [0, \underline{\theta}(\hat{\Phi}^{sb}))$. But, for any given $\underline{\theta}(\hat{\Phi}^{sb})$, Proposition 2 implies that the Legislator can strictly reduce the crime rate by letting the agent talk in some states $\theta < \underline{\theta}(\hat{\Phi}^{sb})$ — i.e.,

$$\int_0^{\underline{\theta}(\hat{\Phi}^{sb})} (q(\theta)\, S_p - u_0) f(\theta)\, d\theta <$$

$$\max_{\underline{\theta},\ s(\theta)\leq\theta}\left\{\int_0^{\underline{\theta}}(q(\theta)S_p-u_0)f(\theta)\,d\theta+\int_{\underline{\theta}}^{\underline{\theta}(\hat{\Phi}^{sb})}\left(Q(s(\theta),\theta)S_p-u_0-Q_\theta(s(\theta),\theta)\frac{F(\underline{\theta}(\hat{\Phi}^{sb}))-F(\theta)}{f(\theta)}\right)f(\theta)\,d\theta\right\}.$$

Hence, a contradiction. It then follows that $\underline{\theta}(\Phi^{sb})>\underline{\theta}_{\Phi}^{sb}$ and that (24) is a necessary condition to identify an internal optimum.

Finally, showing that $\underline{\theta}\left(\Phi^{sb}\right)<\underline{\theta}\left(\Phi^{fb}\right)$, and therefore that $\Phi^{sb}>\Phi^{fb}$, follows from a very simple argument. Let us rewrite the first-order necessary conditions identifying Φ^{fb} and Φ^{sb} respectively, as:

$$(1-p)S_a=\frac{1-F(\underline{\theta}(\Phi^{fb}))}{f(\underline{\theta}(\Phi^{fb}))}S_pQ_\theta(\underline{\theta}(\Phi^{fb}),\underline{\theta}(\Phi^{fb})),\tag{35}$$

and

$$(1-p)S_a+\int_{\underline{\theta}_{\Phi}^{sb}}^{\underline{\theta}(\Phi^{sb})}Q_\theta(s_{\Phi}^{sb}(\theta),\theta)d\theta=\frac{1-F(\underline{\theta}(\Phi^{sb}))}{f(\underline{\theta}(\Phi^{sb}))}S_pQ_\theta(\underline{\theta}(\Phi^{sb}),\underline{\theta}(\Phi^{sb})).\tag{36}$$

Note that the left-hand side of (36) is larger than the left-hand side of (35) because $\underline{\theta}(\Phi^{sb})>\underline{\theta}_{\Phi}^{sb}$ and $Q_\theta(.)>0$. Moreover, note that for any given Φ under **A2** the right-hand of both equations is increasing in Φ — i.e.,

$$\frac{\partial}{\partial\Phi}\left(\frac{1-F(\underline{\theta}(\Phi))}{f(\underline{\theta}(\Phi))}Q_\theta(\underline{\theta}(\Phi),\underline{\theta}(\Phi))\right)>0.$$

Hence, it immediately follows that $\underline{\theta}(\Phi^{sb})<\underline{\theta}(\Phi^{fb})$ so that $\Phi^{sb}>\Phi^{fb}$.

In order to complete the proof we must verify that, given the policy described in the statement of the proposition, neither the agent nor the boss can profitably deviate from the equilibrium where the agent talks in states $\theta<\underline{\theta}(\Phi^{sb})$ and the boss self-reports only if $\theta\geq\underline{\theta}(\Phi^{sb})$.

Consider first the boss. Showing that he cannot gain from talking in the states $\theta<\underline{\theta}(\Phi^{sb})$ is straightforward and it immediately follows from equation (23) and **A1**. Next, suppose that he does not talk in a state $\theta>\underline{\theta}(\Phi^{sb})$, we must show that in this 'off equilibrium' history the following happens: (*i*) the agent will cooperate and, (*ii*) his testimony is such that the boss' deviation is not profitable. In order to do so, consider the allocation $\tilde{s}(\hat{\theta})=\hat{\theta}$ and $\tilde{\phi}=\phi_{\Phi}^{sb}\left(\underline{\theta}(\Phi^{sb})\right)$. Suppose for now that such allocation is incentive compatible (a condition that will be checked ex post), then by construction the boss will not find it profitable to deviate because under **A1** the following is true

$$Q(\theta,\theta)S_p>(1-\Phi^{sb})S_p+\delta\quad\forall\ \ \theta>\underline{\theta}(\Phi^{sb}).$$

We can now show that $\tilde{s}(\hat{\theta})=\hat{\theta}$ and $\tilde{\phi}=\phi_{\Phi}^{sb}\left(\underline{\theta}(\Phi^{sb})\right)$ is indeed incentive compatible — i.e., that in any state $\theta>\underline{\theta}(\Phi^{sb})$ when the boss has (unexpectedly) not self-reported the agent tells the truth — i.e., he cannot profitably deviate neither from mimicking a type $\hat{\theta}>\underline{\theta}(\Phi^{sb})$ nor a type $\hat{\theta}<\underline{\theta}(\Phi^{sb})$. Given the 'off-equilibrium' policy $(\tilde{s}(\hat{\theta}),\tilde{\phi})$, it is immediate to show that mimicking a type $\hat{\theta}>\underline{\theta}(\Phi^{sb})$ is not convenient because both the first- and second-order local incentive constraints are satisfied (which also implies that the global incentive constraint holds). Suppose now that the agent lies by claiming that the state is $\hat{\theta} < \underline{\theta}(\Phi^{sb})$, his utility would then be:

$$u(\hat{\theta}, \theta) = -(1 - \phi_{\Phi}^{sb}(\hat{\theta}))S_a - (1 - Q(s_{\Phi}^{sb}(\hat{\theta}), \theta)), \tag{37}$$

implying that

$$\frac{\partial u(\hat{\theta}, \theta)}{\partial \hat{\theta}} = \dot{\phi}_{\Phi}^{sb}(\hat{\theta})S_a + Q_s(s_{\Phi}^{sb}(\hat{\theta}), \theta)\dot{s}_{\Phi}^{sb}(\hat{\theta}).$$

Note that $Q_{s\theta}(.) > 0$ and $\dot{s}_{\Phi}^{sb}(\hat{\theta}) > 0$ together with the local incentive constraint (8) imply that

$$\dot{\phi}_{\Phi}^{sb}(\hat{\theta})S_a + Q_s(s_{\Phi}^{sb}(\hat{\theta}), \theta)\dot{s}_{\Phi}^{sb}(\hat{\theta}) > \dot{\phi}_{\Phi}^{sb}(\hat{\theta})S_a + Q_s(s_{\Phi}^{sb}(\hat{\theta}), \hat{\theta})\dot{s}_{\Phi}^{sb}(\hat{\theta}) = 0 \quad \forall\, \hat{\theta} < \underline{\theta}(\Phi^{sb}) < \theta.$$

Hence,

$$\frac{\partial u(\hat{\theta}, \theta)}{\partial \hat{\theta}} = \dot{\phi}_{\Phi}^{sb}(\hat{\theta})S_a + Q_s(s_{\Phi}^{sb}(\hat{\theta}), \theta)\dot{s}_{\Phi}^{sb}(\hat{\theta}) > 0 \;\; \forall\, \hat{\theta} < \underline{\theta}(\Phi^{sb}) < \theta.$$

This implies that if the agent mimicks in state $\theta > \underline{\theta}(\Phi^{sb})$ by pretending that the state is $\hat{\theta} \leq \underline{\theta}(\Phi^{sb})$ he will always pretend to be in state $\underline{\theta}(\Phi^{sb})$ — i.e., where the information rent that he gets by reporting a state $\theta' \leq \underline{\theta}(\Phi^{sb})$ is maximal. But then, he could do strictly better from telling the truth and obtain the allocation $\tilde{s}(\theta) = \theta$ and $\tilde{\phi} = \phi_{\Phi}^{sb}\left(\underline{\theta}(\Phi^{sb})\right)$. This can be easily verified by using equation (37); in fact, in this case the agent's expected utility would be:

$$\tilde{u}\,(\theta) = -(1 - \phi_{\Phi}^{sb}(\underline{\theta}(\Phi^{sb})))S_a - (1 - Q(\theta, \theta)),$$

which immediately yields

$$\tilde{u}\,(\theta) > u(\underline{\theta}(\Phi^{sb}), \theta) = -(1 - \phi_{\Phi}^{sb}(\underline{\theta}(\Phi^{sb})))S_a - (1 - Q(\underline{\theta}(\Phi^{sb}), \theta)),$$

since $\theta > \underline{\theta}(\Phi^{sb})$. It then follows that, given the policy described in the statement of the proposition, in a 'off-equilibrium' history where the boss has not self-reported in state $\theta > \underline{\theta}(\Phi^{sb})$ the agent truthfully reveals his type. Hence, the boss' expected sanction in this continuation game would be $Q(\theta, \theta)S_p$, which is clearly above the costs from self-reporting $(1 - \Phi^{sb})S_p + \delta$ by (23) and **A1.**

We are left to show that no agent can profitably deviate when the boss self-reports, this is trivial as in this case they are convicted with certainty and their information is worthless to the prosecutor. Finally, showing that the global incentive constraint holds for all types $\theta \in [\underline{\theta}^{sb}, \underline{\theta}(\Phi^{sb}))$ follows exactly the same steps as those developed in the proof of Proposition 2, and will be thus omitted for simplicity. ■

Proof of Corollary 7: Showing that $s_{\Phi}^{sb}(\theta) > s^{sb}(\theta)$ for all $\theta \in [\underline{\theta}_{\Phi}^{sb}, \underline{\theta}(\Phi^{sb})]$ follows immediately from comparing equation (15) with (22) and $F(\underline{\theta}(\Phi^{sb})) < 1$ since $\underline{\theta}(\Phi^{sb}) < \bar{\theta}$. To show that $\underline{\theta}_{\Phi}^{sb} < \underline{\theta}^{sb}$ consider equations (14) and (21). Note that $\underline{\theta}_{\Phi}^{sb} = \underline{\theta}^{sb}$ for $\Phi^{sb} = 0$. Moreover, for any given Φ let

$$\digamma(\underline{\theta}_{\Phi}^{sb}, \Phi) \equiv (Q(s_{\Phi}^{sb}(\underline{\theta}_{\Phi}^{sb}), \underline{\theta}_{\Phi}^{sb}) - q(\underline{\theta}_{\Phi}^{sb}))S_p - Q_{\theta}(s_{\Phi}^{sb}(\underline{\theta}_{\Phi}^{sb}), \underline{\theta}_{\Phi}^{sb})\frac{F(\underline{\theta}(\Phi)) - F(\underline{\theta}_{\Phi}^{sb})}{f(\underline{\theta}_{\Phi}^{sb})} = 0,$$

note that the Envelope Theorem applied to equation (21) implies:

$$\frac{\partial \underline{\theta}_{\Phi}^{sb}}{\partial \Phi} = -\frac{\mathcal{F}_{\Phi}(\underline{\theta}_{\Phi}^{sb}, \Phi)}{\mathcal{F}_{\underline{\theta}}(\underline{\theta}_{\Phi}^{sb}, \Phi)} = \frac{1}{\mathcal{F}_{\underline{\theta}}(\underline{\theta}_{\Phi}^{sb}, \Phi)} \frac{\partial \underline{\theta}(\Phi^{sb})}{\partial \Phi},$$

where $\mathcal{F}_{\underline{\theta}}(\underline{\theta}_{\Phi}^{sb}, \Phi) > 0$ by concavity of the objective function with respect to $\underline{\theta}$ and $\frac{\partial \underline{\theta}(\Phi^{sb})}{\partial \Phi} < 0$ by (32). Hence, it follows that $\frac{\partial \underline{\theta}_{\Phi}^{sb}}{\partial \Phi} < 0$ implying immediately that $\underline{\theta}_{\Phi}^{sb} < \underline{\theta}^{sb}$. ■

References

[1] Acconcia, A., Immordino, G., Piccolo, S., and Rey P. (2009), "Accomplice-Witnesses and Organized Crime: Theory and Evidence from Italy," CSEF working paper n. 232.

[2] Aubert, C., Kovacic, W., and Rey P. (2006), "The Impact of Leniency and Whistleblowing Programs on Cartels", *International Journal of Industrial Organization*, 24: 1241-1266.

[3] Anderson, A. (1979), *The Business of Organized Crime: A Cosa Nostra Family*, Stanford, The Hoover Institution.

[4] Anselmo, M., and Braucci, M. (2008), *Questa Corte Condanna: Spartacus, il Processo al Clan dei Casalesi*, edited by L'Ancora, Naples.

[5] Baccara M., and Bar-Isaac, H., (2008), "How to Organize Crime," *Review of Economic Studies*, 75: 1039-1067.

[6] Backhaus, J. (1979), "Defending Organized Crime? A note," *Journal of Legal Studies*, 8: 623-631.

[7] Buchanan, J. (1973), "A Defense of Organized Crime?", in *The Economics of Crime and Punishment*, edited by S. Rottenberg, American Enterprise Institute.

[8] Becker, G. (1968), "Crime and Punishment: an Aconomic Approach," *Journal of Political Economy,* 76: 169-217.

[9] Cantone, R. (2008), *Solo per Giustizia*, edited by Mondadori, Milano.

[10] Cassidy, R. M, (2004), ""Soft Words of Hope:" Giglio, Accomplice Witnesses, and the Problem of Implied Inducements", Northwestern University Law Review, 98: 1129-1177.

[11] Chen, Z., and Rey, P. (2007), "On the Design of Leniency Programs," mimeo TSE.

[12] Dickey, J. (2004), *Cosa Nostra: A History of the Sicilian Mafia*, edited by Hodder & Stoughton.

[13] Ehrlich, I. (1973), "Participation in Illegitimate Activities," *Journal of Political Economy,* 81: 521-565.

[14] Falcone, G., (1991), *Cose di Cosa Nostra,* edited by Rizzoli, Milano.

[15] Fiorentini G., and Peltzman, S., (1995), *The Economics of Organized Crime*, Cambridge University Press, Cambridge, and CEPR, 1-30.

[16] Fyfe, N. and Sheptycki J., (2006), "International Trends in the Facilitation of Witness Co-operation in Organized Crime Cases," *European Journal of Criminology*, 3: 319-355.

[17] Gambetta, D. (1992), *La Mafia Siciliana,* edited by Einaudi, Torino.

[18] Garoupa, N. (1997), "The Theory of Optimal Law Enforcement," *Journal of Economic Surveys,* 11: 267-295.

[19] Garoupa, N. (2000), "The Economics of Organized Crime and Optimal Law Enforcement," *Economic Inquiry,* 38: 278-288.

[20] Huber, B., (2001), "Informants Undercover Agents, and Police Informants", in *Towards a European Criminal Law Against Organised Crime,* edited by V., Militello and B. Huber, Freiburg im Breisgau, 109-120.

[21] Kaplow, L., Shavell, S., (1994), "Optimal law enforcement with self-reporting of behavior," *Journal of Political Economy* 102: 583-606.

[22] Konrad, K. A. and Skaperdas, S., (1998), "Extortion," *Economica*, 65: 461-477.

[23] Konrad, K. A. and Skaperdas, S., (1997), "Credible threats in extortion," *Journal of Economic Behavior and Organization*, 33: 23-39.

[24] Kugler, M., Verdier, T. and Zenou Y., (2005), "Organized crime, corruption and punishment," *Journal of Public Economics* 89: 1639-1663.

[25] Jeffries, J.C. and Gleeson, J. (1995), "The federalization of organized crime: advantages of federal prosecution", *Hastings Law Journal*, 46: 1095.

[26] Jennings, W., (1984), "A Note on the Economics of Organized Crime," *Eastern Economic Journal*, 3: 315-321.

[27] Laffont, J.J. and Martimort, D. (2002), *The Theory of Incentives:The Principal-Agent Model*, Princeton University Press.

[28] Laffont, J.J. and Tirole, J. (1993), *A theory of incentives in regulation and procurement*, MIT Press Ltd.

[29] Lodato., S. (2006), *Trent'Anni di Mafia*, edited by Rizzoli, Milano.

[30] Mansour, A., Marceau, N., and Mongrain, S. (2006), "Gangs and Crime Deterrence," *Journal of Law Economics and Organization,* 22: 315-339.

[31] Maskin E., and Riley, J., "Input versus Output Incentive Schemes", *Journal of Public Economics*, 28: 1-23, 1985.

[32] Myerson, R. B. (1981), "Optimal Auction Design," *Mathematics of Operations Research,* 6: 58-73.

[33] Motta, M., and Polo, M. (2003), "Leniency Programs and Cartel Prosecution," *International Journal of Industrial Organization,* 21: 347-379.

[34] Paoli, L. (2003), *Mafia Brotherhoods: Organized Crime, Italian Style*, *New York*, Oxford University Press.

[35] Polo, M., (1995), "Internal Cohesion and Competition among Criminal Organizations", in *The Economics of Organized Crime,* edited by G., Fiorentini and S. Peltzman, Cambridge University Press, Cambridge, and CEPR, 87-109.

[36] Posner, R. (1992), *The Economic Analysis of Law*, 4th Edition. Boston: Little Brown.

[37] Rey, P. (2003), "Toward a Theory of Competition Policy," in Advances in *Economics and Econometrics: Theory and Applications*, Eighth World Congress, M. Dewatripont, L. P. Hansen, S. J. Turnovsky eds, Cambridge University Press.

[38] Schur, G., (1988), "A Father's Dream Come True," *The Pentacle, 8: 13-14.*

[39] Spagnolo, G. (2003), "Optimal Deterrence Mechanisms against Cartels and Organized Crime," mimeo, Mannheim.