Sustaining Cooperation: Community Enforcement vs. Specialized Enforcement*

Daron Acemoglu
MIT

Alexander Wolitzky
MIT

July 2018

Abstract

We introduce the possibility of coercive punishment by specialized enforcers into a model of community enforcement. We assume that, just as regular agents need to be given incentives to cooperate with each other, specialized enforcers need to be given incentives to carry out costly punishments. We fully characterize optimal equilibria in the model. When the specialized enforcement technology is sufficiently effective, cooperation is best sustained by a "one-time enforcer punishment equilibrium", where any deviation by a regular agent is punished only once, and only by enforcers. In contrast, enforcers themselves are disciplined (at least in part) by community enforcement. The reason why there is no community enforcement following deviations by regular agents is that such a response, by reducing future cooperation, would decrease the amount of punishment that enforcers are willing to impose on deviators. Conversely, when the specialized enforcement technology is less effective, optimal equilibria involve a mix of specialized enforcement and community enforcement (which might take the form of "ostracism"). Our results hold both under perfect monitoring of actions and under various types of private monitoring.

Keywords: coercion, community enforcement, cooperation, law enforcement, repeated games, specialized enforcement.

JEL Classification: C73, D72, D74

*We thank Anton Kolotilin and Francesco Nava for insightful and helpful discussions of this paper. For helpful comments, we thank Nageeb Ali, Glenn Ellison, Ben Golub, Matt Jackson, David Levine, Dilip Mookherjee, Juuso Välimäki, Juuso Toikka; seminar participants at Bonn, BU, Cambridge, Duke, Essex, Harvard-MIT, Miami, the Canadian Institute for Advanced Research, the National Bureau of Economic Research, the Stanford Institute for Theoretical Economics, and the Wallis Institute for Political Economy; and the anonymous referees. We thank Kevin Li for excellent research assistance. Wolitzky acknowledges financial support from the NSF and the Sloan Foundation.

1 Introduction

Throughout history, human societies have used a variety of means and practices to foster pro-social behavior among their members. Prominent among these is *decentralized community enforcement*, where deviations from cooperative behavior are discouraged by the threat of withholding future cooperation, or even the threat of the widespread collapse of cooperation throughout society. A large literature in the social sciences, especially in game theory, provides conceptual foundations for this type of enforcement. In small groups, where an individual's behavior can be accurately observed by other members of the community, the threat of exclusion or punishment is a powerful means of supporting cooperation (Axelrod, 1984, Fudenberg and Maskin, 1986, Coleman, 1988, Ostrom, 1990, Greif, 2006). In large groups, where information about past behavior is more limited, cooperation can be supported by *contagion strategies*, which trigger the spread of non-cooperative behavior following a deviation (Kandori, 1992, Ellison, 1994). Furthermore, several prominent examples, such as the cooperative arrangements among the medieval Maghribi traders and their overseas agents (Greif, 1993) and the norms of behavior and compensation between ranchers and landowners in 20th-century Shasta County, California (Ellickson, 1991) demonstrate the practical feasibility of decentralized community enforcement.

In modern societies, however, the basis of cooperative behavior is rather different. Major transgressions are not directly punished by neighbors, nor do they trigger a wave of non-cooperative behavior throughout society. Instead, they are directly punished by specialized law enforcers, including the police, the courts, and other state and non-state institutions. Indeed, following Thomas Hobbes and Max Weber, most social scientists view this type of *specialized enforcement* as desirable, as well as inevitable both in societies with full-fledged states and in those with less developed proto-states (Johnson and Earle, 2000, Flannery and Marcus, 2012). Yet, there exists little formal modeling of the foundations of such specialized enforcement.

The goal of this paper is to develop a model of specialized enforcement, to compare its performance in supporting cooperation with that of community enforcement, and to delineate the conditions under which specialized enforcement emerges as the optimal arrangement for sustaining cooperation.

We consider a model of cooperation within a group of agents. In our baseline model, regular producers randomly match with each other, as well as with specialized enforcers assigned to monitor their relationships. Each producer chooses a level of cooperation (e.g., a contribution to a local public good or an investment in a joint project), which is costly for her but generates benefits for her partners (both the other producers with whom she matches and the enforcers who monitor them). Absent the threat of direct or indirect punishment, a producer would choose zero cooperation. In this model, cooperation can be supported by contagion strategies as in Kandori (1992) and Ellison (1994), where a deviation from pro-social behavior triggers the withdrawal of cooperation throughout the entire community. Cooperation can alternatively be supported by specialized enforcement—in which enforcers coercively punish producers who deviate—provided that enforcers can be given incentives to behave in this way.[1] Cooperation can also be supported by any number of other strategies, including various combinations of community and specialized enforcement involving repentance by the deviator, ostracism of the deviator, and so on. Our question then is not what kinds of strategies can support some cooperation, but rather what strategies support the *maximum possible* level of cooperation (at a *fixed discount factor*).

Our simplest and sharpest results apply under *perfect monitoring*, where each agent observes the entire past history of behavior. We consider both a *small group* version of the model, where all agents in the group interact directly in every period, and a *large group* version, where interactions are determined by random matching. The main results are the same in both cases. The optimal equilibrium always utilizes the specialized enforcement technology in a specific way—using *one-time enforcer punishment strategies*, where enforcers punish a deviant producer as harshly as possible, but only punish once. Enforcers are incentivized to undertake costly punishments through the threat of contagion: if they fail to punish the deviator, this triggers a switch to zero cooperation by all producers.[2] When the specialized enforcement technology is sufficiently effective, following punishment by the enforcers all agents immediately return to equilibrium play, so there is no contagion or withholding of future cooperation. When the specialized enforcement technology is less effective, one-time enforcer punishment must be combined with some form of "community enforcement," where producers withhold cooperation to punish the deviator. In our baseline model, this is optimally achieved via *repentance*, whereby the deviator cooperates at a higher level than other producers in the period immediately following a deviation. If in addition we allow the deviator to be directly excluded from the benefits of cooperation (or "ostracized"), then we show that one-time enforcer punishment strategies are optimally combined with *ostracism* rather than repentance.

The form of our one-time enforcer punishment strategies can be viewed as a stylized representation of how formal and informal incentives interact in modern legal systems. Enforcers' incentives come from the fact that they themselves benefit from societal cooperation (either directly or, in an extension, because the revenues that pay their salaries are generated by such cooperation), and societal cooperation depends on citizens' trust in the integrity of the law enforcement apparatus. If this trust is damaged because enforcers deviate from their expected course of behavior,

[1] In practice, another important problem is ensuring that enforcers do not use their access to violence to expropriate producers. At the level of abstraction of our model, this is similar to the problem of convincing enforcers to choose the appropriate level of punishment in response to transgressions, as we discuss below.

[2] We also show that when enforcers can be directly punished by other enforcers, their deviations trigger both direct punishment and the temporary withholding of cooperation by producers.

societal cooperation collapses, and it is the prospect of such a collapse that incentivizes enforcers.[3] With this interpretation, our results show that societies with more effective specialized enforcement technologies should rely solely on enforcers (or "state institutions") to deter undesirable behavior, while those with less efficient technologies should combine enforcer punishments with community enforcement.

One general implication of our analysis is the optimality of one-time enforcer punishment strategies for supporting cooperation (either by themselves or in combination with repentance or ostracism). The possibility that one-time enforcer punishment strategies (without repentance or ostracism) can be optimal may appear surprising, as one might have conjectured that it would always be better to combine specialized enforcer punishments with decentralized community enforcement—if both coercive punishment and the withdrawal of cooperation are bad for producers, why not use both to provide incentives? The intuition for this result highlights the economic mechanism at the heart of our paper. Adding decentralized punishment to a *given* level of specialized punishment would indeed improve producers' incentives for cooperation. But, crucially, it would also erode the incentives of enforcers to undertake coercive punishment. Enforcers are willing to undertake costly punishments today only because of the future rewards of continued societal cooperation. Hence, if a deviation by a producer also triggered costly community enforcement, then these implicit rewards would be diminished, curtailing the extent of enforcer punishments. This reasoning thus identifies a novel and powerful cost of decentralized punishment: its negative impact on the extent and efficacy of specialized punishment. The reason why it is optimal to punish deviators only once is also interesting: as we show, the gain in efficiency from spreading punishments over time is always more than offset by the reduced willingness of the deviator to return to cooperation during the punishment phase.

The role of specialized punishment by enforcers and the tradeoff between community enforcement and specialized enforcement generalize beyond the perfect monitoring case. First, we show that, for a fairly general class of information structures (including the possibility that each individual observes play only in her own past matches), one-time enforcer punishment strategies outperform pure contagion when either the punishment technology is sufficiently effective or the discount factor is sufficiently large. Because pure contagion is optimal in this environment without the enforcers (Wolitzky, 2013), this result immediately implies that the optimal equilibrium must rely on enforcers to some extent. Second, we establish that, when individuals observe behavior in their partners' most recent matches, one-time enforcer punishment strategies form an optimal equilibrium, provided that the specialized enforcement technology is sufficiently effective and that

[3] More realistically, enforcers may be organized in a hierarchy, where low-level enforcers are incentivized by higher-level enforcers and only the top-level enforcers are incentivized by community enforcement. We discuss such an extension of our model in Section 5.

imperfections in the monitoring structure cannot be used to increase the extent to which enforcers are willing to punish a deviator. This latter requirement can be guaranteed, for example, when enforcers are better informed than producers.[4] We further show that one-time enforcer punishment strategies are also optimal under an additional stability requirement, which postulates that a single deviation by any single individual is not sufficient to start contagion.[5]

Our paper is related to several different lines of research. First, we build on the literature on community enforcement in repeated games, pioneered by Kandori (1992) and Ellison (1994), by introducing costly punishments into this literature. Recent contributions to this literature include Takahashi (2010), Deb (2012), and Deb and González-Díaz (2014). Most closely related to our paper are Wolitzky (2013) and Ali and Miller (2014), which provide conditions under which contagion strategies support the maximum level of cooperation at a fixed discount factor in repeated cooperation games without costly punishments. In contrast, we show that introducing the possibility of costly punishments can radically change the structure of the optimal equilibrium from contagion (grim trigger) strategies to one-time enforcer punishment strategies. Several other papers in this literature emphasize various weaknesses of contagious strategies. Jackson, Rodriguez-Barraquer, and Xu (2012) note that contagion strategies violate a renegotiation-proofness condition and focus instead on equilibria in which social breakdowns are contained following a deviation. Lippert and Spagnolo (2010) and Ali and Miller (2016) show that contagion or permanent exclusion discourages communication about past deviations and argue for equilibria involving temporary exclusion or ostracism. These papers do not consider specialized enforcers and more generally do not investigate optimal equilibria in settings where contagion strategies are suboptimal.[6]

Second, our paper is also related to the literature on optimal penal codes in general repeated games (Abreu, 1988), especially the "stick-and-carrot" equilibria of Abreu (1986). In particular, our one-time enforcer punishment equilibria offer the "stick" of specialized punishment for producers and the "carrot" of continued cooperation for enforcers. However, while in Abreu (1986) stick-and-carrot equilibria are optimal only within the class of pure strategy, strongly symmetric equilibria (i.e., under the restriction that play is symmetric at all histories), we show that one-time enforcer punishment equilibria are globally optimal in our model under perfect monitoring, and we also extend this result to certain classes of imperfect private monitoring. Among other works in related

[4] The superior information of enforcers here might result from communication with producers or from the enforcers' being organized in some institution, such as a law enforcement agency.

[5] We find this requirement attractive because it captures another potential cost of decentralized community enforcement: the danger of contagion being triggered accidentally by trembles or mistaken observations. Indeed, many accounts of cooperation in societies with weak or absent states, such as Lewis's (1994) study of Somalia, emphasize how small transgressions can start major feuds, or even all-out tribal wars. Such accidental contagion would also be triggered in our model under community enforcement if producers trembled with small probability. Under enforcer punishments, however, a similarly costly contagion can occur only if *both* an individual producer trembles *and* an enforcer trembles in response. This makes accidental contagion much less likely under enforcer punishments.

[6] Hirshleifer and Rasmusen (1989) consider a form of ostracism that resembles direct punishment and show how it can support cooperation in the finitely repeated prisoner's dilemma.

environments, Padro-i-Miquel and Yared (2012) consider stick-and-carrot equilibria in a political economy model, and Goldlücke and Kranz (2012) show that stick-and-carrot equilibria are generally optimal in repeated games with transfers.

Third, our work connects to the literature on the economic foundations of the enforcement of laws and norms. Early contributions to this literature, including Ostrom (1990), Greif (1989, 1993), Milgrom, North, and Weingast (1990), Greif, Milgrom, and Weingast (1994), Fearon and Laitin (1996), and Dixit (2003), focused on informal enforcement supported by "reputation" and various ostracism-like arrangements. Dixit (2007) surveys and extends these early frameworks. A particularly relevant contribution by Greif (1994) distinguishes between the "private order" institutions of the Maghribi traders and the "public order" institutions of the rival Genoese traders—which resemble, respectively, our community enforcement and specialized enforcement equilibria—and argues that public order institutions proved more efficient as the scope for trade expanded in the late medieval period.

Other related recent papers include Acemoglu and Verdier (1998), who study how law enforcers matched with pairs of producers can be used to incentivize effort, but must also be discouraged from corruption; Hadfield and Weingast (2012), who model law as a device for coordinating decentralized punishment; Mailath, Morris, and Postlewaite (2017), who develop a model of laws and authority based on cheap talk; Levine and Modica (2016), who consider the problem of designing a specialized enforcement system and emphasize the tradeoff between providing insufficient incentives for cooperation and expending excessive effort in punishment; and Acemoglu and Jackson (2017), who study how social norms can constrain the effectiveness of laws. Two recent papers, Masten and Prüfer (2014) and Aldashev and Zanarone (2017), are especially related because they explore aspects of the trade-off between different types of enforcement. Masten and Prüfer introduce court enforcement in a model similar to Dixit (2003) and analyze the transition from merchant law to court law, while Aldashev and Zanarone compare coercive and non-coercive enforcement in a model with two producers and a state specialized in enforcement. Beyond the differences in emphasis and modeling approach, our paper differs from these analyses by focusing on the globally optimal equilibrium for maximizing cooperation in a repeated game. Finally, in a companion paper (Acemoglu and Wolitzky, 2018) we use a model where punishments are costless but are carried out by "elites" who also engage in production to analyze the emergence of "equality before the law", that is, the equal application of coercive punishments to both normal agent and elites.[7]

Lastly, to the extent that enforcer punishment strategies may be viewed as a type of formal enforcement, our paper relates to the literature on the efficiency of formal versus informal enforce-

[7] There is also an enormous literature on the role of punishments in public good games in experimental economics and evolutionary game theory. Seminal experimental papers include Ostrom, Walker, and Gardner (1992) and Fehr and Gächter (2000, 2002). Seminal theoretical papers include Boyd and Richerson (1992), Sethi and Somanathan (1996), and Henrich and Boyd (2001).

ment of norms and contracts. Theoretical contributions include Kranton (1996) and Kali (1999). Empirical studies of reputation-based contract enforcement include Fafchamps (1996), Clay (1997), Woodruff (1998), McMillan and Woodruff (1999), and Johnson, McMillan, and Woodruff (2002).

The rest of the paper is organized as follows. Section 2 and 3 analyze the small-group and large-group versions of the model, respectively. Section 4 considers private monitoring. Section 5 discusses various modeling issues. Section 6 concludes. Proofs for Sections 2 and 3 are presented in Appendix A. The Online Appendix contains the proofs for Section 4, as well as an extension of our model that allows for ostracism of individual players.

2 Small Group Model: Repeated Prisoner's Dilemma with an Enforcer

We begin with the analysis of a standard repeated game model of cooperation in a small group. This setting is a special case of the large group model considered in the next section, but we present it separately for ease of exposition.

2.1 Environment

There is a group consisting of $k+1$ players. We compare the prospects for cooperation in this group in two situations: first, when all $k+1$ players are *producers*, who can exert effort in providing benefits for their partners; and second, when one of the players is an *enforcer*, who can exert effort in punishing her partners. In each case, we investigate the optimal strategies for supporting cooperation and the resulting level of social welfare. The interpretation is that the group can decide whether to assign one of their number to the role of enforcer—exempting her from production while requiring her to punish deviators—but this decision is not formally modeled as part of the game.

The players take part the following two-stage game in every period $t = 0, 1, 2 \ldots$. The game is a version of the prisoner's dilemma among the producers, with the possibility of costly punishment by the enforcer (if an enforcer is present).

1. *Cooperation Stage:* Each producer i chooses a *level of cooperation* $x_i \in \mathbb{R}_+$. These choices are perfectly observed. Choosing cooperation level x_i costs x_i for player i, and benefits every other player $k \neq i$ by an amount $f(x_i)$, where $f : \mathbb{R}_+ \to \mathbb{R}_+$ is an increasing, strictly concave, bounded, and differentiable function satisfying $f(0) = 0$.[8]

2. *Punishment Stage:* If an enforcer is present, he then chooses a *level of punishment* $y_i \in \mathbb{R}_+$ for each producer i. These choices are also perfectly observed. Choosing punishment level y_i costs the enforcer y_i, and hurts producer i by an amount $g(y_i)$, where $g : \mathbb{R}_+ \to \mathbb{R}_+$ is an

[8] Boundedness is for simplicity and can be replaced by the Inada condition $\lim_{x\to\infty} f'(x) = 0$.

increasing, strictly concave, and differentiable function satisfying $g(0) = 0$. We refer to g as the *specialized enforcement technology.*

In summary, if there is no enforcer, each producer i's stage game payoff is

$$\sum_{i' \neq i} f(x_{i'}) - x_i.$$

If instead there is an enforcer (label him player 1), each producer i's stage game payoff is

$$\sum_{i' \neq 1, i} f(x_{i'}) - x_i - g(y_i),$$

and the enforcer's payoff is

$$\sum_{i \neq 1} (f(x_i) - y_i).$$

Observe that playing $x_i = 0$ ("shirking") is myopically optimal for producer i, and playing $y_i = 0$ for all $i \neq 1$ ("failing to punish") is myopically optimal for the enforcer. Thus, only the shadow of future interactions can incentivize producers to cooperate or incentivize the enforcer to punish.

Note also that the assumptions that f and g are concave imply that there is a technological advantage to spreading out cooperation or punishments over time. Nevertheless, we will show that, while optimal equilibria do spread cooperation over time, they do *not* spread punishments over time. Instead, if an enforcer is present, optimal equilibria always concentrate punishments in a single period.

Players maximize expected discounted payoffs with common discount factor δ. The solution concept is subgame perfect equilibrium (SPE).

2.2 One-Time Enforcer Punishment Strategies and Repentance Strategies

Given a path of play of the repeated game, let x_i^t denote producer i's level of cooperation in period t, and let

$$X_i^t = (1 - \delta) \sum_{\tau=0}^{\infty} \delta^\tau x_i^{t+\tau}$$

denote producer i's present discounted level of cooperation starting in period t. From the perspective of a given equilibrium of the game, x_i^t and X_i^t are (possibly degenerate) random variables. We refer to the quantity $\mathbb{E}\left[X_i^0\right]$ as player i's *average level of cooperation* in a given equilibrium. We say that an equilibrium is the *most cooperative* one if it simultaneously achieves the highest value of $\mathbb{E}\left[X_i^0\right]$ for every player i among all SPE (as we will see, such an equilibrium exists); and we call the corresponding value of $\mathbb{E}\left[X_i^0\right]$ the *maximum level of cooperation.* Note that, by concavity of

f, the most cooperative equilibrium is also the *optimal* equilibrium in terms of utilitarian social welfare—provided that producers choose constant levels of cooperation on path, punishments are not used on path, and the maximum level of cooperation is below the first-best level, x^{FB}, given by $kf'\left(x^{FB}\right) = 1$. This last requirement represents the main case of economic interest, as in most settings the challenge is providing sufficient incentives for cooperation rather than avoiding excessive cooperation.

Our main concern is whether optimal equilibria are based on punishment by enforcers (specialized enforcement) or the withdrawal of future cooperation by producers (community enforcement). We first observe that the problem of maximizing cooperation when all $k+1$ players are producers is a trivial one: in this case, cooperation is maximized by *grim trigger* strategies, where producers always play some $x_i = \hat{x}$ on path and switch to $x_i = 0$ following a deviation. The maximum level of cooperation $\hat{x}$ that can be sustained with grim trigger strategies is given by the unique non-zero solution to the equation

$$\hat{x} = \delta k f\left(\hat{x}\right).$$

This equation simply equates the benefit of a deviation to $x_i = 0$ (the cost-saving of $\hat{x}$) with its cost (the lost benefit of others' future cooperation, which equals $\delta k f\left(\hat{x}\right)$).

Proposition 1 *In the absence of enforcers, grim trigger strategies are optimal, and the maximum level of cooperation is $\hat{x}$.*

Grim trigger strategies constitute an extreme form of community enforcement, since cooperation is incentivized entirely by the threat of the group withdrawing cooperation in the future.

On the other hand, when an enforcer is present, the group can instead rely on various strategies that involve specialized enforcement. An extreme form of specialized enforcement, where following a deviation there is no withdrawal of future cooperation at all, is given by what we call *one-time enforcer punishment strategies.* With these strategies, a producer who deviates is immediately punished by the enforcer. Following this one-time punishment, everyone returns to her normal behavior in the next period. If however the enforcer fails to punish a producer deviation, this triggers the breakdown of cooperation.

Definition 1 *A* one-time enforcer punishment strategy profile *is characterized by a cooperation level x and a punishment level y, and can be represented by the following automaton:*

There are two states, normal and punishment. Play in each state is as follows:

Normal state*: Each producer i plays $x_i = x$. If all producers i play $x_i = x$, then the enforcer plays $y_i = 0$ for all producers i. If instead some producer i plays $x_i \neq x$, then the enforcer plays $y_i = y$ and plays $y_{i'} = 0$ for all producers $i' \neq i$.*[9]

[9] As is standard in repeated games with perfect monitoring, we ignore simultaneous deviations throughout.

Punishment state: *Players always take action* 0 *(producers never cooperate; the enforcer never punishes).*

Players start in the normal state and permanently transition to the punishment state if in any period some producer i *plays* $x_i \neq x$ *and the enforcer then plays* $y_i \neq y$.

The component-wise maximum levels of cooperation and punishment (x^*, y^*) that can be sustained with one-time enforcer punishment strategies are given by the system of equations

$$\begin{aligned} x^* &= g(y^*) \\ y^* &= \frac{\delta}{1-\delta} k f(x^*). \end{aligned} \tag{1}$$

The intuition is that a producer who deviates gains at most x^* (her cost of effort) and loses $g(y^*)$ (the cost of being punished at level y^*), while an enforcer who deviates gains at most y^* and loses $\frac{\delta}{1-\delta} k f(x^*)$ (the future benefit of cooperation at level x^* from k producers). Note that the $\frac{1}{1-\delta}$ term in the formula for y^* reflects the fact that, with one-time enforcer punishment strategies, an enforcer trades off the one-time cost of punishing a deviant producer against the benefit of cooperation in every future period. In contrast, there is no such term in the formula for $\hat{x}$, as under grim trigger strategies a producer trades off the cost of cooperating in every period against the benefit of cooperation in every period.

An alternative form of enforcement, which combines elements of community and specialized enforcement, is given by *(one-time) enforcer punishment plus repentance strategies.* Under these strategies, a producer who deviates is immediately punished by the enforcer, but in addition all other producers reduce their level of cooperation for one period while the deviator "repents" by cooperating at the pre-deviation equilibrium level.

Definition 2 *A (*one-time) enforcer punishment plus repentance strategy profile *is characterized by cooperation levels* x *and* $\underline{x}$ *and a punishment level* y*, and can be represented by the following automaton:*

There are $k+2$ *states: normal, punishment, and* i*-repenting, for each* $i \neq 1$. *Play in these states is as follows:*

Normal state*: Each producer* i *plays* $x_i = x$. *If all producers* i *play* $x_i = x$*, then the enforcer plays* $y_i = 0$ *for all producers* i. *If instead some producer* i *plays* $x_i \neq x$*, then the enforcer plays* $y_i = y$ *and plays* $y_{i'} = 0$ *for all producers* $i' \neq i$.

Punishment state: *Players always take action* 0.

i-repenting state: *Producer* i *plays* $x_i = x$. *Producers* $i' \neq i$ *play* $x_{i'} = \underline{x}$. *If producer* i *plays* $x_i = x$ *and all producers* $i' \neq i$ *play* $x_{i'} = \underline{x}$*, then the enforcer plays* $y_{i'} = 0$ *for all producers (including* i*). If producer* i *plays* $x_i \neq x$*, then the enforcer plays* $y_i = y$ *and plays* $y_{i'} = 0$ *for all* *producers $i' \neq i$. If producer $i' \neq i$ plays $x_{i'} \neq \underline{x}$, then the enforcer plays $y_{i'} = y$ and plays $y_{i''} = 0$ for all producers $i'' \neq i'$.*

Players start in the normal state. In the normal state, if some producer i plays $x_i \neq x$ and the enforcer then plays $y_i = y$, players transition to the i-repenting state. If some producer i plays $x_i \neq x$ and the enforcer plays $y_i \neq y$, then players transition to the punishment state.

In the i-repenting state, players transition to the normal state if producer i plays $x_i = x$ and all producers $i' \neq i$ play $x_i = \underline{x}$. If producer i plays $x_i \neq x$ and the enforcer plays $y_i = y$, then players stay in the i-repenting state. If some producer $i' \neq i$ plays $x_{i'} \neq \underline{x}$ and the enforcer then plays $y_{i'} = y$, players transition to the i'-repenting state. If producer i plays $x_i \neq x$ (resp., some producer $i' \neq i$ plays $x_{i'} \neq \underline{x}$) and the enforcer plays $y_i \neq y$ (resp., $y_{i'} \neq y$), then players transition to the punishment state.

The punishment state is absorbing.

We refer to the special case of enforcer punishment plus repentance strategies with $\underline{x} = 0$ as *enforcer punishment plus full repentance*, and we refer to the case where $\underline{x} > 0$ as *enforcer punishment plus partial repentance*. The maximum levels of cooperation and punishment $(\check{x}, \check{y})$ that can be sustained with enforcer punishment plus full repentance are straightforward to characterize as

$$\begin{aligned} \check{x} &= g(\check{y}) + \delta(k-1) f(\check{x}) \\ \check{y} &= \left(\frac{\delta}{1-\delta}k - \delta(k-1)\right) f(\check{x}). \end{aligned} \tag{2}$$

Intuitively, a producer who deviates gains $\check{x}$ and loses $g(\check{y}) + \delta(k-1) f(\check{x})$ (the cost of being punished at level $\check{y}$ plus the lost benefit of others' cooperation in the next period), while an enforcer who deviates gains $\check{y}$ and loses $\left(\frac{\delta}{1-\delta}k - \delta(k-1)\right) f(\check{x})$ (the benefit of others' future cooperation, taking into account that only the deviant producer cooperates in the very next period).

Note that, under (one-time) enforcer punishment strategies—or enforcer punishment plus repentance strategies—punishments are not used on path and producers choose constant levels of cooperation. This implies that, whenever such a strategy profile sustains the maximum level of cooperation, it is also the optimal equilibrium in terms of utilitarian social welfare, provided the maximum level of cooperation is below the first-best level. If instead such strategies sustain a level of cooperation above the first best, the concavity of the functions f and g makes it possible to reduce cooperation in an incentive compatible manner and exactly achieve the first best.[10]

[10] For example, in the case of one-time enforcer punishments, the utilitarian optimal equilibrium would be characterized by $x = x^{FB}$ and $y = \frac{\delta}{1-\delta} k f\left(x^{FB}\right)$. Concavity of f and g then imply that $x \leq g(y)$ whenever $x^{FB} \leq x^*$.

2.3 Optimal Equilibrium in the Presence of an Enforcer

Our first main result characterizes the most cooperative equilibrium in the presence of an enforcer. The result states that either a one-time enforcer punishment equilibrium or a one-time enforcer punishment plus repentance equilibrium is always optimal.

Define the "efficient" level of punishment, y^E, as follows:[11]

$$y^E = \begin{cases} \infty & \text{if} \quad \lim_{y\to\infty} g'(y) \geq 1 \\ (g')^{-1}(1) & \text{if} \quad g'(0) > 1 > \lim_{y\to\infty} g'(y) \\ 0 & \text{if} \quad g'(0) \leq 1 \end{cases} .$$

With this notation, we can also determine the level of cooperation that can be sustained by enforcer punishment plus partial repentance strategies. In particular, if $y^E < \infty$, define the levels of cooperation $(\tilde{x}, \underline{x})$ under such a strategy profile by the system of equations

$$\begin{aligned} \tilde{x} &= g\left(y^E\right) + \delta(k-1)\left[f(\tilde{x}) - f(\underline{x})\right] \\ y^E &= \left(\frac{\delta}{1-\delta}k - \delta(k-1)\right) f(\tilde{x}) + \delta(k-1) f(\underline{x}) . \end{aligned} \tag{3}$$

The following theorem characterizes the most cooperative equilibrium.

Theorem 1 *If $y^E \geq y^*$, one-time enforcer punishment strategies are optimal, and the maximum level of cooperation is x^*.*

If $y^E \leq \check{y}$, one-time enforcer punishment plus full repentance strategies are optimal, and the maximum level of cooperation is $\check{x}$.

If $y^E \in (\check{y}, y^*)$, one-time enforcer punishment plus partial repentance strategies are optimal, and the maximum level of cooperation is $\tilde{x}$.

One special case of Theorem 1 bears particular emphasis: if $g'(y) \geq 1$ for all $y \in \mathbb{R}_+$ (so that one unit of disutility of effort in punishment incurred by the enforcer always inflicts at least one unit of disutility on a producer), then $y^E = \infty$, so Theorem 1 says that one-time enforcer punishments are optimal. Thus, under this condition, it is optimal for producers' incentives to be provided purely through specialized enforcement.

It is useful to break the intuition for this result into two parts, relating to why the optimal equilibrium can involve only specialized enforcement and why specialized enforcement takes the form of one-time punishment.

Why only specialized enforcement? To give a producer the strongest possible incentive to cooperate, her continuation payoff after a deviation must be made as low as possible. Ideally, her

[11] Intuitively, the "efficient" level of punishment is the one that equates the marginal cost of punishment to its marginal "benefit," which is the disutility imposed on a deviant producer.

continuation payoff would be reduced in two ways: the enforcer would punish her, and other producers would refuse to cooperate with her. However, the enforcer is willing to exert effort in punishing the deviator only if he is subsequently rewarded with cooperation from the producers. Since cooperation benefits both producers and the enforcer, there is no way for producers to reward the enforcer for punishing the deviant producer without also benefiting the deviator herself.[12] Society must then choose between incentivizing the enforcer to punish deviators by immediately restoring cooperation (enforcer punishment strategies), or by reducing enforcer punishments and instead providing incentives by withdrawing cooperation following a deviation (enforcer punishment plus repentance strategies).

We can quantify the tradeoff between incentivizing the enforcer to punish and incentivizing producers by withdrawing cooperation as follows. Consider a history following a producer deviation. The direct effect of reducing another producer's level of cooperation at such a history by one unit is to reduce the deviator's payoff by $f'(x)$ units. This effect increases on-path incentives for cooperation. This direct effect is countered by the indirect effect of reducing the maximum level of punishment the enforcer is willing to impose on the deviator. In particular, reducing the producer's level of cooperation by one unit decreases the amount of punishment the enforcer can be induced to provide by $f'(x)$ units, and each unit of reduced punishment increases the deviator's payoff by $g'(y)$. Thus, the indirect effect reduces on-path incentives for cooperation by $f'(x)g'(y)$. Consequently, the overall impact of withdrawing producer cooperation following a deviation on on-path producer incentives is negative if and only if $g'(y) \geq 1$—that is, if and only if $y \leq y^E$. Therefore, if $y^* \leq y^E$ then, after a producer deviation, it is better to rely solely on enforcer punishments rather than reducing other producers' cooperation levels. Conversely, if $\check{y} \geq y^E$ then reducing other producers' cooperation levels as far as possible is optimal. Finally, if $y^E \in (\check{y}, y^*)$ then it is optimal to reduce other producers' cooperation levels to the point where the efficient punishment level y^E is just incentive-compatible for the enforcer.

It is possible to give a short, heuristic proof of Theorem 1, under the assumption that strategies take the form of enforcer punishment plus full, partial, or no repentance. (However, most of the substance of Theorem 1 is showing that this assumption is without loss.) The problem of choosing an on-path cooperation level x, an off-path cooperation level $\underline{x}$, and a punishment level y to maximize on-path cooperation is

$$\max_{x,\underline{x},y} x$$

[12] In Appendix C, we study how the structure of the optimal equilibrium changes if we allow for *ostracism*—the practice of excluding only deviators from the benefits of cooperation.

subject to

$$
\begin{aligned}
x &\leq g(y) + \delta(k-1)\left[f(x) - f(\underline{x})\right] \quad \text{and} && (4) \\
y &\leq \left(\frac{\delta}{1-\delta}k - \delta(k-1)\right) f(x) + \delta(k-1) f(\underline{x}), && (5)
\end{aligned}
$$

where (4) and (5) are the incentive-compatibility constraints for producers and enforcers, respectively. Using the constraints to substitute for x and y, a necessary condition for optimality is that $\underline{x}$ solves the unconstrained problem

$$
\max_{\underline{x}} g\left(\underbrace{\left(\frac{\delta}{1-\delta}k - \delta(k-1)\right) f(x) + \delta(k-1) f(\underline{x})}_{=y}\right) + \delta(k-1)\left[f(x) - f(\underline{x})\right].
$$

The derivative with respect to $\underline{x}$ equals

$$
\delta(k-1) f'(\underline{x})\left[g'(y) - 1\right].
$$

Hence, if $g'(y) \geq 1$ for all incentive-compatible y (i.e., if $y^E \geq y^*$), then off-path cooperation should be maximized, yielding one-time enforcer punishment. If $g'(y) \leq 1$ when y is chosen to bind (5) with $\underline{x} = 0$ (i.e., $y^E \leq \check{y}$), then off-path cooperation should be minimized, yielding one-time enforcer punishment plus full repentance. Finally, if $g'(y) < 1$ when $\underline{x} = 0$ but $g'(y) > 1$ when $\underline{x} = x$ (i.e., $y^E \in (\check{y}, y^*)$), then $\underline{x}$ should be chosen so that $y = y^E$, yielding one-time enforcer punishment plus partial repentance.

We also emphasize that the result that one-time enforcer punishment strategies are optimal whenever $g'(y) \geq 1$ for all $y \in \mathbb{R}_+$ holds independently of the production technology f. Intuitively, improvements in the production technology increase the greatest level of cooperation that can be sustained with both specialized enforcement and community enforcement, and such improvements cancel out when comparing the two kinds of equilibria. An interesting implication is that, provided the efficiency of the production and punishment technologies are positively related across different societies, Theorem 1 predicts that societies with more effective technologies should rely purely on specialized enforcement, while societies with less effective technologies should rely on a mix of community enforcement and specialized enforcement.

Why are deviators punished only once? One might have conjectured that, to provide the harshest deterrent against a deviation, the enforcer should punish a deviator several times for the same transgression. The reason why this does not occur in the optimal equilibrium is that, with multiple rounds of punishment, the deviator would not be willing to exert as much effort in cooperation during her punishment phase, and the deviator's continuation payoff from being punished once and then returning to full cooperation is weakly lower than her continuation payoff from being punished repeatedly while shirking. (In fact, it is strictly lower, as the deviator's own future cooperation can be used to give the enforcer additional incentives to punish her).

Another way of seeing the intuition is to observe that, in the most cooperative equilibrium, a producer is indifferent between following her equilibrium strategy and following the policy of *always* shirking. If the enforcer were asked to spread the punishment for each individual instance of shirking over multiple periods, this would reduce the total punishment faced by a producer who always shirks, and would therefore reduce the maximum sustainable level of cooperation.

Our results so far show that one-time enforcer punishments are always part of the optimal equilibrium. In particular, society should deploy all available enforcers. However, we also note that enforcer punishments provide very little benefit when the specialized enforcement technology is sufficiently ineffective, so in this case pure community enforcement becomes approximately optimal.

Proposition 2 *For $\varepsilon > 0$, there exists $\eta > 0$ such that if $g'(y) < \eta$ for all $y \in \mathbb{R}_+$ then the grim trigger strategy profile with cooperation level $\hat{x}$ attains within ε of the maximum level of cooperation.*

2.4 The Tradeoff Between Production and Enforcement

We have characterized the optimal equilibrium both in a group consisting of only $k+1$ producers (Proposition 1) and in a group consisting of k producers and a single enforcer (Theorem 1). Given these results, it is straightforward to compare the resulting level of social welfare in the two cases. This answers the question of when a group of producers would gain from designating one of their number as an enforcer.

Theorem 2 *Assume the maximum level of cooperation is below the first-best level, with or without an enforcer. If*

$$g'(0) \leq \frac{1-\delta}{1+\delta k},$$

then utilitarian social welfare in an optimal equilibrium is higher without an enforcer (i.e., with $k+1$ producers, rather than with k producers and one enforcer). Conversely, fixing the values of the other parameters of the model and assuming g' is bounded away from 0, there exists $\bar{\delta} < 1$ (resp., $\alpha < \infty$ or $\bar{k} < \infty$) such that the maximum level of cooperation is higher with an enforcer if $\delta > \bar{\delta}$ (resp., $g'(y) > \alpha$ for all $y \in \mathbb{R}_+$ or $k > \bar{k}$.)

Consequently, if the players are patient, the group is large, or the specialized enforcement technology is effective, then it is optimal to designate an individual as an enforcer. Conversely, if the players are impatient, the group is small, and the specialized enforcement technology is ineffective, it is optimal for all agents to remain producers. The comparative static with respect to δ—wherein higher δ favors specialized enforcement—is the most subtle of these results. It is a consequence of the $\frac{1}{1-\delta}$ term in the formula for y^*, which was explained earlier.

3 Large Group Model: Random Matching

The small group model considered above has the advantage of bringing out as simply as possible the tradeoff between community and specialized enforcement and the possible optimality of one-time enforcer punishment equilibria. However, in large groups it is more realistic to assume that only smaller subsets of the population interact directly in each period, and in general there is also no reason to suppose that each producer is monitored by only a single enforcer. In addition, in large groups it is also more realistic to assume that players can only observe the actions of individuals with whom they interact directly. In this section, we show that exactly the same insights—in fact, essentially the same mathematical results—generalize from small groups to large groups under the assumption of perfect monitoring. Imperfect private monitoring is studied in Section 4.

Specifically, we generalize the small group model of Section 2 as follows. The group now consists of $(k+l)\,n$ players, with $k, l, n \geq 1$. Out of the $(k+l)\,n$ players, kn of them are producers and ln of them are enforcers. The small group model is thus the special case with $n = 1$ and $l = 1$.

Denote the set of producers by P, the set of enforcers by E, and the set of all players by I. In every period $t = 0, 1, 2 \ldots$, the players break into n matches uniformly at random, where each match consists of k producers and l enforcers. Denote the match containing player i by M_i.

The following two-stage game is played simultaneously in each match M.

1. *Cooperation Stage:* Each producer i in match M chooses a level of cooperation $x_i \in \mathbb{R}_+$ before observing the identities of the other players in M.[13] The vector $(i, x_i)_{i \in M \cap P}$ is perfectly observed by all players in M. Choosing cooperation level x_i costs x_i for player i, and benefits every other player $k \neq i$ in M by $f(x_i)$.

2. *Punishment Stage:* Each enforcer $j \in M$ then chooses a level of punishment $y_{ji} \in \mathbb{R}_+$ for each producer $i \in M \cap P$. The vector $(j, i, y_{ji})_{j \in M \cap E, i \in M \cap P}$ is perfectly observed by all players in M. Choosing punishment level y_{ji} costs y_{ji} for player j, and hurts player i by $g(y_{ji})$.

Producer i's stage payoff is thus

$$\sum_{i' \in M_i \cap P \setminus i} f(x_{i'}) - x_i - \sum_{j \in M_i \cap E} g(y_{ji}),$$

[13] We refer to the feature that producers act without knowing their partners' identities as *partial anonymity.* This assumption—which naturally does not arise in the small group model—plays an important role in the large group model, as we discuss in Section 5.

and enforcer j's stage payoff is

$$\sum_{i \in M_j \cap P} \left(f\left(x_i\right) - y_{ji} \right).$$

We refer to the pair $\left((i, x_i)_{i \in M \cap P}, (j, i, y_{ji})_{j \in M \cap E, i \in M \cap P} \right)$ as the *outcome* of match M. Throughout the paper, we maintain the assumption that players perfectly observe the outcomes of their own matches, while varying players' information about the outcomes of other matches. With *perfect monitoring*, players observe the outcomes of all matches at the end of each period. We will also consider two different versions of *private monitoring*—detailed below—where players have less information about what goes on outside their own matches. In all versions of the model, we let h_i^t denote a generic history for player i at the beginning of period t, and we omit the subscript in the perfect monitoring case. The trivial initial history is denoted by h^0. We also denote a generic strategy for player i by σ_i. For example, if player i is a producer then $\sigma_i\left(h_i^t\right) \in \Delta\left(\mathbb{R}_+\right)$ denotes player i's mixed action at history h_i^t.

Players maximize expected discounted payoffs with common discount factor δ. The solution concept is weak perfect Bayesian equilibrium (PBE), with the additional requirement that the equilibrium assessment is derived from a common conditional probability system (Myerson, 1991).[14]

Our goal is again to characterize the most cooperative equilibrium in this game. (The definition of the most cooperative equilibrium is the same as in Section 2, and once again the most cooperative equilibrium is also utilitarian efficient, so long as the maximum level of cooperation is below the first-best level x^{FB}, which is now given by $(k + l - 1) f'\left(x^{FB}\right) = 1$). As in the small group model, contagion strategies, one-time enforcer punishment strategies, and one-time enforcer punishment plus repentance strategies will play a key role. The definitions of all of these strategy profiles are exactly the same as in Section 2. However, the resulting formulas for the maximum level of cooperation sustainable with these strategy profiles must be adjusted to account for the presence of l enforcers in each match and the fact that, following a producer deviation, each enforcer matches with the deviator in the following period with probability $\frac{1}{n}$. The resulting formulas are as follows.

- One-time enforcer punishment strategies:

$$\begin{aligned} x^* &= l g\left(y^*\right) \\ y^* &= \frac{\delta}{1-\delta} k f\left(x^*\right). \end{aligned} \tag{6}$$

[14] This requirement implies that PBE are subgame perfect. We use PBE even for the perfect monitoring version of the model because enforcers' information sets are not roots of proper subgames. Another approach would have been to discretize the action space and use sequential equilibrium. This would lead to the same results, except that with discrete actions the equilibria we characterize would be only approximately rather than exactly optimal.

- One-time enforcer punishment plus full repentance strategies:

$$\begin{aligned} \check{x} &= lg(\check{y}) + \delta(k-1) f(\check{x}) \\ \check{y} &= \left(\frac{\delta}{1-\delta}k - \delta\left(k - \frac{1}{n}\right)\right) f(\check{x}). \end{aligned} \tag{7}$$

- One-time enforcer punishment plus partial repentance strategies

$$\begin{aligned} \tilde{x} &= lg\left(y^E\right) + \delta(k-1)\left[f(\tilde{x}) - f(\underline{x})\right] \\ y^E &= \left(\frac{\delta}{1-\delta}k - \delta\left(k - \frac{1}{n}\right)\right) f(\tilde{x}) + \delta\left(k - \frac{1}{n}\right) f(\underline{x}). \end{aligned} \tag{8}$$

In addition, in this section the definition of the efficient level of punishment y^E is given by

$$y^E = \left\{ \begin{array}{ccc} \infty & \text{if} & \lim_{y\to\infty} g'(y) \geq m \\ (g')^{-1}(m) & \text{if} & g'(0) > m > \lim_{y\to\infty} g'(y) \\ 0 & \text{if} & g'(0) \leq m \end{array} \right\},$$

where $m := \left(\frac{k-1}{k-1/n}\right)\frac{1}{l}$. Note that $m \in \left[\left(\frac{k-1}{k}\right)\frac{1}{l}, \frac{1}{l}\right]$ for all k, l, n.

With these modified definitions, Theorem 1 generalizes verbatim. Recall that all results in the current section concern perfect monitoring.

Theorem 3 *If $y^E \geq y^*$, one-time enforcer punishment strategies are optimal, and the maximum level of cooperation is x^*.*

If $y^E \leq \check{y}$, one-time enforcer punishment plus full repentance strategies are optimal, and the maximum level of cooperation is $\check{x}$.

If $y^E \in (\check{y}, y^*)$, one-time enforcer punishment plus partial repentance strategies are optimal, and the maximum level of cooperation is $\tilde{x}$.

Recalling that $m \leq \frac{1}{l}$ for all k, l, n, Theorem 3 implies that, if $g'(y) \geq \frac{1}{l}$ for all $y \in \mathbb{R}_+$ (so that one unit of disutility of effort in punishment incurred *in total* by the enforcers in a given match always inflicts at least one unit of disutility on a producer), then one-time enforcer punishments are optimal. Thus, as in the small group model, under a mild condition it is optimal for producers' incentives to provided purely through specialized enforcement.

The intuition for Theorem 3 is essentially the same as in the small group model. To understand the precise formula for the constant m, note that the direct effect of reducing another producer's level of cooperation on a deviant producer's payoff is now $\frac{k-1}{kn-1}f'(x)$ (as $\frac{k-1}{kn-1}$ is the probability that the deviator matches with a given producer in any period), while the indirect effect coming through a reduction in the maximum incentive-compatible enforcer punishment is $\frac{l}{n}f'(x)\,g'(y)$ (as

$\frac{1}{n}$ is the probability that a given enforcer matches with a given producer, and there are l enforcers per match). The indirect effect therefore dominates if and only if $g'(y) \geq m$, or $y \leq y^E$.

We also have the following generalization of Theorem 2. This result captures the intuition that, if a social planner has the option to allocate some of the enforcers back to production and the specialized enforcement technology is not very effective, she may prefer to forgo the limited increase in the level of cooperation that these enforcers afford, and also consequently rely on grim trigger strategies strategies.

Theorem 4 *Suppose a social planner can reallocate some of the enforcers back to production, or equivalently chooses k and l (as well as selecting an equilibrium) subject to $k \geq \bar{k}$ and $k + l = s$ to maximize utilitarian social welfare. Suppose also that the maximum level of cooperation is below the first-best level. If*

$$g'(0) \leq \min\left\{\left(\frac{\bar{k}-1}{\bar{k}-\frac{1}{n}}\right)\frac{1}{s-\bar{k}}, \frac{1}{\frac{1}{n}+\frac{\delta}{1-\delta}s}\right\} \tag{9}$$

then the social planner would prefer to have all agents become producers (i.e., set $k = s$) and support cooperation using grim trigger strategies. Conversely, fixing the values of the other parameters of the model and assuming g' is bounded away from 0, there exists $\bar{\delta} < 1$ (resp., $\alpha < \infty$ or $\bar{s} < \infty$) such that the maximum level of cooperation is higher with one enforcer per group than with no enforcers if $\delta > \bar{\delta}$ (resp., $g'(y) > \alpha$ for all $y \in \mathbb{R}_+$ or $s > \bar{k}$.)

Finally, we note that Proposition 2 also applies identically to the large group model. The proof of this result in the appendix thus allows for general l and n.

4 Large Group Model: Private Monitoring

One traditional motivation for studying community enforcement in large groups is the desire to understand how groups can sustain cooperation when individuals have limited information about each other's past behavior. In this more general setting with private monitoring, fully characterizing optimal equilibria at a fixed discount factor (as we have done under perfect monitoring) appears intractable. Nevertheless, we establish two useful results highlighting the robustness of the economic forces we have emphasized so far. First, under *general network monitoring*, where players observe the outcomes of their own matches, as well as possibly the outcomes of some other random matches in the population, we provide conditions under which one-time enforcer punishment strategies outperform pure contagion strategies. Second, we consider a setting with *observable last matches*, where a player observes the outcomes of her own matches and the outcome of each of her current partner's most recent matches. With this information structure, enforcer punishment strategies continue to sustain the same level of cooperation as with perfect monitoring, which implies that they must remain globally optimal (if $y^E \geq y^*$) unless it is possible to sustain *more* cooperation with observable last matches than with perfect monitoring. While this can be possible, we also establish that enforcer punishment equilibria continue to be optimal if enforcers are perfectly informed (which may be a consequence of their organization in an information-sharing institution, such as a police force), or if we impose a requirement of stability in the face of individual trembles.

4.1 General Network Monitoring

The setting considered here is one of general *network monitoring* (e.g., Wolitzky, 2013). At the end of each period t, a *monitoring network* $L_t = (l_{i,j,t})_{i,j \in I \times I}$, $l_{i,j,t} \in \{0,1\}$ is drawn independently from a fixed probability distribution μ on $\{0,1\}^{|I|^2}$. We assume that $\Pr^{\mu}\left((l_{i,j,t})_{i,j \in I \times I}\right) = \Pr^{\mu}\left((l_{\phi(i),\phi(j),t})_{i,j \in I \times I}\right)$ for any permutation $\phi : I \to I$, so the distribution over networks is invariant to relabeling the players. Player i perfectly observes the outcome of match M_0 if and only if $l_{i,j,t} = 1$ for some $j \in M_0$. Otherwise, player i observes nothing about the outcome of match M_0. Assume that $l_{i,i,t} = 1$ with probability one, so players always observe the outcome of their own matches. We compare the performance of contagion strategies and one-time enforcer punishment strategies in this setting.

With contagion strategies, let d_t be the expected number of producers who become *infected* (i.e., enter the punishment state) within t periods of a producer deviation (see the appendix for a formal definition). Intuitively, d_t is the expected number of producers who have observed a producer who has observed a producer who... has observed the deviator within t periods. It follows from standard arguments that the greatest level of cooperation that can be sustained with contagion strategies is given by

$$\hat{x} = (1-\delta) \sum_{t=0}^{\infty} \delta^t \frac{k-1}{kn-1} (d_t - 1) f(\hat{x}).$$

With one-time enforcer punishment strategies, let q_t be the expected number of producers who become infected within t periods of an unpunished producer deviation (once again the details are in the appendix). Note that a player now becomes infected only if *both* a producer and an enforcer in a match she observes are already infected, as only then does she see a producer's failure to cooperate go unpunished. Infection therefore spreads more slowly with enforcer punishment strategies than with contagion strategies, and in particular, q_t is always less than d_t. We will show that an upper bound on the greatest level of cooperation and punishment that can be sustained with enforcer punishment strategies is given by

$$\begin{aligned} x^* &= lg(y^*), \\ y^* &= \sum_{t=0}^{\infty} \delta^t \frac{q_t}{n} f(x^*). \end{aligned}$$

We also show that the resulting strategy profile is indeed part of a PBE whenever x^* is sufficiently high (which holds, for example, if δ is sufficiently high).[15]

We can now inspect the formulas for $\hat{x}$ and x^* and make some basic observations about the relative performance of contagion and enforcer punishment strategies. First, when the specialized enforcement technology is more effective (i.e., g is steeper), enforcer punishment strategies have an advantage over contagion strategies. Second, to the extent that d_t is strictly greater than q_t, contagion strategies have an advantage. Third, as both d_t and q_t converge to kn as $t \to \infty$, this advantage of contagion strategies vanishes when δ is close to 1. Indeed, enforcer punishment strategies have a clear advantage when δ is close to 1, owing to the $(1-\delta)$ term in the definition of $\hat{x}$; the interpretation of this term is the same as in Section 2.[16]

The next theorem formalizes this comparison. The first part shows that, when the specialized enforcement technology is sufficiently effective, enforcer punishment strategies support more cooperation than contagion strategies. The second part establishes the same conclusion when the discount factor δ is sufficiently high.

Theorem 5 *With general network monitoring,*

1. *There exists α such that, if $g'(y) > \alpha$ for all $y \in \mathbb{R}_+$, then one-time enforcer punishment strategies form a PBE strategy profile and support greater cooperation than contagion strategies.*

2. *Assume $\lim_{y\to\infty} g(y) = \infty$.[17] Then there exists $\bar{\delta}$ such that, if $\delta > \bar{\delta}$, then one-time enforcer punishment strategies form a PBE strategy profile and support greater cooperation than contagion strategies.*

We note as well that comparing one-time enforcer punishment strategies and contagion strategies is not as ad hoc at it might seem, as there is a sense in which contagion strategies are optimal among all equilibria in which enforcers never punish. Wolitzky (2013) shows that, under general

[15] To see why such a condition is required, consider the incentives of a producer in the infected state who finds herself with the belief that all of the other producers in her match are in the normal state, while exactly one of the enforcers in her match is in the infected state. If this producer works, she avoids being punished at level y^* by each one of the $l-1$ enforcers in her match in the normal state, but also avoids triggering contagion (because, if she shirked, the infected enforcer's failure to punish her would trigger contagion). When x^* is sufficiently high, this new incentive for cooperation coming from the desire to avoid triggering contagion is necessarily less than the incentive coming from being punished at level y^* by the l^{th} enforcer. In this case (but not otherwise), the fact that the producer is indifferent between working and shirking on path implies that she prefers to shirk when any enforcer is infected.

[16] Presumably, *optimal* equilibria in this setting would take advantage of enforcers' ability to punish while also providing incentives for spreading information faster than one-time enforcer punishment strategies. As providing incentives for strategic communication of this kind is beyond the scope of this paper, we content ourselves with comparing the performance of one-time enforcer punishment strategies and contagion strategies.

[17] The resulting asymmetry between the functions f and g is not essential for this result. If the assumption that f is bounded is relaxed, as discussed in Section 5, the result still holds as long as $\lim_{x\to\infty} f'(x) < 1/l\,(k-1)$, which is consistent with $f = g$ and $\lim_{y\to\infty} g(y) = \infty$.

network monitoring without enforcers, contagion strategies attain the maximum level of cooperation (provided the realized monitoring network is observable). Thus, whenever enforcer punishment strategies outperform contagion strategies, they outperform any equilibrium that does not rely on the enforcers.[18]

4.2 Observable Last Matches

We now turn to the second of the two private monitoring environments we consider: *observable last matches*. This setting, where players observe only the outcomes of their own matches and their current partners' most recent matches, is a natural benchmark and is also tractable enough for us to fully generalize the perfect monitoring results.

One-Time Enforcer Punishment Strategies and Contagion Strategies

We first establish that both enforcer punishment strategies and contagion strategies do exactly as well with observable last matches as they do with perfect monitoring. In particular, any comparison between enforcer punishment strategies and contagion strategies with observable last matches is exactly the same as in the perfect monitoring case.

However, as in the previous section, existence of a enforcer punishment equilibrium requires an additional condition. In what follows, let x^*, y^*, and $\hat{x}$ be defined as in Section 3, and let $\dot{x}$ be the positive solution to $\dot{x} = l\delta (k-1) f(\dot{x})$.

Theorem 6 *With observable last matches,*

1. *If $x^* \geq \dot{x}$, then the one-time enforcer punishment strategy profile with cooperation level x^* and punishment level y^* is a PBE strategy profile. Furthermore, x^* is an upper bound on the level of cooperation in any one-time enforcer punishment equilibrium.*

2. *The contagion strategy profile with cooperation level $\hat{x}$ is a PBE strategy profile. Furthermore, $\hat{x}$ is an upper bound on the level of cooperation in a contagion equilibrium.*

The intuition for this result is simple. Contagion following a producer deviation with contagion strategies, or following an enforcer deviation with enforcer punishment strategies, spreads more

[18] However, recall that our notion of optimality is in terms of supporting a higher level of cooperation. As we have emphasized, this notion corresponds to optimality in terms of utilitarian social welfare *if* this maximum level of cooperation is below the first-best level, but not necessarily otherwise. This caveat is especially important for high discount factor results like part 2 of Theorem 5, as for very high discount factors both the most cooperative one-time enforcer punishment equilibrium and the most cooperative contagion equilibrium are sure to involve an inefficiently high level of cooperation, so the one that supports the higher level of cooperation will actually be worse in terms of welfare. Thus, the main point of Theorem 5 is not that one-time enforcer punishment strategies outperform contagion strategies in the $\delta \to 1$ limit per se, but rather that they outperform contagion strategies for moderately high discount factors where the maximum level of cooperation may still be below the first-best level.

slowly under private monitoring than under perfect monitoring. Nevertheless, the implications for the deviating agent's payoffs are the same as with perfect monitoring, because, when all agents observe the behavior in their partners' last match, the deviator herself always starts suffering the consequences of contagion immediately.

Informed Enforcers

Theorem 6 shows that enforcer punishment strategies can sustain as much cooperation with observable last matches as with perfect monitoring. The question remains whether it is possible to sustain more cooperation with observable last matches than with perfect monitoring, or alternatively if enforcer punishment strategies remain globally optimal with observable last matches (when $y^E > y^*$). In the next subsection, we will see that the former possibility can sometimes arise. That result notwithstanding, we show that, when enforcers have superior information relative to producers, enforcer punishment strategies are indeed globally optimal, under a simple equilibrium refinement in the spirit of partial anonymity. Specifically, we consider the situation with *informed enforcers*, where enforcers perfectly observe all past actions, while producers continue to observe only their partners' most recent matches.

The following condition is in the spirit of the partial anonymity assumption of Section 3.

Non-Discrimination *For every producer i, complete history of play h^t, and pair of players j, k such that $M_j^t = M_k^t$, we have*

$$\mathbb{E}_{h_i^{t+1}}\left[\sigma_i\left(h_i^{t+1}\right)|h^t, i \in M_j^{t+1}\right] = \mathbb{E}_{h_i^{t+1}}\left[\sigma_i\left(h_i^{t+1}\right)|h^t, i \in M_k^{t+1}\right].$$

That is, the distribution over producer i's period-$t+1$ actions is independent of whether i matches with j or k in period $t+1$. This requirement is only imposed for players j and k who are themselves matched at period t, so that producer i's behavior can depend on the outcomes of the period-t matches she observes, but not on which members of those matches she finds herself matched with in period $t+1$. Non-discrimination thus says that a producer's behavior cannot depend on her partners' identities, except insofar as this is informative about past play. Both one-time enforcer punishment strategies and contagion strategies are clearly non-discriminatory.

Theorem 7 *Suppose producers observe their partners' last matches while enforcers are perfectly informed. If $g'(y) \geq m$ for all y, then one-time enforcer punishment strategies with cooperation level x^* and punishment level y^* sustain the maximum level of cooperation among all non-discriminatory equilibria.*[19]

[19] Wth informed enforcers, one-time enforcer punishment strategies constitute a PBE strategy profile even if $x^* < \dot{x}$.

To provide intuition for this result, let us first revisit the case of perfect monitoring. An explanation for why one-time enforcer punishment strategies are optimal with perfect monitoring is that reducing producer j's level of cooperation at a history h_j^{t+1}, after producer i deviates at history h_i^t and is punished by enforcer k, has a direct positive effect on producer i's on-path incentives for cooperation at history h_i^t of

$$\frac{k-1}{kn-1} \Pr\left(h_j^{t+1}|h_i^t\right) \mathbb{E}\left[f'\left(x_j^{t+1}|h_j^{t+1}\right)\right],$$

and has a indirect negative effect of

$$\frac{l}{n}\mathbb{E}\left[\Pr\left(h_j^{t+1}|h_k^t\right)|h_i^t\right] \mathbb{E}\left[f'\left(x_j^{t+1}|h_j^{t+1}\right)\right] \mathbb{E}\left[g'\left(\tilde{y}\right)\right],$$

for some random variable $\tilde{y}$. If monitoring is perfect, or if enforcers always have finer information than producers, then we have

$$\Pr\left(h_j^{t+1}|h_i^t\right) = \mathbb{E}\left[\Pr\left(h_j^{t+1}|h_k^t\right)|h_i^t\right], \tag{10}$$

so the indirect effect outweighs the direct effect whenever $g'(y) \geq m$ for all y. This explains why enforcer punishment strategies are optimal with perfect monitoring (Theorem 3) or with private monitoring with informed enforcers (Theorem 7). However, if monitoring is private and enforcers do not necessarily have finer information than producers, then (10) may fail, and the differing beliefs of enforcers and producers may be exploited to provide stronger on-path incentives than are possible in enforcer punishment strategies. We now construct an example with these features.

A Counterexample: Departures from One-Time Enforcer Punishments

In this subsection, we show by example that, when (10) is not satisfied, it may be possible to support greater cooperation under private monitoring than public monitoring, and this may involve multiple rounds of punishments of a deviator.[20]

Let $n = k = 2$ and $l = 1$. Thus, there are two enforcers and four producers, and every period they randomly split into two groups, each consisting of one enforcer and two producers. Assume that players observe the outcome of their own matches, and that producers—but not enforcers—in addition observe the outcome of each of their partner's most recent matches. This informational edge for the producers is for simplicity; in Appendix B, we sketch a more complicated example without this feature. To complete the description of the physical environment of the example,

[20] Moreover, punishments in this counterexample also take a graduated form, similar to a pattern identified by Ostrom (1990) as an important tool for sustaining cooperation under imperfect information. The intuition here is different from Ostrom's, however: the advantage of graduated punishments in the current setting is that it takes time to build up differences in beliefs among individuals, and these differing beliefs can then be exploited to provide harsher punishments than are possible with perfect monitoring.

assume that $f(x) = 100\sqrt{x}$, $g(y) = y$, and $\delta = .1$. These parameters satisfy our condition for one-time enforcer punishment strategies to be optimal under perfect monitoring ($y^E \geq y^*$).

As we have seen, the highest level of cooperation that can be sustained with one-time enforcer punishment strategies, x^*, is given by $x^* = g\left(\frac{\delta}{1-\delta} 2f(x^*)\right) = \frac{.1}{1-.1} 2\left(100\sqrt{x^*}\right)$, or $x^* \approx 493.8272$.

In contrast, we now describe an equilibrium that sustains a cooperation level of (exactly) 493.830. We call it the *three strikes and you're out* (*3SYO*) equilibrium. In what follows, let $x_1 = 493.830$, $x_2 = 494.102$, $x_3 = 502.058$, and $y = 493.828$.

Producers' strategies:

- On path: play x_1.
- If you play $x < x_1$: play x_2 for one period, then go back to x_1.
- If (i) you play $x' < x_2$ in the period after playing $x < x_1$, and (ii) you matched with the same producer in both of these periods but matched with different enforcers: play x_3 for one period, then go back to x_1. If (i) holds but not (ii): go back to x_1 immediately.
- If you see the same producer play $x < x_1$, then $x' < x_2$, and then $x'' < x_3$, or if you are seen following such a sequence by the same producer, or if you see a producer play $x < x_1$ and see the corresponding enforcer fail to punish her: play 0 forever.

Enforcers' strategies:

- If you see a unique producer play $x < x_1$, punish her at level y. Do not punish anyone if you see two producers deviate.
- If you fail to punish a producer who plays $x < x_1$, or if you see the same producer take actions below x_1 three times in a row, stop punishing forever.

Intuitively, the key difference between the one-time enforcer punishment equilibrium and the 3SYO equilibrium is that, with the latter, if a producer shirks three times in a row and is monitored by the same producer but different enforcers, then after the third time she shirks she is "punished" *both* by the enforcer (who punishes at level y, as usual), *and* by the other producer (who shirks forever, as in a contagion equilibrium). The reason why the enforcer is willing to punish at level y even though the other producer is about to start shirking is that he does not realize that this is what is happening: he has seen the deviator shirk at most once before, so when he sees her shirk again, he thinks this is at most the second straight time she has shirked. He is then certain that the deviator (and the other producer) will return to cooperation in the next period if he punishes, while contagion will start if he does not punish, so he has an incentive to punish. Thus, the 3SYO equilibrium exploits the difference in beliefs between enforcer and producer at such a history to punish the deviator with both coercive punishment and contagion.[21]

Proposition 3 *Under the parameter values presented above, the 3SYO strategy profile is an equilibrium, and therefore one-time enforcer punishment strategies are not optimal.*

Stability

We conclude our analysis by providing another reason why one-time enforcer punishment strategies may be optimal under private monitoring, even when (10) is not satisfied: they are optimal among all equilibria satisfying a simple 1-period stability refinement. While the argument is simple, we believe it is potentially important in light of empirical accounts of how small transgressions can lead to large societal breakdowns in the absence of centralized law enforcement (e.g., Lewis, 1994).

To define this notion of stability, we restrict attention to *deterministic* strategy profiles, defined as profiles where x_i^t and y_{ji}^t are degenerate random variables for all i, j, t.[22]

Definition 3 *A deterministic equilibrium satisfies* Stability *if, whenever a single player i deviates at an on-path history in period t, play returns to the equilibrium path $\left(x_i^\tau, y_{ji}^\tau\right)_{i\in P, j\in E}^{\tau\in\{t+1,\ldots\}}$ in period $t+1$.*

Note that if all players "tremble" with probability ε when choosing their actions, then an equilibrium that fails to satisfy Stability is knocked off its equilibrium path in each period with probability of order ε, while an equilibrium that satisfies Stability is knocked off path with probability of order at most ε^2. In this sense, equilibria that satisfy Stability are more robust to trembles than are equilibria that fail to satisfy this condition.

Theorem 8 *With observable last matches, the one-time enforcer punishment strategy profile with cooperation level x^* and punishment level y^* is the most cooperative deterministic equilibrium satisfying Stability.*

5 Discussion of Model Assumptions

We briefly discuss the role of several key assumptions, indicating how they impact our results and their interpretation, focusing for brevity on the results from Section 3.

[21] The reason why this "extra punishment" at an off-path history allows us to sustain more cooperation on path is as follows: If a producer can be punished "extra hard" after she shirks three times, then she can be asked to work extra hard after she shirks twice. Similarly, if she has to work extra hard after she shirks twice, then she can also be asked to work harder after she shirks once. Finally, if she has to work harder after she shirks once, then she can also be induced to work harder on path.

[22] Equivalently, a determinstic strategy profile is a profile of pure strategies that do not condition on the match realizations.

5.1 Assumptions about the Role of Enforcers

Enforcers can only punish producers: When allowing for multiple enforcers in each match ($l \geq 2$), we have assumed that enforcers can punish only producers and not other enforcers. Changing this assumption by also letting enforcers punish each other would change very little about our results. Specifically, one would redefine one-time enforcer punishment strategies to specify that, if an enforcer fails to punish a deviant producer in period t, then in period $t+1$ there is no production and the deviant enforcer is punished, while cooperation resumes in period $t+2$. This changes the formula for y^* from $\frac{\delta}{1-\delta} kf(x^*)$ to $\frac{\delta}{1-\delta}(kf(x^*) + (l-1)g(y^*))$. (The formula for x^* as a function of y^* stays the same.) With this change, one-time enforcer punishment strategies remain optimal whenever $g'(y) \geq m$ for all $y \in \mathbb{R}_+$. The other parts of Theorem 3 require similar small modifications. As this is a substantive result, we state and prove this modified theorem in Appendix A.

Thus, all that changes when we allow enforcers to punish each other is that they themselves are now incentivized by a mix of withdrawn cooperation and coercive punishment, rather than by the breakdown of cooperation alone. In particular, whether enforcers can punish each other or not does not affect the optimal mode of enforcement for producers, which is our main focus.

An alternative way of extending both the small group and large group models along these lines would be to introduce a *hierarchy* of enforcers with K levels, where "level 1" enforcers can punish producers, "level 2" enforcers can punish level 1 enforcers, and so on. The structure of one-time enforcer punishment equilibria also extends to this setting in a natural way, where each enforcer is incentivized by the threat of punishment from enforcers one level up, and the top-level enforcers are incentivized by the threat of contagion among producers. This variant gives a more realistic model of modern law enforcement: cooperation throughout society does not break down the moment a low-level policeman fails to do his job, but only if this is followed by a breakdown of enforcement at all higher levels.

Partial anonymity: In our large group model, producers choose how much to cooperate before observing their partners' identities, while identities are revealed before enforcers act. Our results also apply exactly if, alternatively, players are completely anonymous and their identities are never revealed. We prefer our baseline assumptions because they emphasize that, even though enforcers have the ability to identify and punish a deviator repeatedly, the optimal equilibrium involves only a single round of punishment.

On the other hand, the assumption that players are anonymous at the cooperation stage plays an important role in our analysis of the large group model. (Of course, this assumption is trivially satisfied in the small group model, as without random matching there is no uncertainty as to one's partners' identities). Without this assumption, it may be possible to partially exclude a deviator from future cooperation by reducing the cooperation level in future matches she is a part of, without simultaneously excluding the enforcers who punished her. This would then allow deviators to be punished more harshly.

We note that all our perfect monitoring results can alternatively be derived by replacing partial anonymity with the requirement that strategies are non-discriminatory (in the section of Section 4.2). Moreover, strongly symmetric strategies (which impose symmetric play at all histories, as in Abreu (1986)) are necessarily non-discriminatory, so without anonymity one-time enforcer punishment equilibria remain optimal in the class of strongly symmetric equilibria.

Separate roles for producers and enforcers: We have assumed that only some agents have the ability to cooperate, and that other, distinct agents have the ability to punish. This implies that the worst continuation play for enforcers is the withdrawal of cooperation, while the best continuation play for enforcers is the most cooperative equilibrium path itself. Both of these features are needed for stick-and-carrot equilibria to be optimal and to take the simple form of one-time enforcer punishment equilibria. If all agents could both cooperate and punish, then the mechanics of the model would be closer to those of Abreu (1986). As in Abreu, stick-and-carrot equilibria would remain optimal in the class of strongly symmetric pure strategy equilibria, while globally optimal equilibria would be more complex. Thus, our assumption that some agents specialize in production or cooperation while others specialize in punishment is a deviation from standard models in a direction that contributes to both realism and tractability.

5.2 Assumptions about Payoffs

Public goods versus bilateral cooperation: We have assumed that the benefits of cooperation are "non-excludable" within a match, and thus have the flavor of a public good. An alternative version of the large group model without this flavor is the following: players match in pairs and do not observe whether their partner is a producer or an enforcer until the end of the period. Thus, cooperation benefits only one's (unique) partner, and at the time she chooses her level of cooperation a producer does not know whether she is matched with another producer (whom she could profitably cheat) or an enforcer (who would punish her if she cheated). All of our results directly translate to this slightly modified setting.

Enforcer payoffs: Yet another interpretation of enforcer payoffs in our model is that enforcers can impose a tax on producers' output, $\sum f(x_i)$, either within their own match or throughout the entire society. If an enforcer's failure to punish a deviant producer leads to reduced cooperation, this then reduces his future payoffs.

Ostracism: In practice, a major tool for sustaining cooperation in small groups is ostracism, or the exclusion of deviators alone from the benefits of societal cooperation (Coleman, 1988, Ostrom, 1990, Ellickson, 1991, Greif, 1993, 2006). The model analyzed so far does not allow for ostracism, because it is not technologically feasible to exclude some players from the benefits of cooperation without excluding everyone. When cooperation corresponds to directed actions (such as simple favors or investments in a bilateral project) rather than undirected actions or the provision of public goods from which all agents benefit, such exclusion becomes a possibility. In Appendix C, we analyze a variant of both the small and large group models where producers have the option to ostracize particular players. We show that all results from these sections apply directly, with the modification that ostracism replaces repentance.

Enforcer misbehavior: In our model enforcer misbehavior takes the form of enforcers' not undertaking costly punishments following a deviation by a producer. Though this is an important consideration in some settings (e.g., motivating law enforcement to pursue powerful individuals, or ensuring that they punish law-breakers who might offer them bribes to avoid such punishment), an equally salient concern is the possibility that enforcers may misuse their positions to expropriate citizens. Introducing this type of misbehavior would not complicate our analysis because our equilibrium construction is already based on giving enforcers the strongest possible incentives to carry out costly punishments. Therefore, if expropriating citizens is as observable as is failing to punish, then the same construction that maximizes enforcers' incentives to punish will minimize their incentives to expropriate.

The specialized enforcement technology: The specialized enforcement technology g measures how much disutility an enforcer must incur to impose a given level of disutility on a producer. This is not to be interpreted as, say, the level of sophistication of a society's instruments of torture, which after all were remarkably advanced even in primitive societies. Rather, it should be interpreted as the cost—and the risk—to enforcers of undertaking the entire process of investigating, pursuing, apprehending, and punishing deviators. To the extent that this cost is less in modern societies than in pre-modern societies, this interpretation suggests a reason why modern societies might tend to rely more on specialized enforcement than community enforcement.

Furthermore, in a natural extension of the large group model where enforcers can monitor multiple matches at the same time, increasing the number of matches monitored by each enforcer would simply scale up the function g. Thus, another reason why specialized enforcement may be more likely in modern societies is that modern technology allows each enforcer to monitor a greater number of interactions at once.[23]

The possibility of transfers and fines: Our results are robust to allowing voluntary monetary transfers from producers, for instance by having deviant producers pay fines to enforcers in lieu of being punished. Indeed, as long as $f'(x) \geq 1$ for all x, it can be checked that our results hold without modification when transfers from producers are allowed. Intuitively, it is inefficient to ask

[23] Other types of monitoring improvements would have more complicated effects, which are beyond the scope of our analysis.

a producer to pay a fine rather than cooperating at a higher level. For example, if producers can pay fines in a separate stage in between the cooperation stage and the punishment stage, they can be asked to do so in equilibrium in lieu of being punished, but this does not increase the maximum level of cooperation, and indeed simply pushes the threat of punishment by enforcers one more step off the equilibrium path. (On the other hand, allowing monetary transfers from enforcers to producers would give enforcers a "cooperative" role, undercutting the separation of roles between producers and enforcers).

6 Conclusion

This paper has introduced a framework for comparing community (private-order) and specialized (public-order) enforcement of pro-social behavior. The key feature of our approach is that we endogenize specialized enforcement by requiring that enforcers have an incentive to carry out the punishment of deviators. We thus require that both community and specialized enforcement are ultimately based on "reputation."

Our main results turn on a novel tradeoff: the withdrawal of future cooperation following a transgression has a positive direct effect on producers' incentives to cooperate, but also a negative indirect effect coming through the erosion of enforcers' incentives to punish. When the specialized enforcement technology is relatively effective, this tradeoff is optimally resolved by going to the extreme of pure enforcer punishments, where the future path of cooperation is completely unaffected by producers' transgressions. All the same, the threat of contagion does play a critical role even under pure enforcer punishments, as in our baseline model it is precisely this threat that gives enforcers the necessary incentives to carry out punishments. A further implication of our analysis is that community enforcement is more likely to emerge in groups with less effective enforcement technologies, while groups with more effective technologies should rely on specialized enforcement. We also illustrate that these results are unchanged when agents have the ability to ostracize (selectively exclude) each other from the benefits of cooperation, and that partial versions of our results remain valid under private monitoring.

The framework introduced in this paper could be developed in several promising directions. First, we have considered the problem of endogenizing the number of specialized enforcers from the perspective of a benevolent social planner (Theorems 2 and 4). One could alternatively analyze the labor market equilibrium of this "occupational choice" problem. Such an exercise would bear some resemblance to the "guns versus butter" tradeoff present in classic models of anarchy, such as Skaperdas (1992), Grossman and Kim (1995), Hirshleifer (1995), and Bates, Greif, and Singh (2002). One could also further extend the model in that direction by allowing "guns" to be used for expropriating others as well as enforcing cooperation.

Second, in a specialized enforcement equilibrium, the enforcers in our model can be interpreted as either a proto-state institution or a non-state institution, such as a mafia. Several scholars, notably Tilly (1985), have argued that states evolve from—or are in fact a form of—private provision of law enforcement. An important question here is when we should expect specialized enforcers to organize in a single institution rather than multiple collectives. While some of our results bear on this question (for example, the results of Section 4.2 on optimal equilibria when enforcers share information with each other), many other interesting questions could be addressed in future work. These include the costs of mafia-like organizations as opposed to states, as well as the dynamics of the process by which proto-states may be transformed into state institutions.

Third, another reason why specialized enforcement may be preferable to community enforcement is the presence of noisy observations, whereby cooperative actions may appear as noncooperative. As briefly discussed in Section 4.2, such noise may make contagion-like strategies prohibitively costly. An analysis of the framework presented here under such richer information structures is an interesting and important area for future work.

Finally, we have only briefly touched on the role of communication and other private actions in supporting specialized enforcement. It would be interesting to analyze more systematically how specialized enforcement (more generally, the legal system) affects the incentives of citizens to cooperate not only with each other but also with state institutions.

Appendix A: Proofs for Sections 2 and 3

Proof of Proposition 1

In the absence of enforcers, a producer's minmax payoff is 0, so grim trigger strategies are an optimal penal code (Abreu, 1988). That the level of cooperation is maximized by a constant path of play then follows from concavity of f. This result is also the special case of Theorem 3 with $l = 0$ and $n = 1$, replacing k with $k + 1$. ■

Proof of Theorem 1

The theorem is the special case of Theorem 3 with $l = 1$ and $n = 1$. ■

Proof of Theorem 2

The theorem is the special case of Theorem 4 with $s = k + 1$, $\bar{k} = k$, and $n = 1$. ■

Proof of Theorem 3

Before proving the theorem, we observe that an enforcer punishment plus partial repentance equilibrium with $y = y^E$ can exist only if $y^E \leq y^*$.

Lemma A1 *If $y^E > y^*$ then the system of equations (8) does not have a solution. That is, an enforcer punishment plus partial repentance equilibrium with $y = y^E$ does not exist.*

Proof. We wish to show that if $y^E > y^*$, then for every $\underline{x} \in [0, x]$ the system of equations

$$\begin{aligned} x &= lg\left(y\right) + \delta\left(k-1\right)\left[f\left(x\right) - f\left(\underline{x}\right)\right] \\ y &= \frac{\delta}{1-\delta} kf\left(x\right) - \delta\left(k - \frac{1}{n}\right)\left[f\left(x\right) - f\left(\underline{x}\right)\right] \end{aligned}$$

does not have a solution with $y = y^E$. Letting $w = 1 - \frac{f(\underline{x})}{f(x)}$, this is equivalent to showing that, for every $w \in [0, 1]$, the system of equations

$$\begin{aligned} x &= lg\left(y\right) + \delta\left(k-1\right) wf\left(x\right) \\ y &= \frac{\delta}{1-\delta} kf\left(x\right) - \delta\left(k - \frac{1}{n}\right) wf\left(x\right) \end{aligned} \tag{A1}$$

does not have a solution with $y = y^E$. We will show that if $y^E > y^*$, then for every $w \in [0, 1]$ the solution to this system has $y \leq y^*$.

To see this, substitute for y to obtain the following equation for x as an implicit function of w:

$$lg\left(\left[\frac{\delta}{1-\delta} k - \delta\left(k - \frac{1}{n}\right) w\right] f\left(x\right)\right) + \delta\left(k-1\right) wf\left(x\right) - x = 0. \tag{A2}$$

Denote the left-hand side of (A2) by $F(x,w)$. Note that $F(x,w)$ is concave in x and satisfies $F(0,w)=0$ for all $w\in[0,1]$, so if $F(x_0,w)=0$ for $x_0>0$ then $\frac{\partial F}{\partial x}|_{(x_0,w)}<0$. Hence, by the implicit function theorem, the solution x to (A2) is differentiable as a function of w, and the sign of $\frac{dx}{dw}$ equals the sign of $\frac{\partial F}{\partial w}$. Next, note that

$$\frac{\partial F}{\partial w}=\delta(k-1)f(x)\left(1-\frac{g'(y)}{m}\right),$$

where y is given by (A1). Therefore, $\frac{dx}{dw}\geq 0$ if and only if $y\geq y^E$, and hence $\frac{dy}{dw}\leq 0$ whenever $y^E>y$ (noting that $\frac{\partial y}{\partial x}\geq 0$ and $\frac{\partial y}{\partial w}\leq 0$). Note also that (A1) coincides with the system defining (x^*,y^*) when $w=0$. Hence, if $y^E>y^*$ then for every $w\geq 0$ the solution to (A1) has $y\leq y^*$. ■

Turning to the proof of the theorem, it is straightforward to check that the three strategy profiles referenced in the theorem are equilibria. It remains to prove that x^* (resp., $\check{x}$, $\tilde{x}$) is an upper bound on each producer's level of cooperation in any PBE when $y^{FB}\geq y^*$ (resp., $y^{FB}\leq\check{y}$, $y^{FB}\in(\check{y},y^*)$). We break the proof into several steps.

Definitions and Preliminary Observations:

Fixing a PBE profile $\sigma=(\sigma_i)_{i\in I}$, let $\underline{u}$ be the infimum continuation payoff of any producer starting from the punishment stage at any history. In addition, let $\operatorname{supp}\sigma_i\left(h^t\right)$ denote the support of producer i's action at history h^t, and let

$$\bar{X}=\sup_{i,h^t,x_i^t\in\operatorname{supp}\sigma_i(h^t)}(1-\delta)\,x_i^t+\delta\mathbb{E}\left[X_i^{t+1}|h^t,x_i^t\right]$$

be the supremum expected present discounted level of cooperation ever taken by any producer at any history.

A preliminary observation is that $\underline{u}>-\infty$ and $\bar{X}<\infty$. To see this, note that, as f is bounded and an enforcer's minmax payoff is 0, there is a finite upper bound $\bar{y}\in\mathbb{R}_+$ on the level of punishment that an enforcer is ever willing to use in equilibrium.[24] Since a producer always has the option of taking action 0 at cost 0, this implies that $\underline{u}\geq -lg(\bar{y})>-\infty$. Given that there is a finite lower bound on $\underline{u}$, it follows that there is a finite upper bound on the level of cooperation that a producer is ever willing to choose in equilibrium, so $\bar{X}<\infty$.

Producer Incentive Compatibility:

[24] By the same argument leading to (A4) below, one such upper bound is $\lim_{x\to\infty}\frac{\delta}{1-\delta}kf(x)$.

A necessary condition for producer i not to deviate to playing $x_i = 0$ at history h^t is that, for all $x_i^t \in \operatorname{supp} \sigma_i \left(h^t\right)$,

$$\begin{aligned} & (1-\delta) \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[\sum_{j \in M_i^{t+\tau} \cap P \backslash i} f\left(x_j^{t+\tau}\right) | h^t, x_i^t\right] - (1-\delta) x_i^t - \delta \mathbb{E}\left[X_i^{t+1} | h^t, x_i^t\right] \\ \geq \; & (1-\delta) \mathbb{E}\left[\sum_{j \in M_i^t \cap P \backslash i} f\left(x_j^t\right) | h^t\right] + \underline{u}, \end{aligned}$$

where $M_i^{t+\tau}$ denotes player i's period-$t+\tau$ match (which is a random variable from the perspective of period t). This is a necessary condition because the left-hand side is an upper bound on player i's equilibrium continuation payoff (as it assumes she is never punished in equilibrium), while the right-hand side is a lower bound on player i's continuation payoff if she deviates (as it assumes she gets her lowest possible continuation payoff).[25] Note that the distribution of x_j^t does not depend on x_i^t, so $\mathbb{E}\left[\sum_{j \in M_i^t \cap P \backslash i} f\left(x_j^t\right) | h^t, x_i^t\right] = \mathbb{E}\left[\sum_{j \in M_i^t \cap P \backslash i} f\left(x_j^t\right) | h^t\right]$, and we can rewrite this necessary condition as

$$(1-\delta) x_i^t + \delta \mathbb{E}\left[X_i^{t+1} | h^t, x_i^t\right] \leq \delta (1-\delta) \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[\sum_{j \in M_i^{t+1+\tau} \cap P \backslash i} f\left(x_j^{t+1+\tau}\right) | h^t, x_i^t\right] - \underline{u}. \qquad \text{(A3)}$$

Using Enforcer Incentive Compatibility to Bound $\underline{u}$:

Letting y_{ki}^t denote enforcer k's punishment action toward player i in period t (which, like x_i^t, is a random variable), a necessary condition for enforcer k not to deviate to playing $y_{ki} = 0$ at history h^t is

$$y_{ki}^t \leq \delta \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[\sum_{j \in M_k^{t+1+\tau} \cap P} f\left(x_j^{t+1+\tau}\right) | h^t\right]. \qquad \text{(A4)}$$

This is a necessary condition because an enforcer's minmax payoff is 0, while her equilibrium continuation payoff is at most $(1-\delta)\left(\delta \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[\sum_{j \in M_k^{t+1+\tau} \cap P} f\left(x_j^{t+1+\tau}\right) | h^t\right] - y_{ki}^t\right)$, as this is her continuation payoff if she does not punish anyone other than player i in period t and never punishes anyone after period t.

Now, producer i's continuation payoff at the punishment stage at history h^t is at least

$$-(1-\delta) \mathbb{E}\left[\sum_{k \in M_i^t \cap E} g\left(y_{ki}^t\right) | h^t\right] + \delta (1-\delta) \mathbb{E}\left[\sum_{j \in M_i^{t+1} \cap P \backslash i} f\left(x_j^{t+1}\right) | h^t\right] + \delta \underline{u},$$

[25] Technically, both expectations in this expression should also be conditioned on the event $j \in M_i^{t+\tau} \cap C \backslash i$. However, because identities are concealed at the point where producers choose their actions, the distribution of $x_j^{t+\tau}$ conditional on this event equals its unconditional distribution. We therefore omit this conditioning throughout the proof.

as a producer always has the option of playing $x_i = 0$ in period $t+1$. Therefore, there exists a producer i and a history h^t such that

$$\underline{u} \geq -(1-\delta)\mathbb{E}\left[\sum_{k\in M_i^t\cap E} g\left(y_{ki}^t\right)|h^t\right] + \delta(1-\delta)\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap P\backslash i} f\left(x_j^{t+1}\right)|h^t\right] + \delta\underline{u},$$

or equivalently $\underline{u} \geq -\mathbb{E}\left[\sum_{k\in M_i^t\cap E} g\left(y_{ki}^t\right)|h^t\right] + \delta\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap P\backslash i} f\left(x_j^{t+1}\right)|h^t\right]$. In particular, by (A4) and the observation that the quantity $\mathbb{E}\left[\sum_{j\in M_k^{t+1+\tau}\cap P} f\left(x_j^{t+1+\tau}\right)|h^t\right]$ is the same for all $k \in E$, we have

$$\underline{u} \geq -lg\left(\delta\sum_{\tau=0}^{\infty}\delta^\tau\mathbb{E}\left[\sum_{j\in M_k^{t+1+\tau}\cap P} f\left(x_j^{t+1+\tau}\right)|h^t\right]\right) + \delta\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap P\backslash i} f\left(x_j^{t+1}\right)|h^t\right]. \tag{A5}$$

Bounding $\underline{u}$ in terms of $\bar{X}$:

By the definition of $\bar{X}$, for every producer j, history h^{t+1}, and level of cooperation $x_j^{t+1} \in \operatorname{supp}\sigma_j\left(h^{t+1}\right)$, we have

$$x_j^{t+1} \leq \frac{1}{1-\delta}\bar{X} - \delta\sum_{\tau=0}^{\infty}\delta^\tau\mathbb{E}\left[x_j^{t+2+\tau}|h^{t+1}, x_j^{t+1}\right]. \tag{A6}$$

We now consider three cases:

Case 1: This case applies if, after replacing x_j^{t+1} on the right-hand side of (A5) with its upper bound in (A6) for all j, the resulting argument of g is less than y^E.

Case 2: This case applies if, after replacing x_i^{t+1} on the right-hand side of (A5) with its upper bound in (A6) and replacing x_j^{t+1} with 0 for all $j \in P\backslash i$, the resulting argument of g is greater than y^E.

Case 3: This case applies when Cases 1 and 2 do not apply. Note that, in this case, there exists a unique value for the term $\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap P\backslash i} f\left(x_j^{t+1}\right)|h^t\right]$ such that

$$\begin{aligned}\delta\mathbb{E}\left[f\left(\frac{1}{1-\delta}\bar{X} - \delta\sum_{\tau=0}^{\infty}\delta^\tau\mathbb{E}\left[x_i^{t+2+\tau}|h^{t+1}, x_i^{t+1}\right]\right)|h^t\right] + \delta^2\sum_{\tau=0}^{\infty}\delta^\tau\mathbb{E}\left[f\left(x_i^{t+2+\tau}\right)|h^t\right] \\ + \delta\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap P\backslash i} f\left(x_j^{t+1}\right)|h^t\right] + \delta^2\sum_{\tau=0}^{\infty}\delta^\tau\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap P\backslash i} f\left(x_j^{t+2+\tau}\right)|h^t\right] = y^E.\end{aligned}$$

Call this value f^E.

We now argue that, in Case 1, the bound (A5) can be relaxed to

$$\underline{u} \geq -lg\left(\frac{\delta}{1-\delta}kf\left(\bar{X}\right)\right) + \delta(k-1)f\left(\bar{X}\right). \tag{A7}$$

Similarly, we argue that in Case 2, (A5) can be relaxed to

$$\underline{u} \geq -lg\left(\left(\frac{\delta}{1-\delta}k - \delta\left(k - \frac{1}{n}\right)\right) f\left(\bar{X}\right)\right), \tag{A8}$$

and in Case 3, (A5) can be relaxed to

$$\underline{u} \geq -lg\left(y^E\right) + \delta\left(k-1\right) f\left(\bar{X}\right) - \frac{\delta}{1-\delta} k f\left(\bar{X}\right) + y^E. \tag{A9}$$

Start with Case 1. For every $j \in P \backslash i$ and $k \in E$, we have $\Pr\left(j \in M_k^{t+1}\right) = \frac{1}{n}$ and $\Pr\left(j \in M_i^{t+1}\right) = \frac{k-1}{kn-1}$, so the derivative of the right-hand side of (A5) with respect to x_j^{t+1} equals $-\delta l \frac{1}{n} g'\left(y\right) f'\left(x_j^{t+1}\right) + \delta \frac{k-1}{kn-1} f'\left(x_j^{t+1}\right)$ for some number y. In Case 1, we have $y \leq y^E$ whenever $x_{j'}^{t+1}$ is below its its upper bound in (A6) for all j', so this derivative is non-positive. Thus, by the fundamental theorem of calculus, replacing x_j^{t+1} with its upper bound in (A6) for all j relaxes (A5). The resulting lower bound equals

$$\begin{gathered} -lg\left(\begin{array}{c} \delta \mathbb{E}\left[\sum_{j \in M_k^{t+1} \cap P} f\left(\frac{1}{1-\delta}\bar{X} - \delta \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[x_j^{t+2+\tau} | h^{t+1}, x_j^{t+1}\right]\right) | h^t\right] \\ +\delta^2 \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[\sum_{j \in M_k^{t+2+\tau} \cap P} f\left(x_j^{t+2+\tau}\right) | h^t\right] \end{array}\right) \\ +\delta \mathbb{E}\left[\sum_{j \in M_i^{t+1} \cap P \backslash i} f\left(\frac{1}{1-\delta}\bar{X} - \delta \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[x_j^{t+2+\tau} | h^{t+1}, x_j^{t+1}\right]\right) | h^t\right]. \end{gathered} \tag{A10}$$

We next derive an upper bound on the argument of g in (A10). Letting

$$X_j\left(h^{t+1}\right) = \left(1-\delta\right) \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[x_j^{t+2+\tau} | h^{t+1}\right],$$

by the concavity of f and Jensen's inequality we have

$$\begin{gathered} \delta \mathbb{E}\left[\sum_{j \in M_k^{t+1} \cap P} f\left(\frac{1}{1-\delta}\bar{X} - \delta \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[x_j^{t+2+\tau} | h^{t+1}, x_j^{t+1}\right]\right) | h^t\right] \\ +\delta^2 \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[\sum_{j \in M_k^{t+2+\tau} \cap P} f\left(x_j^{t+2+\tau}\right) | h^t\right] \\ \leq \delta \mathbb{E}\left[\sum_{j \in M_k^{t+1} \cap P} f\left(\frac{1}{1-\delta}\left(\bar{X} - \delta X_j\left(h^{t+1}\right)\right)\right) | h^t\right] + \frac{\delta^2}{1-\delta} \mathbb{E}\left[\sum_{j \in M_k^{t+2} \cap P} f\left(X_j\left(h^{t+1}\right)\right) | h^t\right]. \end{gathered}$$

Next, again by the concavity of f, the maximum of $\delta f\left(\frac{1}{1-\delta}\left(\bar{X} - \delta X_j\left(h^{t+1}\right)\right)\right) + \frac{\delta^2}{1-\delta}\mathbb{E} f\left(X_j\left(h^{t+1}\right)\right)$ over $X_j\left(h^{t+1}\right) \leq \bar{X}$ is attained at $X_j\left(h^{t+1}\right) = \bar{X}$ for all j and h^{t+1}. This gives an upper bound on the argument of g in (A10) of $\delta k f\left(\frac{1}{1-\delta}\left(\bar{X} - \delta \bar{X}\right)\right) + \frac{\delta^2}{1-\delta} k f\left(\bar{X}\right) = \frac{\delta}{1-\delta} k f\left(\bar{X}\right)$. On the other hand,

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap P\backslash i} f\left(\frac{1}{1-\delta}\bar{X}-\delta\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[x_j^{t+2+\tau}|h^{t+1},x_j^{t+1}\right]\right)|h^t\right] \\
\geq\ &\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap P\backslash i} f\left(\frac{1}{1-\delta}\bar{X}-\frac{\delta}{1-\delta}\bar{X}\right)\right]=(k-1)\,f\left(\bar{X}\right).
\end{aligned}
$$

Combining these observations, we see that (A10) is lower-bounded by $-lg\left(\frac{\delta}{1-\delta}kf\left(\bar{X}\right)\right)+\delta\left(k-1\right)f\left(\bar{X}\right)$. This yields (A7).

Next, consider Case 2. In this case, replacing x_i^{t+1} with its upper bound in (A6) and replacing x_j^{t+1} with 0 for all $j\in P\backslash i$ relaxes (A5). The resulting lower bound equals

$$
-lg\left(\begin{array}{c}\frac{1}{n}\left[\delta\mathbb{E}\left[f\left(\frac{1}{1-\delta}\bar{X}-\delta\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[x_i^{t+2+\tau}|h^{t+1},x_i^{t+1}\right]\right)|h^t\right]+\delta^2\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[f\left(x_i^{t+2+\tau}\right)|h^t\right]\right] \\ +\delta^2\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[\sum_{j\in M_k^{t+2+\tau}\cap P\backslash i}f\left(x_j^{t+2+\tau}\right)|h^t\right]\end{array}\right). \tag{A11}
$$

As we have seen,

$$
\delta\mathbb{E}\left[f\left(\frac{1}{1-\delta}\bar{X}-\delta\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[x_i^{t+2+\tau}|h^{t+1},x_i^{t+1}\right]\right)|h^t\right]+\delta^2\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[f\left(x_i^{t+2+\tau}\right)|h^t\right]\leq\frac{\delta}{1-\delta}f\left(\bar{X}\right).
$$

In addition, by concavity of f, Jensen's inequality, and the definition of $\bar{X}$,

$$
\delta^2\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[\sum_{j\in M_k^{t+2+\tau}\cap P\backslash i}f\left(x_j^{t+2+\tau}\right)|h^t\right]\leq\frac{\delta^2}{1-\delta}\left(k-\frac{1}{n}\right)f\left(\bar{X}\right).
$$

As $\frac{1}{n}\frac{\delta}{1-\delta}+\frac{\delta^2}{1-\delta}\left(k-\frac{1}{n}\right)=\frac{\delta}{1-\delta}k-\delta\left(k-\frac{1}{n}\right)$, this yields (A8).

Finally, consider Case 3. Here, replacing x_i^{t+1} with its upper bound in (A6) and replacing $\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap P\backslash i}f\left(x_j^{t+1}\right)|h^t\right]$ with f^E relaxes (A5). The resulting lower bound equals

$$
\begin{aligned}
&-lg\left(y^E\right)+\delta f^E \\
=\ &-lg\left(y^E\right)-\delta\mathbb{E}\left[f\left(\frac{1}{1-\delta}\bar{X}-\delta\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[x_i^{t+2+\tau}|h^{t+1},x_i^{t+1}\right]\right)|h^t\right]-\delta^2\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[f\left(x_i^{t+2+\tau}\right)|h^t\right] \\
&-\delta^2\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap P\backslash i}f\left(x_j^{t+2+\tau}\right)|h^t\right]+y^E.
\end{aligned}
$$

As above, this bound can be relaxed to

$$
-lg\left(y^E\right)-\frac{\delta}{1-\delta}f\left(\bar{X}\right)-\frac{\delta^2}{1-\delta}\left(k-1\right)f\left(\bar{X}\right)+y^E,
$$

which equals the right-hand side of (A9).

Finishing the Proof:

To finish the proof, we show that the maximum level of cooperation is always either x^*, $\check{x}$, or $\tilde{x}$, and that if the maximum level of cooperation is x^* (resp., $\check{x}$, $\tilde{x}$), then y^E must be weakly greater than y^* (resp., weakly less than $\check{y}$, in between $\check{y}$ and y^*). This implies the desired result.

To begin, fix a sequence of equilibria converging to the maximum level of cooperation for producer 1 (say).

Suppose that, for all $\varepsilon > 0$, there exists an equilibrium along the sequence yielding within ε of the maximum level of cooperation for which Case 1 applies for some producer i and history h^t. Combining (A3) and (A7), it follows that, in this equilibrium for every player i, history h^t, and level of cooperation $x_i^t \in \operatorname{supp} \sigma_i\left(h^t\right)$,

$$\begin{aligned}
& (1-\delta)\, x_i^t + \delta \mathbb{E}\left[X_i^{t+1} | h^t, x_i^t\right] \\
\leq \;& lg\left(\frac{\delta}{1-\delta} k f\left(\bar{X}\right)\right) - \delta (k-1) f\left(\bar{X}\right) + \delta (1-\delta) \sum_{\tau=0}^{\infty} \delta^\tau \left[\sum_{j \in M_i^{t+1+\tau} \cap P \backslash i} f\left(x_j^{t+1+\tau}\right) | h^t, x_i^t\right] \\
\leq \;& lg\left(\frac{\delta}{1-\delta} k f\left(\bar{X}\right)\right) - \delta (k-1) f\left(\bar{X}\right) + \delta (k-1) f\left(\bar{X}\right) = lg\left(\frac{\delta}{1-\delta} k f\left(\bar{X}\right)\right).
\end{aligned}$$

As $\bar{X} = \sup_{i, h^t, x_i^t \in \operatorname{supp} \sigma_i(h^t)} (1-\delta)\, x_i^t + \delta \mathbb{E}\left[X_i^{t+1} | h^t, x_i^t\right]$, we have $\bar{X} \leq lg\left(\frac{\delta}{1-\delta} k f\left(\bar{X}\right)\right)$. By the definition of x^*, this implies that $\bar{X} \leq x^*$. Furthermore, as $\mathbb{E}\left[X_i^0 | h^0\right] \leq \bar{X}$, we have $\mathbb{E}\left[X_i^0 | h^0\right] \leq x^*$. Thus, in this case x^* is an upper bound on each player's maximum equilibrium level of cooperation. As we have seen that x^* is supportable with one-time enforcer punishment strategies, this means that x^* equals each player's maximum level of cooperation. Finally, if one-time enforcer punishment strategies are optimal then $y^E \geq y^*$, as otherwise we could support more cooperation in enforcer punishment plus partial repentance strategies with $\underline{x} = x - \eta$ for sufficiently small $\eta > 0$ (recalling from the proof of Lemma A1 that $\frac{dx}{dw} > 0$ if $y > y^E$).

Next, suppose that, for all $\varepsilon > 0$, there exists an equilibrium along the sequence yielding within ε of the maximum level of cooperation for which Case 2 applies for some producer and history. Combining (A3) and (A8), we have, for every player i, history h^t, and level of cooperation $x_i^t \in \operatorname{supp} \sigma_i\left(h^t\right)$,

$$\begin{aligned}
& (1-\delta)\, x_i^t + \delta \mathbb{E}\left[X_i^{t+1} | h^t, x_i^t\right] \\
\leq \;& lg\left(\left(\frac{\delta}{1-\delta} k - \delta\left(k - \frac{1}{n}\right)\right) f\left(\bar{X}\right)\right) + \delta (1-\delta) \sum_{\tau=0}^{\infty} \delta^\tau \left[\sum_{j \in M_i^{t+1+\tau} \cap P \backslash i} f\left(x_j^{t+1+\tau}\right) | h^t, x_i^t\right] \\
\leq \;& lg\left(\left(\frac{\delta}{1-\delta} k - \delta\left(k - \frac{1}{n}\right)\right) f\left(\bar{X}\right)\right) + \delta (k-1) f\left(\bar{X}\right).
\end{aligned}$$

As above, this gives $\bar{X} \leq lg\left(\left(\frac{\delta}{1-\delta}k - \delta\left(k - \frac{1}{n}\right)\right) f\left(\bar{X}\right)\right) + \delta\left(k-1\right) f\left(\bar{X}\right)$. By the definition of $\check{x}$, this implies that $\bar{X} \leq \check{x}$, and hence $\mathbb{E}\left[X_i^0|h^0\right] \leq \check{x}$. Thus, in this case $\check{x}$ is an upper bound on each player's maximum equilibrium level of cooperation, and hence enforcer punishment plus repentance is optimal. And, if these strategies are optimal, then $y^E \leq \check{y}$ as otherwise we could support more cooperation in one-time enforcer punishment plus partial repentance strategies with $\underline{x} = \eta$ for sufficiently small $\eta > 0$ (recalling from the proof of Lemma A1 that $\frac{dx}{dw} < 0$ if $y < y^E$).

Finally, suppose that, for all $\varepsilon > 0$, there exists an equilibrium along the sequence yielding within ε of the maximum level of cooperation for which Case 3 applies for some producer and history. Combining (A3) and (A9), we have, for every player i, history h^t, and level of cooperation $x_i^t \in \operatorname{supp} \sigma_i\left(h^t\right)$,

$$\begin{aligned}
& \left(1-\delta\right) x_i^t + \delta \mathbb{E}\left[X_i^{t+1}|h^t, x_i^t\right] \\
\leq \quad & lg\left(y^E\right) - \delta\left(k-1\right) f\left(\bar{X}\right) + \frac{\delta}{1-\delta} k f\left(\bar{X}\right) - y^E + \delta\left(1-\delta\right) \sum_{\tau=0}^{\infty} \delta^\tau \left[\sum_{j \in M_i^{t+1+\tau} \cap P \setminus i} f\left(x_j^{t+1+\tau}\right) |h^t, x_i^t\right] \\
\leq \quad & lg\left(y^E\right) + \frac{\delta}{1-\delta} k f\left(\bar{X}\right) - y^E.
\end{aligned}$$

This gives $\bar{X} \leq lg\left(y^E\right) + \frac{\delta}{1-\delta} k f\left(\bar{X}\right) - y^E$. By the definition of $\tilde{x}$, this implies that $\bar{X} \leq \tilde{x}$, and hence $\mathbb{E}\left[X_i^0|h^0\right] \leq \tilde{x}$. Thus, now $\tilde{x}$ is an upper bound on each player's maximum equilibrium level of cooperation, and one-time enforcer punishment plus partial repentance is optimal. By Lemma A1, such an equilibrium exists only if $y^E \leq y^*$. Finally, these strategies can be optimal only if $y^E \geq \check{y}$, as since $\tilde{x} \geq \check{x}$ we have

$$\begin{aligned}
y^E \quad &= \quad \left(\frac{\delta}{1-\delta}k - \delta\left(k - \frac{1}{n}\right)\right) f\left(\tilde{x}\right) + \delta\left(k - \frac{1}{n}\right) f\left(\underline{x}\right) \geq \left(\frac{\delta}{1-\delta}k - \delta\left(k - \frac{1}{n}\right)\right) f\left(\tilde{x}\right) \\
&\geq \quad \left(\frac{\delta}{1-\delta}k - \delta\left(k - \frac{1}{n}\right)\right) f\left(\check{x}\right) = \check{y}.
\end{aligned}$$

■

Proof of Theorem 4

Per-match social welfare with k producers per match and on-path cooperation x is $k\left(\left(s-1\right) f\left(x\right) - x\right)$. Thus, assuming that the maximum level of cooperation is below the first-best level, a sufficient condition for setting $k = s$ to maximize social welfare is that the maximum (per producer) level of cooperation is maximized at $k = s$. The maximum level of cooperation when $k = s$ is given by the solution to $\hat{x} = \delta\left(s-1\right) f\left(\hat{x}\right)$. On the other hand, if $k < s$ and $g'\left(y\right) \leq \left(\frac{k-1}{k-\frac{1}{n}}\right) \frac{1}{s-k}$ for all y (which is guaranteed by the assumptions that $g'\left(y\right) \leq \left(\frac{\bar{k}-1}{\bar{k}-\frac{1}{n}}\right) \frac{1}{s-\bar{k}}$ for all y and $k \geq \bar{k}$), then by Theorem 3 the maximum level of cooperation is given by the solution to

$$x = (s-k)\, g\left(\frac{\delta}{n} f(x) + \frac{\delta^2}{1-\delta} k f(x)\right) + \delta (k-1) f(x). \tag{A12}$$

Thus, the planner prefers that all agents become producers if (A12) is maximized at $k = s$. In turn, a sufficient condition for this to hold is that the derivative of the right-hand side of (A12) with respect to k is non-negative for all x. This derivative equals

$$(s-k)\, g'\left(\frac{\delta}{n} f(x) + \frac{\delta^2}{1-\delta} k f(x)\right) \frac{\delta^2}{1-\delta} f(x) - g\left(\frac{\delta}{n} f(x) + \frac{\delta^2}{1-\delta} k f(x)\right) + \delta f(x).$$

As the first term is positive, a sufficient condition for the whole derivative to be positive is $\delta f(x) \geq g\left(\frac{\delta}{n} f(x) + \frac{\delta^2}{1-\delta} k f(x)\right)$ for all x, or, letting $z = \delta f(x)$, $z \geq g\left(\left(\frac{1}{n} + \frac{\delta}{1-\delta} k\right) z\right)$ for all z. Since $k \leq s$, a sufficient condition for this is $z \geq g\left(\left(\frac{1}{n} + \frac{\delta}{1-\delta} s\right) z\right)$ for all z. Finally, since g is concave with $g(0) = 0$, a sufficient condition is $g'(0) \leq \frac{1}{\frac{1}{n} + \frac{\delta}{1-\delta} s}$. Hence, if (9) holds, then the planner prefers that all agents become producers.

For the second part of the theorem, it suffices to show that $\check{x}$ (with $k = s-1$ and $l = 1$) is greater than $\hat{x}$. By the same argument as above, a sufficient condition for this to hold is $z < g\left(\left(\frac{1}{n} + \frac{\delta}{1-\delta} k\right) z\right)$ for all z. Finally, this condition is satisfied if $g'(y)$ is sufficiently large for all $y \in \mathbb{R}_+$, or if g' is bounded away from 0 and either δ is sufficiently close to 1 or k is sufficiently large. ■

Proof of Proposition 2

We prove the proposition for arbitrary l and n, so that it covers the large group setting as well.

Formally, we show that, for all $\varepsilon > 0$, there exists $\eta > 0$ such that $\hat{x} + \varepsilon$ is an upper bound on the maximum level of cooperation whenever $g'(y) < \eta$ for all y.

By (A5), the definition of $\bar{X}$ (from the proof of Theorem 3), and Jensen's inequality,

$$\underline{u} \geq -lg\left(\frac{\delta}{1-\delta} k f(\bar{X})\right). \tag{A13}$$

Combining (A3) and (A13) yields that, for every player i, history h^t, and level of cooperation $x_i^t \in \operatorname{supp} \sigma_i(h^t)$,

$$(1-\delta)\, x_i^t + \delta \mathbb{E}\left[X_i^{t+1} | h^t, x_i^t\right] \leq lg\left(\frac{\delta}{1-\delta} k f(\bar{X})\right) + \delta (k-1) f(\bar{X}).$$

Again by the definition of $\bar{X}$, whenever $g'(y) < \eta$ for all y we have

$$\bar{X} \leq \left(\frac{\delta}{1-\delta} kl\eta + \delta (k-1)\right) f(\bar{X}).$$

Recall that $\hat{x} = \delta (k-1) f(\hat{x})$. In addition, since $\delta (k-1) f(x) - x$ is concave and crosses 0 from above at $x = \hat{x}$, there exists $\rho > 0$ such that $\delta (k-1) f'(\hat{x}) < 1 - \rho$. Hence, as f is concave, for all $\varepsilon > 0$ we have

$$\begin{aligned}
\left(\frac{\delta}{1-\delta} kl\eta + \delta (k-1)\right) f(\hat{x}+\varepsilon) &\leq \left(\frac{\delta}{1-\delta} kl\eta + \delta (k-1)\right) \left(f(\hat{x}) + \varepsilon f'(\hat{x})\right) \\
&\leq \left(\frac{\delta}{1-\delta} kl\eta + \delta (k-1)\right) \left(f(\hat{x}) + \varepsilon \frac{1-\rho}{\delta (k-1)}\right) \\
&= \hat{x} + \varepsilon (1-\rho) + \frac{\delta}{1-\delta} kl\eta \left(f(\hat{x}) + \varepsilon \frac{1-\rho}{\delta (k-1)}\right).
\end{aligned}$$

For sufficiently small $\eta > 0$, this is less than $\hat{x} + \varepsilon$. Thus, $\left(\frac{\delta}{1-\delta} kl\eta + \delta (k-1)\right) f(x) - x$ crosses 0 to the left of $\hat{x} + \varepsilon$, so $\bar{X} \leq \hat{x} + \varepsilon$. Finally, as $\mathbb{E}\left[X_i^0 | h^0\right] \leq \bar{X}$, we have $\mathbb{E}\left[X_i^0 | h^0\right] \leq \hat{x} + \varepsilon$. Therefore, $\hat{x} + \varepsilon$ is an upper bound on the maximum equilibrium level of cooperation. ■

Theorem 3 when Enforcers Can Punish Each Other

Suppose the enforcers can punish each other. The modified definition of one-time enforcer punishment strategies is as follows: On path, producers cooperate at level x^*. If a producer deviates, all enforcers in her match punish her at level y^*, and play returns to the equilibrium path next period. If an enforcer j deviates, then in the next period all producers shirk, all enforcers matched with j next period punish him at level y^*, and j himself randomly punishes another enforcer at level $\frac{\delta}{1-\delta} \left(kf(x^*) + (l-1) g(y^*)\right)$. Finally, define (x^*, y^*) by the system of the equations

$$\begin{aligned}
x^* &= lg(y^*), \\
y^* &= \frac{\delta}{1-\delta} \left(kf(x^*) + (l-1) g(y^*)\right).
\end{aligned}$$

Theorem 9 *Suppose enforcers can punish other enforcers in addition to producers. If $g'(y) \geq m$ for all $y \in \mathbb{R}_+$, then one-time enforcer punishment strategies are optimal and the maximum level of cooperation is x^*.*

Proof (sketch). The argument is similar to the proof of Theorem 3. In particular, fixing a PBE $\sigma = (\sigma_i)_{i \in I}$, let $\bar{y}$ be the greatest action in the support of any enforcer's equilibrium strategy at any history. As an enforcer can always take action 0, $-\delta (l-1) g(\bar{y})$ is now a lower bound on each enforcer's equilibrium continuation payoff at any history. Equation (A4) ("enforcer incentive compatibility") then becomes

$$\bar{y} \leq \delta \sum_{\tau=0}^{\infty} \delta^t \mathbb{E}\left[\sum_{j \in M_k^{t+1+\tau} \cap P} f\left(x_j^{t+1+\tau}\right) | h^t\right] + \frac{\delta}{1-\delta} (l-1) g(\bar{y}).$$

Carrying the new $\frac{\delta}{1-\delta}(l-1)\,g(\bar{y})$ term throughout the proof of Theorem 3, we obtain the bounds

$$\begin{aligned}\bar{X} &\leq lg(\bar{y}), \\ \bar{y} &\leq \frac{\delta}{1-\delta}\left(kf\left(\bar{X}\right)+(l-1)\,g(\bar{y})\right)\end{aligned}$$

in Case 1 of the proof (which always applies when $g'(y) \geq m$ for all $y \in \mathbb{R}_+$). With the modified definition of (x^*, y^*), this implies that $\bar{X} \leq x^*$. Thus, x^* is an upper bound on each player's maximum level of cooperation. ■

References

[1] Abreu, Dilip. "Extremal equilibria of oligopolistic supergames." *Journal of Economic Theory* 39 (1986): 191-225.

[2] Abreu, Dilip. "On the theory of infinitely repeated games with discounting." *Econometrica* 56 (1988): 383-396.

[3] Acemoglu, Daron, and Matthew O. Jackson. "Social norms and the enforcement of laws." *Journal of the European Economic Association* 15 (2017): 245-295.

[4] Acemoglu, Daron, and Thierry Verdier. "Property rights, corruption and the allocation of talent: A general equilibrium approach." *Economic Journal* 108 (1998): 1381-1403.

[5] Acemoglu, Daron, and Alexander Wolitzky. "A Theory of Equality Before the Law." *Working paper* (2018).

[6] Aldashev, Gani, and Giorgio Zanarone. "Endogenous enforcement institutions." *Journal of Development Economics* 128 (2017): 49-64.

[7] Ali, S. Nageeb, and David A. Miller. "Enforcing cooperation in networked societies." *Working paper* (2014).

[8] Ali, S. Nageeb, and David A. Miller. "Ostracism and forgiveness." *American Economic Review* 106 (2016): 2329-2348.

[9] Axelrod, Robert. *The Evolution of Cooperation*, Basic Books, 1984.

[10] Bates, Robert, Avner Greif, and Smita Singh. "Organizing violence." *Journal of Conflict Resolution* 46 (2002): 599-628.

[11] Boyd, Robert, and Peter J. Richerson. "Punishment allows the evolution of cooperation (or anything else) in sizable groups." *Ethology and Sociobiology* 13 (1992): 171-195.

[12] Clay, Karen. "Trade without law: private-order institutions in Mexican California." *Journal of Law, Economics, and Organization* 13 (1997): 202-231.

[13] Coleman, James S. "Social capital in the creation of human capital." *American Journal of Sociology* (1988): S95-S120.

[14] Deb, Joyee. "Cooperation and Community Responsibility: A Folk Theorem for Repeated Matching Games with Names," mimeo (2012).

[15] Deb, Joyee, and Julio González-Diaz. "Community Enforcement beyond the Prisoner's Dilemma," mimeo (2014).

[16] Dixit, Avinash. "Trade expansion and contract enforcement." *Journal of Political Economy* 111 (2003): 1293-1317.

[17] Dixit, Avinash K. *Lawlessness and economics: alternative modes of governance*. Princeton University Press, 2007.

[18] Ellickson, Robert C. *Order without law: How neighbors settle disputes.* Harvard University Press, 1991.

[19] Ellison, Glenn. "Cooperation in the prisoner's dilemma with anonymous random matching." *Review of Economic Studies* 61 (1994): 567-588.

[20] Fafchamps, Marcel. "The enforcement of commercial contracts in Ghana." *World Development* 24 (1996): 427-448.

[21] Fearon, James D., and David D. Laitin. "Explaining interethnic cooperation." *American Political Science Review* 90 (1996): 715-735.

[22] Fehr, Ernst, and Simon Gächter. "Cooperation and punishment in public goods experiments." *American Economic Review* 90 (2000): 980-994.

[23] Fehr, Ernst, and Simon Gächter. "Altruistic punishment in humans." *Nature* 415 (2002): 137-140.

[24] Flannery, Kent, and Joyce Marcus. *The creation of inequality: how our prehistoric ancestors set the stage for monarchy, slavery, and empire.* Harvard University Press, 2012.

[25] Fudenberg, Drew, and Eric Maskin. "The folk theorem in repeated games with discounting or with incomplete information." *Econometrica* 54 (1986): 533-554.

[26] Goldlücke, Susanne, and Sebastian Kranz. "Infinitely repeated games with public monitoring and monetary transfers." *Journal of Economic Theory* 147 (2012): 1191-1221.

[27] Greif, Avner. "Reputation and coalitions in medieval trade: evidence on the Maghribi traders." *Journal of Economic History* 49 (1989): 857-882.

[28] Greif, Avner. "Contract enforceability and economic institutions in early trade: The Maghribi traders' coalition." *American Economic Review* (1993): 525-548.

[29] Greif, Avner. "Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies." *Journal of Political Economy* 102 (1994): 912-950.

[30] Greif, Avner. *Institutions and the path to the modern economy: Lessons from medieval trade.* Cambridge University Press, 2006.

[31] Greif, Avner, Paul Milgrom, and Barry R. Weingast. "Coordination, commitment, and enforcement: The case of the merchant guild." *Journal of Political Economy* (1994): 745-776.

[32] Grossman, Herschel I., and Minseong Kim. "Swords or plowshares? A theory of the security of claims to property." *Journal of Political Economy* 103 (1995): 1275-1288.

[33] Hadfield, Gillian K., and Barry R. Weingast. "What is law? A coordination model of the characteristics of legal order." *Journal of Legal Analysis* 4 (2012): 471-514.

[34] Henrich, Joseph, and Robert Boyd. "Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas." *Journal of Theoretical Biology* 208 (2001): 79-89.

[35] Hirshleifer, David, and Eric Rasmusen. "Cooperation in a repeated prisoners' dilemma with ostracism." *Journal of Economic Behavior & Organization* 12 (1989): 87-106.

[36] Hirshleifer, Jack. "Anarchy and its breakdown." *Journal of Political Economy* 103 (1995): 26-52.

[37] Jackson, Matthew O., Tomas Rodriguez-Barraquer, and Xu Tan. "Social capital and social quilts: Network patterns of favor exchange." *American Economic Review* 102.5 (2012): 1857-1897.

[38] Johnson, Allen W. and Timoty Earle. *The evolution of human societies: from foraging group to agrarian state*. Stanford University Press, 2000.

[39] Johnson, Simon, John McMillan, and Christopher Woodruff. "Courts and relational contracts." *Journal of Law, Economics, and Organization* 18 (2002): 221-277.

[40] Kali, Raja. "Endogenous business networks." *Journal of Law, Economics, and Organization* 15 (1999): 615-636.

[41] Kandori, Michihiro. "Social norms and community enforcement." *Review of Economic Studies* 59 (1992): 63-80.

[42] Kranton, Rachel E. "Reciprocal exchange: a self-sustaining system." *American Economic Review* 86 (1996): 830-851.

[43] Levine, David K., and Salvatore Modica. "Peer discipline and incentives within groups." *Journal of Economic Behavior & Organization* 123 (2016): 19-30.

[44] Lewis, Ioan M. *Blood and Bone: The Call of Kinship in Somali Society*, Red Sea Press, 1994.

[45] Lippert, Steffen, and Giancarlo Spagnolo. "Networks of relations and word-of-mouth communication." *Games and Economic Behavior* 72 (2011): 202-217.

[46] Mailath, George J., Stephen Morris, and Andrew Postlewaite. "Laws and authority." *Research in Economics* 71 (2017): 32-42.

[47] Masten, Scott E., and Jens Prüfer. "On the evolution of collective enforcement institutions: communities and courts." *Journal of Legal Studies* 43 (2014): 359-400.

[48] McMillan, John, and Christopher Woodruff. "Interfirm relationships and informal credit in Vietnam." *Quarterly Journal of Economics* 114 (1999): 1285-1320.

[49] Milgrom, Paul R., and Douglass C. North. "The role of institutions in the revival of trade: The law merchant, private judges, and the champagne fairs." Economics & Politics 2.1 (1990): 1-23.

[50] Myerson, Roger B. *Game theory*. Harvard University Press, 1991.

[51] Ostrom, Elinor. *Governing the commons: The evolution of institutions for collective action.* Cambridge University Press, 1990.

[52] Ostrom, Elinor, James Walker, and Roy Gardner. "Covenants with and without a sword: Self-governance is possible." *American Political Science Review* 86 (1992): 404-417.

[53] Padró i Miquel, Gerard, and Pierre Yared. "The Political Economy of Indirect Control." *Quarterly Journal of Economics* 127 (2012): 947-1015.

[54] Sethi, Rajiv, and Eswaran Somanathan. "The evolution of social norms in common property resource use." *American Economic Review* 86 (1996): 766-788.

[55] Skaperdas, Stergios. "Cooperation, conflict, and power in the absence of property rights." *American Economic Review* 82 (1992): 720-739.

[56] Takahashi, Satoru. "Community enforcement when players observe partners' past play." *Journal of Economic Theory* 145 (2010): 42-62.

[57] Tilly, Charles. "War making and state making as organized crime" in *Bringing the State Back* In, Peter B. Evans, Dietrich Rueschemeyer, Theda Skocpol eds. Cambridge: Cambridge University Press, 1985.

[58] Tirole, Jean. "A theory of collective reputations (with applications to the persistence of corruption and to firm quality)." *Review of Economic Studies* 63 (1996): 1-22.

[59] Wolitzky, Alexander. "Cooperation with network monitoring." *Review of Economic Studies* 80 (2013): 395-427.

[60] Woodruff, Christopher. "Contract enforcement and trade liberalization in Mexico's footwear industry." *World Development* 26.6 (1998): 979-991.

Appendix B (Not-For-Publication): Proofs for Section 4

Proof of Theorem 5

There are two steps. First, we derive a sufficient condition for one-time enforcer punishment strategies with cooperation level x^* and punishment level y^* to form a PBE strategy profile (Lemma B1). We then show that, under the conditions of the theorem, this sufficient condition is satisfied, and one-time enforcer punishment strategies support greater cooperation than contagion strategies.

To state the sufficient condition for existence, we require some notation. Define the set $D(\tau, t, i)$ recursively by

$$\begin{aligned} D(\tau, t, i) &= \emptyset \text{ if } \tau \leq t, \\ D(t+1, t, i) &= P \cap \{j : \exists k \in I \text{ such that } l_{j,k,t} = 1 \text{ and } k \in M_i\}, \\ D(\tau+1, t, i) &= P \cap \left\{ \begin{array}{c} j : \exists k, k' \in P \text{ such that} \\ l_{j,k,\tau} = 1 \text{ and } k' \in M_k \cap D(\tau, t, i) \end{array} \right\} \text{ if } \tau \geq t+1. \end{aligned}$$

Under contagion strategies, if producer i deviates in period t, then $D(\tau, t, i)$ is the set of producers in the infected phase in period τ. Note that the probability distribution of $D(\tau, t, i)$ is the same as the probability distribution of $D(\tau - t) \equiv D(\tau - t, 0, 1)$ for all $i \in P$ and $\tau \geq t$. Let $d_t \equiv \mathbb{E}[|D(t)|]$.

Similarly, define the set $Q(\tau, t, i)$ recursively by

$$\begin{aligned} Q(\tau, t, i) &= \emptyset \text{ if } \tau \leq t, \\ Q(t+1, t, i) &= \{j : \exists k \in I \text{ such that } l_{j,k,t} = 1 \text{ and } k \in M_i\}, \\ Q(\tau+1, t, i) &= \left\{ \begin{array}{c} j : \exists k, k', k'' \in I \text{ such that} \\ l_{j,k,\tau} = 1,\ k' \in M_k \cap P \cap Q(\tau, t, i),\ k'' \in M_k \cap E \cap Q(\tau, t, i) \end{array} \right\} \text{ if } \tau \geq t+1. \end{aligned}$$

Under one-time enforcer punishment strategies, if producer i deviates in period t and the corresponding enforcer $j \in M_i$ fails to punish her, then $Q(\tau, t, i)$ is the set of players in the infected phase in period τ. Note that the probability distribution of $Q(\tau, t, i)$ is the same as the probability distribution of $Q(\tau - t) \equiv Q(\tau - t, 0, 1)$ for all i and $\tau \geq t$. Let $q_t \equiv \mathbb{E}[|Q(t) \cap P|]$.

Finally, define the set $Z(\tau, t, i)$ by

$$\begin{aligned} Z(\tau, t, i) &= \emptyset \text{ if } \tau \leq t, \\ Z(t+1, t, i) &= \{j : \exists k \in I \text{ such that } l_{j,k,t} = 1 \text{ and } k \in M_i\}, \\ Z(\tau+1, t, i) &= \left\{ \begin{array}{c} j : \exists k, k' \in I \text{ such that} \\ l_{j,k,\tau} = 1 \text{ and } k' \in M_k \cap Z(\tau, t, i) \end{array} \right\} \text{ if } \tau \geq t+1. \end{aligned}$$

The set $Z(\tau, t, i)$ is the set of "infected" players in period τ if an infection process starts in period t in match M_i and spreads through *both* producers and enforcers (rather than only through producers, as is the case with contagion strategies). Let $Z(t) \equiv Z(t, 0, i)$ and $z_t \equiv \mathbb{E}[Z(t) \cap P]$.

Note that the distribution of $|D(t)|$ first-order stochastically dominates the distribution of $|Q(t) \cap P|$, as for every realization of the monitoring technology there are more infected producers with contagion strategies than with one-time enforcer punishment strategies. Similarly, the distribution of $|Z(t) \cap P|$ first-order stochastically dominates the distribution of $|D(t)|$. In particular, $z_t \geq d_t \geq q_t$ for all t.

The formulas for $\hat{x}$ and x^* as functions of d_t and q_t are given in the text. In addition, let

$$\dot{x} = l(1-\delta) \sum_{t=0}^{\infty} \delta^t \frac{k-1}{kn-1} (z_t - 1) f(\dot{x}).$$

Thus, if $l = 1$, then $\dot{x}$ is the greatest level of cooperation that could be sustained with contagion strategies if contagion spread through the process $Z(t)$ rather than $D(t)$. Otherwise, $\dot{x}$ is the level of cooperation that could be sustained when the benefits of cooperation lost through contagion are scaled up by a factor of l.

Our sufficient condition for existence is as follows:

Lemma B1 *If $x^* \geq \dot{x}$, then the one-time enforcer punishment strategy profile with cooperation level x^* and punishment level y^* is a PBE strategy profile.*

Proof. Let the off-path beliefs be that any zero-probability action is viewed as a deviation, rather than a response to an earlier deviation. We check sequential rationality first in the normal state and then in the infected state.

Given our specification of off-path beliefs, whenever a player is in the normal state, she believes that all of her opponents are also in the normal state. Hence, playing $x_i = x^*$ is optimal for producers, as deviating can save a cost of at most x^* but incurs a punishment of $lg(y^*) = x^*$. In addition, if an enforcer deviates when he is supposed to play $y_{ji} = 0$, this incurs an instantaneous cost but yields no future benefit. Finally, if an enforcer deviates when he is supposed to play $y_{ji} = y^*$, this saves a cost of at most y^* and leads to lost future benefits of $\sum_{t=0}^{\infty} \delta^t \frac{q_t}{n} f(x^*) \geq y^*$.

It remains to consider players' incentives in the infected state. For enforcers, note that whenever an enforcer is in the infected state, he believes that at least k producers are also in the infected state (namely, the producers with whom he was matched in the period when he became infected). As these producers will never cooperate, deviating from $y_{ji} = 0$ to $y_{ji} = y^*$ (which is the only tempting deviation) incurs a cost of y^* and yields future benefits worth strictly less than $\sum_{t=0}^{\infty} \delta^t \frac{q_t}{n} f(x^*) < x^*$. So enforcers' off-path play is optimal.

Finally, for a producer i in period t, the only tempting deviation is to $x_{i,t} = x^*$. If every enforcer in M_i^t is in the normal state, then every player enters period $t+1$ in the same state regardless of producer i's choice of $x_{i,t}$. Hence, in this case, producer i is indifferent between playing $x_{i,t} = 0$ (and incurring a punishment of $lg(y^*) = x^*$) and playing $x_{i,t} = x^*$. On the other hand, suppose at least one enforcer in M_i^t is in the infected state. If there is also at least one other producer $j \in M_i^t \backslash \{i\}$ in the infected state, then again every player enters period $t+1$ in the same state regardless of $x_{i,t}$, and in this case producer i strictly prefers playing $x_{i,t} = 0$ to playing $x_{i,t} = x^*$ (as now playing $x_{i,t} = 0$ incurs a punishment of at most $(l-1)\, g(y^*)$). The remaining case is when at least one enforcer in M_i^t is in the infected state, but all other producers $j \in M_i^t \backslash \{i\}$ are in the normal state. In this case, producer i can slow the spread of contagion by playing $x_{i,t} = x^*$, and we must verify that she does not have an incentive to do so.

To see this, let r_τ denote the number of producers who *do* enter the infected state by period τ when producer i plays $x_{i,\tau} = 0$ for all $\tau \in \{t, \dots, \tau\}$, but *do not* enter the infected state by period τ when producer i plays $x_{i,t} = x^*$ and $x_{i,\tau} = 0$ for all $\tau \in \{t+1, \dots, \tau\}$. The difference between producer i's continuation payoff from playing $x_{i,t} = x^*$ as opposed to $x_{i,t} = 0$ (and subsequently playing $x_{i,\tau} = 0$) is then equal to $(1-\delta)\sum_{\tau=0}^{\infty} \delta^\tau \frac{k-1}{kn-1} \mathbb{E}[r_\tau]\, f(x^*)$. As producer i may be punished for playing $x_{i,t} = 0$ by at most $l-1$ enforcers (as we are assuming that at least one of her enforcers is infected), to show that playing $x_{i,t} = 0$ is optimal, it suffices to show that

$$x^* \geq (l-1)\, g(y^*) + (1-\delta) \sum_{\tau=0}^{\infty} \delta^\tau \frac{k-1}{kn-1} \mathbb{E}[r_\tau]\, f(x^*). \tag{B1}$$

We will show below that $\mathbb{E}[r_\tau] \leq z_\tau$ for all τ. This will imply (B1) because, recalling that $\dot{x} = l(1-\delta)\sum_{\tau=0}^{\infty} \delta^\tau \frac{k-1}{kn-1} z_\tau f(\dot{x})$ by definition, the fact that $x^* \geq \dot{x}$ implies that $x^* \geq l(1-\delta)\sum_{\tau=0}^{\infty} \delta^\tau \frac{k-1}{kn-1} z_\tau f(x^*)$, and then $\mathbb{E}[r_\tau] \leq z_\tau$ for all τ implies that this lower bound exceeds $l(1-\delta)\sum_{\tau=0}^{\infty} \delta^\tau \frac{k-1}{kn-1} \mathbb{E}[r_\tau]\, f(x^*)$. Combining this bound with $x^* = lg(y^*)$ yields (B1).

We now show that $\mathbb{E}[r_\tau] \leq z_\tau$ for all τ. For any subset of players $S \subseteq I$, define $Q(\tau, t, S)$ by

$$\begin{aligned} Q(\tau, t, S) &= \emptyset \text{ if } \tau < t \\ Q(t, t, S) &= S \\ Q(\tau+1, t, S) &= \left\{ \begin{array}{c} j : \exists k, k', k'' \in I \text{ such that} \\ l_{j,k,\tau} = 1, k' \in M_k \cap P \cap Q(\tau, t, S), k'' \in M_k \cap E \cap Q(\tau, t, S) \end{array} \right\} \text{ if } \tau \geq t. \end{aligned}$$

Note that if $\tilde{Q} \ni i$ is the set of infected players at the beginning of period t, then when i plays $x_{i,t'} = 0$ for all $t' \in \{t, \dots, \tau\}$ the set of infected players at the beginning of period τ is $Q\left(\tau, t, \tilde{Q}\right)$, while when i plays $x_{i,t} = x^*$ and $x_{i,t'} = 0$ for all $t' \in \{t+1, \dots, \tau\}$ the set of infected players at the beginning of period τ is $Q\left(\tau, t+1, Q\left(t+1, t, \tilde{Q} \backslash \{i\}\right) \cup \{i\}\right)$. Hence,

$$r_\tau = \left| \left(Q\left(\tau, t, \tilde{Q}\right) \backslash Q\left(\tau, t+1, Q\left(t+1, t, \tilde{Q} \backslash \{i\}\right) \cup \{i\}\right)\right) \cap P \right|.$$

We show that, for all $\tau > t$ and for every subset of players $S \ni i$,

$$Q(\tau, t, S) \backslash Q(\tau, t+1, Q(t+1, t, S \backslash \{i\}) \cup \{i\}) \subseteq Z(\tau, t, \{i\}). \tag{B2}$$

This implies that $\mathbb{E}[r_\tau] \leq \mathbb{E}[|Z(\tau,t,i)| \cap P] = z_\tau$, completing the proof.

The proof of (B2) is by induction on τ. For the $\tau = t+1$ case, we have

$$Q(t+1,t,S) \backslash Q(t+1,t+1,Q(t+1,t,S\backslash\{i\}) \cup \{i\}) = Q(t+1,t,S) \backslash (Q(t+1,t,S\backslash\{i\}) \cup \{i\}),$$

which is clearly contained in $Z(t+1,t,\{i\})$. Suppose now that (B2) holds for some $\tau > t$. We then have

$$Q(\tau,t,S) \subseteq Q(\tau,t+1,Q(t+1,t,S\backslash\{i\}) \cup \{i\}) \cup Z(\tau,t,\{i\}),$$

so

$$\begin{aligned} Q(\tau+1,t,S) &= Q(\tau+1,\tau,Q(\tau,t,S)) \\ &\subseteq Q(\tau+1,\tau,Q(\tau,t+1,Q(t+1,t,S\backslash\{i\}) \cup \{i\}) \cup Z(\tau,t,i)). \end{aligned}$$

On the other hand,

$$Q(\tau+1,t+1,Q(t+1,t,S\backslash\{i\}) \cup \{i\}) = Q(\tau+1,\tau,Q(\tau,t+1,Q(t+1,t,S\backslash\{i\}) \cup \{i\})).$$

Hence,

$$\begin{aligned} & Q(\tau+1,t,S) \backslash Q(\tau+1,t+1,Q(t+1,t,S\backslash\{i\}) \cup \{i\}) \\ \subseteq\ & Q(\tau+1,\tau,Q(\tau,t+1,Q(t+1,t,S\backslash\{i\}) \cup \{i\}) \cup Z(\tau,t,i)) \\ & \backslash Q(\tau+1,\tau,Q(\tau,t+1,Q(t+1,t,S\backslash\{i\}) \cup \{i\})), \end{aligned}$$

and this set is clearly contained in $Z(\tau+1,\tau,Z(\tau,t,\{i\})) = Z(\tau+1,t,\{i\})$. Thus, (B2) holds for $\tau+1$, so by induction it holds for all $\tau > t$. ∎

We now complete the proof of Theorem 5. Recall that $\dot{x} \geq \hat{x}$. Thus, by Lemma B1, to show that one-time enforcer punishment strategies form a PBE strategy profile and support greater cooperation than contagion strategies, it suffices to show that $x^* \geq \dot{x}$.

For the first part of the theorem, note that $lg\left(\sum_{t=0}^{\infty} \delta^t \frac{q_t}{n} f(x)\right) - x$ crosses 0 from above at $x = x^*$. Therefore,

$$lg'\left(\sum_{t=0}^{\infty} \delta^t \frac{q_t}{n} f(x^*)\right) \sum_{t=0}^{\infty} \delta^t \frac{q_t}{n} f'(x^*) < 1.$$

Thus, if $g'(y) \geq \bar{g}$ for all y then $f'(x^*) < \frac{n}{l\bar{g}\sum_{t=0}^{\infty}\delta^t q_t}$. As f is increasing, concave, and bounded, $f'(x^*)$ is positive, non-increasing, and goes to 0 as $x^* \to \infty$. Hence, $x^* \to \infty$ as $\bar{g} \to \infty$. On the other hand, $\dot{x}$ is finite and independent of $\bar{g}$, so if $\bar{g}$ is sufficiently high then $x^* > \dot{x}$.

For the second part, as z_t can never exceed kn, $\dot{x}$ is bounded from above by the greatest solution to $x = l(k-1)f(x)$, which is finite as f is bounded. On the other hand, $\lim_{t\to\infty} q_t = kn$, so if $\lim_{y\to\infty} g(y) = \infty$ then $\lim_{\delta\to 1} lg\left(\sum_{t=0}^{\infty} \delta^t \frac{q_t}{n} f(x)\right) - x = \infty$ for all $x > 0$. Hence, x^*, the greatest root of $lg\left(\sum_{t=0}^{\infty} \delta^t \frac{q_t}{n} f(x)\right) - x$, goes to infinity as $\delta \to 1$. ■

Proof of Theorem 6

For one-time enforcer punishment strategies, the proof is similar to the proof of Lemma B1, with the following differences:

First, the reason why punishments of up to $\frac{\delta}{1-\delta} kf(x^*) = y^*$ are incentive-compatible for enforcers is that, if an enforcer in the normal state deviates when he is supposed to play $y_{ji} = y^*$, he then loses all future benefits of cooperation (recall that a player in the normal state believes that all of her opponents are also in the normal state). To see this, note that his partners in the next period will observe his deviation and will therefore play $x_i = 0$, and—since the enforcer will then be in the infected state—he will play $y_{ji} = 0$. Hence, his partners in the period after next will also play $x_i = 0$, and so on.

Second, for a producer in the infected state, the only tempting deviation is to $x_i = x^*$. This is clearly unprofitable if all of her enforcers are in the normal state, as it incurs a cost of $lg(y^*) = x^*$ and does not affect her continuation payoff. If, instead, at least one of her enforcers is in the infected state, then it is unprofitable because the fact that $x^* \geq \dot{x}$ implies that $x^* \geq l\delta(k-1) f(x^*)$, and therefore $x^* \geq (l-1) g(y^*) + \delta(k-1) f(x^*)$.

The proof for contagion strategies is simpler and is omitted. ■

Proof of Theorem 7

For existence, the only difference from the proof of Theorem 6 is that, with informed enforcers, whenever a producer is in the infected state she believes that every enforcer in her match is also in the infected state. The condition that $x^* \geq \dot{x}$ is thus no longer required.

The proof that x^* is an upper bound on each producer's level of cooperation with informed enforcers follows the proof of Theorem 3, with one key additional step. In particular, if we follow the proof of Theorem 3 while conditioning on private rather than public histories where appropriate, as well as conditioning on the realizations of the matching technology, we arrive at the following analogue of inequality (A5):

$$\begin{aligned} \underline{u} \quad \geq \quad & -l\mathbb{E}\left[g\left(\delta \sum_{\tau=0}^{\infty} \delta^{\tau} \sum_{j \in M_k^{t+1+\tau} \cap P} \mathbb{E}\left[f\left(x_j^{t+1+\tau}\right) | h_k^t, j \in M_k^{t+1+\tau}\right]\right) | h_i^t\right] \\ & +\delta\mathbb{E}\left[\sum_{j \in M_i^{t+1} \cap P \setminus i} f\left(x_j^{t+1}\right) | h_i^t, j \in M_i^{t+1}\right]. \end{aligned}$$

(Here, h_k^t is the private history of the enforcer k in match M_i^t, which equals h^t as enforcers are perfectly informed.) As enforcers are perfectly informed, and thus in particular have finer information than producers, we have

$$\mathbb{E}_{h_j^{t+1}}\left[\sigma_j^{t+1}\left(h_j^{t+1}\right)|h_i^t, j \in M_i^{t+1}\right] = \mathbb{E}_{h^t}\left[\mathbb{E}_{h_j^{t+1}}\left[\sigma_j^{t+1}\left(h_j^{t+1}\right)|h^t, j \in M_i^{t+1}\right]|h_i^t\right].$$

By non-discrimination,

$$\mathbb{E}_{h_j^{t+1}}\left[\sigma_j^{t+1}\left(h_j^{t+1}\right)|h^t, j \in M_i^{t+1}\right] = \mathbb{E}_{h_j^{t+1}}\left[\sigma_j^{t+1}\left(h_j^{t+1}\right)|h^t, j \in M_k^{t+1}\right].$$

Hence, we have

$$\begin{aligned} \underline{u} \quad \geq \quad & -l\mathbb{E}\left[g\left(\delta\sum_{\tau=0}^{\infty}\delta^{\tau}\sum_{j\in M_k^{t+1+\tau}\cap P}\mathbb{E}\left[f\left(x_j^{t+1+\tau}\right)|h^t, j \in M_k^{t+1+\tau}\right]\right)|h_i^t\right] \\ & +\delta\mathbb{E}\left[\mathbb{E}\left[\sum_{j\in M_k^{t+1}\cap P\setminus i} f\left(x_j^{t+1}\right)|h^t, j \in M_k^{t+1}\right]|h_i^t\right]. \end{aligned} \tag{B3}$$

For any producer $j^* \neq i$ and any history $h_{j^*}^{t+1}$, replacing $\sigma_{j^*}\left(h_{j^*}^{t+1}\right)$ with $\sigma_{j^*}\left(h_{j^*}^{t+1}\right)+\varepsilon$ in the right-hand side of (B3) and differentiating with respect to ε yields

$$\begin{aligned} & -\delta\frac{l}{n}\mathbb{E}\left[f'\left(x_{j^*}^{t+1}|h_{j^*}^{t+1}\right)\right]\Pr\left(h_{j^*}^{t+1}|h_i^t\right)\mathbb{E}\left[g'\left(\delta\sum_{\tau=0}^{\infty}\delta^{\tau}\sum_{j\in M_k^{t+1+\tau}\cap P}\mathbb{E}\left[f\left(x_j^{t+1+\tau}\right)|h^t, j \in M_k^{t+1+\tau}\right]\right)|h_i^t\right] \\ & +\delta\frac{k-1}{kn-1}\mathbb{E}\left[f'\left(x_{j^*}^{t+1}|h_{j^*}^{t+1}\right)\right]\Pr\left(h_{j^*}^{t+1}|h_i^t\right). \end{aligned}$$

The assumption that $g'(y) \geq m$ for all y implies that the derivative is non-positive for all $h_{j^*}^{t+1}$. Thus, a lower bound on (B3) is obtained by setting x_j^{t+1} equal to its upper bound in (A6) for all j, as in Case 1 of the proof of Theorem 3. The remainder of the argument is identical to the proof of Theorem 3. ■

Proof of Proposition 3

Start with incentive compatibility for the enforcers. Whenever an enforcer is asked to punish, he believes that cooperation will return to x_1 forever if he punishes, while contagion will start if he fails to punish. Thus, his incentive compatibility condition is

$$y \leq \frac{\delta}{1-\delta}2f(x_1) = \frac{.1}{1-.1}2\left(100\sqrt{493.830}\right) \approx 493.8286.$$

So he is willing to punish at level $y = 493.828$ whenever he sees a producer deviation.

Now turn to incentive compatibility for the producers. We start with incentives to exert effort x_1, x_2, and x_3, rather than deviating and choosing effort 0.

If a producer shirks when she is supposed to play x_3, she is punished at level y and starts contagion with probability $\frac{1}{3}$ (while returning to her equilibrium payoff of $f(x_1) - x_1$ with probability $\frac{2}{3}$). Her contagion payoff is at most $\frac{2}{3}f(x_1)$, as this would be her payoff if the contagion never reached the other two producers and she was never punished in the future. Thus, a sufficient condition for incentive compatibility here is

$$x_3 \leq y + \frac{\delta}{3}\left(f(x_1) - x_1 - \frac{2}{3}f(x_1)\right) = 493.828 + \frac{.1}{3}\left(\frac{100\sqrt{493.830}}{3} - 493.830\right) \approx 502.0584.$$

So $x_3 = 502.058$ is incentive-compatible.

If a producer shirks when she is supposed to play x_2, she is punished at level y and faces the additional "punishment" of having to play x_3 rather than x_1 when matched with the same producer partner tomorrow. Thus, her incentive compatibility condition is

$$x_2 \leq y + \frac{\delta}{3}(x_3 - x_1) = 493.828 + \frac{.1}{3}(502.058 - 493.830) \approx 494.1023.$$

So $x_2 = 494.102$ is incentive-compatible.

Finally, if a producer shirks when she is supposed to play x_1, she is punished at level y and also has to play x_2 rather than x_1 tomorrow with probability $\frac{1}{3}$. Thus, her incentive compatibility condition is

$$x_1 \leq y + \frac{\delta}{3}(x_2 - x_1) = 493.828 + \frac{.1}{3}(494.102 - 493.830) \approx 493.8371.$$

So $x_1 = 493.830$ is incentive-compatible.

We also have to check a couple more incentive compatibility conditions. In particular, we have to show that a producer does not want to deviate to playing x_1 rather than x_2 or x_3 (which avoids direct punishment but still risks starting contagion); and we have to show that a producer is willing to go through with contagion (play $x = 0$) when she is supposed to.

A sufficient condition for playing x_3 rather than x_1 when x_3 is called for is:

$$x_3 - x_1 \leq \frac{\delta}{3}\left(\frac{f(x_1)}{3} - x_1\right) \approx 8.2304.$$

As $x_3 - x_1 = 8.228$, this is satisfied.

A sufficient condition for playing x_2 rather than x_1 when x_2 is called for is:

$$x_2 - x_1 \leq \frac{\delta}{3}(x_3 - x_1) \approx 0.2743$$

As $x_2 - x_1 = 0.272$, this is also satisfied.

As for the incentives to carry out contagion, note that the only reasons to work today are to avoid punishment and to encourage others to work in the future. Working today makes enforcers less likely to enter the contagion phase (i.e., stop punishing), which is bad for producers. So an upper bound on the present value of reduced punishments from working is the value of reducing punishments in the current period only. As at least one other producer is also in the contagion phase, this value of reduced punishments is at most $\frac{2}{3}y$ (as enforcers do not punish when both producers shirk). Finally, working can prevent another producer from entering the contagion phase only if this is the third straight time that the same matches have realized, and the two infected producers are in difference matches. In this case, the other producer in the other match will get infected no matter what in the current period, so working keeps at most one other producer out of the contagion phase. An upper bound on the value of this is $\frac{1}{3}\delta f(x_1)$. Hence, a sufficient condition for carrying out contagion is

$$x_1 \geq \delta \frac{f(x_1)}{3} + \frac{2}{3}y \approx 403.2930.$$

Since $x_1 = 493.830$, carrying out contagion is incentive-compatible.

We now describe how the example would have to be modified if enforcers also observed the outcomes of their partners' most recent matches. The reason why some modification is needed is that the counterexample rests on there being some history where, if a producer shirks, the other producer in her match knows that this is the third straight time she has shirked, while the enforcer does not. If enforcers observe the outcomes of their partners' last histories, then a "three strikes and you're out" strategy profile is not enough to generate such a history. For example, if we label the two enforcers A and B, if a producer's match history is ABA then enforcer A will see that she shirked three times in a row.

To restore the counterexample, consider a "five strikes and you're out" strategy profile, where contagion starts only if a producers shirks five times in a row *and* her match history for the first four matches is either AABB or BBAA. With such a match history, the fifth enforcer the producer meets surely will not know that she shirked five times in row. ■

Proof of Theorem 8

One-time enforcer punishment strategies are clearly deterministic and satisfy Stability. It remains only to show that x^* is an upper bound on each player's maximum level of cooperation in any deterministic equilibrium satisfying Stability.

Under Stability, a necessary condition for producer i not to have a profitable one-shot deviation in period t in a deterministic equilibrium with equilibrium path $\left(x_i^\tau, y_{ji}^\tau\right)_{i\in P, j\in E}^{\tau\in\{0,1,\ldots\}}$ is

$$\begin{aligned}
&\sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[\sum_{j\in M_i^{t+\tau}\cap P\setminus i} f\left(x_j^{t+\tau}\right)\right] - \sum_{\tau=0}^{\infty} \delta^\tau x_i^{t+\tau} \\
\geq\ & -\mathbb{E}\left[\sum_{k\in M_i^t\cap E} g\left(y_{ki}^t\right)|x_i^t = 0\right] + \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[\sum_{j\in M_i^{t+\tau}\cap P\setminus i} f\left(x_j^{t+\tau}\right)\right] - \delta \sum_{\tau=0}^{\infty} \delta^\tau x_i^{t+1+\tau},
\end{aligned}$$

or equivalently

$$x_i^t \leq \mathbb{E}\left[\sum_{k \in M_i^t \cap E} g\left(y_{ki}^t\right) | x_i^t = 0\right]. \tag{B4}$$

Next, under Stability, (A4) becomes

$$y_{ki}^t \leq \delta \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[\sum_{j \in M_k^{t+1+\tau} \cap P} f\left(x_j^{t+1+\tau}\right)\right]. \tag{B5}$$

Let $\bar{x} = \sup_{i,t} x_i^t$, and note that $\bar{x} < \infty$ as in the proof of Theorem 3. Combining (B4) and (B5) then yields $\bar{x} \leq lg\left(\frac{\delta}{1-\delta} kf\left(\bar{x}\right)\right)$. By definition of x^*, this implies that $\bar{x} \leq x^*$. Hence, x^* is an upper bound on each player's maximum equilibrium level of cooperation. ∎

Appendix C (Not-For-Publication): Ostracism

We now show how ostracism can be introduced into the small and large group models of Sections 2 and 3. We demonstrate that this does not affect our main results on the role of specialized enforcement and the optimality of one-time enforcer punishment strategies, which are now combined with ostracism.

C1 Small Group Model

In the small group model of Section 2, suppose that, when producer i chooses her period t cooperation level x_i^t, she also chooses an *ostracism level* $w_i^t \in [0,1]$, as well as a set of players to ostracize. Players ostracized by producer i receive benefit only

$$\left(1 - w_i^t\right) f\left(x_i^t\right)$$

from producer i's cooperation in period t, while players not ostracized by producer i receive benefit

$$\left(1 - \gamma w_i^t\right) f\left(x_i^t\right),$$

where $\gamma \in [0,1]$ is a parameter measuring the efficiency cost of ostracism. Thus, if $\gamma = 1$ then, as in the main model, producer i must always provide the same benefit to every other player; while if $\gamma = 0$ then producer i can completely exclude player j from the benefits of her actions without reducing the benefit to player k.

When ostracism is available, there is a simple way of improving on the performance of enforcer punishment plus repentance strategies: in the period following a producer deviation, rather than having all other producers stop cooperating, have them continue to cooperate while fully ostracizing the deviator. This leads to the following definition of enforcer punishment plus ostracism strategies.

Definition 4 *A* one-time enforcer punishment plus ostracism strategy profile *is characterized by cooperation level x, a punishment level y, and an ostracism level w, and can be represented by the following automaton:*

There are $k+2$ states: normal, punishment, and i-ostracizing, for each $i \neq 1$. Play in these states is as follows:

Normal state*: Each producer i plays $x_i = x$. If all producers i play $x_i = x$, then the enforcer plays $y_i = 0$ for all producers i. If, instead, some producer i plays $x_i \neq x$, then the enforcer plays $y_i = y$ and plays $y_{i'} = 0$ for all producers $i' \neq i$.*

Punishment state: *Players always take action 0.*

i-ostracizing state: *Producer i plays $x_i = x$. Producers $i' \neq i$ play $x_{i'} = x$ and ostracize player*

i at level w. If all producers play x, then the enforcer plays $y_{i'} = 0$ for all producers i'. If producer i' plays $x_{i'} \neq x$, then the enforcer plays $y_{i'} = y$ and plays $y_{i''} = 0$ for all producers $i'' \neq i'$.

Players start in the normal state. In the normal state, if some producer i plays $x_i \neq x$ and the enforcer then plays $y_i = y$, players transition to the i-ostracizing state. If some producer i plays $x_i \neq x$ and the enforcer plays $y_i \neq y$, then players transition to the punishment state.

In the i-ostracizing state, players transition to the normal state if all producers play x. If some producer i' plays $x_{i'} \neq x$ and the enforcer plays $y_{i'} = y$, then players transition to the i'-ostracizing state. If some producer i' plays $x_{i'} \neq x$ and the enforcer plays $y_{i'} \neq y$, then players transition to the punishment state.

The punishment state is absorbing.

In analogy to repentance strategies, *full ostracism* will refer to the special case where $w = 1$, and *partial ostracism* to the case where $w \in (0, 1)$. The maximum levels of cooperation and punishment that can be sustained with enforcer punishment plus full ostracism strategies can now be characterized as

$$\begin{aligned} \check{x} &= g(\check{y}) + \delta(k-1) f(\check{x}) \\ \check{y} &= \left(\frac{\delta}{1-\delta} k - \delta(k-1)\gamma\right) f(\check{x}) . \end{aligned}$$

Notice that the difference between this system of equations and the system (2) is the presence of the γ term, reflecting the fact that the future benefits from cooperation lost to the enforcer when he fails to punish a deviant producer are now equal to $\left(\frac{\delta}{1-\delta} k - \delta(k-1)\gamma\right) f(\check{x})$ rather than $\left(\frac{\delta}{1-\delta} k - \delta(k-1)\right) f(\check{x})$.

When ostracism is unavailable, withdrawing a unit of benefit from future cooperation for a deviant producer also withdraws one unit of benefit for the enforcer. With ostracism, it is now possible to withdraw a unit of benefit for the deviator while withdrawing only γ units of benefit for the enforcer. The relevant notion of the "efficient" level of punishment is thus given by

$$y^E(\gamma) = \left\{ \begin{array}{ccc} \infty & \text{if} & \lim_{y\to\infty} g'(y) \geq \frac{1}{\gamma} \\ (g')^{-1}\left(\frac{1}{\gamma}\right) & \text{if} & g'(0) > \frac{1}{\gamma} > \lim_{y\to\infty} g'(y) \\ 0 & \text{if} & g'(0) \leq \frac{1}{\gamma} \end{array} \right\} .$$

This suggests that the system of equations analogous to (3) now becomes

$$\begin{aligned} \tilde{x} &= g\left(y^E(\gamma)\right) + \delta(k-1) w f(\tilde{x}) \\ y^E(\gamma) &= \left(\frac{\delta}{1-\delta} k - \delta(k-1)\gamma w\right) f(\tilde{x}) . \end{aligned} \tag{C1}$$

The generalization of Theorem 1 when ostracism is available is as follows.

Theorem C1 *Suppose the players can employ ostracism with efficiency cost γ.*

If $y^E(\gamma) \geq y^*$, one-time enforcer punishment strategies are optimal, and the maximum level of cooperation is x^*.

If $y^E(\gamma) \leq \check{y}$, one-time enforcer punishment plus full ostracism strategies are optimal, and the maximum level of cooperation is $\check{x}$.

If $y^E(\gamma) \in (\check{y}, y^*)$, one-time enforcer punishment plus partial ostracism strategies are optimal, and the maximum level of cooperation is $\tilde{x}$.

Proof. The theorem is the special case of Theorem C2 with $l = 1$ and $n = 1$. ■

In particular, if $g'(y) \geq \frac{1}{\gamma}$ for all $y \in \mathbb{R}_+$ then it is optimal to rely on one-time enforcer punishment strategies and avoid using ostracism altogether. If, instead, $g'(y) \leq \frac{1}{\gamma}$ for all $y \in \mathbb{R}_+$, then it is optimal to combine enforcer punishments with ostracism.

C2 Large Group Model

Consider now a version of the large group model of Section 3 where players can employ ostracism with an efficiency cost of γ. Differently from the small group setting where the deviator could be ostracized directly, in the large group model this type of ostracism is not feasible because producers choose their level of cooperation without knowing their partners' identities ("partial anonymity"). As a result, ostracism will now take the form of all producers (except the deviator) ostracizing all other producers following a deviation. Put differently, ostracism in the large group setting will take the form of "group ostracism," rather than the "individual ostracism" in the small group setting.

In what follows, let

$$\begin{aligned} \check{x} &= lg(\check{y}) + \delta(k-1) f(\check{x}) \\ \check{y} &= \left(\frac{\delta}{1-\delta} k - \delta\left(k - \frac{1}{n}\right)\gamma\right) f(\check{x}). \end{aligned}$$

Also, let

$$y^E(\gamma) = \left\{ \begin{array}{ccc} \infty & \text{if} & \lim_{y\to\infty} g'(y) \geq m/\gamma \\ (g')^{-1}(m/\gamma) & \text{if} & g'(0) > m/\gamma > \lim_{y\to\infty} g'(y) \\ 0 & \text{if} & g'(0) \leq m/\gamma \end{array} \right\},$$

and let

$$\begin{aligned} \tilde{x} &= lg\left(y^E(\gamma)\right) + \delta(k-1) w f(\tilde{x}) \\ y^E(\gamma) &= \left(\frac{\delta}{1-\delta} k - \delta\left(k - \frac{1}{n}\right)\gamma w\right) f(\tilde{x}). \end{aligned}$$

The definition of (x^*, y^*) is unchanged. In addition, for the reasons we have just explained, we modify the definition of one-time enforcer punishment plus ostracism strategies to specify that, in the period following a producer deviation, all producers except the deviator ostracize *all other* *producers*, rather than only the deviator.

With these modifications (and the modified interpretation of ostracism), Theorem C1 generalizes verbatim to large groups.

Theorem C2 *Suppose the players can employ ostracism with efficiency cost γ.*

If $y^E(\gamma) \geq y^*$, one-time enforcer punishment strategies are optimal, and the maximum level of cooperation is x^*.

If $y^E(\gamma) \leq \check{y}$, one-time enforcer punishment plus ostracism strategies are optimal, and the maximum level of cooperation is $\check{x}$.

If $y^E(\gamma) \in (\check{y}, y^*)$, one-time enforcer punishment plus partial ostracism strategies are optimal, and the maximum level of cooperation is $\tilde{x}$.

Proof. It is straightforward to see that the strategy profiles described in the theorem are equilibria. We prove that x^* (resp., $\check{x}$) is an upper bound on the maximum level of cooperation when $y^E \geq y^*(\gamma)$ (resp., $y^E \leq \check{y}(\gamma)$). The argument closely parallels the proof of Theorem 3, so we omit some details.

We first observe that a one-time enforcer punishment plus partial ostracism equilibrium with $y = y^E$ can exist only if $y^E \leq y^*$.

Lemma C1 *If $y^E > y^*$ then the system of equations (C1) does not have a solution. That is, an enforcer punishment plus partial ostracism equilibrium with $y = y^E$ does not exist.*

Proof. Analogous to the proof of Lemma 3. ■

Turning to the proof of the theorem, the substance of the result is again that x^* (resp., $\check{x}$, $\tilde{x}$) is an upper bound on each producer's level of cooperation in any PBE when $y^{FB} \geq y^*$ (resp., $y^{FB} \leq \check{y}$, $y^{FB} \in (\check{y}, y^*)$). We break the proof into steps analogous to the proof of Theorem 3.

Using Incentive Compatibility to Bound $\underline{u}$:

For each producer i and player j, let

$$\tilde{f}_{ji}\left(x_i^t, w_i^t\right) = \left\{ \begin{array}{c} f\left(x_i^t\right) \text{ if } i \text{ does not ostracize anyone in period } t \\ \left(1 - w_i^t \gamma\right) f\left(x_i^t\right) \text{ if } i \text{ ostracizes someone other than } j \text{ in period } t \\ \left(1 - w_i^t\right) f\left(x_i^t\right) \text{ if } i \text{ ostracizes } j \text{ in period } t \end{array} \right\}.$$

As in the proof of Theorem 3, a necessary condition for producer i not to deviate to playing $x_i = 0$ at history h^t is that, for all $x_i^t \in \operatorname{supp} \sigma_i\left(h^t\right)$,

$$(1-\delta) x_i^t + \delta \mathbb{E}\left[X_i^{t+1} | h^t, x_i^t\right] \leq \delta (1-\delta) \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[\sum_{j \in M_i^{t+1+\tau} \cap P \setminus i} \tilde{f}_{ij}\left(x_j^{t+1+\tau}, w_j^{t+1+\tau}\right) | h^t, x_i^t\right] - \underline{u}. \tag{C2}$$

Similarly, a necessary condition for enforcer k not to deviate to playing $y_{ki}=0$ at history h^t is

$$y_{ki}^t \leq \delta \sum_{\tau=0}^{\infty} \delta^{\tau} \mathbb{E}\left[\sum_{j \in M_k^{t+1+\tau} \cap P} \tilde{f}_{kj}\left(x_j^{t+1+\tau}, w_j^{t+1+\tau}\right) |h^t \right]. \tag{C3}$$

Combining these conditions as in the proof of Theorem 3, there must exist a producer i and a history h^t such that

$$\underline{u} \geq -lg\left(\delta \sum_{\tau=0}^{\infty} \delta^{\tau} \mathbb{E}\left[\sum_{j \in M_k^{t+1+\tau} \cap P} \tilde{f}_{kj}\left(x_j^{t+1+\tau}, w_j^{t+1+\tau}\right) |h^t \right]\right) + \delta \mathbb{E}\left[\sum_{j \in M_i^{t+1} \cap P \setminus i} \tilde{f}_{ij}\left(x_j^{t+1}, w_j^{t+1}\right) |h^t \right]. \tag{C4}$$

Bounding $\underline{u}$ in terms of $\bar{X}$:

By the definition of $\bar{X}$, for every producer j, history h^{t+1}, and level of cooperation $x_j^{t+1} \in \operatorname{supp} \sigma_j\left(h^{t+1}\right)$, we have

$$x_j^{t+1} \leq \frac{1}{1-\delta}\bar{X} - \delta \sum_{\tau=0}^{\infty} \delta^{\tau} \mathbb{E}\left[x_j^{t+2+\tau} | h^{t+1}, x_j^{t+1}\right]. \tag{C5}$$

Now, in order to reduce the right-hand side of (C4) as much as possible, it is optimal to replace x_j^{t+1} with its upper bound in (C5) for all j. This follows because the right-hand side of (C4) is decreasing in the $\tilde{f}_{kj}$ terms and increasing in the $\tilde{f}_{ij}$ terms, so it is minimized when, for a given set of values of the $\tilde{f}_{kj}$ terms, the $\tilde{f}_{ij}$ terms are reduced as far as possible. This is achieved when only the producers (and not the enforcers) are ostracized in period t, in which case

$$\begin{aligned}
\tilde{f}_{kj}\left(x_j^{t+1}, w_j^{t+1}\right) &= \left(1-\gamma w_j^{t+1}\right) f\left(x_j^{t+1}\right), \\
\tilde{f}_{ij}\left(x_j^{t+1}, w_j^{t+1}\right) &= \left(1-w_j^{t+1}\right) f\left(x_j^{t+1}\right) = \frac{1-w_j^{t+1}}{1-\gamma w_j^{t+1}} \tilde{f}_{kj}\left(x_j^{t+1}\right).
\end{aligned}$$

It is now clear that, for a fixed value of $\tilde{f}_{kj}\left(x_j^{t+1}\right) \leq f\left(\frac{1}{1-\delta}\bar{X} - \delta \sum_{\tau=0}^{\infty} \delta^{\tau} \mathbb{E}\left[x_j^{t+2+\tau} | h^{t+1}, x_j^{t+1}\right]\right)$, $\tilde{f}_{ij}\left(x_j^{t+1}\right)$ is minimized by setting $x_j^{t+1} = \frac{1}{1-\delta}\bar{X} - \delta \sum_{\tau=0}^{\infty} \delta^{\tau} \mathbb{E}\left[x_j^{t+2+\tau} | h^{t+1}, x_j^{t+1}\right]$ and setting w_j^{t+1} to attain the given value of $\tilde{f}_{kj}\left(x_j^{t+1}\right)$.

After replacing x_j^{t+1} with its upper bound in (C5) for all j, we consider three cases analogous to those in the proof of Theorem 3.

Case 1: This case applies if, after setting $w_j^{t+1}=0$ for all j in (C4), the resulting argument of g is less than y^E.

Case 2: This case applies if, after setting $w_i^{t+1}=0$ and setting $w_j^{t+1}=1$ for all $j \in P \setminus i$ in (C4), the resulting argument of g is greater than y^E.

Case 3: This case applies when Cases 1 and 2 do not apply. Note that, in this case, there exists a unique value of w such that

$$\delta\mathbb{E}\left[f\left(\frac{1}{1-\delta}\bar{X}-\delta\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[x_i^{t+2+\tau}|h^{t+1},x_i^{t+1}\right]\right)|h^t\right]+\delta^2\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[f\left(x_i^{t+2+\tau}\right)|h^t\right]$$
$$+\delta w\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap P\backslash i}f\left(x_j^{t+1}\right)|h^t\right]+\delta^2\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap P\backslash i}f\left(x_j^{t+2+\tau}\right)|h^t\right]=y^E.$$

Call this value w^*.

We now argue that, in Case 1, the bound (C4) can be relaxed to

$$\underline{u}\geq -lg\left(\frac{\delta}{1-\delta}kf\left(\bar{X}\right)\right)+\delta\left(k-1\right)f\left(\bar{X}\right). \tag{C6}$$

Similarly, we argue that in Case 2 (C4) can be relaxed to

$$\underline{u}\geq -lg\left(\left(\frac{\delta}{1-\delta}k-\delta\left(k-\frac{1}{n}\right)\gamma\right)f\left(\bar{X}\right)\right), \tag{C7}$$

and in Case 3 (C4) can be relaxed to

$$\underline{u}\geq -lg\left(y^E\right)+\delta\left(k-1\right)f\left(\bar{X}\right)-\frac{k-1}{k-\frac{1}{n}}\frac{1}{\gamma}\left(\frac{\delta}{1-\delta}kf\left(\bar{X}\right)-y^E\right). \tag{C8}$$

The argument for Case 1 is exactly as in the proof of Theorem 3.

For Case 2, replacing x_j^{t+1} with its upper bound in (C5) for all j and setting $w_j^{t+1}=1$ for all $j\in P\backslash i$ relaxes (C4). The resulting lower bound equals

$$-lg\left(\begin{array}{c}\frac{1}{n}\left[\begin{array}{c}\delta\mathbb{E}\left[f\left(\frac{1}{1-\delta}\bar{X}-\delta\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[x_i^{t+2+\tau}|h^{t+1},x_i^{t+1}\right]\right)|h^t\right]\\+\delta^2\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[f\left(x_i^{t+2+\tau}\right)|h^t\right]\end{array}\right]\\+\delta\left(1-\gamma\right)\mathbb{E}\left[\sum_{j\in M_k^{t+1}\cap P\backslash i}f\left(\frac{1}{1-\delta}\bar{X}-\delta\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[x_j^{t+2+\tau}|h^{t+1},x_j^{t+1}\right]\right)|h^t\right]\\+\delta^2\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[\sum_{j\in M_k^{t+2+\tau}\cap P\backslash i}f\left(x_j^{t+2+\tau}\right)|h^t\right]\end{array}\right). \tag{C9}$$

As we have seen,

$$\delta\mathbb{E}\left[f\left(\frac{1}{1-\delta}\bar{X}-\delta\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[x_i^{t+2+\tau}|h^{t+1},x_i^{t+1}\right]\right)|h^t\right]$$
$$+\delta^2\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[f\left(x_i^{t+2+\tau}\right)|h^t\right]\leq\frac{\delta}{1-\delta}f\left(\bar{X}\right).$$

By the same argument, for any $j \in P \backslash i$,

$$\begin{aligned}\delta (1-\gamma) \mathbb{E}\left[f\left(\frac{1}{1-\delta} \bar{X} - \delta \sum_{\tau=0}^{\infty} \delta^{\tau} \mathbb{E}\left[x_j^{t+2+\tau} | h^{t+1}, x_j^{t+1} \right] \right) | h^t \right] \\ + \delta^2 \sum_{\tau=0}^{\infty} \delta^{\tau} \mathbb{E}\left[f\left(x_j^{t+2+\tau} \right) | h^t \right] \leq \left(\delta (1-\gamma) + \frac{\delta^2}{1-\delta} \right) f\left(\bar{X} \right),\end{aligned}$$

and therefore

$$\begin{aligned}\delta (1-\gamma) \mathbb{E}\left[\sum_{j \in M_k^{t+1} \cap P \backslash i} f\left(\frac{1}{1-\delta} \bar{X} - \delta \sum_{\tau=0}^{\infty} \delta^{\tau} \mathbb{E}\left[x_j^{t+2+\tau} | h^{t+1}, x_j^{t+1} \right] \right) | h^t \right] \\ + \delta^2 \sum_{\tau=0}^{\infty} \delta^{\tau} \mathbb{E}\left[\sum_{j \in M_k^{t+2+\tau} \cap P \backslash i} f\left(x_j^{t+2+\tau} \right) | h^t \right] \leq \left(k - \frac{1}{n} \right) \left(\delta (1-\gamma) + \frac{\delta^2}{1-\delta} \right) f\left(\bar{X} \right).\end{aligned}$$

Combining these inequalities, and noting that $\frac{1}{n} \frac{\delta}{1-\delta} + \left(k - \frac{1}{n} \right) \left(\delta (1-\gamma) + \frac{\delta^2}{1-\delta} \right) = \frac{\delta}{1-\delta} k - \delta \left(k - \frac{1}{n} \right) \gamma$, this yields (C7).

Finally, consider Case 3. Here, (C4) equals

$$\begin{aligned} & -lg\left(y^E \right) + \delta w^* \mathbb{E}\left[\sum_{j \in M_i^{t+1} \cap P \backslash i} f\left(x_j^{t+1} \right) | h^t \right] \\ = & -lg\left(y^E \right) - \delta \mathbb{E}\left[f\left(\frac{1}{1-\delta} \bar{X} - \delta \sum_{\tau=0}^{\infty} \delta^{\tau} \mathbb{E}\left[x_i^{t+2+\tau} | h^{t+1}, x_i^{t+1} \right] \right) | h^t \right] - \delta^2 \sum_{\tau=0}^{\infty} \delta^{\tau} \mathbb{E}\left[f\left(x_i^{t+2+\tau} \right) | h^t \right] \\ & -\delta^2 \sum_{\tau=0}^{\infty} \delta^{\tau} \mathbb{E}\left[\sum_{j \in M_i^{t+1} \cap P \backslash i} f\left(x_j^{t+2+\tau} \right) | h^t \right] + y^E.\end{aligned}$$

As above, this bound can be relaxed to

$$-lg\left(y^E \right) - \frac{\delta}{1-\delta} f\left(\bar{X} \right) - \frac{\delta^2}{1-\delta} (k-1) f\left(\bar{X} \right) + y^E,$$

which equals the right-hand side of (C8).

Finishing the Proof:

As in the proof of Theorem 3, we finally show that the maximum level of cooperation is always either x^*, $\check{x}$, or $\tilde{x}$, and that if the maximum level of cooperation is x^* (resp., $\check{x}$, $\tilde{x}$) then y^E must be weakly greater than y^* (resp., weakly less than $\check{y}$, in between $\check{y}$ and y^*).

Fix a sequence of equilibria converging to the maximum level of cooperation for producer 1.Suppose that, for all $\varepsilon > 0$, there exists an equilibrium along the sequence yielding within ε of the maximum level of cooperation for which Case 1 applies for some producer i and history h^t. As in the proof of Theorem 3, it follows that one-time enforcer punishment strategies are optimal. Finally, this implies that $y^E \geq y^*$, as otherwise we could support more cooperation in enforcer punishment plus partial ostracism strategies for small $w > 0$. (This follows because $\frac{dx}{dw} > 0$ if $y > y^E$, which follows by the same argument that shows $\frac{dx}{d\eta} > 0$ if $y > y^E$ in Lemma A1.)

Next, suppose that, for all $\varepsilon > 0$, there exists an equilibrium along the sequence yielding within ε of the maximum level of cooperation for which Case 2 applies for some producer and history. Combining (C2) and (C7), we have, for every player i, history h^t, and level of cooperation $x_i^t \in \operatorname{supp} \sigma_i\left(h^t\right)$,

$$
\begin{aligned}
& (1-\delta)\, x_i^t + \delta \mathbb{E}\left[X_i^{t+1} | h^t, x_i^t\right] \\
\leq \;& lg\left(\left(\frac{\delta}{1-\delta}k - \delta\left(k - \frac{1}{n}\right)\gamma\right) f\left(\bar{X}\right)\right) + \delta\,(1-\delta)\sum_{\tau=0}^{\infty}\delta^{\tau}\left[\sum_{j \in M_i^{t+1+\tau} \cap P \backslash i} f\left(x_j^{t+1+\tau}\right) | h^t, x_i^t\right] \\
\leq \;& lg\left(\left(\frac{\delta}{1-\delta}k - \delta\left(k - \frac{1}{n}\right)\gamma\right) f\left(\bar{X}\right)\right) + \delta\,(k-1)\, f\left(\bar{X}\right).
\end{aligned}
$$

As above, this gives $\bar{X} \leq lg\left(\left(\frac{\delta}{1-\delta}k - \delta\left(k - \frac{1}{n}\right)\gamma\right) f\left(\bar{X}\right)\right) + \delta\,(k-1)\, f\left(\bar{X}\right)$. By the definition of $\check{x}$, this implies that $\bar{X} \leq \check{x}$, and hence $\mathbb{E}\left[X_i^0 | h^0\right] \leq \check{x}$. Thus, in this case $\check{x}$ is an upper bound on each player's maximum equilibrium level of cooperation, and hence enforcer punishment plus ostracism is optimal. And, if these strategies are optimal, then $y^E \leq \check{y}$, as otherwise we could support more cooperation in enforcer punishment plus partial ostracism strategies with w less than but sufficiently close to 1. (This follows because $\frac{dx}{dw} < 0$ if $y < y^E$, which follows by the same argument that established $\frac{dx}{d\eta} < 0$ if $y < y^E$ in Lemma A1.)

Finally, suppose that, for all $\varepsilon > 0$, exists an equilibrium along the sequence yielding within ε of the maximum level of cooperation for which Case 3 applies for some producer and history. Combining (C2) and (C8), we have, for every player i, history h^t, and level of cooperation $x_i^t \in \operatorname{supp} \sigma_i\left(h^t\right)$,

$$
\begin{aligned}
& (1-\delta)\, x_i^t + \delta \mathbb{E}\left[X_i^{t+1} | h^t, x_i^t\right] \\
\leq \;& lg\left(y^E\right) - \delta\,(k-1)\, f\left(\bar{X}\right) + \frac{k-1}{k-\frac{1}{n}}\frac{1}{\gamma}\left(\frac{\delta}{1-\delta}kf\left(\bar{X}\right) - y^E\right) \\
& + \delta\,(1-\delta)\sum_{\tau=0}^{\infty}\delta^{\tau}\left[\sum_{j \in M_i^{t+1+\tau} \cap P \backslash i} f\left(x_j^{t+1+\tau}\right) | h^t, x_i^t\right] \\
\leq \;& lg\left(y^E\right) + \frac{k-1}{k-\frac{1}{n}}\frac{1}{\gamma}\left(\frac{\delta}{1-\delta}kf\left(\bar{X}\right) - y^E\right).
\end{aligned}
$$

This gives $\bar{X} \leq lg\left(y^E\right) + \frac{k-1}{k-\frac{1}{n}}\frac{1}{\gamma}\left(\frac{\delta}{1-\delta}kf\left(\bar{X}\right) - y^E\right)$. By the definition of $\tilde{x}$, this implies that $\bar{X} \leq \tilde{x}$, and hence $\mathbb{E}\left[X_i^0 | h^0\right] \leq \tilde{x}$. Thus, now $\tilde{x}$ is an upper bound on each player's maximum equilibrium level of cooperation, and enforcer punishment plus partial repentance is optimal. By Lemma C1, such an equilibrium exists only if $y^E \leq y^*$. Finally, these strategies can be optimal only if $y^E \geq \check{y}$, as since $\tilde{x} \geq \check{x}$ we have

$$\begin{aligned} y^E &= \left(\frac{\delta}{1-\delta}k - \delta\left(k - \frac{1}{n}\right)w^E\gamma\right)f\left(\tilde{x}\right) \geq \left(\frac{\delta}{1-\delta}k - \delta\left(k - \frac{1}{n}\right)\gamma\right)f\left(\tilde{x}\right) \\ &\geq \left(\frac{\delta}{1-\delta}k - \delta\left(k - \frac{1}{n}\right)\gamma\right)f\left(\check{x}\right) = \check{y}. \end{aligned}$$

■