Sustaining Cooperation: Community Enforcement vs. Specialized Enforcement*

Daron Acemoglu
MIT

Alexander Wolitzky
MIT

February 2016

Abstract

We introduce the possibility of direct punishment by specialized enforcers into a model of community enforcement. Specialized enforcers need to be given incentives to carry out costly punishments. Our main result shows that, 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 actions, 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 ineffective, optimal equilibria do punish deviations by regular agents with community enforcement. Our results hold both under perfect monitoring of actions and under various types of private monitoring.

Keywords: cooperation, community enforcement, law enforcement, repeated games
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 Toikka, and seminar participants at Bonn, BU, Cambridge, 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.

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, 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, 1990) 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 contagion toward 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 best arrangement for sustaining cooperation.

We consider a model of cooperation within a group of agents. In our baseline model, regular citizens ("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 directly punish producers who deviate—provided that enforcers can be given the proper 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. 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. In this case, we establish a simple condition on the effectiveness of the specialized enforcement (direct punishment) technology under which the maximum level of cooperation is sustained by a simple form of specialized enforcement—*one-time enforcer punishment strategies*—wherein all punishment for producer deviations takes place instantaneously and is carried out by enforcers, and there is no contagion or withholding of future cooperation. In our baseline setting, enforcers are incentivized to undertake such costly punishments because, unlike deviations by the producers themselves, deviations by enforcers trigger contagion among the producers.[2] Conversely, when direct punishments are ineffective, the maximum level of cooperation is maintained by community enforcement. The model thus predicts that—for example—societies with more effective technologies in both production and punishment should rely on specialized enforcement, while societies with less effective technologies should rely on community enforcement.

The form of our specialized enforcement equilibrium can be viewed as a stylized representation of modern state-society relations. 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 the enforcers representing the state deviate from their expected course of behavior, societal cooperation collapses, and it is the prospect of such a collapse that incentivizes enforcers.

The optimality of one-time enforcer punishment strategies in our model may appear surprising, as one might have conjectured that it would be better to combine specialized enforcer punishments with decentralized community enforcement: if both direct punishment and the withdrawal of cooperation are bad for producers, why not use both to provide incentives? The intuition for this result helps highlight the novel economic mechanism at the heart of our paper. Adding decentralized

[1] In practice, another important problem is ensuring that enforces do not use their access to violence to expropriate others. 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.

punishment—for example, some degree of contagion—to a *given* level of specialized punishment would indeed improve producers' incentives for cooperation. And yet, crucially, it would also erode the incentives of enforcers to undertake direct punishment: enforcers are willing to undertake costly punishments today only because of the future reward of continued societal cooperation. Hence, if a deviation by a producer also triggered partial or full contagion, then this implicit reward for enforcers would be diminished, undermining the power of enforcer punishments. This reasoning thus identifies a novel—and in our setting, quite powerful—cost of decentralized punishment: its negative impact on the extent and efficacy of specialized punishment.

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 in only 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 (though we do not completely characterize the optimal equilibrium in this more general environment).

Second, we establish that when individuals observe behavior in their partners' most recent matches, one-time enforcer punishment strategies form an optimal equilibrium unless the group can somehow benefit from the imperfections in the monitoring structure in this setting. While we show by example that this is in fact possible in general, we also provide two natural settings in which one-time enforcer punishment strategies are indeed optimal. First, this is the case if enforcers are better informed than producers.[3] Second, 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. 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 a small probability. Under enforcer punishments, however, a similarly costly contagion can occur only if *both* individual producers tremble *and* an enforcer trembles in response. This makes accidental contagion much less likely under enforcer punishments.

We also consider one main extension of our model, where we allow producers to directly

[3] 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.

exclude—or *ostracize*—each other from the benefits of cooperation, while continuing to provide benefits to enforcers. This extension softens the key tradeoff in our model between withdrawing future cooperation from deviators and using the promise of future cooperation to give enforcers incentives to punish. Nonetheless, the optimal arrangement in this setting either remains one-time enforcer punishment strategies or becomes a combination of these strategies with the short-term ostracism of deviant producers. The simple structure of optimal equilibria is thus largely preserved in the presence of ostracism.

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 (2014a), 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 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 (2014b) show that contagion 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.[4]

Second, our paper is also related to the literature on optimal penal codes in general repeated games (Abreu, 1988, see Mailath and Samuelson, 2006, for a survey), 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 only optimal within the class of 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 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.

[4] 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.

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; Mailath, Morris, and Postlewaite (2007), who develop a model of laws and authority based on cheap talk; Levine and Modica (2014), who study the problem of designing a specialized enforcement technology to sustain group cooperation and emphasize the tradeoff between providing insufficient incentives for cooperation and expending excessive effort in punishment; and Acemoglu and Jackson (2014) who analyze the converse problem to the one in this literature—studying how social norms can constrain the effectiveness of laws. Two recent papers, Masten and Prüfer (2014) and Aldashev and Zanarone (2015), also 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. Neither paper, nor any other of which we are aware, considers whether one-time enforcer punishments are part of an optimal equilibrium.

Finally, 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 enforcement 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 presents the model. Section 3 presents our main result on the optimality of specialized enforcement under perfect monitoring. Section 4 discusses various modeling issues. Section 5 presents our results with private monitoring. Section 6 extends the model to allow for ostracism. Section 7 concludes. Omitted proofs and additional results are included in the Appendix.

2 Model

We first introduce the basic environment and information structure, and then define the key concepts of contagion strategies and one-time enforcer punishment strategies.

We note at the outset that we consider a simple model in order to bring out our main insights on the relationship between community enforcement and specialized enforcement as cleanly as possible. In Section 4, we show that many of our simplifying assumptions are irrelevant for our main results.

2.1 Environment

There is a population of $(k+l)\,n$ players, with $k, l, n \geq 1$. Out of the $(k+l)\,n$ players, kn of them are producers (or "citizens") and ln of them are enforcers (or "police"). Denote the set of producers by C, the set of enforcers by P, and the set of all players by N. 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. The game is a version of the prisoner's dilemma among the producers, with the possibility of costly punishment by the enforcers.

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.[5] The vector $(i, x_i)_{i \in M \cap C}$ is then 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 $f\,(x_i)$, where $f : \mathbb{R}_+ \to \mathbb{R}_+$ is an increasing, concave, bounded, and differentiable function satisfying $f\,(0) = 0$.

2. *Punishment Stage:* Each enforcer $j \in M$ then chooses a *level of punishment* $y_{ji} \in \mathbb{R}_+$ for each producer $i \in M \cap C$. The vector $(j, i, y_{ji})_{j \in M \cap P, i \in M \cap C}$ is perfectly observed by all players in M. Choosing punishment level y_{ji} costs y_{ji} for player j, and hurts player i by $g\,(y_{ji})$, where $g : \mathbb{R}_+ \to \mathbb{R}_+$ is an increasing and differentiable function satisfying $g\,(0) = 0$. We refer to g as the *specialized enforcement technology.*

Thus, producer i's stage payoff is

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

and enforcer j's stage payoff is

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

[5] We refer to the feature that producers act without knowing their partners' identities as *partial anonymity.* This assumption plays an important role in our results, as we discuss further in Section 4.

Note that playing $x_i = 0$ ("shirking") is myopically optimal for producer i, and playing $y_{ji} = 0$ for all $i \in M_j \cap C$ ("failing to punish") is myopically optimal for enforcer j. Thus, only the shadow of future interactions can incentivize producers to cooperate and enforcers to punish.

For some secondary results, we also consider a variation with "money burning," in which producers (resp., enforcers) are also allowed to publicly destroy some of their *own* utility at the same time that they choose their level of cooperation (resp., punishment). The only difference between this type of money burning and cooperation (resp., punishment) is that cooperation (resp., punishment) also benefits (resp., hurts) other players. We explicitly indicate below which of our results involve the possibility of money burning.

We refer to the pair $\left((i, x_i)_{i \in M \cap C}, (j, i, y_{ji})_{j \in M \cap P, i \in M \cap C}\right)$ 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 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 of player i's at the beginning of period t, where we omit the subscript in the perfect monitoring case. The trivial initial history is denoted by h^0. We also denote a generic strategy of player i's 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 with perfect monitoring is subgame perfect equilibrium (SPE). The solution concept with private monitoring is weak perfect Bayesian equilibrium (PBE), with the additional requirement that the equilibrium assessment is derived from a common conditional probability system (Myerson, 1991).[6]

Note that the assumption that f is concave implies that there is a technological advantage to spreading out cooperation over time. Similarly, g can be concave, so there can also be a technological advantage to spreading out punishments over time (though we do not impose in our main analysis that g is concave or convex). Nevertheless, we will show that, while optimal equilibria do spread cooperation over time, they do *not* spread punishments over time. Instead, optimal equilibria either do not use punishments at all or concentrate them in a single period.

[6] 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.

2.2 Contagion Strategies and One-Time Enforcer Punishment 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 $\mathbb{E}\left[X_i^0|h^0\right]$ 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 $\mathbb{E}\left[X_i^0|h^0\right]$ for every player i among all equilibria (with the solution concept of SPE for the perfect monitoring version and PBE for the private monitoring version); and we call the corresponding value of $\mathbb{E}\left[X_i^0|h^0\right]$ the *maximum level of cooperation.* Note that, by concavity of f, the most cooperation equilibrium is also the *optimal* equilibrium in terms of utilitarian social welfare if 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 $(k+l-1)\,f'\left(x^{FB}\right)=1$. This is indeed the main case of economic interest, as the main problem in most settings is providing sufficient incentives for cooperation rather than avoiding excessive levels of 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). An extreme form of community enforcement, where enforcers are inactive, is given by contagion strategies.

Definition 1 *A* contagion *(or* grim trigger*)* strategy profile *is characterized by a cooperation level $x \in \mathbb{R}_+$ and is represented by the following 2-state automaton:*

Producers have two states, normal and infected. Producers play $x_i = x$ in the normal state and play $x_i = 0$ in the infected state. Producers start in the normal state, and permanently transition to the infected state if they observe the outcome of a match (including their own) in which some producer $i' \in C$ plays $x_{i'} \neq x$.

A contagion strategy profile involves no punishment by enforcers.

On the other hand, an extreme form of specialized enforcement, where there is no withdrawal of future cooperation at all following a deviation, is given by what we call *one-time enforcer punishment strategies.* With these strategies, a producer who deviates is immediately punished by the enforcers in her match. Following this one-time punishment, everyone returns to her normal behavior next period. If however an enforcer fails to punish a producer deviation, then this triggers contagion, and eventually all producers choose $x_i = 0$.

Definition 2 *A* one-time enforcer punishment strategy profile *is characterized by a cooperation level x and a punishment level y, and is represented by the following 2-state automaton:*

All players have two states, normal and infected. Play in the two states is as follows:

Normal state*: Producer i plays $x_i = x$. If all producers $i \in M_j \cap C$ play $x_i = x$, then enforcer j plays $y_{ji} = 0$ for all $i \in M_j \cap C$. If instead some producer $i \in M_j \cap C$ plays $x_i \neq x$, then enforcer j plays $y_{ji} = y$ for one of the producers $i \in M_j \cap C$ who played $x_i \neq x$—choosing arbitrarily if more than one producer i played $x_i \neq x$—and plays $y_{ji'} = 0$ for all $i' \in M_j \cap C \backslash i$.*

Infected state: *Players always take action* 0 *(producers never cooperate; enforcers never punish).*

Players start in the normal state, and permanently transition to the infected state if they observe the outcome of a match (including their own) in which some producer i plays $x_i \neq x$ and *some enforcer $j \in M_i \cap P$ then plays $y_{ji'} \neq y$ for all $i' \in M_j \cap C$.*

With perfect monitoring, everyone transitions to the infected state simultaneously under one-time enforcer punishment strategies (or contagion strategies). With private monitoring, transitions to the infected state follow a more general process of contagion, described below. Note that we have specified that if two producers deviate simultaneously, the corresponding enforcer is supposed to punish one of them.

Note also that under both contagion and one-time enforcer punishment strategies, punishments are not used on path and producers choose constant levels of cooperation. This implies that whenever a contagion equilibrium or a one-time punishment equilibrium sustains the maximum level of cooperation, it is also the optimal equilibrium in terms of utilitarian social welfare (provided that, as noted above, this maximum level of cooperation is below the first-best level).

3 Perfect Monitoring

We now characterize the most cooperative equilibrium under perfect monitoring.

Let (x^*, y^*) be the greatest solution to the following system of equations:

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

Intuitively, x^* and y^* are the greatest levels of cooperation and punishment that can be sustained with one-time enforcer punishment strategies. With these strategies, a producer who deviates gains at most x^* (her cost of effort) and loses $lg(y^*)$ (the cost of being punished at level y^* by l enforcers), while an enforcer who deviates gains at most y^* and loses $\frac{\delta}{1-\delta} kf(x^*)$ (the future benefit of cooperation at level x^* from k producers).

Similarly, let $\hat{x}$ be the greatest solution to the equation

$$\hat{x} = \delta (k-1) f(\hat{x}).$$

Intuitively, $\hat{x}$ is the greatest level of cooperation that can be sustained with contagion strategies.

Finally, define the parameter

$$m \equiv \frac{(k-1)n}{(kn-1)l} \in [0,1].$$

Note that m depends only on the number of producers and enforcers in the population, and not on the production and punishment technologies f and g or the discount factor δ.

Our main result for the perfect monitoring version of the model is the following:

Theorem 1 *With perfect monitoring,*

1. *If $g'(y) \geq m$ for all $y \in \mathbb{R}_+$, then the one-time enforcer punishment strategy profile with cooperation level x^* and punishment level y^* is the most cooperative equilibrium. Moreover, if g is concave, then for every $x < x^*$ the one-time enforcer punishment strategy profile with cooperation level x and punishment level $y = \frac{\delta}{1-\delta} kf(x)$ is also a SPE.*

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

The first part of Theorem 1 is our most important result: one-time enforcer punishment strategies attain the maximum level of cooperation whenever the specialized enforcement technology, g, is sufficiently effective.[7] It is useful to break the intuition for this result into two parts, relating to why the optimal equilibrium involves 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 continuation payoff would be reduced in two ways: enforcers would punish her, and other producers would refuse to cooperate with her. However, enforcers are only willing to exert effort in punishing the deviator if they are subsequently rewarded with cooperation from the producers. Since cooperation benefits both producers and enforcers, there is no way for producers to reward enforcers for punishing the deviant producer without also benefiting the deviator herself.[8] Society must

[7] The relevance of the observation that any cooperation level below x^* can also be supported in a one-time enforcer punishment equilibrium is simply that such equilibria remain optimal even when x^* is above x^{FB}. Unlike our other results, this observation requires that g is concave.

[8] In Section 6, we study how the structure of the optimal equilibrium changes if we allow for *ostracism*—the practice of excluding deviators from the benefits of cooperation.

then choose between incentivizing enforcers to punish deviators by subsequently restoring cooperation (one-time enforcer punishment strategies), or giving up on enforcer punishments and instead providing incentives by withdrawing cooperation following a deviation (contagion strategies).

We can quantify the tradeoff between incentivizing enforcers 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 $\frac{k-1}{kn-1}f'(x)$ units (here $\frac{k-1}{kn-1}$ is the probability that the deviator matches with a given producer in any period). This effect increases on-path incentives for cooperation. This direct effect is countered by the indirect effect of reducing the maximum level of punishment an 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 each enforcer can be induced to provide by $\frac{1}{n}f'(x)$ units (as $\frac{1}{n}$ is the probability that the enforcer in question matches with a given producer in any period), and each unit of reduced punishment increases the deviator's payoff by $lg'(y)$ (as the deviator is punished by l enforcers). Thus, the indirect effect of reducing on-path incentives for cooperation is $\frac{l}{n}f'(x)g'(y)$. Consequently, the overall impact of withdrawing producer cooperation following a deviation on on-path producer incentives is negative if and only if $g'(y) \geq \frac{(k-1)n}{(kn-1)l} = m$. Therefore, if $g'(y) \geq m$ for all y, after a producer deviation, it is better to rely solely on enforcer punishments (and return to full cooperation thereafter) rather than reducing other producers' cooperation levels.[9]

Why are deviators punished only once? One might have conjectured that, to provide the harshest deterrent against a deviation, the enforcers should punish a deviator several times for the same transgression (recall that enforcers observe producers' identities, so multiple rounds of punishment are feasible). 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 just as low as her continuation payoff from being punished repeatedly while shirking (and in fact it is strictly lower, as the deviator's own future cooperation can be used to give enforcers additional incentives to punish her).[10]

The proof of Theorem 1 (in Appendix A) also shows that if $g'(y)$ is *strictly* greater than m for all y, then the one-time enforcer punishment strategy profile with cooperation level x^* and punishment

[9] The derivative condition $g'(y) \geq m$ for all $y \in \mathbb{R}_+$ can also be weakened to the slope condition $\frac{g(f(\bar{x}))-g(y)}{f(\bar{x})-y} \geq m$ for all $y \leq f(\bar{x})$, where $\bar{x}$ is defined as the greatest level of cooperation ever played in any SPE. With this modification, the first case in Theorem 1 can apply even if g satisfies the Inada condition $\lim_{y\to\infty} g'(y) = 0$.

[10] Another way of seeing the intuition is noting that, in the most cooperative equilibrium, a producer is indifferent between following her equilibrium strategy and following the policy of *always* shirking. If enforcers 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.

level y^* is essentially the *unique* most cooperative equilibrium. Specifically, any equilibrium that supports cooperation level x^* for each producer must have the following features:

1. Each producer i plays $x_i = x^*$ at every on-path history.
2. If a single producer i deviates to $x_i = 0$ at an on-path history, she is punished at level y^*, and the path of play then returns to all producers' playing x^* forever.
3. If a single producer i deviates to $x_i = 0$ at an on-path history and an enforcer $j \in M_i \cap P$ deviates to $y_{ji} = 0$, then all producers stop cooperating forever.
4. The path of continuation play where all producers play x^* forever is always supported by the threat of punishment at level y^*—which in turn is always supported by the threat of all producers' withdrawing cooperation—even when this continuation path starts at an off-path history.[11]

Theorem 1 holds whether or not money burning is available. In addition, when money burning is available, the lower bound on g' in part 1 of the theorem is tight:

Proposition 1 *With perfect monitoring, if $g'(y) < m$ for all $y \in \mathbb{R}_+$ and money burning is available, then one-time enforcer punishment strategies are not optimal.*

Proposition 1 shows that, if $g'(y) < m$ for all y, then optimal equilibria must involve some form of community enforcement (at least when money burning is available), meaning that deviations will be punished not only by specialized enforcement but also by the withholding of future cooperation. The details and a discussion of the role of money burning are provided in Section 6.

However, there is a sense in which Proposition 1 overstates the case for hybrid forms of community enforcement. 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 pure contagion strategies.

Theorem 2 *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*

$$g'(y) \leq \min\left\{\frac{\left(\bar{k}-1\right)n}{\left(\bar{k}n-1\right)l}, \frac{1}{\frac{1}{n}+\frac{\delta}{1-\delta}s}\right\} \text{ for all } y \in \mathbb{R}_+ \tag{1}$$

[11] There are, however, other equilibria which support cooperation level x^* and differ from single punishment strategies in various inessential ways. For example, equilibrium play after multiple simultaneous deviations can be specified arbitrarily, and there is some flexibility in specifying play after deviations to actions that are less tempting than $x_i = 0$ and $y_{ji} = 0$.

then the social planner would prefer to have all agents become producers (i.e., set $k = s$) and support cooperation using contagion strategies.

Thus, except when δ is very large (in which case the right-hand side of (1) goes to 0), in much of the region covered by Proposition 1, pure contagion is the optimal arrangement once one takes into account the social cost of allocating individuals to enforcement roles.

It is also straightforward to show that, conversely, if $g'(y)$ is sufficiently large for all y then it is optimal to allocate some individuals to enforcement and rely on one-time punishment strategies.[12]

Finally, we remark that the conclusion of Theorem 1 that one-time enforcer punishment equilibria are optimal if the specialized enforcement technology g is sufficiently effective, and that community enforcement is optimal if this technology is ineffective, is *independent* of the production technology f. Intuitively, improvements in the production technology increase the greatest level of cooperation that can be sustained in both one-time enforcer punishment equilibria and contagion equilibria in the same manner, so improvements in the production technology cancel out when comparing the two kinds of equilibria. An interesting implication is that, provided the efficiency of production and punishment technologies are positively related across different societies, Theorem 1 predicts that societies with more effective technologies should rely on one-time enforcer punishment strategies, while societies with less effective technologies should rely on contagion strategies.

4 Discussion of Model Assumptions

Before turning to the private monitoring version of the model, we briefly discuss several of our underlying assumptions, indicating how our results and their interpretation do or do not depend on the various assumptions.

4.1 Assumptions about the Role of Enforcers

Enforcers can only punish producers: We have assumed that enforcers can punish producers but 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

[12] The condition for this to be the case is

$$s\left[(s-1)\, f(\hat{x}) - \hat{x}\right] < k^*\left[(s-1)\, f\left(x^*\left(k^*\right)\right) - x^*\left(k^*\right)\right],$$

where $\hat{x} = \delta\,(s-1)\, f(\hat{x})$, $x^*(k) = (s-k)\, g\left(\frac{\delta}{1-\delta} k f\left(x^*(k)\right)\right)$, and k^* is the welfare-maximizing number of enforcers under specialized enforcement, given by $k^* = \text{argmax}_{k\in\{1,\ldots,s\}}\, k\left[(s-1)\, f\left(x^*(k)\right) - x^*(k)\right]$. For example, in the linear case $g(y) = \alpha y$, this condition implicitly defines a lower bound on α, as the right-hand side of the inequality is non-decreasing in α and goes to infinity as $\alpha \to \infty$.

period $t+2$. With this modified definition, Theorem 1 applies verbatim, with the sole modification that the formula for y^* changes from $\frac{\delta}{1-\delta}kf(x^*)$ to $\frac{\delta}{1-\delta}(kf(x^*)+(l-1)g(y^*))$, while the formula for x^* as a function of y^* remains $lg(y^*)$. (See the Appendix for a formal statement and proof).

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 direct 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 concern.

An alternative way of extending the model 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 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 state-society relations: 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 baseline 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, instead, 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. Without this assumption, it may be possible to exclude a deviator from future cooperation without simultaneously excluding the enforcers who punished her, which would let deviators be punished more harshly. In Section 6, we discuss an extension where the use of ostracism makes this type of selective exclusion possible.

This assumption of partial anonymity could also be replaced by one of several alternatives. In Section 5, we introduce the notion of non-discriminatory strategies, in which an individual's behavior does not depend on her partners' identities, except insofar as this is informative of past play. The point of our partial anonymity assumption is to guarantee that strategies are non-discriminatory, and all our perfect monitoring results can be derived by replacing partial anonymity with the requirement that strategies are non-discriminatory. We also note that the strongly symmetric strategies of Abreu (1986) (which impose symmetric play at all histories) are necessarily non-discriminatory, so without anonymity one-time enforcer punishment equilibria are also optimal in the class of strongly symmetric equilibria. Finally, as we discuss in Section 6, one-time enforcer punishment equilibria are also optimal even without anonymity when $n = 1$.

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 under the assumption of strong symmetry, 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.

4.2 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 (but see Section 6 below). An alternative 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 setup.

Bounded benefits from cooperation: We have assumed that f is bounded. This is for simplicity and can be replaced, with no change in any of our results, with the weaker assumptions that (1) $\lim_{x\to\infty} (k-1) f(x) - x = -\infty$, and (2) there exists $x_0 \in \mathbb{R}_+$ such that $lg\left(\frac{\delta}{1-\delta} kf(x)\right) - x < 0$ for all $x > x_0$. These assumptions hold if, for example, g' is bounded and f satisfies the Inada condition $\lim_{x\to\infty} f'(x) = 0$.

Enforcer payoffs: In addition to the public good and anonymous bilateral cooperation interpretations of our model, a third way to motivate why enforcers' payoffs depend on producers' levels of cooperation is to assume that the enforcers can impose a tax on the producers' output $\sum_{i\in C} f(x_i)$ of up to some maximum rate $\tau < 1$, with the proceeds split equally among themselves. If an enforcer's failure to punish a deviant producer leads to contagion, each enforcer's future benefits then decline because the tax revenues that determine their payoffs dry up.

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 enforcers' misusing 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. This cost was presumably much greater in early societies—where the weapons, technology, and organization available to law enforcement and to criminals were much closer to each other—than it is today.

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 Theorem 1 holds without modification when transfers from producers are allowed. Intuitively, it is inefficient to ask 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" instrument, which would undercut the separation of roles between producers and enforcers.

5 Private Monitoring

The perfect monitoring version of our model has the advantage of bringing out our main insights on the optimal structure of specialized enforcement in a particularly simple way. Yet, perfect monitoring is generally not a satisfactory assumption for studying community enforcement, as one of the main motivations for analyzing these settings is to understand how groups can sustain cooperation when individuals have limited information about each other's past behavior. We therefore turn to the question of whether one-time enforcer punishment equilibria remain optimal under private monitoring. Our answer is nuanced but generally positive: one-time enforcer punishment equilibria still outperform contagion equilibria in a wide range of private monitoring environments, and in some environments they remain optimal among all equilibria satisfying some natural conditions.

More specifically, we first provide conditions under which one-time enforcer punishment equilibria outperform contagion equilibria 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 then 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. This informational assumption is inconsistent with players' remaining completely anonymous at the point where producers take actions, so it sometimes gives players *more* information than does the perfect monitoring model of Section 3. With this information structure, one-time enforcer punishment strategies continue to sustain the same level of cooperation as with perfect monitoring, which implies that they must remain globally optimal unless it is possible to sustain *more* cooperation with observable last matches than with perfect monitoring. While we show that this is in fact possible, we also establish that one-time enforcer punishment equilibria continue to be optimal either 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.

5.1 General Network Monitoring

The setting considered here is one of general *network monitoring* (Wolitzky, 2013). At the end of each period t, a *monitoring network* $L_t = (l_{i,j,t})_{i,j \in N \times N}$, $l_{i,j,t} \in \{0,1\}$ is drawn independently from a fixed probability distribution μ on $\{0,1\}^{|N|^2}$. We assume that $\Pr^{\mu}\left((l_{i,j,t})_{i,j\in N\times N}\right) = \Pr^{\mu}\left((l_{\phi(i),\phi(j),t})_{i,j\in N\times N}\right)$ for any permutation $\phi : N \to N$, 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 analyze 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 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 (e.g., Wolitzky, 2013) that the greatest level of cooperation that can be sustained with contagion strategies, $\hat{x}$, is given by the greatest solution to

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

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 being presented 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 one-time 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 one-time enforcer punishment strategies is given by the greatest solution to

$$\begin{aligned} x^* &= lg\left(y^*\right), \\ y^* &= \sum_{t=0}^{\infty} \delta^t \frac{1}{n} q_t f\left(x^*\right). \end{aligned}$$

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).[13]

We can now inspect the formulas for $\hat{x}$ and x^* and make some basic observations about the relative performance of contagion and one-time enforcer punishment strategies. First, when the specialized enforcement technology is more effective (i.e., g is steeper), one-time enforcer punishment equilibria have an advantage over contagion strategies. Second, to the extent to which 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, one-time enforcer punishment strategies have a clear advantage when δ is close to 1, owing to the $(1-\delta)$ term in the definition of $\hat{x}$. This term reflects the fact that with contagion strategies, a producer who chooses $x = 0$ forever can only lose the future benefit of others' cooperation once, while with one-time enforcer punishment strategies, a producer who shirks forever is punished in every period, and each punishment is calibrated so as to be as costly for the enforcer as losing all future benefits of cooperation would be.[14]

The next theorem formalizes this comparison. The first part shows that, when the specialized enforcement technology is sufficiently effective, one-time enforcer punishment equilibria support more cooperation than do contagion equilibria. The second part establishes that the same conclu-

[13] 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 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.

[14] 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.

sion holds when the discount factor δ is sufficiently high.

Theorem 3 *With general network monitoring,*

1. *There exists $\bar{g}$ such that if $g'(y) \geq \bar{g}$ 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 that $\lim_{y\to\infty} g(y) = \infty$.*[15] *Then there exists $\bar{\delta}$ such that if $\delta \geq \bar{\delta}$ then one-time enforcer punishment strategies form a PBE strategy profile and support greater cooperation than contagion strategies.*

We also note that comparing one-time enforcer punishment equilibria and contagion equilibria is not as ad hoc at it might seem, as there is a sense in which contagion equilibria are optimal among all equilibria in which enforcers never punish. In particular, Wolitzky (2013) shows that under general network monitoring without enforcers, contagion strategies attain the maximum level of cooperation (provided that the realized monitoring network is observable). Thus, whenever one-time enforcer punishment equilibria outperform contagion equilibria, they outperform any equilibrium that does not rely on the enforcers.[16]

5.2 Observable Last Matches

We now turn to the second of the two private monitoring environments we consider: the *observable last matches* setting. This is the setting where players observe only the outcomes of their own matches and their current partners' most recent matches. In addition to being a natural benchmark, as we will see this setting is also quite tractable.

5.2.1 One-Time Enforcer Punishment Equilibria and Contagion Equilibria

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

[15] 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 4, the result still holds as long as $\lim_{x\to\infty} f'(x) < 1/l(k-1)$, which is consistent with $f = g$ and $\lim_{y\to\infty} g(y) = \infty$.

[16] 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 3, 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 whichever supports a higher level of cooperation will actually be worse in terms of welfare. Thus, the main point of Theorem 3 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.

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

Theorem 4 *With observable last matches,*

1. *If $x^* \geq \tilde{x}$, 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 one-time enforcer punishment strategies, spreads more 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.

5.2.2 Informed Enforcers

Theorem 4 shows that one-time enforcer punishment equilibria 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 one-time enforcer punishment equilibria remain globally optimal with observable last matches (when $g'\left(y\right) \geq m$ for all y). In the next subsection, we will see that the former possibility can sometimes arise. That result notwithstanding, we now show that, when enforcers have superior information to producers (i.e., have finer information sets), one-time enforcer punishment equilibria are indeed globally optimal, under a simple equilibrium refinement in the spirit of anonymity. Specifically, we consider the situation with *informed enforcers*, where enforcers perfectly observe all past actions, while producers have more limited information.

The following condition is in the spirit of the partial anonymity assumption of Section 2:

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. As noted in Section 4, partial anonymity implies non-discrimination. Both one-time enforcer punishment strategies and contagion strategies are clearly non-discriminatory.

Theorem 5 *Suppose that producers observe their partners' last matches and 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.*[17]

To provide intuition for this result, let us first revisit the case of perfect monitoring. An explanation for why one-time enforcer punishment equilibria 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

$$\Pr\left(h_j^{t+1}|h_i^t\right) = \mathbb{E}\left[\Pr\left(h_j^{t+1}|h_k^t\right)|h_i^t\right], \tag{2}$$

so the indirect effect outweighs the direct effect whenever $g'(y) \geq m = \frac{(k-1)n}{(kn-1)l}$ for all y. This explains why one-time enforcer punishment equilibria are optimal with perfect monitoring (Theorem 1) or with private monitoring with informed enforcers (Theorem 5). On the other hand, if monitoring is private and enforcers may not have finer information than producers, then (2) may fail, and the differing beliefs of enforcers and producers may be exploited to provide stronger on-path incentives than are possible in one-time enforcer punishment equilibria. We construct an example with these features in the next subsection.

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

5.2.3 A Counterexample: Graduated Punishments

In this section, we show by example that when (2) is not satisfied, it may be possible to support greater cooperation under private monitoring than public monitoring and thus improve on one-time enforcer punishment equilibria. Interestingly, such greater cooperation is supported by *graduated punishments*, 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.

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 match. This informational edge for the producers is for simplicity; in the Appendix, we sketch a more complicated example without this feature. To complete the description of the physical environment of the example, assume that $f(x) = 100\sqrt{x}$, $g(y) = y$, and $\delta = .1$. These parameters satisfy our condition for one-time enforcer punishment equilibria to be optimal under perfect monitoring ($g'(y) \geq m$ for all 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 construct an equilibrium that sustains a cooperation level of (exactly) 493.830. We call it the "three strikes and you're out" (3SYO) equilibrium.

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 you play $x' < x_2$ in the period after after playing $x < x_1$, *and* you are monitored by the same producer in both of these periods, *but* you are monitored by different enforcers: play x_3 for one period, then go back to x_1. Otherwise, go back to x_1 immediately. (We define x_1, x_2, x_3 below.)
- 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 belief difference between enforcer and producer at such a history to punish a deviator with both direct punishment and contagion. The reason why this "extra punishment" at an off-path history allow 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.

Formally, we establish the following result in the Appendix:

Proposition 2 *The 3SYO strategy profile is an equilibrium when* $x_1 = 493.830$*,* $x_2 = 494.102$*,* $x_3 = 502.058$*, and* $y = 493.828$*. Consequently, one-time enforcer punishment equilibria are not optimal.*

5.2.4 Stability

We conclude our treatment of the model with private monitoring by providing another reason why one-time enforcer punishment equilibria may be optimal under private monitoring, even when (2) is not satisfied. In particular, we show that one-time enforcer punishment equilibria are optimal among all equilibria satisfying a simple 1-period stability refinement. While the argument is simple, we believe that 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.[18]

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 $\left(x_i^\tau, y_{ji}^\tau\right)_{i \in C, j \in P}^{\tau \in \{t+1,\ldots\}}$ in period $t+1$.*

Note that if all players "trembled" 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 on the order of 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 6 *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.*

Note that, unlike Theorem 1, Theorem 6 does not require the assumption that $g'(y) \geq m$ for all y to establish that one-time enforcer punishment strategies are optimal. This is because community enforcement strategies always violate Stability, regardless of g.

6 Ostracism

An important way in which many groups sustain cooperation among their members in practice is ostracism, or the exclusion of deviators alone from the benefits of societal cooperation (Ostrom, 1990, Ellickson, 1991, Greif, 1993, 2006). The main model analyzed so far does not allow 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 providing public goods, such exclusion becomes a possibility. To study how ostracism interacts with specialized enforcement, we now consider a variant of our baseline model in which producers can exclude one another other from the benefits of cooperation. We consider both a version of the model where we retain the assumption of partial anonymity (so that producers can only be ostracized as a group) and a version where we relax it (which makes ostracism of individual producers possible).

The main message of this section is that our results generalize to an environment that incorporates ostracism. This is despite the fact that producers' ability to ostracize each other while

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

still cooperating with enforcers works against the tradeoff underlying our main model, as now a deviant producer can be directly punished by enforcers while simultaneously facing the withdrawal of cooperation from other producers. Furthermore, the structure of optimal equilibria in this environment is quite interesting: producers' ability to use ostracism makes the conditions for one-time enforcer punishment equilibria to be optimal more restrictive in an intuitive way, but when pure one-time enforcer punishments are not optimal, optimal equilibria simply combine one-time enforcer punishments with ostracism.

6.1 Group Ostracism

We first consider the perfect monitoring version of our model with the modification that players have the ability to *ostracize* each other: when producer i chooses her level of cooperation x_i^t, she can also choose to exclude one or more of the other producers and enforcers in M_i^t from the benefits of this cooperation. In this subsection, this choice must be made subject to our usual partial anonymity requirement. That is, a producer cannot discriminate which other producers and enforcers in her match she wishes to ostracize, as she does not observe their identities (for example, if she decides to ostracize one producer, then one of the other producers is chosen at random to be the recipient of this ostracism). If in period t producer i ostracizes some player j but not player k, then player j receives 0 benefit from player i's cooperation, while player k receives benefit

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

for some $\gamma \in [0,1]$. Thus, the parameter γ measures the efficiency cost of employing ostracism: if $\gamma = 1$ then, as in our main model, producer i cannot provide any benefit to player k without also benefiting player j, while if $\gamma = 0$ then player i can completely exclude player j from the benefits of her actions without reducing the benefit to player k. Note that the $\gamma = 1$ case corresponds to our main model with the possibility of money burning, as when $\gamma = 1$ "cooperating" while ostracizing anyone is equivalent to burning money.

We will show that in this version of our model, *one-time enforcer punishment plus group ostracism (EPGO) strategies* may be optimal. Formally, an EPGO strategy profile is characterized by a cooperation level x and punishment level y, and is represented by the following 4-state automaton, with normal, infected, other-ostracism and self-ostracism states. Play in these states is as follows:

Normal state: Producer i plays $x_i = x$ and does not ostracize anyone. If all producers $i \in M_j \cap C$ play $x_i = x$—or if multiple producers $i \in M_j \cap C$ play $x_i \neq x$—then enforcer j plays $y_{ji} = 0$ for all $i \in M_j \cap C$. If instead a unique producer $i \in M_j \cap C$ plays $x_i \neq x$, then enforcer j plays $y_{ji} = y$ and plays $y_{ji'} = 0$ for all $i' \in M_j \cap C \backslash i$.

Infected state: Players always take action 0.

Self-ostracism state: Producer i plays $x_i = x$ and does not ostracize anyone.

Other-ostracism state: Producer i plays $x_i = x$ and ostracizes all other producers in her match (so her cooperation only benefits the enforcers). Enforcers play exactly as in the normal state.

As usual, all players starts in the normal state. In the normal state or the self- or other-ostracism state, all players transition to the normal state unless a unique producer i plays $x_i \neq x$. If producer i plays $x_i \neq x$ and all enforcers $j \in M_i \cap P$ play $y_{ji} = y$, player i transitions to the self-ostracism state and all other players transition to the other-ostracism state. If producer i plays $x_i \neq x$ and some enforcer $j \in M_i \cap P$ plays $y_{ji} \neq y$, all players transition to the infected state. The infected state is absorbing.

Note that under EPGO strategies, enforcer behavior and state transitions are conditioned only on each producer's level of cooperation and not on her choice of whom to ostracize. This does not cause incentive problems for producers, since, given her level of cooperation, a producer is indifferent as to whether or not she ostracizes anyone. For example, when $\gamma = 1$ and producer i is in the self-ostracism state, producer i cooperates at level x while other producers burn x utils each.

Let $\check{x}$ and $\check{y}$ be the greatest solution to the following system of equations:

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

Intuitively, $\check{x}$ and $\check{y}$ are the greatest levels of cooperation and punishment that are sustainable in one-time enforcer punishment plus anonymous ostracism strategies. Specifically, a producer who deviates gains $\check{x}$ and loses $lg\left(\check{y}\right) + \delta\left(k-1\right) f\left(\check{x}\right)$ (one period of punishment plus one period of lost cooperation), while an enforcer who fails to punish a deviating producer gains $\check{y}$ and loses $\left(\frac{\delta}{1-\delta} - \frac{nk-1}{nk}\delta\gamma\right) kf\left(\check{x}\right)$ (all future benefits of cooperation, noting that benefits from all but one producer next period are only $\left(1-\gamma\right) f\left(\check{x}\right)$ instead of the usual $f\left(\check{x}\right)$).

We show the following result:

Theorem 7 *With anonymous perfect monitoring, when (group) ostracism is available,*

1. *If $g'\left(y\right) \geq \frac{m}{\gamma}$ for all $y \in \mathbb{R}_+$, then the one-time enforcer punishment strategy profile with cooperation level x^* and punishment level y^* is the most cooperative equilibrium.*

2. *If $g'\left(y\right) \leq \frac{m}{\gamma}$ for all $y \in \mathbb{R}_+$, then the EPGO strategy profile with cooperation level $\check{x}$ and punishment level $\check{y}$ is the most cooperative equilibrium.*

Theorem 7 makes two main points. First, the necessary condition for pure one-time enforcer punishment strategies to be optimal becomes more restrictive when ostracism is more effective: the lower bound of m from Theorem 1 must simply be scaled up by $\frac{1}{\gamma}$. Second, when the converse condition is satisfied, a version of one-time enforcer punishment equilibria remain optimal, but now these strategies must be combined with group ostracism.

In addition to showing how our results change with γ, Theorem 7 is of independent interest in demonstrating conditions under which strategies involving ostracism are optimal. While numerous empirical and case studies have argued for the practical importance of ostracism for sustaining cooperation (e.g., Ostrom, 1990), in most game-theoretic models of cooperation in groups ostracism strategies do no better than contagion strategies. Two notable exceptions are Lippert and Spagnolo (2011) and Ali and Miller (2014b), who study community enforcement games with private monitoring and show that ostracism can have advantages over contagion in terms of providing incentives for communication about others' past behavior.[19] In contrast, Theorem 7 concerns a setting with perfect information—where there is no need for communication—and shows that ostracism strategies can be essentially uniquely optimal.

Finally, it is also the case that EPGO strategies can still support cooperation level $\check{x}$ under private monitoring with observable last matches. Thus, with observable last matches, the comparison between one-time enforcer punishment strategies and EPGO strategies is exactly the same as under perfect monitoring.

6.2 Individual Ostracism with $n = 1$

As we have emphasized at several points throughout the paper, letting producers observe their partners' identities at the time they take actions could substantially complicate the analysis. When $n = 1$, however, this would be completely immaterial, as producers know their partners' identities anyway. (For example, in the $n = 1$ case of our baseline model, one-time enforcer punishment strategies remain optimal when $g'(y) \geq m$ for all y, even without "anonymity.") In this section, we show that when we introduce targeted ostracism into the model without anonymity in the $n = 1$ case, the structure of optimal equilibria is essentially identical to that under anonymity, except that now it is possible to avoid ostracizing innocent producers.

We now define a *one-time enforcer punishment plus individual ostracism (EPIO) strategy profile* as a profile characterized by a cooperation level x and punishment level y, and represented by the following $k + 2$-state automaton:

Normal state and infected state: Same as with EPGO strategies.

i-ostracism state (one for each $i \in C$)*:* Producer i plays $x_i = x$ and does not ostracize anyone. Producer $i' \neq i$ plays $x_{i'} = x$ and ostracizes player i only. Enforcers play exactly as in the normal state.

[19] Tirole (1996) shows how collective reputation or community responsibility systems, which also resemble group ostracism, might encourage cooperation in certain settings.

Play starts in the normal state. In the normal state or one of the ostracism states, all players transition to the normal state unless a unique producer i plays $x_i \neq x$. If producer i plays $x_i \neq x$ and all enforcers $j \in M_i$ play $y_{ji} = y$, play transitions to the i-ostracism state. If producer i plays $x_i \neq x$ and some enforcer $j \in M_i \cap P$ plays $y_{ji} \neq y$, play transitions to the infected state. The infected state is absorbing.

Theorem 8 *With non-anonymous perfect monitoring and $n = 1$, when ostracism is available,*

1. *If $g'(y) \geq \frac{m}{\gamma}$ for all $y \in \mathbb{R}_+$, then the one-time enforcer punishment strategy profile with cooperation level x^* and punishment level y^* is the most cooperative equilibrium.*

2. *If $g'(y) \leq \frac{m}{\gamma}$ for all $y \in \mathbb{R}_+$, then both the EPGO and EPIO strategy profiles with cooperation level $\check{x}$ and punishment level $\check{y}$ sustain the maximum level of cooperation.*

Thus, perhaps surprisingly, with non-anonymous perfect monitoring and $n = 1$, EPGO strategies (which ostracize innocent producers as well as the deviator) remain optimal, but EPIO strategies (which only ostracize the deviator) perform equally well. This is because whether innocent producers are ostracized or not is irrelevant for on-path incentives for cooperation: all that matters is whether the deviator herself is ostracized, and this is possible whether or not players are anonymous. Nonetheless, EPIO strategies seem to more closely resemble ostracism as it is used in practice, so it is interesting to note that these strategies can indeed be optimal. Most importantly, the same forces that make specialized enforcement optimal remain operational even in the presence of the ability to ostracize specific individuals.

7 Conclusion

This paper introduces 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 citizens' incentives to cooperate, but also a negative indirect effect coming through the erosion of enforcers' incentives to punish. Whenever the specialized enforcement technology is either uniformly effective or uniformly ineffective, this tradeoff is optimally resolved by going to either the extreme of pure community enforcement (where specialized enforcers are completely inactive) or pure enforcer punishments (where the future path of cooperation is completely unaffected by producers' transgressions). Yet, the threat of contagion does play a role even under pure enforcer punishments—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 societies with less effective production and enforcement technologies, while societies with more effective technologies should rely on specialized enforcement.

The framework introduced in this paper could be developed in several promising directions. First, our analysis takes the number of specialized enforcers in society as given (except briefly in Theorem 2). As we have characterized optimal equilibria for any number of producers and enforcers, endogenizing the number of enforcers in our model is conceptually straightforward. 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), and one could further extend the framework in that direction by allowing "guns" to be used for expropriating others as well as enforcing cooperation. Because of the pivotal role of specialized enforcers in sustaining cooperation, the insights of this analysis are likely to be different from those of existing models of anarchy.

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, including Tilly (1985), have argued that states evolve out of—or are, in fact, a form of—the 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 5.2.2 on optimal equilibria when all 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 5.2.4, 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 (or, more generally, laws) impact the incentives of citizens to cooperate not only with each other but also with state institutions themselves.

Appendix: Proofs and Additional Results

Proof of Theorem 1

Optimality of One-Time Enforcer Punishment Strategies when $g'(y) \geq m$ for all y

It is straightforward to see that the one-time enforcer punishment strategy profile with cooperation level $x \leq x^*$ and punishment level $y = \frac{\delta}{1-\delta} kf(x)$ is a SPE: In the normal state, a producer who deviates saves an effort cost of at most x and receives punishment $lg(y) = lg\left(\frac{\delta}{1-\delta} kf(x)\right)$, which equals x if $x = x^*$ and is weakly greater than x if $x < x^*$ and g is concave. Similarly, in the normal state an enforcer who deviates saves a cost of at most y and loses future benefits of cooperation worth $\frac{\delta}{1-\delta} kf(x) \geq y$. Finally, incentives in the infected state are trivial.

The main part of the theorem is thus showing that x^* is an upper bound on each producer's level of cooperation in any SPE when $g'(y) \geq m$ for all y. We break the proof into several steps.

Definitions and Preliminary Observations:

Fixing a SPE $\sigma = (\sigma_i)_{i \in N}$, 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.[20] 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 in particular $\bar{X} < \infty$.

Producer Incentive Compatibility:

A necessary condition for producer i not to deviate to playing $x_i = 0$ at history h^t is that, for

[20] By the same argument leading to (A2) below, one such upper bound is $\lim_{x \to \infty} \frac{\delta}{1-\delta} kf(x)$.

all $x_i^t \in \operatorname{supp} \sigma_i \left(h^t\right)$,

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

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).[21] Note that the distribution of x_j^t does not depend on x_i^t, so $\mathbb{E}\left[\sum_{j \in M_i^t \cap C \backslash i} f\left(x_j^t\right) | h^t, x_i^t\right]=\mathbb{E}\left[\sum_{j \in M_i^t \cap C \backslash i} f\left(x_j^t\right) | h^t\right]$, and we can rewrite this necessary condition as

$$(1-\delta) x_i^t+\delta \mathbb{E}\left[X_i^{t+1} | h^t, x_i^t\right] \leq \delta(1-\delta) \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[\sum_{j \in M_i^{t+1+\tau} \cap C \backslash i} f\left(x_j^{t+1+\tau}\right) | h^t, x_i^t\right]-\underline{u}. \tag{A1}$$

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

$$y_{ki}^t \leq \delta \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[\sum_{j \in M_k^{t+1+\tau} \cap C} f\left(x_j^{t+1+\tau}\right) | h^t\right]. \tag{A2}$$

This is a necessary condition because an enforcer's minmax payoff is 0, while her equilibrium continuation payoff is at most $(1-\delta)\left(\delta \sum_{\tau=0}^{\infty} \delta^\tau \mathbb{E}\left[\sum_{j \in M_k^{t+1+\tau} \cap C} f\left(x_j^{t+1+\tau}\right) | h^t\right]-y_{ki}^t\right)$, 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

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

as a producer always has the option of playing $x_i=0$ in period $t+1$. Therefore, there exist a

[21] 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.

producer i and a history h^t such that

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

or equivalently $\underline{u} \geq -\mathbb{E}\left[\sum_{k\in M_i^t\cap P} g\left(y_{ki}^t\right)|h^t\right] + \delta\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap C\setminus i} f\left(x_j^{t+1}\right)|h^t\right]$. In particular, by (A2) and the observation that the quantity $\mathbb{E}\left[\sum_{j\in M_k^{t+1+\tau}\cap C} f\left(x_j^{t+1+\tau}\right)|h^t\right]$ is the same for all $k \in P$, we have

$$\underline{u} \geq -lg\left(\delta\sum_{\tau=0}^{\infty}\delta^\tau\mathbb{E}\left[\sum_{j\in M_k^{t+1+\tau}\cap C} f\left(x_j^{t+1+\tau}\right)|h^t\right]\right) + \delta\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap C\setminus i} f\left(x_j^{t+1}\right)|h^t\right]. \tag{A3}$$

Bounding $\bar{X}$ when $g'(y) \geq m$:

Suppose that $g'(y) \geq m$ for all $y \in \mathbb{R}_+$. We argue that (A3), together with the definition of $\bar{X}$, implies that

$$\underline{u} \geq -lg\left(\frac{\delta}{1-\delta}kf\left(\bar{X}\right)\right) + \delta(k-1)f\left(\bar{X}\right). \tag{A4}$$

To see this, note that, 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

$$x_j^{t+1} \leq \frac{1}{1-\delta}\bar{X} - \delta\sum_{\tau=0}^{\infty}\delta^\tau\mathbb{E}\left[x_j^{t+2+\tau}|h^{t+1}, x_j^{t+1}\right]. \tag{A5}$$

Next, if in computing the right-hand side of (A3) we replace the random variable $\sigma_j\left(h^{t+1}\right)$ with $\sigma_j\left(h^{t+1}\right)+\varepsilon$ for some constant $\varepsilon > 0$ (i.e., shift $\sigma_j\left(h^{t+1}\right)$ up by ε realization-by-realization), the derivative of the resulting expression with respect to ε equals

$$-\delta\frac{l}{n}\mathbb{E}\left[f'\left(x_j^{t+1}\right)|h^t\right]g'\left(\delta\sum_{\tau=0}^{\infty}\delta^\tau\mathbb{E}\left[\sum_{j\in M_k^{t+1+\tau}\cap C} f\left(x_j^{t+1+\tau}\right)|h^t\right]\right) + \delta\frac{k-1}{kn-1}\mathbb{E}\left[f'\left(x_j^{t+1}\right)|h^t\right]. \tag{A6}$$

As $g'(y) \geq m$ for all y, this derivative is non-positive for all j and h^{t+1}. Hence, a lower bound on the right-hand side of (A3) is obtained by setting x_j^{t+1} equal to its upper bound in (A5) for all j. The resulting lower bound equals

$$\begin{aligned}&-lg\left(\begin{array}{c}\delta\mathbb{E}\left[\sum_{j\in M_k^{t+1}\cap C} f\left(\frac{1}{1-\delta}\bar{X} - \delta\sum_{\tau=0}^{\infty}\delta^\tau\mathbb{E}\left[x_j^{t+2+\tau}|h^{t+1}, x_j^{t+1}\right]\right)|h^t\right]\\ +\delta^2\sum_{\tau=0}^{\infty}\delta^\tau\mathbb{E}\left[\sum_{j\in M_k^{t+2+\tau}\cap C} f\left(x_j^{t+2+\tau}\right)|h^t\right]\end{array}\right)\\ &+\delta\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap C\setminus i} f\left(\frac{1}{1-\delta}\bar{X} - \delta\sum_{\tau=0}^{\infty}\delta^\tau\mathbb{E}\left[x_j^{t+2+\tau}|h^{t+1}, x_j^{t+1}\right]\right)|h^t\right].\end{aligned} \tag{A7}$$

We next derive an upper bound on the argument of g in (A7). Letting

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

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

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

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 (A7) of $\delta kf\left(\frac{1}{1-\delta}\left(\bar{X}-\delta\bar{X}\right)\right)+\frac{\delta^2}{1-\delta}kf\left(\bar{X}\right)=\frac{\delta}{1-\delta}kf\left(\bar{X}\right)$. On the other hand,

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

Combining these observations, we see that (A7) 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 (A4).

We can now complete the proof of the main part of the theorem. Combining (A1) and (A4), we have, for every player i, history h^t, and level of cooperation $x_i^t\in\operatorname{supp}\sigma_i\left(h^t\right)$,

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

As $\bar{X}=\sup_{i,h^t,x_i^t\in\operatorname{supp}\sigma_i(h^t)}\left(1-\delta\right)x_i^t+\delta\mathbb{E}\left[X_i^{t+1}|h^t,x_i^t\right]$, we have $\bar{X}\leq lg\left(\frac{\delta}{1-\delta}kf\left(\bar{X}\right)\right)$. By the definition of x^*, this implies that $\bar{X}\leq x^*$. Finally, as $\mathbb{E}\left[X_i^0|h^0\right]\leq\bar{X}$, we have $\mathbb{E}\left[X_i^0|h^0\right]\leq x^*$. Thus, x^* is an upper bound on each player's maximum equilibrium level of cooperation.

Near-Optimality of Contagion Strategies when $g'(y)$ is Small

It is straightforward to see that the contagion strategy profile with cooperation level $\hat{x}$ is a SPE. It remains to show that, for all $\varepsilon>0$, there exists $\eta>0$ such that $\hat{x}+\varepsilon$ is an upper bound on each player's maximum level of cooperation whenever $g'(y) < \eta$ for all y.

By (A3), the definition of $\bar{X}$, and Jensen's inequality, we have

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

Combining (A1) and (A8) yields that, for every player i, history h^t, and level of cooperation $x_i^t \in \operatorname{supp} \sigma_i\left(h^t\right)$,

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

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

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

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

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

For sufficiently small $\eta > 0$, this is less than $\hat{x} + \varepsilon$. Thus, $\left(\frac{\delta}{1-\delta} k l \eta + \delta (k-1)\right) f(x) - x$ 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 each player's maximum equilibrium level of cooperation.

Proof of Proposition 1

When money burning is available, the model is almost identical to the model of Section 6 in the $\gamma = 1$ case, and the conclusion of the proposition follows from the conclusion of Theorem 7 when $\gamma = 1$. The sole difference between the models is that is that the model with money burning lets a producer cooperate and burn money simultaneously. The proof of Theorem 7 is unaffected by this possibility, so Proposition 1 follows from the proof of Theorem 7.

Proof of Theorem 2

Per-match social welfare with k producers per match and on-path cooperation x is $k\left((s-1) f(x) - x\right)$. 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 $\hat{x}=\delta(s-1)f(\hat{x})$. On the other hand, if $k<s$ and $g'(y)\leq\frac{(k-1)n}{(kn-1)l}$ for all y (which is guaranteed by the assumptions that $g'(y)\leq\frac{(\bar{k}-1)n}{(\bar{k}n-1)l}$ for all y and $k\geq\bar{k}$), then by Theorem 7 the maximum level of cooperation (even if money burning is available) is given by the solution to

$$x=(s-k)\,g\left(\frac{\delta}{n}f(x)+\frac{\delta^2}{1-\delta}kf(x)\right)+\delta(k-1)f(x). \tag{A9}$$

Thus, the planner prefers that all agents become producers if (A9) is maximized at $k=s$. In turn, a sufficient condition for this is that the derivative of the right-hand side of (A9) with respect to k is non-negative for all x. This derivative equals

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

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

Theorem 1 when Enforcers Can Punish Each Other

Suppose that the enforcers can punish each other. Then we consider the following modified definition of one-time enforcer punishment strategies: 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 instead 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 one of his partners at level $\frac{\delta}{1-\delta}(kf(x^*)+(l-1)g(y^*))$ (or alternatively he burns this level of utility, if money burning is allowed). Finally, let (x^{**},y^{**}) be greatest solution of 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 (1′) *When enforcers can punish other enforcers, Theorem 1 applies verbatim with this modified definition of one-time enforcer punishment strategies.*

Proof (sketch). Start with the first part of the theorem: optimality of one-time enforcer punishment strategies when $g'(y) \geq m$ for all y. It is straightforward to check that the above strategy profile is a SPE, so it remains to show that x^{**} is an upper bound on each producer's level of cooperation in any SPE. The argument is similar to the proof of Theorem 1. In particular, fixing a SPE $\sigma = (\sigma_i)_{i \in N}$, 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 (A2) ("enforcer incentive compatibility") then becomes

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

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

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

By 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, as desired.

For the second part of the theorem (near-optimality of contagion strategies when $g'(y)$ is small), note that if $g'(y) < \eta$ for all y then

$$\bar{y} \leq \frac{\delta}{1-\delta}\left(kf\left(\bar{X}\right) + (l-1) g(\bar{y})\right) \leq \frac{\delta}{1-\delta}\left(kf\left(\bar{X}\right) + (l-1)\eta\bar{y}\right),$$

or equivalently $\bar{y} \leq \frac{\delta}{1-\delta-\delta(l-1)\eta} kf\left(\bar{X}\right)$. We then obtain the bound

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

As $\lim_{\eta \to 0} \frac{\delta k l \eta}{1-\delta-\delta(l-1)\eta} = 0$ for all $\delta < 1$, this bound suffices for the result, by the same argument as in the proof of Theorem 1. ■

Proof of Theorem 3

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 1). 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

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

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 $i \in C$ and $\tau \geq t$. Let $d_t \equiv \mathbb{E}[|D(t)|]$.

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

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

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 $q_t \equiv \mathbb{E}[|Q(t) \cap C|]$.

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

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

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 $z_t \equiv \mathbb{E}[Z(t) \cap C]$.

Note that the distribution of $|D(t)|$ first-order stochastically dominates the distribution of $|Q(t) \cap C|$, 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 $|Z(t) \cap C|$ 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

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

Thus, if $l = 1$, then $\tilde{x}$ is the greatest level of cooperation that could be sustained with contagion strategies if contagion spread through the process $Z\left(t\right)$ rather than $D\left(t\right)$. Otherwise, $\tilde{x}$ 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 1 *If $x^* \geq \tilde{x}$, 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 $lg\left(y^*\right) = x^*$. 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 $\sum_{t=0}^{\infty}\delta^t \frac{1}{n} q_t f\left(x^*\right) \geq y^*$.

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 $\sum_{t=0}^{\infty}\delta^t \frac{1}{n} q_t f\left(x^*\right) < x^*$. 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 $lg\left(y^*\right) = x^*$) 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 $\left(l-1\right) g\left(y^*\right)$). 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 $\tau \in \{t, \ldots, \tau\}$, 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 $\tau \in \{t+1, \ldots, \tau\}$. 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

$$x^* \geq (l-1) g(y^*) + (1-\delta) \sum_{\tau=0}^{\infty} \delta^\tau \frac{k-1}{kn-1} \mathbb{E}[r_\tau] f(x^*). \tag{A10}$$

We will show below that $\mathbb{E}[r_\tau] \leq z_\tau$ for all τ. This will imply (A10) because, recalling that $\tilde{x} = l(1-\delta) \sum_{\tau=0}^{\infty} \delta^\tau \frac{k-1}{kn-1} z_\tau f(\tilde{x})$ by definition, the fact that $x^* \geq \tilde{x}$ 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 $l(1-\delta) \sum_{\tau=0}^{\infty} \delta^\tau \frac{k-1}{kn-1} \mathbb{E}[r_\tau] f(x^*)$. Finally, the fact that $x^* = lg(y^*) \geq \tilde{x}$ also implies that $g(y^*) \geq (1-\delta) \sum_{\tau=0}^{\infty} \delta^\tau \frac{k-1}{kn-1} z_\tau f(x^*)$, so we have (A10).

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

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

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 $t' \in \{t, \ldots, \tau\}$ 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 $t' \in \{t+1, \ldots, \tau\}$ 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,

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

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

$$Q(\tau, t, S) \backslash Q(\tau, t+1, Q(t+1, t, S \backslash \{i\}) \cup \{i\}) \subseteq Z(\tau, t, \{i\}). \tag{A11}$$

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

The proof of (A11) 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 (A11) 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, (A11) holds for $\tau+1$, so by induction it holds for all $\tau > t$. ■

We now complete the proof of Theorem 3. Recall that $\tilde{x} \geq \hat{x}$. Thus, by Lemma 1, 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 $x^* \geq \tilde{x}$.

For the first part of the theorem, note that $lg\left(\sum_{t=0}^{\infty}\delta^t\frac{1}{n}q_t f(x)\right) - x$ crosses 0 from above at $x = x^*$. Therefore,

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

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, $\tilde{x}$ is finite and independent of $\bar{g}$, so if $\bar{g}$ is sufficiently high then $x^* > \tilde{x}$.

For the second part, as z_t can never exceed kn, $\tilde{x}$ 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 $\lim_{\delta\to 1} lg\left(\sum_{t=0}^{\infty}\delta^t\frac{1}{n}q_t f(x)\right) - x = \infty$ for all $x > 0$. Hence, x^*, the greatest root of $lg\left(\sum_{t=0}^{\infty}\delta^t \frac{1}{n} q_t f(x)\right) - x$, goes to infinity as $\delta \to 1$.

Proof of Theorem 4

For one-time enforcer punishment strategies, the proof is similar to the proof of Lemma 1, with the following differences:

First, the reason why punishments of up to $\frac{\delta}{1-\delta} k f(x^*) = y^*$ 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 $x^* \geq \tilde{x}$ 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 5

For existence, the only difference from the proof of Theorem 4 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 $x^* \geq \tilde{x}$ 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 1, with one key additional step. In particular, if we follow the proof of Theorem 1 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 (A3):

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

(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

$$\begin{aligned}\underline{u} \quad \geq \quad & -l\mathbb{E}\left[g\left(\delta\sum_{\tau=0}^{\infty}\delta^{\tau}\sum_{j\in M_k^{t+1+\tau}\cap C}\mathbb{E}\left[f\left(x_j^{t+1+\tau}\right)|h^t, j\in M_k^{t+1+\tau}\right]\right)|h_i^t\right]\\ & +\delta\mathbb{E}\left[\mathbb{E}\left[\sum_{j\in M_k^{t+1}\cap C\backslash i}f\left(x_j^{t+1}\right)|h^t, j\in M_k^{t+1}\right]|h_i^t\right]. \qquad \text{(A12)}\end{aligned}$$

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 (A12) and differentiating with respect to ε yields

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

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 (A12) is obtained by setting x_j^{t+1} equal to its upper bound in (A5) for all j, as in the proof of Theorem 1. The remainder of the argument is identical to the proof of Theorem 1.

Proof of Proposition 2

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

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

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 ($x = 0$) 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 6

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 $\left(x_i^\tau, y_{ji}^\tau\right)_{i\in C, j\in P}^{\tau\in\{0,1,...\}}$ is

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

or equivalently

$$x_i^t \leq \mathbb{E}\left[\sum_{k\in M_i^t\cap P} g\left(y_{ki}^t\right)|x_i^t=0\right]. \tag{A13}$$

Next, under Stability, (A2) becomes

$$y_{ki}^t \leq \delta\sum_{\tau=0}^{\infty}\delta^\tau \mathbb{E}\left[\sum_{j\in M_k^{t+1+\tau}\cap C} f\left(x_j^{t+1+\tau}\right)\right]. \tag{A14}$$

Let $\bar{x} = \sup_{i,t} x_i^t$, and note that $\bar{x} < \infty$ as in the proof of Theorem 1. Combining (A13) and (A14) 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.

Proof of Theorem 7

It is straightforward to see that the EPGO strategy profile with cooperation level $\check{x}$ and punishment level $\check{y}$ is a SPE: a producer who deviates gains at most $\check{x}$ and loses $g\left(\check{y}\right) + \delta\left(k-1\right)f\left(\check{x}\right)$, while an enforcer who deviates gains at most $\check{y}$ and loses $\left(\frac{\delta}{1-\delta} - \frac{nk-1}{nk}\delta\gamma\right)kf\left(\check{x}\right)$. (More precisely, a producer's incentive compatibility constraint in the normal or ostracism phase is $\left(k-1\right)f\left(\check{x}\right) - \check{x} \geq \left(1-\delta\right)\left(\left(k-1\right)f\left(\check{x}\right) - lg\left(\check{y}\right)\right) - \delta\check{x} + \delta^2\left(k-1\right)f\left(\check{x}\right)$, or $\check{x} \leq lg\left(\check{y}\right) + \delta\left(k-1\right)f\left(\check{x}\right)$, which holds by definition of $\check{x}$.) Similarly, the one-time enforcer punishment strategy profile with cooperation level x^* and punishment level y^* remains a SPE when ostracism is possible. The main part of the proof is showing that x^* (resp., $\check{x}$) is an upper bound on each producer's level of cooperation in any SPE when $g'\left(y\right) \geq \frac{m}{\gamma}$ (resp., $\leq \frac{m}{\gamma}$) for all y. The argument is parallel to the proof of Theorem 1.

Fix a SPE $\sigma = \left(\sigma_i\right)_{i\in N}$, and let $\underline{u}$ and $\bar{X}$ be defined as in the proof of Theorem 1. As in the proof of Theorem 1, a necessary condition for producer i not to deviate to playing $x_i = 0$ at history h^t is that, for all $x_i^t \in \operatorname{supp}\sigma_i\left(h^t\right)$,

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

For each producer i and player j, let

$$\tilde{x}_{ij}^{t}=\left\{\begin{array}{c}x_i^t \text{ if } i \text{ does not ostracize anyone in period } t\\ (1-\gamma)\,x_i^t \text{ if } i \text{ ostracizes someone other than } j \text{ in period } t\\ 0 \text{ if } i \text{ ostracizes } j \text{ in period } t\end{array}\right\}.$$

Then, arguing as in the proof of Theorem 1, a necessary condition for enforcer i not to deviate to playing $y_{ik}=0$ in period t is

$$y_{ik}^{t}\leq\delta\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[\sum_{j\in M_i^{t+1+\tau}\cap C}f\left(\tilde{x}_{ji}^{t+1+\tau}\right)|h^t\right].$$

Furthermore, again as in the proof of Theorem 1,

$$\underline{u}\geq -lg\left(\delta\sum_{\tau=0}^{\infty}\delta^{\tau}\mathbb{E}\left[\sum_{j\in M_k^{t+1+\tau}\cap C}f\left(\tilde{x}_{jk}^{t+1+\tau}\right)|h^t\right]\right)+\delta\mathbb{E}\left[\sum_{j\in M_i^{t+1}\cap C\setminus i}f\left(\tilde{x}_{ji}^{t+1}\right)|h^t\right]. \qquad \text{(A15)}$$

Bounding $\bar{X}$ when $g'(y)\geq\frac{m}{\gamma}$:

Suppose that $g'(y)\geq\frac{m}{\gamma}$ for all y. We claim that, for any distribution over x_j^{τ}, the right-hand side of (A15) is minimized when producer j never ostracizes anyone. If $\tau>t+1$, then this is immediate, as if producer j ostracizes anyone the only effect of this on the right-hand side of (A15) is to decrease $\tilde{x}_{jk}^{\tau}$. If instead $\tau=t+1$, then ostracizing player i increases the right-hand side of (A15) by

$$l\left(g\left(a\right)-g\left(b\right)\right)-\delta\frac{k-1}{kn-1}f\left(x_j^{t+1}\right), \qquad \text{(A16)}$$

for some numbers a and b that differ by $\delta\frac{1}{n}\gamma f\left(x_j^{t+1}\right)$. In particular, (A16) is no less than

$$\inf_{y\in\mathbb{R}_+}\frac{l}{n}\delta\gamma f\left(x_j^{t+1}\right)g'\left(y\right)-\delta\frac{k-1}{kn-1}f\left(x_j^{t+1}\right).$$

Hence, if $g'(y)\geq\frac{m}{\gamma}$ for all y, then ostracizing player i increases the right-hand side of (A15) by a positive number, so the right-hand side of (A15) is minimized when producer j never ostracizes another producer. A similar argument shows that the right-hand side of (A15) is minimized when producer j never ostracizes an enforcer, so the right-hand side of (A15) is minimized when producer j never ostracizes anyone. Given this observation, the fact that $g'(y)\geq\frac{m}{\gamma}\geq m$ for all y implies that a lower bound on the right-hand side of (A15) is obtained by setting x_j^{t+1} equal to its upper bound in (A5) for all j, as in the proof of Theorem 1. As in the proof of Theorem 1, this implies (A4), and combining (A1) and (A4) then implies that x^* is an upper bound on each player's maximum equilibrium level of cooperation.

Bounding $\bar{X}$ when $g'(y)\leq\frac{m}{\gamma}$:

Suppose now that $g'(y) \leq \frac{m}{\gamma}$ for all y. The converse of the argument for the $g'(y) \geq \frac{m}{\gamma}$ case then implies that the right-hand side of (A15) is minimized when each producer $j \neq i$ ostracizes all of the producers (but none of the enforcers) in period $t+1$ and never ostracizes anyone in periods $\tau > t+1$, while producer i herself never ostracizes anyone. This gives a lower bound on the right-hand side of (A15) of

$$\underline{u} \geq -lg \left(\begin{array}{c} \delta \mathbb{E}\left[\sum_{j \in M_k^{t+1} \cap C} \left(1 - \gamma 1_{\{j \neq i\}}\right) f\left(x_j^{t+1}\right) | h^t\right] \\ + \delta \sum_{\tau=1}^{\infty} \delta^{\tau} \mathbb{E}\left[\sum_{j \in M_k^{t+1+\tau} \cap C} f\left(x_j^{t+1+\tau}\right) | h^t\right] \end{array} \right), \tag{A17}$$

where $1_{\{\cdot\}}$ denotes the indicator function. By definition of $\bar{X}$, concavity of f, and Jensen's inequality, this implies the following:[22]

$$\underline{u} \geq -lg\left(\left(\frac{\delta}{1-\delta} - \frac{nk-1}{nk}\delta\gamma\right) kf\left(\bar{X}\right)\right). \tag{A18}$$

Combining (A1) and (A18), we have, for every player i, history h^t and level of cooperation $x_i^t \in \operatorname{supp} \sigma_i\left(h^t\right)$,

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

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

$$\bar{X} \leq lg\left(\left(\frac{\delta}{1-\delta} - \frac{nk-1}{nk}\delta\gamma\right) kf\left(\bar{X}\right)\right) + \delta(k-1) f\left(\bar{X}\right).$$

By definition of $\check{x}$, this implies that $\bar{X} \leq \check{x}$. Finally, as $\mathbb{E}\left[X_i^0 | h^0\right] \leq \bar{X}$, we have $\mathbb{E}\left[X_i^0 | h^0\right] \leq \check{x}$. Thus, $\check{x}$ is an upper bound on each player's maximum equilibrium level of cooperation.

Proof of Theorem 8

It is straightforward to see that the EPGO strategy profile with cooperation level $\check{x}$ and punishment level $\check{y}$ remains a SPE with non-anonymous perfect monitoring. In addition, the argument that EPIO strategy profile with cooperation level $\check{x}$ and punishment level $\check{y}$ is also a SPE is identical, because each player faces the same incentives at every history (whether or not an innocent producer

[22] More specifically, by the concavity of f and Jensen's inequality, if the right-hand side of (A17) is less than $-lg\left(\left(\frac{\delta}{1-\delta} - \frac{nk-1}{nk}\delta\gamma\right) kf\left(\bar{X}\right)\right)$ and $X_j^{t+1} \leq \bar{X}$, then it must be that $X_j^{t+2} > \bar{X}$. As $X_j^{\tau} \leq \bar{X}$ for all τ by the definition of $\bar{X}$, it follows that the right-hand side of (A17) is not less than $-lg\left(\left(\frac{\delta}{1-\delta} - \frac{nk-1}{nk}\delta\gamma\right) kf\left(\bar{X}\right)\right)$.

i' is ostracized in the i-ostracism state affects her payoffs in that state, but not her incentives). Finally, the argument that x^* (resp., $\check{x}$) is an upper bound on each producer's level of cooperation in any SPE when $g'(y) \geq \frac{m}{\gamma}$ (resp., $\leq \frac{m}{\gamma}$) for all y is essentially the same as in the anonymous case. In particular, the key way in which anonymity is used in the proof of Theorem 7 (or Theorem 1) is that expectations of producers' levels of cooperation need not be conditioned on their realized matches. When there is only one match, this is always the case.

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." mimeo, 2014.

[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] Aldashev, Gani, and Giorgio Zanarone. "Endogenous enforcement institutions." mimeo, 2015.

[6] Ali, S. Nageeb, and David A. Miller. "Enforcing cooperation in networked societies." mimeo, 2014a.

[7] Ali, S. Nageeb, and David A. Miller. "Ostracism." mimeo, 2014b.

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

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

[10] Clay, Karen. "Trade without law: private-order institutions in Mexican California." *Journal of Law, Economics, and Organization* 13.1 (1997): 202-231.

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

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

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

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

[15] Ellickson, Robert C. *Order without law: How neighbors settle disputes.* Harvard University Press, 2009.

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

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

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

[19] 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.

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

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

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

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

[24] 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.

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

[26] 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.

[27] 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.

[28] Hirshleifer, David, and Eric Rasmusen. "Cooperation in a repeated prisoners' dilemma with ostracism." *Journal of Economic Behavior & Organization* 12.1 (1989): 87-106.

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

[30] 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.

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

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

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

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

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

[36] Levine, David K. and Salvatoe Modica, "Peer discipline and incentives within groups," mimeo.

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

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

[39] Mailath, George J., Stephen Morris, and Andrew Postlewaite. "Maintaing authority," mimeo.

[40] Mailath, George J., and Larry Samuelson. *Repeated games and reputations.* Oxford. Oxford University Press 2006.

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

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

[43] 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.

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

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

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

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

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

[49] 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.

[50] 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.

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

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