Euro. Jnl of Applied Mathematics (2016), *vol.* 27, *pp.* 317–337. © Cambridge University Press 2015
doi:10.1017/S0956792515000649

Crime, punishment, and evolution in an adversarial game†

MICHAEL MCBRIDE[1], RYAN KENDALL[2], MARIA R. D'ORSOGNA[3] and MARTIN B. SHORT[4]‡

[1]*Department of Economics, University of California, Irvine, CA, USA*
email: *mcbride@uci.edu*
[2]*Department of Economics, University of Southern California, Los Angeles, CA, USA*
email: *rakendal@usc.edu*
[3]*Department of Mathematics, California State University, Northridge, CA, USA*
email: *dorsogna@csun.edu*
[4]*School of Mathematics, Georgia Institute of Technology, Atlanta, GA, USA*
email: *mbshort@math.gatech.edu*

(*Received 10 December 2014, revised 28 October 2015; accepted 4 November 2015; first published online 21 December 2015*)

We examine the game theoretic properties of a model of crime first introduced by Short *et al.* (2010 *Phys. Rev. E* **82**, 066114) as the SBD Adversarial Game. We identify the rationalizable strategies and one-shot equilibria under multiple equilibrium refinements. We further show that SBD's main result about the effectiveness of defecting-punishers ("Informants") in driving the system to evolve to the cooperative equilibrium under an imitation dynamic generalizes to a best response dynamic, though only under certain parameter regimes. The nature of this strategy's role, however, differs significantly between the two dynamics: in the SBD imitation dynamic, Informants are sufficient but not necessary to achieve the cooperative equilibrium, while under the best response dynamic, Informants are necessary but not sufficient for convergence to cooperation. Since a policy of simply converting citizens to Informants will not guarantee success under best response dynamics, we identify alternative strategies that may help the system reach cooperation in this case, e.g., the use of moderate but not too severe punishments on criminals.

Key words: adversarial game theory, mathematical criminology, social modeling, best response dynamics, cooperative behavior

1 Introduction

While various mechanisms can sustain cooperative behaviour in social dilemmas, numerous theoretical models, case studies, and experimental data have shown the punishment of defectors to be particularly effective (e.g., Ostrom 1990; Ledyard 1995; Fehr and Gächter 2000; Henrich and Boyd 2001; Cinyabuguma *et al.* 2005; Boyd *et al.* 2010; Fudenberg

† The material herein is based upon work supported by the Army Research Office Award No. #W911NF-11-1-0332. McBride and Short were also supported by the Air Force Office of Scientific Research Award No. FA9550-10-1-0569.

‡ Corresponding author. 686 Cherry Street, Atlanta, GA, 30332-0160, USA, mbshort@math.gatech.edu.

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

and Pathak 2010; Takahashi 2010; Chaudhuri 2011; Xiao and Hauser 2011). Given that the punishment of defectors is itself a second-order social dilemma, recent attention has been given to how cooperative and punishing behaviour coevolve in evolutionary settings (e.g., Fehr and Fischbacher 2004; Carpenter 2007; Hauert *et al.* 2007; Levine and Pesendorfer 2007; Bowles and Gintis 2011). Most of these works examine the emergence of cooperative and punishing behaviours in the standard social dilemma games, such as the Prisoner's Dilemma, Public Good Games, and Common Property Resource Games. The lack of cooperation in these settings is driven by the tension between personal and collective interests: players maximize their personal interest which, in turn, has an indirect and negative effect on the collective interest.

Standard social dilemmas pervade much of social and political life, but individuals in many societies—especially those with weak governments or state capacity—face dilemmas that are more adversarial in nature than standard social dilemmas. Such settings have been largely ignored in the literature, yet an important exception is Short *et al.* (2010), hereafter SBD. In their work, the authors introduce an evolutionary game in which two actors are selected at random, one placed in a potential criminal role, the other placed in a potential victim role. The former decides whether or not to steal from the latter, and the latter decides whether or not to report this theft to authorities, if it takes place. Whether a report leads to conviction, which directly affects the expected net benefit of reporting a crime, is probabilistically increasing in the population's overall proclivity to cooperate with authorities (i.e., proclivity to punish criminals). SBD modify the traditional cooperate and defect strategies to account for punishing behaviour, and then show that under a specific imitation dynamic, the presence of a novel strategy type—dubbed the Informant—that both commits crimes and cooperates with authorities, is highly influential in driving the system away from the "Dystopian"state of high crime and toward the efficient, no-crime, "Utopian" steady state. Specifically, the presence of Informants is a sufficient, but not necessary, condition for achieving the Utopian state.

Further experimental studies of this model have been performed (D'Orsogna *et al.* 2013), and the results of these experiments do indicate that the Informant strategy is in fact critical in achieving a low-crime steady state. However, differing from the SBD results, the experiments seem to show that Informants are *necessary* (and possibly *sufficient*) for the low-crime state to flourish; those experiments with Informant as an available strategy always tend toward a Utopian state, while those without Informants available always tend toward a Dystopian state. Though these experiments structure the game in essentially the same way as in the SBD model, subjects are free to choose strategies at will from amongst those that are available. Hence, the update rules used in SBD may not accurately capture all of the aspects of the decision making process that subjects use during the experiments.

To perhaps better understand the experimental results referenced above, this paper more closely examines the game theoretic properties of the SBD Adversarial Game. We consider the Adversarial Game worthy of study for multiple reasons. It provides an original depiction of an adversarial setting but also presents a simple formalization of the punishment of defectors (criminals) that explicitly captures the positive externality in punishment present in many societies: civilians play a crucial role in the self-regulation of pro-social norms (e.g., Sampson and Groves 1989; Skogan 1990; Bursik and Grasmick 1993), but fear of retaliation may lead to disengagement from law enforcement and

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

the proliferation of criminal behaviour (e.g., Gambetta 1988; Beittel 2009). Furthermore, the SBD model uses a simple, stylized schema of strategy types, including the novel Informant type, that is of independent interest. Finally, though the game's design is directly influenced by features of crime and punishment in disorganized societies, the model's simple structure has the potential to illuminate our understanding of the emergence of cooperative behaviour in a more general way.

In this paper, we recast the SBD Adversarial Game in a classic game theoretic setting. A specific goal is to assess the robustness of SBD's main finding regarding the influential role of Informants in the evolution of cooperation. From a broader perspective, we would like to identify what the Adversarial Game can teach us about the evolution of cooperation. As part of that goal, we seek to identify exactly how the SBD game compares with other, more commonly studied social dilemma games. We thus present both static and evolutionary analyses. We first examine the one-shot Adversarial Game and fully characterize the set of rationalizable strategies and (Bayesian) Nash Equilibria. As a bridge between the static and evolutionary analyses, we then identify which equilibria survive two equilibrium refinements: Evolutionary Stable Strategy and Trembling Hand Perfect Equilibrium (THPE). We finally turn to evolutionary analysis. Whereas SBD assume a particular imitation dynamic, we examine the evolutionary path of cooperative behaviour under a simple best response dynamic.

Two main findings and two main lessons emerge from the analysis. Unlike other one-shot social dilemma games in which the unique equilibrium is inefficient, the one-shot Adversarial Game has multiple equilibria, one of which is efficient—Utopia. In effect, the specific form of the positive externality in reporting (i.e., punishing) converts the second-order public goods problem into a second-order coordination game, which in turn converts the overall game into a non-standard coordination game. A key lesson is that an institution that converts the second-order public goods problem from a social dilemma into a coordination game can foster crime-deterring punishment. We also show that SBD's main result regarding the power of Informants to foster a low-crime society generalizes to the case of best response dynamics, though in a much different way than in the imitation dynamics of SBD. That is, we find the availability of the Informant strategy to be a necessary, but insufficient, condition to maintain a low-crime state, which will only occur given appropriate parameters and initial conditions. This finding suggests a modification to SBD's policy recommendation of converting defectors into Informants to improve overall cooperation, under best response dynamics. Our analysis identifies settings under which this policy is more likely or less likely to work. This finding is also in better agreement with experimental results than the original SBD model.

We emphasize that this Adversarial Game does not manifest the second-order free-rider problem in which the collective level of punishment on defectors is itself a type of public good problem, as in Fehr and Gächter (2000). Rather, the second-order punishment game is a coordination game in which the expected net benefits of punishing are increasing in the overall proportion of others that also punish. Thus, this Adversarial Game is closer in spirit to social dilemma models, such as Ostrom *et al.* (1992) and Boyd *et al.* (2010), in which punishers must solve a coordination problem. Our findings thus complement prior theoretical work by showing how coordinated punishments can enforce cooperative behaviour in a directly adversarial setting.

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

Our work also complements prior work in the economics literature on crime. In his review of literature, Ehrlich (1996) distinguishes two main strands of recent work: market models of crime and optimal crime control policies. Our paper resides in the former, though we do offer suggestions regarding the latter. Prior models have considered the equilibrium patterns of crime in static settings (e.g., Fender 1999), while others have considered the instability of equilibria (e.g., Eaton and Wen 2008). Our paper differs by considering an stylized, evolutionary setting, and looking at the evolution and steady states of crime and reporting. In this way, we draw more direct comparisons to the evolution of cooperation literature.

2 The Adversarial Game

Consider a population of expected-payoff maximizing actors, each labelled by an index $i = 1, ..., N$, and N, the total population size, is large and even. Each actor has an endowment of 1 and must next choose one of the four following strategies:

$$\begin{aligned} P &= \text{“Paladin”} &&= \{\text{not steal, report}\}, \\ A &= \text{“Apathetic”} &&= \{\text{not steal, not report}\}, \\ I &= \text{“Informant”} &&= \{\text{steal, report}\}, \\ V &= \text{“Villain”} &&= \{\text{steal, not report}\}. \end{aligned}$$

Let the number of players in total who choose a particular strategy $\ell \in \{P, A, I, V\}$ be denoted as N_ℓ. After choosing their strategies, two of the actors, say i and j, are chosen at random (uniformly) and paired for an interaction. All other actors are *bystanders*. Each bystander receives payoff 1 no matter what others do, yet as bystanders their strategies may affect the payoffs for i and j.

The two selected actors, i and j, are randomly assigned (uniformly) into different roles: one is assigned the first-mover role of “potential criminal,”and the other is assigned the second-mover role of “potential victim.” Here, the terms “first-mover” and “second-mover” differ from the sense that they are often used in, for example, Stackelberg games. In Stackelberg games, the “first-mover” must commit to her strategy before the “second-mover” must commit to his; this allows the second-mover to optimize against a known (but probably mixed) strategy. In our case, the first- and second-movers must both choose their strategies simultaneously, before any interaction occurs. We use the term “first-mover” merely to signify that this player’s strategy choice determines what happens in the first phase of the game, while the term “second-mover” signifies that this player’s strategy choice determines what happens in the second phase of the game; see below for details.

After this pairing, the selected players’ strategies are carried out in the following way. If the first mover does not steal, then each keeps his or her initial endowment of 1 regardless of the second mover’s report strategy. If the first mover steals, then δ is taken from the second mover, and $\alpha\delta$ (with $0 < \alpha < 1$) is given to the first mover. Hence, an amount $(1 - \alpha)\delta$ is destroyed (forgone productive activity by the first mover, etc.) in this process, making theft socially inefficient. When the first mover steals, then the second mover’s report-decision plays a role in determining the outcome. If no reporting occurs, final payoffs are $(1 + \alpha\delta,\ 1 - \delta)$. If reporting occurs, the probability of a conviction is

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

given by $r \equiv (N_P + N_I)/N$. Notice that r, which we shall refer to as the "report rate", is the proportion of actors that chose report as part of their strategy. If conviction occurs, then the first mover returns δ to the victim and pays an additional punishment cost θ, ending with a payoff of $1 + \alpha\delta - \delta - \theta$. In this case, the victim is fully reimbursed to receive payoff 1. If no conviction occurs, the criminal keeps $1 + \alpha\delta$, and the victim pays an additional cost ε (due to a loss in reputation or to retaliation) for a final payoff of $1 - \delta - \varepsilon$. Finally, we define $s \equiv (N_I + N_V)/N$, the proportion of actors that chose steal as part of their strategy, as the "steal rate".

The Paladin, Apathetic, Informant, and Villain labels, which are taken directly from SBD, are meant to convey the essence of the strategies. Paladins act as pure cooperators by not stealing and reporting criminals; Apathetics disengage from society by neither stealing nor reporting (first-order cooperator, second-order defector); Informants commit crimes but also punish criminals (first-order defector, second-order cooperator), a strategy that plays a strong role in SBD's original analysis; and Villains act as pure defectors by both committing and not reporting crimes. We emphasize that these labels do not necessarily match how those terms are used in everyday discourse (if indeed they are used every day). Rather, they are meant to convey the essence of the four possible strategies in this Adversarial Game. We also note that strategies are chosen at the start of a period and not at information sets within the period. Given the sequential nature of the game, it is natural to consider choices by information set, yet the single choice at the start of the round allows us to consider off-path choices and matches SBD's original analysis.

SBD's original Adversarial Game differs from our version in a few ways, two of which are worthy of note and relate to SBD's primary motivation to examine evolutionary dynamics of cooperative behaviour in adversarial settings. The first difference is that our formulation above is of a one-shot game, whereas their set-up is an evolutionary model with repeated periods. Our one-shot game mimics their stage game in that in each period two actors are selected at random and matched, and all other actors are bystanders for that period. The one-shot game is of sufficient interest and, as we shall see, serves as a useful benchmark for our later evolutionary analysis. The second difference is that strategies in SBD's version are either inherited or adopted via a pseudo-imitation dynamic (described in detail in Section 4 of this paper) rather than being selected by payoff maximization (best responding). Strategy revision via imitation is commonly assumed in evolutionary settings where the primary interest is the evolution of behaviour over time (see Sandholm 2010). We consider a best response dynamic instead of an imitation dynamic because it is a dynamic more closely tied to standard (non-evolutionary) game theoretic analysis and because, as will be shown, it will generate different evolutionary paths of behaviour.

The Adversarial Game defined above is a simultaneous game of imperfect information, but it could alternatively be defined as a sequential move game in which the potential criminal makes the steal decision knowing his or her role, and the victim and bystanders make their report decisions after observing a crime. The potential criminal would choose from the strategy set {steal, not steal}, and the others would choose from the set {report, not report}, thus decoupling the steal and report decisions from the overall strategy plan as defined with the simultaneous structure. This sequential structure may more accurately reflect the timing of real-life steal and report decisions; however, we use the simultaneous structure for three reasons. First, having all actors make strategy decisions before knowing

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

their roles makes decisions salient for all actors rather than just those in the matched pair. If bystanders make their report decisions knowing that they are bystanders, then their actions have no relevance to their payoffs. Second, the simultaneous structure better reflects the notion of societal conditions that partly motivates the model. Actors live in a societal setting that has certain properties that pre-date and influence the effectiveness of criminal acts. The committed behaviour of all actors, including bystanders, reflects their latent support, or lack thereof, for punishing criminals that is realized upon the commission of a crime. Third, SBD assume strategy-types are determined at the end of the prior period rather than sequentially within each period. In choosing two-dimensional strategies at the start of the period, we can, when moving to the dynamics, track the full evolution of strategy types, which is a fundamental feature of SBD's original analysis.

Unlike other social dilemma games, the one-shot Adversarial Game has a single potential deviant. This feature better reflects the inherent asymmetry of criminal behaviour, where at a given point in time only one actor may be in a position to take advantage of others or be victimized. The Adversarial Game also formalizes a stylized form of punishment for deviators directly into the basic game by allowing the victim to immediately challenge deviant behaviour. Moreover, as with some types of actual criminal behaviour, the victim has more to gain than the bystanders when the criminal is punished because he or she is reimbursed. The victim also has even more to lose because an unsuccessful challenge results in an additional private cost. However, the expected result of attempting to punish a deviant depends on the larger societal characteristics, i.e., the behaviour of the bystanders. Bystanders, though not directly affected by the realized crime, do indirectly foster or inhibit deviant behaviour by influencing the likelihood of successful punishment. Thus, in a very simple way, the Adversary Game captures the positive externalities inherent in the punishment of deviants. The more supportive the environment, the larger the expected benefit of attempting to punish, and the smaller the expected cost. We discuss the importance of this feature of the model in more detail after our analysis of the one-shot game.

3 Static analysis

3.1 Best-response functions

An actor's report decision is only relevant if he or she is a victimized second mover, which occurs with probability $\approx s/N$ (approximate because, if the second mover in this case has strategy V or I, the first mover is a victimizer with probability $s - 1/N$, rather than s); with probability $\approx (1-s)/N$, this report decision is irrelevant. Conditional on the decision being relevant, the expected payoff for reporting is higher than when not reporting if

$$\begin{aligned} r(1) + (1-r)(1-\delta-\varepsilon) &> 1-\delta \Rightarrow \\ r &> \frac{\varepsilon}{\varepsilon+\delta} \equiv R. \end{aligned} \tag{3.1}$$

Hence, the best report decision is to report when $r > R$, not report when $r < R$, and either report or not report when $r = R$.

An actor's steal decision is only relevant if he or she is the first mover, which occurs with probability $1/N$; otherwise, the steal decision is irrelevant. Conditional on the

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

decision being relevant, the probability that the second mover will report the crime is $\approx r$ (approximate because, if the first mover in this case has strategy I, the second mover reports with probability $r - 1/N$, rather than r). Hence, the expected payoff for stealing is higher than when not stealing if

$$r\left[r\left(1+\alpha\delta-\delta-\theta\right)+\left(1-r\right)\left(1+\alpha\delta\right)\right]+\left(1-r\right)\left(1+\alpha\delta\right) \;>\; 1 \Rightarrow$$
$$r \;<\; \sqrt{\frac{\alpha\delta}{\delta+\theta}} \equiv S. \quad (3.2)$$

Hence, the best steal decision is to steal when $r < S$, not steal when $r > S$, and either steal or not steal when $r = S$.

Altogether, the actor's *ex ante* best response function, assuming $S \neq R$, is thus

$$BR_i\left(s,r\right) = \begin{cases} \{P,A\}, & \text{if } s=0 \text{ and } r>S, \\ \{P,A,I,V\}, & \text{if } s=0 \text{ and } r=S, \\ \{I,V\}, & \text{if } s=0 \text{ and } r<S, \\ P, & \text{if } s>0 \text{ and } r>S \text{ and } r>R, \\ A, & \text{if } s>0 \text{ and } S<r<R, \\ I, & \text{if } s>0 \text{ and } R<r<S, \\ V, & \text{if } s>0 \text{ and } r<S \text{ and } r<R, \\ \{P,A\}, & \text{if } s>0 \text{ and } S<r=R, \\ \{P,I\}, & \text{if } s>0 \text{ and } R<r=S, \\ \{A,V\}, & \text{if } s>0 \text{ and } r=S<R, \\ \{I,V\}, & \text{if } s>0 \text{ and } r=R<S. \end{cases} \quad (3.3)$$

This function is depicted graphically in Figure 1.

3.2 Rationalizable strategies

The game-theoretic notion of rationalizability identifies the set of strategies that will be played when players have common knowledge that every player will play a best response (Fudenberg and Tirole 1996). This set is identified via iterative logic. If one actor will never play one strategy because it is never a best response, then the other players know that fact and take it into account when identifying their own best responses. Moreover, the first player knows that the other players take that fact into account when making her own choice, and this fact is known by the other players, and so on.

One way to identify the set of rationalizable strategies is to construct an infinite chain of justification, at each step of which one player justifies her own best response by a conjecture of what the other players do, and then at the next step, the other players justify those conjectured strategies by their own conjectures, and so on. If such a chain can be constructed in which each player's conjectured strategy is a best response to a conjecture that is also a best response, then the strategies are rationalizable.

Proposition 1 *Fix δ, α, ε, and θ. Then, every pure strategy is rationalizable.*

Proof Consider strategy P. P is a best response to a conjecture that all others choose P, which is a best response for each actor to the conjecture that all others choose P, and so

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

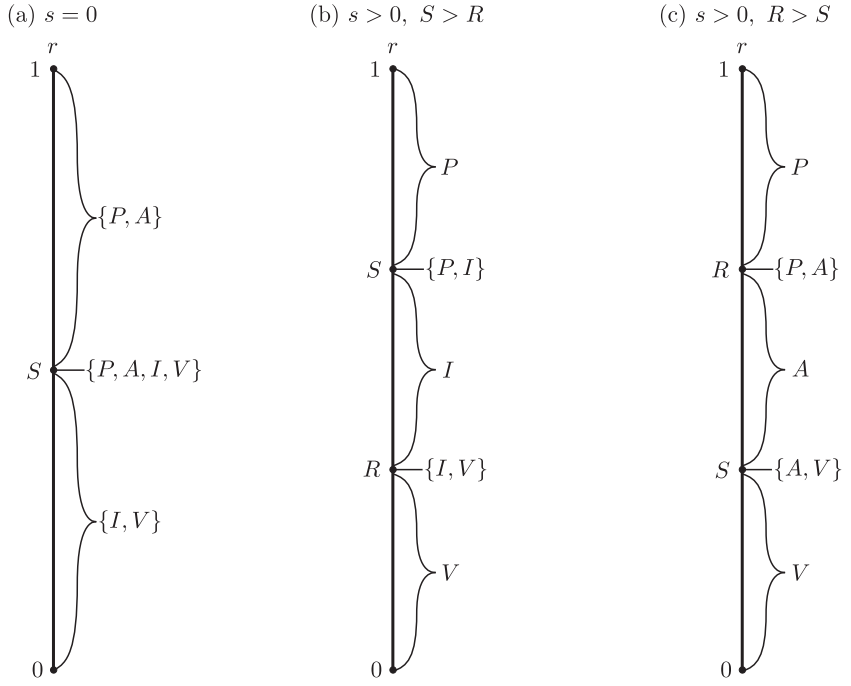


FIGURE 1. A graphical representation of the best response function $BR_i(s, r)$ of equation (3.3) under the three cases: (a) $s = 0$, (b) $s > 0,\ S > R$, and (c) $s > 0,\ R > S$.

on. An infinite chain of justification can thus be created in which each actor believes all others choose P at each step, thus making P rationalizable.

Next consider strategy A. A is a best response to a conjecture that all others choose P, which we know is rationalizable from above. An infinite chain of justification can thus be created for A.

Now consider strategy V. V is a best response to a conjecture that all others choose V, which is a best response for each actor to the conjecture that all others choose V, and so on. An infinite chain of justification can thus be created in which each actor believes all others choose V at each step, thus making V rationalizable.

Finally, consider strategy I. I is a best response if all others choose A, which we know is rationalizable from above. An infinite chain of justification can thus be created for I. □

As evident by this result, common knowledge of rationality does not alone restrict the set of potentially observable behaviours.

3.3 Perfect Bayesian equilibrium

The standard solution concept for (Bayesian) games in which the actors act simultaneously not knowing the move by Nature is the Bayesian Nash Equilibrium (BNE) concept. Our second result identifies the BNE of the Adversarial Game.

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

Proposition 2 *Fix δ, α, ε, and θ.*

(a) *The set of pure BNE consists of the following:*

(i) *symmetric profile "Utopia", in which each actor chooses P;*

(ii) *symmetric profile "Dystopia", in which each actor chooses V;*

(iii) *any asymmetric "Semi-Utopia" profile, in which a fraction z of actors choose P, the remaining fraction $(1-z)$ of actors choose A, and $S \leqslant z < 1$;*

(iv) *and the asymmetric "Semi-Dystopia" profile, in which the fraction R of actors choose I and the remaining fraction $(1-R)$ of actors choose V.*

(b) *The sum of expected utilities is maximized in Utopia and Semi-Utopia.*

Proof (a) 1. We first show that the profiles listed are equilibria.

(a-i) Utopia with every actor choosing P implies $(s = 0,\ r = 1)$. From the best response function, we see that P is a best response to $(s = 0,\ r = 1)$ for each actor. Hence, Utopia is a pure, symmetric BNE.

(a-ii) Dystopia with every actor choosing V implies $(s = 1,\ r = 0)$. From the best response function, we see that V is a best response to $(s = 1,\ r = 0)$ for each actor. Hence, Dystopia is a pure, symmetric BNE.

(a-iii) Consider Semi-Utopia with $S \leqslant z < 1$. It follows that $(s = 0,\ r = z \geqslant S)$. From the best response function, we see that both P and A are best responses. With each actor choosing a best response, it follows that Semi-Utopia is an asymmetric, pure BNE.

(a-iv) Consider Semi-Dystopia with the fraction R of players choose I and the fraction $(1-R)$ of players choose V, so that $(s = 1,\ r = R)$. From the best response function, each Informant and Villain is playing a pure best response. Hence, this Semi-Dystopia is an asymmetric, pure BNE.

2. It is straightforward to show that for every other $(s,\ r)$ combination, at least one actor is strictly better off in expectation by changing strategy. The cases left to consider are as follows:

(i) $(s = 0,\ r < S)$, where non-criminals are strictly better off in expectation by stealing;

(ii) $(s > 0,\ r > S)$, where criminals are strictly better off in expectation by not stealing;

(iii) $(s > 0,\ R < r \leqslant S)$, where non-reporters are strictly better off in expectation by reporting;

(iv) $(s > 0,\ 0 < r < R)$, where reporters are strictly better off in expectation not reporting;

(v) $(0 < s < 1,\ r = 0$ or $r = R < S)$, where non-criminals are strictly better off in expectation by stealing.

(b) Social utility is lost anytime a crime occurs, and crime occurs in expectation if and only if at least one actor selects a V or I strategy. Hence, Utopia and Semi-Utopia have maximized sum of utilities. □

The Utopia and Dystopia labels are taken from SBD, though the social configurations they apply these labels to are slightly different. SBD use Utopia to refer to any state in

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

which no crime occurs, thus combining our definition of Utopia as well as our definition of Semi-Utopia under one label. They refer to both Utopia and Semi-Utopia as Utopia because the distinction was not important given their dynamic analysis. However, our analysis below finds the distinction to be important. SBD use Dystopia to refer to the state in which no punishing occurs, as we do here. However, the SBD Dystopian state includes both Villains and Apathetics, while ours includes only Villains. There is no analogue of our Semi-Dystopia in the SBD dynamics.

Two technical matters regarding Proposition 2(b) are of note. First, the sum of utilities is maximized in any setting with no crime on the equilibrium path, so it is possible that the sum of*realized* utilities can be maximized with some actors choosing I or V. This occurs when none of those crime committing actors is selected to be the first mover. However, whenever I or V are chosen by at least one actor and before roles are realized, there is a non-zero probability that an inefficient outcome will occur. Second, if a fraction z of actors choose P and the remaining fraction $(1-z)$ choose A, then the sum of utilities will be maximized with any z, $0 \leqslant z \leqslant 1$. In effect, there are efficient, non-equilibrium Semi-Utopia strategy profiles when $0 \leqslant z < R$.

3.4 Equilibrium refinements

As seen above, the Adversarial Game has multiple equilibria. Without any further assumptions about how strategies are selected, it is not clear which, if any, of the equilibria would be chosen. One way to approach this equilibrium selection problem is to perform an evolutionary analysis. However, before turning to such analysis, we consider equilibrium refinements in the one-shot game, which can identify which equilibria are most likely to be played and thereby identify which equilibria we expect may arise in our evolutionary analysis.

Although we find many asymmetric Semi-Utopia to be equilibria in the one-shot game, we do not find many asymmetric Semi-Dystopia equilibria. The reason for this result is that the symmetric Utopia with all Paladins is a weak BNE while the symmetric Dystopia with all Villains is a strict BNE. To see this, observe that if all players other than player i are Paladins, then i is indifferent between being a Paladin and an Apathetic, while if all others are Villains, then i's unique best response is to be a Villain. This fact has implications for refining the set of equilibria.

We consider here two refinements. The first is the notion of an Evolutionarily Stable Strategy (ESS), which is a strategy that, if adopted by all actors, cannot be invaded by a mutant (Sandholm 2010). Recognizing that the BNE of the Adversarial Game is effectively a Nash Equilibrium, we here consider as an ESS in this Bayesian Game a strategy whose expected utilities satisfy the utility conditions for ESS. The second refinement is THPE in which no actor wants to change his or her strategy given the small chance that others will deviate from their intended strategies (Fudenberg and Levine 1996). As with ESS, we apply the THPE concept using expected payoffs rather than realized payoffs.

Proposition 3 *Fix δ, α, ε, and θ.*

(a) *V is the only pure ESS.*

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

(b) *Dystopia and Utopia are the only pure THPE.*

Proof

(a) Let $U_i(x, y)$ be the expected payoff to actor i when he or she chooses x and all other $N-1$ actors choose y. By definition, a pure ESS is a symmetric pure equilibria in which all actors choose strategy x such that (i) $U_i(x, x) \geqslant U_i(y, x)$ and (ii) if $U_i(x, x) = U_i(y, x)$, then $U_i(x, y) > U_i(y, y)$, for any strategy $y \neq x$. From Proposition 2(a), the only two symmetric equilibria that may potentially be ESS are all choose V (Dystopia) and all choose P (Utopia).

Consider Utopia. From the best response function, it is evident that $U_i(P, P) = U_i(A, P)$ but $U_i(P, A) = U_i(A, A)$. Hence, P is not an ESS. Now consider Dystopia. From the best response function, it is clear that $U_i(V, V) > U_i(y, V)$ for $y = \in \{P, A, I\}$. Hence, V (Dystopia) is an ESS.

(b) A THPE is a Nash Equilibrium with certain properties, one of those being that no weakly dominated pure strategy can be played in a THPE (see Fudenberg and Tirole 1996). From Proposition 2(a), the only equilibria to consider are Dystopia, Utopia, and Semi-Utopia. From the best response function: V is the unique best response when all others choose V, so Dystopia is a THPE; P is the unique best response when s is close to 0 and r is close to 1, so Utopia is a THPE; but A is weakly dominated by P in Semi-Utopia, so Semi-Utopia is not a THPE.

□

That Dystopia withstands typical revisions follows from it being a strict equilibrium. Conversely, all-Paladin Utopia cannot withstand invasions by Apathetics, but it can withstand small mutations or errors in which crime occurs.

3.5 Other social dilemma games and the reporting externality

It is readily apparent from Propositions 1 and 2 that one fundamental difference between the Adversarial Game and standard social dilemma games (Prisoners Dilemma, Public Good, Common Property Resource) is that whereas a social dilemma game typically has a single equilibrium that is inefficient, the Adversarial Game has multiple equilibria, one of which is efficient. Indeed, upon closer inspection, though the Adversarial Game is a social dilemma game in spirit and purpose, it is actually an N-actor, four-strategy coordination game with efficient Utopia and inefficient Dystopia as the two focal equilibria. The Adversarial Game is transformed from a social dilemma into a non-standard coordination game via the reporting externality which makes punishment (reporting) a best response when enough others also punish (we refer to it as a non-standard coordination game because coordination is successful on only two of the four strategies). The second-order punishment problem is not a typical social dilemma but rather a coordination game that has enough force to transform the first-order social dilemma, indeed, the entire game, into a coordination game.

It is instructive to revisit SBD's interpretation of the model with this new insight in mind. SBD explain that the reporting reflects a willingness to cooperate with authorities.

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

		Player 2 C	Player 2 D
Player 1	C	2,2,0	0,3,0
	D	3,0,0	1,1,0*

(A) Reporting rate $r = 0$ (Player 3)

		Player 2 C	Player 2 D
Player 1	C	2,2,0*	0,2,0
	D	0,2,0	0,0,0

(B) Reporting rate $r = 1$ (Player 3)

FIGURE 2. Modified prisoner's dilemma game.

That is, the setting is one in which there already exists in place a third-party institutional framework that can leverage society members' willingness to cooperate into effective punishment of criminal behaviour. The larger lesson is that if such institutions can be developed, then cooperation can be sustained even in short horizon interactions. This begs the question of how such institutions can be developed in the first place; however, it is evident that such institutions fundamentally change the incentives of potential criminals when there is sufficient societal support.

In principle, any social dilemma game can be converted from a social dilemma to a modified coordination game through a similar mechanism. Consider the following example. Suppose the Adversarial Game payoffs were replaced with a simultaneous move Prisoners Dilemma Game payoffs and then appropriately modified. Also suppose that whether defectors are reported and punished depends only on the choice of a single bystander, actor 3, whose actions represent society's support or lack thereof for punishing defectors. The bystander can choose the reporting rate to be either $r = 0$ or $r = 1$, and his or her payoff does not depend on this choice. The left matrix in Figure 2 depicts the typical Prisoners Dilemma payoffs when the bystander does not report or punish defectors. The payoffs in the right matrix correspond to when there is full reporting and punishment. The payoffs are calculated by subtracting 3 from each defector and adding 2 to each victim; the defector pays back the 2 lost by the defection, which goes to the victim, and then pays an additional 1 as punishment.

In this modified Prisoners Dilemma Game, as with the Adversarial Game, we have defection as the unique best response when there is no reporting, and we have cooperation as the unique best response with full reporting. The two equilibria of Dystopia and Utopia (denoted by * in the matrices) are the only two pure Nash Equilibria in the game. More generally, if actors coordinate their reporting, then the incentives to defect will be overcome by the threat of punishment. These two equilibria reveal it to be a game of coordination.

4 Evolutionary analysis with best response dynamics

SBD's main purpose is to examine the evolution of behaviour with long-run repetition of the Adversarial Game. They assume that strategies switch according to a special imitation dynamic. Say that an actor is a "loser" in a period if that actor's payoff is strictly less than the initial endowment. Clearly, a first mover will only be a loser if he or she was an Informant or Villain that was convicted after being reported against, while a second mover is a loser if he or she is a victimized Apathetic or Villain or a victimized Paladin or Informant that unsuccessfully challenged after being victimized. The "winner" is the other player in that period. SBD assume that only losers switch strategies and that they

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

do so by (possibly imperfectly) mimicking one of the players in that round. The choice of which player to mimic is made with a probability proportional to the players' payoffs for that round. This probability implies a certain amount of inertia in the system. However, a caveat is that any loser who chooses to mimic the second player always becomes a non-criminal type, directly mimicking the second player's reporting strategy only. The idea behind this is that the loser in this case has decided not to mimic the criminal (which would certainly cause him or her to become a criminal type), and has therefore implicitly decided not to commit criminal acts. In short, the SBD dynamic is best described as a modified imitation dynamic with inertia.

SBD show that, with a deterministic version of their imitation dynamic, the Adversarial Game always converges to either Utopia/Semi-Utopia or their form of Dystopia (consisting of mostly Villains with some Apathetics, due to the imperfect mimicking), and that whether the system converges to one or the other depends strongly on the presence of Informants in the initial population. If there are no Informants in the initial population, then the system converges to Dystopia unless a large fraction of the initial population are Paladins; if there are any Informants, then the system converges to Utopia/Semi-Utopia. With a stochastic imitation dynamic, the system converges to Utopia/Semi-Utopia with quickly increasing probability as the initial number of Informants increases. In short, increasing the second-order cooperation (punishment of defectors) among defectors themselves has a powerful effect on the system's resting state, so much so that any number of initial Informants is sufficient to eventually bring the system to Utopia.

It remains to be seen whether this striking result holds for other strategy revision dynamics. We here consider the best response dynamic, which is an important alternative to the imitation dynamic (Sandholm 2010). In an imitation dynamic, the actor adopts the strategy of another actor, usually with the likelihood of adoption increasing in the performance of the actor potentially copied. Inertia takes two forms: only one or a small number of actors are allowed to switch strategies in a given round, and each of those potential switchers will switch with probability less than 1. In a best response dynamic, the actor that switches does so with more foresight: he or she switches to a strategy that is a best response to the current strategies of the other actors. As with an imitation dynamic, a best response dynamic usually incorporates inertia, with only one or a few actors allowed to switch in a given period.

There are many forms that a best response dynamic can take. Three main characteristics are the rate of inertia, how choice is made among competing best responses, and whether decision errors are allowed during strategy switches. We consider a Best Response Dynamic, denoted SBD-BRD, that selects on these three characteristics to maintain a closeness in spirit to the SBD imitation dynamic. The intent is to find the closest best response dynamic analogue to the SBD imitation dynamic, thus allowing for the sharpest comparison of results when switching the dynamic. We then consider an important variation on the dynamic to check robustness.

Definition 1 SBD-BRD is the following switching protocol:

(a) *At the end of period t, with probability* $q > 0$ *one of the two selected actors is chosen to switch strategy.*

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

(b) *Conditional on being selected to switch, the actor switches to a best response to the population strategy profile of period t, and if more than one best response exists then one is selected at random (uniformly).*

Note how SBD-BRD handles inertia, multiple best responses, and switching errors. It intentionally mimics the degree of inertia in SBD's imitation dynamic: there is a chance no actor switches strategy; at most one actor switches strategy; and that actor is one of the two matched actors. SBD's assumption that the loser is the one to switch matches the spirit of imitation dynamics because a loser would want to imitate a winner but not vice versa. However, in the spirit of the best response dynamic, potentially any actor may see benefits in switching, so we allow either of the two actors to be the one to switch. SBD's imitation dynamic does not have a natural analogue to the issue of multiple best responses in the best response dynamic, so without guidance from the SBD dynamic we here adopt a conventional assumption that the actor mixes equally among best responses. SBD's imitation dynamic also does not include switching errors, and SBD-BRD makes a similar assumption.

Proposition 4 *Fix δ, α, ε, and θ. Under SBD-BRD, there are two possible long-term behaviours:*

(a) *the system evolves to Dystopia, or*

(b) *the system maintains a level of reporting $r \gtrsim S > R$.*

Proof Here, we assume that $|R - S| \gg 1/N$, and define the following regimes of (s,r) space, which cover all possible values:

(i) $s > 0,\ r < R$,

(ii) $s = 0,\ r \geqslant S$,

(iii) $s = 0,\ r < S$,

(iv) $s > 0,\ r \geqslant R$.

(a) First, consider the case $S < R$, with initial conditions in regime (i). Here, the best response is to not report, so that each strategy update will either leave r unchanged or will decrease r. There is a non-zero probability that these sequences of updates will also leave $s > 0$. Conditional on s remaining non-zero, then, the system is guaranteed to eventually evolve to a point in which $r < S$, in which case V is the unique best response, leading the system to eventually reach Dystopia and stay there forever. If s becomes zero before this can happen, the system switches to regime (ii).

Now, consider the case $S < R$ with initial conditions in regime (ii). If $r > S$, either P or A is a best response, so each update will keep $s = 0$ and may increase, decrease, or maintain r. With probability 1, then, the system will eventually evolve through neutral drift to either a point in regime (iii) or to the specific point $s = 0,\ r = S$. If $r = S$, all strategies are best responses, so eventually an update will occur that either creates a criminal and puts the system back in regime (i), maintains $s = 0$ and causes

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

$r > S$ (still in regime (ii)), or puts the system into regime (iii). Hence, the eventual end result of beginning in regime (ii) is to either enter regime (i) or regime (iii).

Let us consider, then, $S < R$ with initial conditions in regime (iii). Here, the best responses are V and I, so the next update will necessarily put the system into regime (i) above.

Next, consider the case $S < R$, with initial conditions in regime (iv). If $r = R$, either P or A is a best response, so there is a non-zero probability that r decreases below R, leading to regime (i). Otherwise, the unique best response is P, causing r to either increase or stay the same each round, and causing s to decrease or stay the same each round. Eventually, then, the system will evolve to a state in which $s = 0,\ r > R > S$, which is regime (ii).

Hence, when $S < R$, initial conditions in regimes (ii)–(iv) will all eventually (though possibly indirectly) lead to regime (i), which leads to the absorbing Dystopia state with a non-zero probability, and otherwise leads back to regime (ii). Therefore, when $S < R$ the system will eventually evolve into, and forever remain in, the Dystopian state, regardless of initial conditions.

Finally, consider the case $R < S$ with initial conditions in regime (i). The best response here is V, causing the system to evolve to Dystopia with probability 1.

(b) Consider now $R < S$, with initial conditions in regime (iv). If $r = R$, both I and V are best responses, so the system may evolve to regime (i) and, therefore, to Dystopia. However, there is a non-zero chance that the system instead evolves such that $r > R$. With r strictly greater than R, reporting is always the best response, so r will only increase or maintain. Furthermore, if $r < S$, we are guaranteed to maintain $s > 0$, as I is the unique best response. Therefore, the system will eventually evolve to a point with $s > 0,\ r \geqslant S$. At this point, P is certainly a possible best response (usually the unique best response), so the system will eventually evolve into regime (ii).

Initial conditions in regime (ii) behave the same way when $R < S$ as when $R > S$, though this behaviour may lead to entering different subsequent regimes, due to the switching of the order of R and S. That is, initial conditions (ii) will eventually lead through neutral drift either to regime (iv), specifically with $S - \frac{1}{N} \leqslant r$, or regime (iii), specifically with $S - \frac{1}{N} \leqslant r < S$.

Now we turn to $R < S$ with initial conditions in regime (iii). Here, the best response is I or V, so the next update will cause us to leave regime (iii) as s becomes non-zero. But, there are three possibilities for how this will happen. First, if $r \geqslant R + \frac{1}{N}$, this update will put us into regime (iv). This is guaranteed to happen if we entered regime (iii) from regime (ii), as mentioned above, and will specifically lead to regime (iv) with $r \geqslant S - \frac{2}{N}$. If $r < R - \frac{1}{N}$, this update will put us into regime (i), and lead to Dystopia. If $R - \frac{1}{N} \leqslant r < R + \frac{1}{N}$, the system may evolve into either (iv) or (i).

We can therefore summarize these results in the following way. For $R < S$, if the system is ever in a state in which $r \geqslant S - \frac{2}{N}$, r is guaranteed never to fall out of that region again. In this case, the system will maintain an r in this range while cycling between regimes (iv) and (ii) (potentially with very brief stops in (iii) between).

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

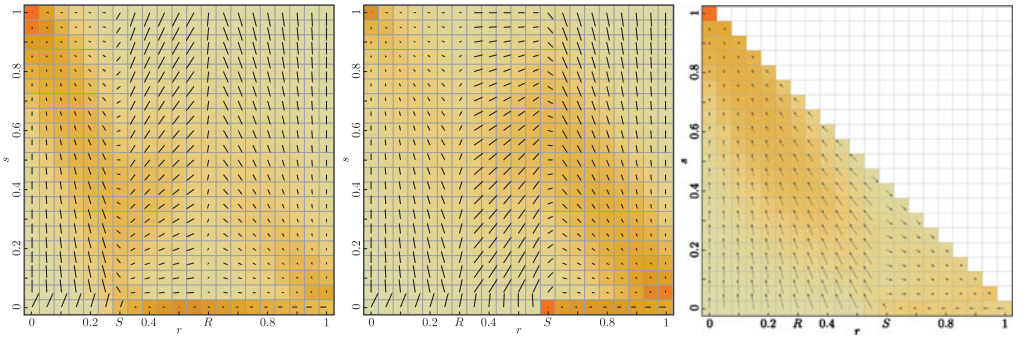


FIGURE 3. Mean flow fields in (r, s) space for the cases $R > S$ (left), $S > R$ (centre), and $S > R$ but with no Informants (right), using $N = 20$. Colours represent the total amount of time spent at a cell, ranging from light gray (little time) to bright orange (much time), assuming a uniform initial distribution on the number of players of each strategy type. For $R > S$, all trajectories eventually end in Dystopia. For $S > R$ (centre), we have treated $r = S$ as an absorbing point (all trajectories passing through that point actually cycle near it indefinitely), so that all trajectories end either there or at Dystopia. With Informants disallowed (right), however, all trajectories end in Dystopia.

Furthermore, once this cycling begins, there will never be more than 1 Villain in the system at a time, and never more than $SN + 1$ Informants (the typical number will be much less than this). Finally, the only initial conditions that are guaranteed to not lead to this long term behaviour (and are guaranteed to enter Dystopia) are those in regime (i) or those in the subset of regime (iii) with $r < R - \frac{1}{N}$; essentially those initial conditions with $r \lesssim R$. All other initial conditions are either guaranteed to end in this cycling (when $r \gtrsim R$) or may lead to either this cycling or to Dystopia (when $r \approx R$).

□

The above argument is explored graphically in Figure 3. Here, we have plotted the mean flow fields in (r, s) space for the two cases $R > S$ and $S > R$, using $N = 20$. These mean fields assume a uniform distribution on the initial number of players of each strategy. The background colours range from light gray to bright orange, and represent the total amount of time spent in each of the available cells, from very little to very much. For the case $R > S$ (left), we see that the trajectories are separated along the line $r = R$ when $s > 0$, but that the $s = 0$ line serves as a funnel from the region $r > R$ to the region $r < R$. When $r < R$, the trajectories inevitably lead to Dystopia, where they remain forever (hence the brightest orange square on the figure). For the case $S > R$ (centre), we also see a separation around $r = R$, but the line $s = 0$ no longer funnels trajectories from above this separation to below it. Here, we have treated $r = S$, $s = 0$ as an absorbing point (hence the brightest orange on that plot), because any trajectory passing through it will in fact never settle down to a single point, and will forever cycle near that region.

Proposition 4 reveals that the SBD result about the influential role of Informants does generalize in the case of SBD-BRD, as long as $R < S$, though the role that the Informants play differs from SBD to best response. In SBD, Informants are a sufficient but not

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

necessary condition to drive the system to Utopia/Semi-Utopia. In SBD-BRD, on the other hand, Informants are a necessary but not sufficient condition to keep the system near Utopia/Semi-Utopia. This result is formalized in the following proposition, which considers a modified form of SBD-BRD.

Proposition 5 *Fix δ, α, ε, and θ. Under SBD-BRD, but with the strategy I now disallowed, the system will always evolve to Dystopia in the long term.*

Proof With I disallowed, we must now reconsider our Best Response function $BR_i(s,r)$. For all entries that list I as a non-unique best response, we may simply remove I. This leaves only the special case $R < r < S$, for which I is the unique best response. Note, therefore, that if $S < R$, removing I from our possible list of strategies does not change the Best Response function in any qualitative way, so that the result above—that if $S < R$ the system will always evolve to Dystopia—still holds. Also, note that if I is disallowed, the inequality $s \leqslant 1 - r$ must hold, since in this case $s = V/N$ and $r = P/N$, and $V = N - A - P$.

If $R < S$, we must determine the new best response for $R < r < S$. Since $r < S$, V is superior to A, and since $R < r$, P is superior to A. Hence, either V or P is the best response, or both. The expected payoffs for P and V are

$$1 - \frac{s}{2}(1-r)(\delta+\varepsilon) \text{ and } 1 - \frac{s}{2}\delta + \frac{1}{2}\left[\alpha\delta - r^2(\delta+\theta)\right],$$

respectively. These payoffs are equal when $s = Q(r)$, with

$$Q(r) = \frac{(\delta+\theta)(S^2-r^2)}{(\delta+\varepsilon)(r-R)}.$$

Thus, if $s > Q(r)$, P is the best response, if $s < Q(r)$, V is the best response, and if $s = Q(r)$, both P and V are the best response.

The function $Q(r)$ is defined on $s \in (R,S]$, is decreasing and concave up with slope guaranteed to satisfy $Q'(r) \leqslant -2$, and $Q(S) = 0$ while $Q(s)$ diverges as $r \to R$. Furthermore, there is a critical value of r, call it r_*, where the curve $s = Q(r)$ intersects the line $s = 1 - r$ (the maximum value s can attain for any given r), and $R < r_* < S$. Therefore, V is always the unique best response if $r < r_*$ since in this case s cannot possibly be greater than $Q(r)$. We now introduce two new regimes for initial conditions, based upon $Q(s)$:

(v) $s > 0,\ r \geqslant Q(s)$,

(vi) $s > 0,\ r < Q(s)$.

Now, consider $R < S$ with initial conditions in regime (v). If $r = Q(s)$, the next update may either remain in regime (v) with $r > Q(s)$, put the system into regime (vi), or put the system into regime (ii). For $r > Q(s)$, P is the unique best response, so the system will eventually evolve to a state in regime (ii).

Consider $R < S$ with initial conditions in regime (ii). Here, through neutral drift the system will eventually evolve into one of three other regimes: regime (v), specifically with $s = \frac{1}{N},\ r = S$; regime (iii), specifically with $S - \frac{1}{N} \leqslant r < S$; or regime (vi), specifically

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

with $s = \frac{1}{N}$, $r = S - \frac{1}{N}$. To see this last possibility, it is important to remember that $Q'(r) \leqslant -2$ and $Q(S) = 0$, so that the point specified certainly resides within regime (vi) and not regime (v).

Consider next $R < S$ with initial conditions in regime (iii). Here, the unique best response is V, so the next step will bring the system out of this regime and into either regime (vi), which is always a possibility, or regime (v), which is only a possibility if $S - \frac{1}{2N} \leqslant r < S$. As above, this limitation on the possibility of bringing the system into regime (v) relies on the facts that $Q'(r) \leqslant -2$ and $Q(S) = 0$.

Finally, consider $R < S$ with initial conditions in regime (vi). Here, V is the unique best response, so each update that changes r or s is guaranteed to either decrease r while increasing s (if a Paladin becomes a Villain) or maintain r while increasing s (if an Apathetic becomes a Villain). If $r \geqslant r_*$, then there is a non-zero probability that the system will undergo a sufficient number of these latter updates to bring it back into regime (v). But, there is also a non-zero probability that a sequence of updates occurs whereby the system remains in regime (vi) and attains $r < r_*$. Below this point, regime (v) can never be re-entered, so the system evolves to Dystopia.

Hence, if $R < S$ and I is disallowed, all initial conditions will eventually bring the system into regime (vi), which leads to Dystopia with a non-zero probability. Therefore, all initial conditions are guaranteed to eventually evolve to Dystopia, where the system will then remain forever. □

The above argument is illustrated in Figure 3 on the right.

5 Conclusion

SBD's evolutionary game showed that Informants alone may drive a society from Dystopia to Semi-Utopia. As a consequence, its primary policy implication is that recruiting Informants from the general population may be an overall valid approach in helping reduce crime. SBD's result stems from the primary role Informants play in their imitation dynamics, where Informants render Dystopia unstable to small perturbations. In SBD's case, then, Dystopia may be reached under a set of initial conditions containing no Informants, but any infinitesimal deviation from this will drive the system to Semi-Utopia in the long run. This is a universal result, in the sense that it holds across parameter space, as long as Informants are present. It becomes natural then to ask what is the best way to recruit Informants from the general population—and hasten the transition towards Semi-Utopia—given specific costs associated with different player's current strategies and histories. This question is explored in Short *et al.* (2012).

Our current analysis presents a more nuanced picture: we find that under the best response dynamics, the role Informants play depends on parameter choices, and that their mere presence does not necessarily drive the system to Utopia as they did in SBD's original work. In particular, when $S < R$, Informants do not play a central role because the strategy of committing crimes and reporting to authorities is never a best response. On the other hand, when $R < S$, the Informant strategy plays a pivotal role in evolving the system toward, and then maintaining, a level of reporting $r \gtrsim S$, as long as the initial reporting level $r_0 \gtrsim R$. Thus, the availability of I as a possible strategy is the only

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

channel allowing the system to reach and maintain a state very near Semi-Utopia, making Informants a necessary, but not sufficient, criteria for the emergence of a low-crime state.

As discussed earlier, simply converting citizens to Informants in the context of best response dynamics will not necessarily guarantee the system to evolve towards Utopia as it did in SBD's dynamics. Carefully choosing parameters, however, may allow the system to make this transition; specifically, if parameters are chosen so that $R < S$. In terms of the original model parameters, the $R < S$ constraint implies that several conditions must be met in order for the low-crime state to emerge. Let us first note that we may assume α, δ, and ε to be fixed parameters intrinsic to the game between victim and victimizer, that cannot be adjusted by law enforcement. We may instead allow the degree of punishment θ to be a variable parameter that authorities may fine tune at will. In terms of θ, then, the constraint $R < S$ implies that

$$\theta < \frac{\delta}{\varepsilon^2}\left[\alpha(\varepsilon+\delta)^2 - \varepsilon^2\right]. \tag{5.1}$$

In other words, the degree of punishment imparted on criminals must not be too onerous, so as not to discourage the strategy I from being chosen. Hence, one policy implication of the current work is that punishments should not be made too harsh, lest Semi-Utopia be made unobtainable.

Furthermore, note that equation 5.1 can be made true only if $\alpha > \varepsilon^2/(\varepsilon+\delta)^2$. This implies, though, that for a given $\alpha < 1$, criminals may always choose a set of parameters δ and ε that violates this second inequality, by choosing

$$\varepsilon \geqslant \frac{\delta\sqrt{\alpha}}{1-\sqrt{\alpha}}, \quad \delta \leqslant 1-\sqrt{\alpha}\ ; \tag{5.2}$$

the latter condition arises from the constraint that $\delta + \varepsilon \leqslant 1$. Thus, if retaliation against witnesses is too large, the system has no chance of reaching a low-crime state, regardless of θ. Of course, this raises another game theoretic challenge, this time for criminals: to guarantee a crime friendly society, criminals must restrain themselves to stealing no more than $\delta = 1 - \sqrt{\alpha}$, but each criminal is individually motivated to choose a δ as large as possible, i.e., 1. Hence, it seems plausible that only in societies where criminals themselves are highly organized can such a situation arise, and if each criminal is acting purely in his own best interest, the authorities should be able to arrange for punishment levels that promote a Semi-Utopian state.

Of course, whether an imitation, best response, or some other dynamic best captures actual decisional strategies is an empirical question. Are players perfectly rational and capable of choosing a best response in each one-shot game, perhaps allowing for an initial learning phase, or do they rather learn by imitation and by adaptation? In relation to the experiments already performed (D'Orsogna *et al.* 2013), the best response dynamic leads to an Informant role that is closer to reality—specifically, that Informants are necessary to achieve Utopia. However, we note that when experimental conditions disallow the Informant and a Dystopian state follows, the Dystopian state still contains a significant minority of players who choose to be Paladins, despite the fact that Paladins are certainly not the best response to the system configuration. Hence, it appears that while some

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

players may in fact be payoff maximizing and are approximating the best response dynamic, others may be attempting to maximize a different utility, perhaps one that includes concepts such as fairness and justice. It would thus be interesting to further analyse, using the experimental data, whether the actual test subjects followed either of these two dynamics—best response or imitation—or some other decisional process when faced with the choice of committing and reporting crimes. We leave this as future work.

References

[1] Beittel, J. (2009) *Mexico's Drug Related Violence*, Congressional Research Service, Washington, D.C.

[2] Bowles, S. & Gintis, H. (2011) *A Cooperative Species: Human Reciprocity and its Evolution*, Princeton University Press, Princeton, NJ.

[3] Boyd, R., Gintis, H. & Bowles, S. (2010) Coordinated punishment of defectors Sustains cooperation and can proliferate when rare. *Science* **328**, 617–620.

[4] Bursik, R. & Grasmick, H. (1993) *Neighborhoods and Crime: Dimensions of Effective Community Control,* Lexington Books, New York.

[5] Carpenter, J. (2007) Punishing free-riders: How group size affects mutual monitoring and the provision of public goods. *Games Econ. Behav.* **60**, 31–51.

[6] Chaudhuri, A. (2011) Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Exp. Econ.* **14**, 47–83.

[7] Cinyabuguma, M., Page, T. & Putterman, L. (2005). Cooperation under the threat of expulsion in a public goods experiment. *J. Public Econ.* **89**, 1421–1435.

[8] D'Orsogna, M. R., Kendall, R., McBride, M. & Short, M. B. (2013) Criminal defectors lead to the emergence of cooperation in an experimental, adversarial game. *PloS One* **8**, e61458.

[9] Eaton, B. C. & Wen, J. (2008) Myopic deterrence policies and the instability of equilibria. *J. Econ. Behav. Organ.* **65**, 609–624.

[10] Ehrlich, I. (1996) Crime, punishment, and the market for offenses. *J. Econ. Perspect.* **10**, 43–67.

[11] Fehr, E. & Fischbacher, U. (2004) Third-party punishment and social norms. *Evol. Human Behav.* **25**, 63–87.

[12] Fehr, E. & Gächter, S. (2000) Cooperation and punishment in public goods experiments. *Am. Econ. Rev.* **90**, 980–994.

[13] Fender, J. (1999) A general equilibrium model of crime and punishment. *J. Econ. Behav. Organ.* **39**, 437–453.

[14] Fudenberg, D. & Levine, D. (1996) *Game Theory,* MIT Press, Cambridge, MA.

[15] Fudenberg, D. & Pathak, P. (2010) Unobserved punishment supports cooperation. *J. Public Econ.* **94**, 78–86.

[16] Gambetta, D. (1988) *Trust: Making and Breaking Cooperative Relations,* Blackwell, Oxford, UK.

[17] Hauert, C. *et al.* (2007) Via freedom to Coercion: The emergence of costly punishment. *Science* **316**, 1905–1907.

[18] Henrich, J. & Boyd, R. (2001) Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *J. Theor. Biol.* **208**, 79–89.

[19] Ledyard, J. (1995) Public goods: Some experimental results. In: J. Kagel & A. Roth (editors), *Handbook of Experimental Economics*, Princeton University Press, Princeton, NJ, pp. 111–194.

[20] Levine, D. & Pesendorfer, W. (2007) The evolution of cooperation through imitation. *Games Econ. Behav.* **58**, 293–315.

[21] Ostrom, E. (1990) *Governing the Commons: The Evolution of Institutions for Collective Action,* Cambridge University Press, New York, NY.

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press

[22] Ostrom, E., Walker, J. & Gardner, R. (1992) Covenants with and without a sword: Self-governance is possible. *Am. Political Sci. Rev.* **86**, 404–417.
[23] Sampson, R. & Groves, W. (1989) Community structure and crime: Testing social-disorganization theory. *Am. J. Sociol.* **94**, 774–802.
[24] Sandholm, W. (2010) *Population Games and Evolutionary Dynamics,* MIT Press, Cambridge, MA.
[25] Short, M. B., Brantingham, P. J. & D'Orsogna, M. R. (2010) Cooperation and punishment in an adversarial game: How defectors pave the way to a peaceful society. *Phys. Rev. E* **82**, 066114.
[26] Short, M. B., Pitcher, A. & D'Orsogna, M. R. (2013) External conversions of player strategy in an evolutionary game: A cost-benefit analysis through optimal control. *Eur. J. Appl. Math.* **24**, 131–159.
[27] Skogan, W. (1990) *Disorder and Decline: Crime and the Spiral of Decay in American Neighborhoods,* University of California Press, Los Angeles, CA.
[28] Takahashi, S. (2010) Community enforcement when players observe partners' past play. *J. Econ. Theory* **145**, 42–62.
[29] Xiao, E. & Houser, D. (2011) Punish in public. *J. Public Econ.* **95**, 1006–1017.

https://doi.org/10.1017/S0956792515000649 Published online by Cambridge University Press