

Heiner, Ronald Asher

Working Paper

Robust Evolution Of Contingent Cooperation In Pure One-Shot Prisoners' Dilemmas. Part I: Vulnerable Contingent Participators Versus Stable Contingent Cooperators

CSLE Discussion Paper, No. 2002-09

Provided in Cooperation with:

Saarland University, CSLE - Center for the Study of Law and Economics

Suggested Citation: Heiner, Ronald Asher (2002) : Robust Evolution Of Contingent Cooperation In Pure One-Shot Prisoners' Dilemmas. Part I: Vulnerable Contingent Participators Versus Stable Contingent Cooperators, CSLE Discussion Paper, No. 2002-09, Universität des Saarlandes, Center for the Study of Law and Economics (CSLE), Saarbrücken

This Version is available at:

<https://hdl.handle.net/10419/23123>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

ROBUST EVOLUTION OF CONTINGENT COOPERATION IN PURE ONE–SHOT PRISONERS’ DILEMMAS

Part I: Vulnerable Contingent Participators Versus Stable Contingent Cooperators

Center for the Study of Law and Economics Discussion Paper 2002-09

Part II: Evolutionary Dynamics & Testable Predictions

Center for the Study of Law and Economics Discussion Paper 2002-10

Ronald A. Heiner
James Buchanan Center For Political Economy
George Mason University
MSN 1D3 Carow Hall
Fairfax, VA 22030
email: rheiner@gmu.edu

(September 2002)

ABSTRACT

ROC curves from the signal detection literature are used in an evolutionary analysis of one-shot and repeated prisoners’ dilemmas: showing if there is any discounting of future payoffs, or any cost of searching for an additional partner, then cooperative players who contingently participate – in terms of who to play with or when to exit – cannot survive when most other players unconditionally defect; even when contingent participators only interact with themselves by perfectly detecting their own type.

However, quite different results hold for players who act contingently, not in terms of whether to play or exit, but rather in terms of how to act with any given partner. There is a form of contingent cooperation in one-shot prisoners’ dilemmas (called CD behavior) that will robustly evolve through any payoff monotonic process, such as replicator dynamics. That is, whenever CD-players can detect their own type better than pure chance, they are guaranteed to evolve from any initial population – eventually to a unique evolutionarily stable population composed entirely of contingent cooperators – provided the fear payoff difference is less than the sum of greed and cooperation payoff differences.

The adaptive capabilities just described hold for *pure* one-shot prisoners’ dilemmas : meaning no repeated interactions or pairings in any generation are involved; no information or third party reports about past behavior are involved, all signal information arises only from symptoms detected after two strangers meet for the first time; and no subjective preferences for altruism, fairness, equity, reciprocity, or morality affect the raw evolutionary dynamics.

Testable predictions are also derived that agree with a large body of experimental data built up since the prisoners dilemma was first introduced in 1950. They describe how the CD-players’ equilibrium probability of cooperating changes: depending on the relative size of fear, greed, and cooperation payoff differences; and depending on the players’ history of communication, especially when face-to-face discussion is involved.

(JEL C60, C62, C72, C92, D80)

KEYWORDS: prisoners’ dilemma, cooperation, Nash equilibrium, evolutionary stability, replicator dynamics, signal detection, ROC curves, experiment, testable predictions

**ROBUST EVOLUTION OF CONTINGENT COOPERATION
IN PURE ONE-SHOT PRISONERS' DILEMMAS, PART I:**

Vulnerable Contingent-Participants Versus Stable Contingent-Cooperators

Center for the Study of Law and Economics Discussion Paper 2002-09

Ronald A. Heiner ¹

Various models show *contingent participants* can evolve in prisoners' dilemmas: by avoiding play with certain partners, or by exiting after starting play.² However, contingent participants are vulnerable to their share in the population. For example, Axelrod (1984) avoids this problem by assuming tit-for-tat players arrive in clusters; so they interact more often with themselves than otherwise implied by their frequency of the population.

However, suppose individual players cannot last in isolation from others, but instead must eventually search out partners and actually play with them in order to survive. Suppose players also discount future payoffs obtained from successive rounds of play, or there exists a cost of searching for new partners. Part I of this two part paper shows that contingent participants cannot succeed in a population composed mostly of always defecting players; even if they only interact with themselves by perfectly detecting their own type.

On the other hand, quite different results hold if players act contingently; not in terms of whether to participate, but rather in terms of how to act with any given partner: there is a form of contingent cooperation in one-shot prisoners' dilemmas that is always stable against invasion by other behavior, including always defect behavior; and will robustly evolve from any initial population. These results are shown by using principles from the signal detection literature and ROC curves to analyze the following situation:³

- A Players from a larger population are randomly matched into pairs, called *Adam* and *Eve*, who are strangers knowing nothing about each other's past actions or reputations. Their payoffs satisfy the prisoners' dilemma ordering, $T > R > P > S$; where $(T - R)$, $(R - P)$, $(P - S)$ are called the *greed*, *cooperation*, and *fear* payoff differences. Once randomly matched, Adam and Eve interact in two stages. In stage one, they meet for the first time, enabling them to detect an 'x' or 'y' signal from each other. These two signals result from each player dividing a continuous space of symptoms received from its partner into two categories: one category defined as an x-signal, and the other defined as a y-signal.
- B Neither player can choose the specific symptoms that will be detected by the other player. Instead, the likelihood of each player detecting symptoms from each other is governed by a pair of overlapping density functions whose shape and position are caused by the type of internal motivations determining how each player will react to the symptoms categorized as either an x or y signal by the receiving player.
- C In stage two, all communication between Adam and Eve ends, and they play a one-shot simultaneous

¹ I thank greatly colleagues Dieter Schmidtchen and Max Albert. The following persons also provided helpful discussion: Robert Axelrod, Robert Boyd, James Buchanan, Robert Frank, Herbert Gintis, Werner Güth, Jürgen Eichberger, Bruno Frey, Roland Kirstein, Robert Nelson, Charles Plott, Vernon Smith, and Jörgen Weibull.

² Axelrod & Hamilton 1981; Axelrod 1984; Frank 1988; Gauthier 1986; Orbell et. al. 1984; Schuessler 1989; Vanberg & Congleton 1992, 2001.

³ See Green & Swets 1966; Swets 1961, 1964; Egan 1975; plus other references in the text.

prisoners' dilemma while isolated from each other: meaning they must each react to their detected signal by choosing to cooperate C or defect D without knowing each other's chosen action. Each player can be programmed to react according to one of four mappings from signal x or y into actions C or D, denoted CC, CD, DC, DD. Each player can also be programmed to randomly choose action C or D independent of signal x or y, denoted C λ D; where λ versus $1 - \lambda$ equals the probability of choosing action C versus D.

- D Players cannot engage in any kind of “assortative interaction” or “clustering” that would cause the probability of interacting with another player to deviate from that generated by random matching over the whole population. Finally, Adam and Eves' interaction ends after they each choose an action and receive a payoff; and no information about their actions or payoffs becomes available to anyone else.

Conditions A – D imply always defecting DD behavior remains dominant over CC and C λ D behavior, as expected from traditional theory. However, DD behavior is not dominant over both signal-contingent strategies (CD, DC). Instead, Section IV of Part I shows that contingently cooperating CD behavior is a *strict* Nash equilibrium plus *evolutionarily stable*, whenever CD-players can detect their own type better than pure chance.

However, this does not rule out other stable equilibria, say with mostly DD-players. So Part II of this paper shows that no other stable equilibria exist: because CD behavior has robust adaptive capabilities guaranteeing it will evolve from any initial population. That is, in every case where CD-players can detect their own type better than pure chance, and the fear payoff difference is less than the sum of greed and cooperation payoff differences, any initial population will evolve – either directly through any payoff monotonic process (such as replicator dynamics), or combined with arbitrarily small perturbations from unstable equilibria – eventually to a *unique evolutionarily stable population* composed entirely of CD-players.

The adaptive capabilities just described hold independent of any argument about whether CD behavior is or is not ‘optimal’ according to a definition such as Bayesian rationality. They also hold for *pure* one-shot prisoners' dilemmas: meaning no repeated interactions or pairings in any generation are involved; no information or third party reports about past behavior are involved, all signal information arises only from symptoms detected after two randomly matched strangers meet for the first time; and no subjective preferences for altruism, fairness, reciprocity, or morality affect the raw evolutionary dynamics.⁴

Testable predictions are also derived in Part II that agree with experimental data built up since the prisoners' dilemma was first introduced in 1950. They describe how the CD-players' probability of cooperating changes: depending on the relative size of fear, greed, and cooperation payoff differences; and depending on the players' history of communication, especially when face-to-face discussion is involved.

I. SIGNAL DETECTION PRINCIPLES & ROC CURVES

ROC curves from the signal detection literature are drawn in the *unit square*, as shown in Figure 1a. The vertical axis shows the probability of detecting a signal (denoted x) indicating certain events (denoted Q) have happened instead of not happened (denoted ~Q); represented by the conditional probability, $r = p(x|Q)$. Alternatively, signal x may still arise when events ~Q have happened instead of Q; represented on the horizontal axis by the conditional probability, $w = p(x|\sim Q)$. So the letters (r, w) represent the probability of signal x ‘*rightly*’ versus ‘*wrongly*’ indicating that events Q instead of ~Q have happened.

Figures 1a & 1b About Here

A wide variety of ROC experiments have been conducted: including tasks like detecting a signal masked by background noise; detecting cancer tissue with x-ray or MRI pictures; or observing speech and body language to detect another person's preferences or beliefs, including lie detection. Across these different tasks, experiments during the last 40 years have shown that both human and animal detection behavior conforms to basic patterns shown by the ROC curves in Figure 1a.⁵ So let us describe how these curves are generated.

⁴ So no subjective factors are involved as for example analyzed by, Trivers 1971; Geanakoplos, Pearce, & Stacchetti 1989; Rabin 1993; Falk & Fischbacher 1998; Fehr & Schmidt 1999; Bolton & Ockenfels 2000; and Gintis 2000.

⁵ See footnote 3; plus Lusted 1968; Ben-Shakhar et. al. 1986; Getty et. al. 1988.

A. *Dividing Symptoms Into Signal Categories With Overlapping Density Functions*

Figure 1b shows two density functions: one generated by events Q happening, and the other generated by complementary events $\sim Q$ happening. The variable ‘ z ’ on the horizontal axis represents a continuous set of detectable symptoms whose likelihood depends on the two density functions caused by events Q versus $\sim Q$.

Notice particular z -symptoms can arise under either density function, and so cannot be uniquely associated with events Q or $\sim Q$ happening. But the z -symptoms can be partitioned into two disjoint subsets separated by a boundary in Figure 1b, denoted by the letter ‘ b ’. These two subsets can then be categorized as two different signals: where signal x versus y represents the subset to the *left* versus *right* of the boundary. Signal x is the subset whose area under the density function generated by Q is at least as great as the area generated by $\sim Q$. In Figure 1b, this subset extends to the *left* of b . The areas under the two density functions to the left of boundary b equal the two signal probabilities along an ROC curve in Figure 1a: $r = p(x|Q) \geq w = p(x|\sim Q)$.

In this paper, we are interested in ROC curves for a person detecting symptoms related to the *internal motivations* of another person. Such motivations include standard economic factors such as a player’s subjective preferences, beliefs, and an evaluation criterion for combining preferences and beliefs, like expected utility. Appendix A describes an example involving ‘*feedback questioning*’ that generates both verbal and body-language symptoms that are causally linked to the internal motivations of another person.⁶

B. *Other Game Theory Models With Signaling Between Players*

The above signal detection environment is quite different than that used to analyze “cheap talk” communication between players; by Wärneryd (1993), Kim & Sobel (1995), Bhaskar (1998), and Banerjee & Weibull (2000). For example, Banerjee & Weibull assume each player can choose a particular signal to be sent to another player, who is also assumed to perfectly identify that signal from other signals that might be sent.

However, the above signaling environment implies it is impossible for a player to target particular symptoms for sending to another player: because a player’s internal motivations are causally linked to a density function spread over a range of detectable symptoms. So there exist no specific symptoms – such as a “green beard”, “secret handshake”, or particular spoken words – that are uniquely associated with a given type of internal motivations. Moreover, particular symptoms can be categorized by the receiving player as either an x or y signal depending on their location relative the boundary b in Figure 1b.

These properties imply Adam and Eve cannot choose the specific symptoms that will be sent to each other, nor can they choose whether these symptoms will be categorized as an x or y signal by the receiving player. Adam and Eve thereby know the x or y signal received from each other, but not the signal each other receives from themselves; nor how each other will react to their received signal. As noted earlier, this type of signaling environment is the typical case in experiments during the last 40 years over a wide variety of detection tasks. So let us further describe the ROC curves in Figure 1a.⁷

C. *ROC Curves Ranging From Pure Chance To Perfect Detection Skill*

The *diagonal* line in Figure 1a represents feasible signal probabilities resulting from pure chance detection. This happens when the two density functions in Figure 1b coincide exactly: so the signal probabilities (r, w) always equal each other no matter where the boundary between the x and y signals is located.⁸

If the two density functions in Figure 1b do not coincide, then shifting the boundary b to the right will raise

⁶ Both human and animal subjects in experiments typically are not aware of the concepts and variables described above; such as the signal probabilities (r, w) and ROC curves in Figure 1a, or the density functions and boundary point b in Figure 1b. Instead, they may be only aware of the relative frequency of events Q versus $\sim Q$, or the costs/benefits of correct versus mistaken detections. So subjects do not need to know the theoretical elements in Figures 1(a, b) to behave according to an ROC curve.

⁷ The cheap talk models just discussed ignore set up costs of meeting and talking to receive signals between two players. If we similarly ignore set up costs of meeting and talking to detect symptoms described by an ROC curve, then such detection has no effect on the mapping from chosen actions into received payoffs. Such detection is then likewise a form of cheap talk. However, it is quite different than the cheap talk assumed in the models just discussed.

⁸ An example would be trying to detect another person’s internal motivations with an externally randomized signal, such as rolling a multi-sided die. The (r, w) probabilities from such signals are necessarily *equal*, and thus on the diagonal line in Figure 1a.

the r probability initially faster than w . But the area under the density function generating the r probability will at some point rise slower than the other area generating the w probability: causing (r, w) to move along a *convex* path above the diagonal line in Figure 1a; called an ROC curve in the signal detection literature.

Perfect detection represents the limit where the ROC curve coincides with the left and top sides of the unit square; when the two density functions in Figure 1b have no overlap at all. So the probability of rightly detecting another person's internal motivations can be raised all the way to 1 without producing any positive probability of wrongly doing so – located at the top-left corner of the unit square where $(r, w) = (1, 0)$.

The ROC curves in Figure 1a are referred to as NN-ROC curves. The letters 'NN' indicate detection errors arise from two *normal* density functions in Figure 1b. Numerous experiments have shown NN-ROC curves can explain observed detection behavior across a wide range of tasks; including auditory & visual signals, body language signals, pattern recognition in medical & psychological diagnosis, and so on.⁹ So normally distributed errors are assumed in detecting symptoms related to the internal motivations of another player.

D. *Properties of NN-ROC Curves Needed For Later Analysis*

The signal categories x and y are hereafter referred to as simply signals x and y , in order to simplify subsequent discussion. This does not mean players are dealing with only two signals from each other. Rather, each player must divide a continuous space of detectable symptoms into two disjoint subsets in an attempt to determine which of two overlapping density functions generated the specific symptoms received from another player.

Now consider another key principle: that both human and animal subjects can vary the boundary between the x and y signals in Figure 1b – thereby causing their signal probabilities (r, w) to shift between $(0, 0)$ and $(1, 1)$ along an ROC curve. This pattern has been observed across a variety of different detection tasks, whenever experiments vary the benefits and costs of making false positive or false negative detection mistakes.¹⁰

Note also from the definitions of r and w , that shifting (r, w) closer to $(0, 0)$ means signal x is less likely to be detected regardless of whether it rightly or wrongly indicates another player's type of internal motivations. A player is thus “*more cautiously*” detecting when its pair of signal probabilities (r, w) shifts closer to $(0, 0)$.

Cautious detection implies another property for convex NN-ROC curves (beyond pure chance detection): shifting (r, w) closer to $(0, 0)$ will cause the odds favoring rightly instead of wrongly detecting another player's motivation type to rise arbitrarily high – thereby causing the ratio of signal probabilities r/w to rise toward infinity as (r, w) drops to $(0, 0)$. So we have the following property:

$$\frac{r}{w} \rightarrow \infty \quad \text{as} \quad (r, w) \rightarrow (0, 0) \quad (1)$$

Note that even though the ratio of rightly to wrongly detecting rises arbitrarily high in statement (1), this property does *not* involve any extreme or unusual situation, nor any special signal like a ‘secret handshake’. Instead, it is the ordinary result of cautious detection in numerous experiments with both human and animal subjects; achievable with no unusual skill beyond pure chance detection – by simply shifting (r, w) along a convex NN-ROC curve that bows above the diagonal line in Figure 1a.

One reason for describing ROC curves and principles from the signal detection literature is to show numerous experiments already demonstrate the signal detection properties used in later analysis are the typical result of ordinary people in ordinary circumstances; thereby achievable without any special detection skills nor any special signals like secret handshakes, and so on.

II. APPLICATION TO PRISONERS' DILEMMAS WITH CONTINGENT PARTICIPATION

Consider Axelrod's (1984) analysis of prisoners' dilemmas with repeated play. Appendix B also analyzes one-shot play; including Frank's (1988) analysis. To begin, assume the population includes two or more different types of players: *AllD*–players who defect in every successive round with their partners; and *tit-for-tat* (*TFT*) players who cooperate in the first round, and then do whatever their partner did in the previous round.

Axelrod wanted to show *TFT* players could evolve even within a hostile environment composed mostly of *AllD*–players. So he assumed a form of non-random interaction (1984, chapter 3) in which the *TFT*–players arrive in “clusters”. An example might be kin relationships, where parents propagate children whose local environment

⁹ See footnote 5; plus Saxe 1985; Swets 1973, 1988; and Swets & Pickett 1982.

¹⁰ See for example Swets 1988 article in *Science*; plus references in footnote 9.

contains mostly near genetic relatives.

However, Axelrod did not model how clustering occurs. Interpretations about kin selection also give no basis for explaining how cooperation could evolve in prisoners' dilemmas with partners *not* linked by a family bond. So let us apply signal detection analysis to the repeated play environment studied by Axelrod.

To do so, let $H(\text{AllD})$ and $H(\text{TFT})$ denote histories that will cause a player such as Eve to be programmed by her internal motivations to behave according to AllD or TFT. So events Q versus $\sim Q$ in Figures 1a and 1b are now interpreted as histories $H(\text{TFT})$ versus $H(\text{AllD})$: where these two types of histories may also affect the likelihood of a player such as Adam detecting a signal x instead of y from his partner Eve. So the r and w signal probabilities in Figure 1a now become:

$$r = p[x | Q = H(\text{TFT})] \quad \& \quad w = p[x | \sim Q = H(\text{AllD})] \quad (2)$$

Assume TFT players *contingently participate*: meaning they will start playing only if they detect signal x from their partner; otherwise, they will refuse to start playing. Such players are denoted TFT_{CP} ; where the subscript 'CP' means '*contingently participating*'. Alternatively, the AllD-players will always (non-contingently) start repeated play with whomever they are matched – provided the other player chooses to participate – called AllD_{NP} -players; where the subscript 'NP' means '*non-contingently participating*'.

Let ' c ' denote the cost of engaging in each successive matchup in search of a partner from whom the x signal is detected, where $c \geq 0$ holds. Also let $0 \leq \delta \leq 1$ denote the per period discount factor for valuing future payoffs, including delays from refusing to play with one's current partner. So a delay of one period from not playing one's current partner results in one period discounted cost of engaging in the next matchup, $\delta c \geq 0$

We can then study the limiting case most favorable to TFT_{CP} -players against AllD_{NP} -players, depending on their population frequencies, denoted ξ and $1 - \xi$: namely, when the TFT_{CP} -players perfectly detect whether their current partner is another TFT_{CP} -player versus an AllD_{NP} -player.

To analyze this case, we must calculate the difference between the two players' expected payoffs from playing against themselves and each other, denoted $E(\text{TFT}_{\text{CP}})$ and $E(\text{AllD}_{\text{NP}})$, when perfect detection holds, so that $(r, w) = (1, 0)$. The details are shown in Part 2 of Appendix B; which yields the following results :

$$(r, w) = (1, 0) \text{ implies } E(\text{AllD}_{\text{NP}}) - E(\text{TFT}_{\text{CP}}) = \frac{(1-x)\left(\frac{P}{1-d}\right) - xdc}{1-dx} - \frac{x\left(\frac{R}{1-d}\right) - (1-x)dc}{1-d(1-x)} \quad (3a)$$

So as ξ drops to 0, we obtain the following inequality :

$$E(\text{AllD}_{\text{NP}}) - E(\text{TFT}_{\text{CP}}) = \frac{P}{1-d} + \frac{dc}{1-d} > 0 \quad (3b)$$

The positive difference in the above inequality implies the evolutionary success of the TFT_{CP} -players against AllD_{NP} -players is not robust to beginning with a small frequency in the total population, regardless of how skillfully they might distinguish themselves from AllD_{NP} -players.

Next generalize this result by considering any player who will '*sometimes cooperate*' at least once in a prisoners' dilemma, and who contingently participates by detecting an x or y signal from each partner; denoted SC_{CP} . SC_{CP} thus includes TFT_{CP} , plus a variety of other examples; such as "*tit-for-two-tats*", "*testor*", or, "*retaliator*" from Axelrod's computer tournaments (1984, chapter 2) .

Alternatively, we can consider any player who will cooperate at least once with their partners, but who will also exit after n rounds with an AllD_{NP} -player. These players are denoted Ex_n -players; where n denotes the stage at which repeated defection will cause such a '*contingent-exitor*' to stop playing. An example of Ex_n for $n = 1$, is the '*prudent moral*' strategy studied by Vanberg & Congleton 1992, 2001; that starts play by cooperating against any partner, and then exits after the first defection by its partner.

Parts 2 & 3 of Appendix B show a similar relationship holds as (3a, b) above; when either SC_{CP} or Ex_n competes in a population composed mostly of AllD_{NP} . That is, it doesn't matter how SC_{CP} -players behave once they start playing, provided they will cooperate at least once against an AllD_{NP} -player. Likewise, it doesn't matter how the Ex_n -players' decide to exit, provided they will cooperate a least once with an AllD_{NP} -player before exiting. All potential types of SC_{CP} or Ex_n players are vulnerable to their share in a population of mostly AllD_{NP} -players.

Notice the definitions of (AllD_{NP} , SC_{CP} , Ex_n) players imply they will choose respectively (D, C, C) in one-shot games, where $n = 1$. So the above results hold for both one-shot and repeated games, as follows :

THEOREM 1 (Vulnerability of SC_{CP} and Ex_n Players To Their Share In The Population)

Any discounting of future payoffs ($\delta < 1$) along with any nonnegative search cost ($c \geq 0$) in one-shot or repeated prisoners' dilemmas implies any type of SC_{CP} or Ex_n players are inherently vulnerable to their share in a population with mostly of $AllD_{NP}$ -players: no matter how skillful SC_{CP} or Ex_n players might be at avoiding or exiting participation with $AllD_{NP}$ -players.

III. REPLACING CONTINGENT-PARTICIPATION WITH CONTINGENT-COOPERATION

Given the vulnerability of contingent participators to their share in the population, one might expect that cooperative behavior would be even more vulnerable if no exit from playing with always defecting players was possible. However, there is a form of contingent cooperation in one-shot prisoners' dilemmas – not in terms of whether to participate, but rather in terms of *how* to play any given partner – whose survival does not depend on its share in the population, nor on its ability to distinguish itself from other types of behavior. The next objective is to describe this type of contingently cooperative behavior, and show why the above statement holds.

A. Feasible Strategies & Signal Detection Probabilities

Start by representing strategies as mappings from detected signals into feasible actions. There are four non-random mappings represented by 1st and 2nd letters for the C or D action chosen in response to detecting signal x versus y: CC, meaning to *always cooperate* regardless of signal x or y; CD, meaning to *contingently-cooperate* if and only if signal x instead of y is detected from the other player; DC, meaning the opposite strategy of *contingently-defecting* if and only if signal x instead of y is detected; and DD, meaning to *always defect* regardless of signal x or y. Each player can also choose a *mixed strategy*, denoted $C\lambda D$ for $\lambda \in (0, 1)$; that randomly chooses C or D independent of signal x or y, with probability $\lambda = p(C)$ versus $1 - \lambda = p(D)$.

So there are five types of signal response strategies, (CC, CD, DC, DD, $C\lambda D$). The four strategies besides CD are denoted $\sim CD = \{CC, DC, DD, C\lambda D\}$. The Greek letter ξ is used with a subscript to represent each type's frequency in the whole population; and the population frequency profile is denoted with no subscript, $\xi = (\xi_{CC}, \xi_{CD}, \xi_{DC}, \xi_{DD}, \xi_{C\lambda D})$. Let $H(CD)$ and $H(\sim CD)$ denote histories that will cause a player such as Eve to be programmed either to contingently cooperate CD, to behave according to one of the other signal-response strategies $\sim CD$.

Next analyze a player such as Adam detecting an x or y signal causally influenced by the type of histories $H(CD)$ versus $H(\sim CD)$ also influencing his partner Eve's behavior. Accordingly, the Q versus $\sim Q$ events in Figure 1b now represent histories $H(CD)$ versus $H(\sim CD)$, and the r and w signal probabilities in Figure 1a now become:

$$r = p[x | Q = H(CD)] \quad \& \quad w = p[x | \sim Q = H(\sim CD)] \quad (4)$$

So r versus w indicates the probability of signal x *rightly* versus *wrongly* indicating that Adam's partner Eve has been programmed by past events to contingently react to her signal from him according to strategy CD, instead of being programmed to react differently by $H(\sim CD)$.

Recall that neither player can choose the specific symptoms detected by the other player. Instead, the likelihood of each player detecting specific symptoms from each other is governed by a pair of overlapping density functions whose shape and position relative to each other are caused by the type of internal motivations resulting from histories $H(CD)$ versus $H(\sim CD)$.

Also recall CD-players cooperate if and only if they detect a signal positively correlated with their own type (signal x instead of y); while DC-players follow the reverse strategy. Since both players react contingently, they both have independently chosen (r, w) signal probabilities; subscripted by 'CD' and 'DC' to distinguish them.

B. Linear Versus Quadratic Expected Payoffs

Table 1 shows the $5^2 = 25$ expected payoff formulas resulting from each player competing against its own type, or one of the other four types. Notice the cells in Table 1 may contain a 3rd line representing a *quadratic* payoff term; meaning an expression with a multiple of two signal or action probabilities, such as $(r_{CD})^2$, λ^2 , or λw_{CD} . In all such cases, the probability multiples are themselves multiplied by the difference between the fear and greed payoff differences, $(P - S) - (T - R)$. So quadratic payoff terms arise only when players' fear and greed payoff differences are *not* equal – otherwise expected payoffs are *linear* instead of quadratic.

Table 1 & Figure 2a About Here

IV. DD BEHAVIOR IS NOT DOMINANT & CD BEHAVIOR IS EVOLUTIONARILY STABLE

Figure 2a shows an example with linear expected payoffs over the subpopulation with only CD and DD players, called the CD/DD subpopulation. Two expected payoff lines are shown, which depend on the frequency of CD-players, denoted ξ_{CD} . These lines are weighted averages of expected payoff formulas in the CD and DD columns of Table 1. Note that linear expected payoffs simplifies the quadratic expected payoff for two CD-players against each other, to the formula $P + r_{CD}(R - P)$; shown on the right hand vertical scale where $\xi_{CD} = 1$.

Notice also the interior *crossover threshold* where the CD and DD players' expected payoff lines intersect, denoted $\xi_{CD/DD}^0$. CD-players will outperform DD-players if and only if their population frequency ξ_{CD} exceeds the crossover threshold. Such an interior crossover point implies CD-players do better against themselves than DD-players do against them; as implied by the ranking of the two expected payoff formulas on the right-hand vertical scale of Figure 2a, $P + r_{CD}(R - P) > P + w_{CD}(T - P)$. This in turn implies DD behavior is *not* a dominant strategy in the CD/DD subpopulation.

This same conclusion applies to the whole population. To see this, compare the expected payoff formulas in the (CC, DD, C λ D) columns of Table 1. Doing so implies the DD-players' expected payoff always exceeds that of CC players, and the C λ D players for any $0 < \lambda < 1$. Thus, DD behavior is dominant within the CC/DD/C λ D subpopulation; as expected from traditional analysis.

However, DD behavior is not dominant when players' strategies also include the two signal contingent strategies (CD, DC). Instead, CD behavior can be a *strict* Nash equilibrium. This follows from the CD column in Table 1: by noticing a CD-player's expected payoff against its own type *strictly* exceeds the expected payoffs of the other four types against CD-players, when their signal probabilities (r_{CD} , w_{CD}) are near (1, 0). This follows because at $(r_{CD}, w_{CD}) = (1, 0)$, the expected payoffs for (CC, CD, DC, DD, C λ D) against CD-players equal respectively :

$$[S; \quad R; \quad S + r_{DC}(P - S) \leq P; \quad P; \quad P - \lambda(P - S) < P] \quad (5)$$

Since payoff R strictly exceeds payoffs P and S, the above profile of expected payoffs immediately implies an exclusive population of only CD players is *evolutionarily stable* whenever the signal probabilities (r_{CD} , w_{CD}) are sufficiently close to (1, 0). However, these expected payoff comparisons might require the CD-players' to have near perfect detection skill in order for them to shift (r_{CD}, w_{CD}) near (1, 0).

But a much stronger result holds: *CD behavior is guaranteed to be evolutionarily stable whenever the CD-players cautiously detect their own type by shifting their signal probabilities (r_{CD} , w_{CD}) sufficiently close to (0, 0) along a convex NN-ROC curve – that may be arbitrarily close to the diagonal line in Figure 1a corresponding to pure chance detection.* This result holds for linear expected payoffs, and for quadratic expected payoffs, provided the fear payoff difference is less than the sum of greed and cooperation payoff differences. It is formalized next, and proved in Appendix C.

THEOREM 2 (*Evolutionary Stability Of CD-Players With Any Detection Skill Beyond Pure Chance*)

Let the CD-players detect according to a convex NN-ROC curve that bows above the diagonal line in Figure 1a, and assume $(P - S) < (T - R) + (R - P)$. Then there exists a positive interval extending from (0, 0), such that any signal probabilities (r_{CD} , w_{CD}) along this interval imply: DD behavior is not dominant over the remaining population with (CC, CD, DC, C λ D) players; and a population with only CD-players is evolutionarily stable against invasion by other players, including DD-players.

Consider Ken Binmore's comment (1994, page 173): *The game theoretic equivalent of squaring the circle consists of justifying the use of a strongly dominated strategy in the one-shot Prisoners' Dilemma.* In other words, suppose an analysis of one-shot prisoners' dilemmas is offered that implies always defect behavior strictly dominates other possible behavior, while also showing other possible behavior exists that can still invade always defect behavior. As Binmore suggests, such analysis would be just as impossible as successfully squaring the circle.

Notice Theorem 2 immediately bypasses such impossibility: because it implies CD-players cautiously detecting their own type will prevent DD behavior from being dominant over all other types of signal-response behavior. And once a unique dominant strategy is no longer possible, then other Nash equilibria may exist. Moreover, if such a Nash equilibrium is strict, then it will also be evolutionarily stable; as implied for CD behavior.

Albert Tucker devised the original prisoners' dilemma (in 1950) partly as a counter example to his graduate student, John Nash's equilibrium concept for non-zero sum games. Yet the Nash equilibrium has subsequently come to be almost universally used in strategic analysis despite mostly falsified predictions in one-shot prisoners' dilemma

experiments. The stable Nash equilibrium shown in Theorem 2 thus vindicates the Nash equilibrium's ability to explain behavior even in one-shot prisoners' dilemmas; especially if CD behavior implies testable predictions about the frequency of cooperation that agree with a large body of data from one-shot and finitely-repeated prisoners' dilemma experiments published since 1950. Such predictions are developed in Part II of this paper.

V. CONCLUSION

Part I of this paper used ROC curves in an evolutionary analysis of prisoners' dilemmas. Doing so implies the following: if there is any discounting of future payoffs, or any cost of searching for an additional partner, then cooperative players who contingently participate – in terms of who to play with or when to exit – cannot survive in a population containing mostly players who always defect. This vulnerability holds even when contingent participants only interact with themselves by perfectly detecting their own type.

However, quite different results hold if players act contingently; not in terms of whether to participate, but rather in terms of how to act with any given partner. There is a form of cooperation in one-shot prisoners dilemmas that reacts contingently to symptoms detected from its partner (called CD behavior): where such behavior is a *strict* Nash equilibrium plus evolutionarily stable against invasion by other types of signal-response behavior, including always defect DD behavior; provided CD-players can detect their own type better than pure chance, and the fear payoff difference is less than the sum of both greed and cooperation payoff differences.

These implications do not rule out other stable equilibria, say with mostly DD-players. So Part II of this paper shows no other stable equilibria exist: because CD behavior has robust adaptive capabilities guaranteeing it will evolve in one-shot prisoners dilemmas from any initial population through any payoff monotonic process (such as replicator dynamics). The reader is invited to continue with Part II: to see why robust evolution of CD behavior must happen, and why it implies certain testable predictions, as noted at the end of the last section.

APPENDIX A (*Detection Through Feedback Questioning*)

Suppose Adam and Eve are involved in the original prisoner's dilemma story, and they talk with each other about an upcoming crime. Adam asks questions that might prompt Eve to respond in a way that generates symptoms related to her type of internal motivations; as indicated by the z-axis in Figure 1b. Since Adam asks Eve questions in order to get Eve to respond to them as feedback, we can say Adam is involved in "*feedback questioning*", and similarly call the symptoms he receives, "*feedback symptoms*".

Feedback symptoms might be "*endogenous*", meaning they are *causally related* to the other player's internal motivations. Such symptoms may thus be relevant for Adam and Eve to detect in a way that "*exogenous*" symptoms are not; meaning symptoms obtained through an external process, such as signals generated from a bingo cage: because the latter signals have nothing to do with whatever might cause a player to respond to them in a particular way. Hence the reason for calling such signals *exogenous* or *external* in the first place.

Eve must determine which endogenous symptoms she will respond to by choosing action C instead of D. Statistical decision theory implies she must divide the range of possible symptoms (spread across the z-axis in Figure 1b) into two disjoint subsets. One subset is categorized as a "*favorable*" signal 'x' and the other subset is categorized as an "*unfavorable*" signal 'y' — so that particular symptoms categorized as favorable instead of unfavorable can be mapped into choosing action C instead of D respectively.

For example, suppose Adam intends to always defect DD but tries to fool Eve by saying he will cooperate if he detects a favorable x signal of her willingness to respond in a similar contingent manner to feedback symptoms from himself. However, Eve is interested in the *combined package* of verbal and non-verbal symptoms resulting from Adam responding to her questions: because the likelihood of such combined feedback symptoms is causally related to Adam's type of internal motivations.

So the reason for feedback questioning is to strengthen causal links to Adam's internal motivations through detecting combined verbal-plus-body-language symptoms. That is why close-range, "*face-to-face*" interaction is required to have the greatest access to the causal links involved. Other ways of communicating such as with

telephones, video conferencing, and email will be noticeably less effective: because they restrict the ability to receive combined symptoms otherwise observable in Adam's direct presence; thereby degrading the causal links required to receive endogenous symptoms that more reliably indicate Adam's internal motivations.

It is also likely that both Adam and Eve will take for granted that feedback questioning is imperfect at eliciting symptoms related to his or her future strategy decisions. Such mistakes arise because the same symptoms can be generated from different overlapping density functions caused by different types of internal motivations that might be influencing how Adam reacts to Eve's questions.

Nevertheless, suppose Adam could train himself to maintain completely steady non-verbal symptoms of all types. For example, suppose Adam was involved in a game of poker, and trained himself to display a steady facial expression (a "poker face") intended to prevent other players from detecting whether he will bluff or not if dealt a poor hand, and so on. So other players such as Eve may interpret Adam's poker face as a detectable symptom that trying to forecast his betting strategy is especially uncertain, and thereby fruitless to attempt in the first place.

Accordingly, otherwise "uninformative" (poker face) symptoms can themselves be interpreted by Eve as an unfavorable signal (y instead of x) that may cause Eve to defect instead of cooperate with Adam. So regardless of whether Adam tries to control his verbal and non-verbal feedback or not, there always exists symptoms that are relevant from Eve's point of view, in order to justify defecting instead of cooperating with Adam.

APPENDIX B (*Proofs for Contingent Participation With One-Shot Or Repeated Play*)

PART 1 (*Proofs For One-Shot Prisoners' Dilemmas*)

First consider one-shot prisoners' dilemmas, and let g represent the probability that a Cooperator who *contingently participates*, denoted C_{CP} , is able to play its current partner; where a C_{CP} player will cooperate if and only if it detects the x signal instead of the y signal from its partner. Let $r = p[x|H(C)]$ and $w = p[x|H(D)]$ denote the probabilities of the x signal either 'rightly' or 'wrongly' indicating that one's partner has been programmed by a history of past events to choose C instead of D .

The other type of player is a Defector who *non-contingently participates*, denoted D_{NP} . The frequencies ξ versus $1 - \xi$ determine the probability of being matched with a C_{CP} -player versus D_{NP} -player. Finally recall that ' δ ' denotes the factor used to discount future payoffs; and ' c ' denotes the cost searching for an additional partner.¹¹

Now let $\alpha = p(x)$ denote the probability of a C_{CP} -player detecting the x signal, either rightly or wrongly, depending on whether its partner is a C_{CP} -player or a D_{NP} -player; so that,

$$\alpha = p(x) = \xi r + (1 - \xi)w \quad (B1a)$$

Then let the posterior probability of one's partner being a C_{CP} -player, conditional on detecting the x signal from that

$$\text{partner, be denoted, } \beta = p(C_{CP} | x) = \frac{\alpha r}{\alpha r + (1 - \alpha)w} \quad (B1b)$$

The complementary chance that one's partner is a D_{NP} -player instead of C_{CP} -player after having detected the x signal equals, $1 - \beta = p(D_{NP} | x)$. There are three requirements for a C_{CP} -player to end up actually playing its current partner: first it must detect the x signal with probability α (otherwise it will refuse to play); and then its current partner must also be willing to play, either if it is a non-contingently participating D_{NP} -player (encountered with probability with probability $1 - \beta$ after detecting the x signal), or if it is a C_{CP} -player who has also detected an x signal.

Consequently, the probability g that a C_{CP} -player is able to play its current partner satisfies,

$$g = \alpha[(1 - \beta) + \beta r] = \alpha[1 - \beta(1 - r)] = p[C\text{-player plays its current partner}] \quad (B2)$$

¹¹ This is the formulation considered by Frank (1988, page 60). However, he does not allow discounting of future payoffs; nor the possibility of a search cost resulting from C_{CP} -players encountering new partners until they detect signal x instead of y from their current partner. Frank eliminates such costs by assuming players can immediately switch to a "work alone" option (1988, page 261); allowing them to survive with no search or delay costs from otherwise not playing until they encounter a partner from whom the x signal is detected. This assumption denies the possibility that players cannot survive indefinitely in isolation from others, but instead must search out partners and actually play with them in order to survive.

A C_{CP} -player's expected payoff conditional on being able to play its current partner is denoted $E^0(C_{CP})$, and is computed as follows:

$$E^0(C_{CP}) = \frac{a b r}{g} \bullet E(C_{CP} / C_{CP}) + \frac{a(1-b)}{g} \bullet E(C_{CP} / D_{NP}) \quad (B3a)$$

$$= \frac{a b r}{g} \bullet R + \frac{a(1-b)}{g} \bullet S = \frac{b r}{1-b(1-r)} \bullet R + \frac{1-b}{1-b(1-r)} \bullet S \quad (B3b)$$

At the start of each matchup, a C_{CP} -player has a probability g of actually playing its current partner: if it detects the x signal from its partner, and its partner is also willing to play. If so, then it will receive its expected payoff contingent on actually playing, $E^0(C_{CP})$. Alternatively, it has a probability $(1-g)$ of not playing its current partner, because it doesn't detect the x signal and/or its partner is not willing to play.

If the latter happens during the current period, then the C_{CP} -player must search for a new partner in the next period, and thereby pay a discounted cost of δc in order to restart the whole process with a new randomly matched partner in the next period – where its total expected payoff at that point is the same as it was at the start of the last matchup. So a C_{CP} -player's expected payoff, denoted $E(C_{CP})$, satisfies the following relationship that recursively determines its value depending on δ , g , and $E^0(C_{CP})$.

$$E(C_{CP}) = g E^0(C_{CP}) + (1-g) [\delta E(C_{CP}) - \delta c] \quad (B4a)$$

Then rearrange terms to solve for $E(C_{CP})$, obtaining:

$$E(C_{CP}) = \frac{g E^0(C_{CP}) - (1-g) \delta c}{1 - \delta(1-g)} \quad (B4b)$$

Next use a similar procedure to determine a D_{NP} -player's expected payoff depending on its expected payoff if it is able to actually play its current partner, and on the probability of playing or not playing with its current partner. The above definitions imply the probability of a D_{NP} -player being able to actually play its current partner, denoted q , depends on the chance of either being matched with another D_{NP} -player (who will always non-contingently play), or being matched with a C_{CP} -player who also detects the x signal, as follows:

$$q = p(\text{D-player plays its current partner}) = (1 - \xi) + \xi w = 1 - \xi(1 - w) \quad (B5a)$$

The expected payoff of a D_{NP} -player conditional on it actually playing its current partner then becomes,

$$E^0(D_{NP}) = E(D_{NP} | \text{plays its current partner}) = \frac{1-x}{q} \bullet E(D_{NP} / D_{NP}) + \frac{xw}{q} \bullet E(D_{NP} / C_{CP}) \quad (B5b)$$

$$= \frac{1-x}{q} \bullet P + \frac{xw}{q} \bullet T \quad (B5c)$$

Similar to (B4b), a D_{NP} -player's expected payoff across all future possibilities where it may or may not be able to play its current partner, then becomes:

$$E(D_{NP}) = \frac{q E^0(D_{NP}) - (1-q) \delta c}{1 - \delta(1-q)} \quad (B6)$$

By subtracting the formulas (B4b) and (B6) for $E(C_{CP})$ and $E(D_{NP})$, and using the formulas (B2) and (B5a) for g and q , the difference $E(C_{CP}) - E(D_{NP})$ becomes:

$$E(D_{NP}) - E(C_{CP}) = \frac{[1-x(1-w)]E^0(D_{NP}) - x(1-w)\delta c}{1 - \delta x(1-w)} - \frac{a[1-b(1-r)]E^0(C_{CP}) - (1-a[1-b(1-r)])\delta c}{1 - \delta(1-a[1-b(1-r)])} \quad (B7a)$$

Notice that as the C_{CP} -players' signal probabilities converge to perfect detection, $(r, w) \rightarrow (1, 0)$, then $E^0(C_{CP}) = R$ and $E^0(D_{NP}) = P$. And so, using the formulas (B1a–b) for α and β , (B7a) converges to the next difference

formula, as the C_{CP} -players' signal probabilities converge to perfect detection, $(r, w) = (1, 0)$:

$$E(D_{NP}) - E(C_{CP}) = \frac{(1-x)P - xd c}{1-dx} - \frac{xR - (1-x)d c}{1-d(1-x)} \quad (B7b)$$

Next take the limit of the above difference as ξ drops to 0, to obtain the following inequality:

$$E(D_{NP}) - E(C_{CP}) = P + \frac{dc}{1-d} > 0 \quad (B7c)$$

Inequality (B7c) is strictly positive for all nonnegative $c \geq 0$, and for all nonnegative discount factors up to the limit where the denominator in (B7c) is well defined (all $0 \leq \delta < 1$). This implies the evolutionary success of the C_{CP} -players against D_{NP} -players is not robust to beginning with a small frequency in the total population, regardless of how skillfully they might distinguish themselves from D_{NP} -players, and regardless of how small the cost of being matched up with a new partner might be.

However, a contrary result follows if we consider a special case assumed by Frank (1988, page 60): in which the search cost for an additional partner c is set equal to 0, and no discounting of future payoffs takes place (so that $\delta = 1$) before allowing ξ to converge to 0 in (B7b). If these limiting assumptions are imposed before ξ drops to 0, along with assuming perfect detection, $(r, w) = (1, 0)$; then the difference formula (B7b) becomes,

$$E(D_{NP}) - E(C_{CP}) = \frac{(1-x)P}{1-x} - \frac{xR}{x} = P - R < 0 \quad \text{for all } 0 < \xi \leq 1 \quad (B7d)$$

Notice the reversal to a negative difference, $E(D_{NP}) - E(C_{CP}) < 0$. This reversal implies C_{CP} -players will outperform D_{NP} -players regardless of their own population frequency ξ from 0 to 1. However, this result is misleading because it only holds at the limit where both $c = 0$ and $\delta = 1$; but not near the limit: because we just showed the reverse positive difference must hold for any values $c \geq 0$ and/or $\delta < 1$, as ξ drops to 0.

PART 2 (Proofs for Repeated Play With TFT or Sometimes Cooperating, SC-Players)

The repeated play analysis involved TFT_{CP} and $AllD_{NP}$ players, as defined in Section II of the main text. It is similar to the analysis just presented in Part 1, except for the formulas (B3a-b) and (B5c-d) used to calculate expected payoffs from actually playing a current partner. This is because choosing to play one's current partner will start an indefinitely repeated game with either TFT_{CP} or $AllD_{NP}$, instead of only one round of choosing C or D actions. So, recalling the one-period discount parameter is denoted δ , we have the following formulas.

$$E(TFT_{CP} | TFT_{CP}) = R/(1-\delta) \quad E(TFT_{CP} | AllD_{NP}) = S + \delta P/(1-\delta) \quad (B8a)$$

$$E(AllD_{NP} | TFT_{CP}) = T + \delta P/(1-\delta) \quad E(AllD_{NP} | AllD_{NP}) = P/(1-\delta) \quad (B8b)$$

Then substituting these formulas in place of the expected payoff formulas for C_{CP} and D_{NP} players in equations (B3a-b) and (B5c-d), gives the following:

$$\begin{aligned} E^0(TFT_{CP}) &= E(TFT_{CP} | \text{plays its current partner}) \\ &= \frac{a b r}{g} \bullet E(TFT_{CP} / TFT_{CP}) + \frac{a(1-b)}{g} \bullet E(TFT_{CP} / AllD_{NP}) \end{aligned} \quad (B9a)$$

$$= \frac{b r}{1-b(1-r)} \bullet \frac{R}{1-d} + \frac{1-b}{1-b(1-r)} \bullet \left(S + \frac{d P}{1-d} \right) \quad (B9b)$$

$$\begin{aligned} E^0(AllD_{NP}) &= E(AllD_{NP} | \text{plays its current partner}) \\ &= \frac{x w}{q} \bullet E(AllD_{NP} / TFT_{CP}) + \frac{1-x}{q} \bullet E(AllD_{NP} / AllD_{NP}) \end{aligned} \quad (B9c)$$

$$= \frac{x w}{q} \bullet \left(T + \frac{d P}{1-d} \right) + \frac{1-x}{q} \bullet \frac{P}{1-d} \quad (B9d)$$

Then substitute formulas (B9b) and (B9d) in place of the expected payoffs $E^0(C_{CP})$ and $E^0(D_{NP})$ used in equations (B4b) and (B6) above, to calculate the TFT_{CP} and $AllD_{NP}$ players' expected payoffs, denoted $E(TFT_{CP})$ and $E(AllD_{NP})$; and rearrange terms to obtain the following difference formula, similar to (B7a) above :

$$E(AllD_{NP}) - E(TFT_{CP}) = \frac{[1-x(1-w)]E^0(AllD_{NP}) - x(1-w)dc}{1-dx(1-w)} - \frac{a[1-b(1-r)]E^0(TFT_{CP}) - (1-a[1-b(1-r)])dc}{1-d(1-a[1-b(1-r)])} \quad (B10a)$$

So assuming perfect detection, $(r, w) = (1, 0)$, gives the same difference as statement (3a) of the main text,

$$E(AllD_{NP}) - E(TFT_{CP}) = \frac{(1-x)\left(\frac{P}{1-d}\right) - xdc}{1-dx} - \frac{x\left(\frac{R}{1-d}\right) - (1-x)dc}{1-d(1-x)} \quad (B10b)$$

So as ξ drops to 0, we obtain the following inequality :

$$E(AllD_{NP}) - E(TFT_{CP}) = \frac{P}{1-d} + \frac{dc}{1-d} > 0 \quad (B10c)$$

So as described in the main text, $AllD_{NP}$ -players are guaranteed to outperform TFT_{CP} -players as their population ξ frequency drops to 0, even in the limiting case where they can perfectly detect and thereby limit interaction to playing only with themselves; and even when the cost of searching for another partner c also drops to 0.

Next generalize the last result by substituting any repeated play strategy in place of TFT_{CP} ; that also contingently participates, and which '*sometimes cooperates*' at least once against $AllD_{NP}$ (denoted SC_{CP} in the main text) – provided the SC_{CP} -players' expected payoff conditional on starting to play a current partner, denoted $E^0(SC_{CP})$, is *finite* as ξ drops to 0. Again assuming $(r, w) = (1, 0)$, the following statements are essentially the same as (B10b-c):

$$E(AllD_{NP}) - E(SC_{CP}) = \frac{(1-x)\left(\frac{P}{1-d}\right) - xdc}{1-dx} - \frac{x E^0(SC_{CP}) - (1-x)dc}{1-d(1-x)} \quad (B11a)$$

$$\text{where (B11a) converges to: } \frac{P}{1-d} + \frac{dc}{1-d} > 0 \quad \text{as } \xi \rightarrow 0 \quad (B11b)$$

So we have the same conclusion; namely,

THEOREM B1

Any discounting of future payoffs ($0 \leq \delta < 1$) implies any type of SC_{CP} player in one-shot or repeated games is vulnerable to its share in a population composed mostly of $AllD_{NP}$ players; where this vulnerability holds no matter how skillful a SC_{CP} player might be at avoiding participation with $AllD_{NP}$ players, and even when there is a zero cost of searching for another partner ($c = 0$).

PART 3 (Proofs For Contingently Exiting Players)

Next consider any '*contingently exiting*' player: who always starts playing; but will also exit after n rounds of continual defection by its partner, and who will cooperate at least once before exiting – denoted in the main text by Ex_n -players. An example for $n = 1$ is the '*prudent moral*' strategy (Vanberg & Congleton 1992): that always starts play by cooperating, and then exits after the first defection by its partner.

A contingent exitor's payoff during the periods of continued play with an $AllD$ -player before exiting after the n^{th} round is denoted V_n , and satisfies the following inequality :

$$\sum_{t=0}^{t=n-1} d^t P > V_n = \sum_{t=0}^{t=n-1} d^t (P \text{ or at-least-one-} S) \quad \text{for all } n \geq 1 \quad (B12a)$$

The strict equality just above follows because any type of strategy against an AllD_{NP}-player will receive either payoff P or S on each round of play. So $P > S$ implies any player who cooperates at least once with an AllD_{NP}-player will necessarily receive lower expected payoff over $n \geq 1$ rounds of play.

Consider a population with an ξ versus $1 - \xi$ frequency of Ex_n -players versus AllD_{NP}-players. The the Ex_n -players will then have an ξ probability of being matched with their own type, receiving an expected payoff denoted by $V^0 = E(Ex_n | Ex_n)$. Otherwise, they have a $1 - \xi$ chance of playing against an AllD_{NP}-player: receiving an expected payoff V_n defined above in (B12a), before exiting n -rounds of play later, and then paying a discounted search cost equal to, $\delta^n c \geq 0$, before starting play with a new partner encountered from the larger population.

So the expected payoff of a Ex_n player over all cases of encountering either themselves or AllD_{NP}-players in the future, denoted $V = E(Ex_n)$, satisfies the relationship: $E(Ex_n) = \xi V^0 + (1 - \xi)[V_n - \delta^n c + \delta^n E(Ex_n)]$. Rearranging the last formula obtains:

$$E(Ex_n) = \frac{\mathbf{x} V^0 + (1 - \mathbf{x})[V_n - \mathbf{d}^n c]}{1 - \mathbf{d}^n (1 - \mathbf{x})} \quad (\text{B12b})$$

Let the AllD_{NP}-players' expected payoff be denoted by, $U = E(\text{AllD}_{NP})$. Also let U_n denote the expected payoff received by an AllD_{NP}-player when matched with a Ex_n -player, over the n -rounds of play before the latter player exits the relationship; causing the AllD_{NP}-player to pay a discounted search cost that equals $\delta^n c \geq 0$, before starting play with a new partner.

The AllD_{NP}-players' expected payoff satisfies a similar relationship as $E(Ex_n)$ in statement (B12a), and so we

$$\text{also have: } E(\text{AllD}_{NP}) = \xi [U_n - \delta^n c + \delta^n E(\text{AllD}_{NP})] + (1 - \xi) \left(\frac{P}{1 - \mathbf{d}} \right) \quad (\text{B13a})$$

$$E(\text{AllD}_{NP}) = \frac{\mathbf{x}[U_n - \mathbf{d}^n c] + (1 - \mathbf{x}) \left(\frac{P}{1 - \mathbf{d}} \right)}{1 - \mathbf{d}^n \mathbf{x}} \quad (\text{B13b})$$

$$E(\text{AllD}_{NP}) - E(Ex_n) = \frac{\mathbf{x}[U_n - \mathbf{d}^n c] + (1 - \mathbf{x}) \left(\frac{P}{1 - \mathbf{d}} \right)}{1 - \mathbf{d}^n \mathbf{x}} - \frac{\mathbf{x} V^0 + (1 - \mathbf{x})[V_n - \mathbf{d}^n c]}{1 - \mathbf{d}^n (1 - \mathbf{x})} \quad (\text{B13c})$$

Then take the limit of (B13c) as the Ex_n players' population frequency ξ drops to 0, to see how well they will do in an almost exclusive population of AllD-players; thereby obtaining:

$$E(\text{AllD}_{NP}) - E(Ex_n) = \left(\frac{P}{1 - \mathbf{d}} - \frac{V_n}{1 - \mathbf{d}^n} \right) + \frac{\mathbf{d}^n c}{1 - \mathbf{d}^n} > 0 \quad (\text{B13d})$$

Next consider the following identity that follows from a standard proof by induction:

$$\frac{1}{1 - \mathbf{d}} \equiv \frac{\sum_{t=0}^{t=n-1} \mathbf{d}^t}{1 - \mathbf{d}^n} \quad \text{for all } n \geq 1 \quad (\text{B14a})$$

Then recalling the inequality for V_n given in (B12a) above, the difference expression in the curved brackets () on the right-hand side of (B13d) is necessarily positive, by the following sequence of inequalities implied by identity (B14a) just above:

$$\frac{P}{1 - \mathbf{d}} \equiv \frac{\sum_{t=0}^{t=n-1} \mathbf{d}^t P}{1 - \mathbf{d}^n} > \frac{\sum_{t=0}^{t=n-1} \mathbf{d}^t (P \text{ or at-least-one-} S)}{1 - \mathbf{d}^n} \equiv \frac{V_n}{1 - \mathbf{d}^n} \quad \text{for all } n \geq 1 \quad (\text{B14b})$$

The last result implies the strict inequality in (B13d) is guaranteed to hold for any type of contingently exiting player Ex_n . So we have the same qualitative results as before:

THEOREM B2

Any discounting of future payoffs ($0 \leq \delta < 1$) along with any nonnegative search cost ($c \geq 0$) implies any type

of Ex_n player is vulnerable to its share in a larger population composed mostly of AllD players.

APPENDIX C (Proof of Theorem 2)

The (CD, DC, CλD) players' signal and action probabilities are denoted by $p = (r_{CD}, w_{CD}, r_{DC}, w_{DC}, \lambda)$; and payoffs are denoted $Z = (T, R, P, S)$. Players' strategies are also indexed from 1 to 5, by $(s^1, s^2, s^3, s^4, s^5) = (CC, CD, DC, DD, CλD)$. Let $E(s^m | s^k, p, Z)$ denote an s^m -player's expected payoff when it plays against an s^k -player; depending on the signal and action probabilities p , and depending on payoffs Z .

It is sufficient to prove that a CD-player's expected payoff against its own type will *strictly* exceed each of the other four players' expected payoffs when matched against a CD-player; provided (r_{CD}, w_{CD}) is shifted sufficiently close to $(0, 0)$ along a convex NN-ROC curve. To do so, look at the CD column in Table 1 of the main text, and notice the first comparison between CD and CC is immediate; because their expected payoffs in this column converge respectively to (P, S) as $(r_{CD}, w_{CD}) \rightarrow (0, 0)$.

The other comparisons involve additional algebra, but are otherwise straight forward. Let the next expression denote the expected payoff difference between player s^m versus s^n when they both play against player s^k :

$$\Delta(s^m, s^n | s^k, p, Z) = E(s^m | s^k, p, Z) - E(s^n | s^k, p, Z) \quad \text{for } m \neq n, k = 1, \dots, 5$$

Now compare CD with DD, again using the expected payoffs in the CD column of Table 1. Algebraic manipulation yields the following inequality for strictly positive $(r_{CD}, w_{CD}) > (0, 0)$:

$$\frac{\Delta(CD, DD | CD, p, Z)}{r_{CD}} > 0 \quad \text{if and only if} \quad \frac{1}{r_{CD} / w_{CD}} \quad (C1a)$$

$$r_{CD} [(P - S) - (T - R)] + (T - P)(1 - r_{CD} / w_{CD}) - (P - S) > 0 \quad (C1b)$$

Then recall $r_{CD}/w_{CD} \rightarrow \infty$ as $(r_{CD}, w_{CD}) \rightarrow (0, 0)$ along any convex NN-ROC curve that bows above the diagonal line in Figure 1a, by statement (1) in Section I. So the expression (C1b) converges to, $(T - P) - (P - S)$ as $(r_{CD}, w_{CD}) \rightarrow (0, 0)$; where $(T - P) = (T - R) + (R - P)$. Since the denominator r_{CD} of (C1a) is positive as $(r_{CD}, w_{CD}) \rightarrow (0, 0)$ along a convex NN-ROC curve, then the ratio $\Delta(CD, DD | CD, p, Z)/r_{CD}$ must become positive as $(r_{CD}, w_{CD}) \rightarrow (0, 0)$, because $(T - P) = (T - R) + (R - P) > (P - S)$ was already assumed. This establishes $\Delta(CD, DD | CD, p, Z) > 0$ must hold as $(r_{CD}, w_{CD}) \rightarrow (0, 0)$ along a convex NN ROC curve.

Next compare CD with DC. Algebraic manipulation of expected payoffs in the CD column of Table 1 yields the following inequality for strictly positive $(r_{CD}, w_{CD}) > (0, 0)$:

$$\Delta(CD, DC | CD, p, Z)/r_{CD} > 0 \quad \text{if and only if} \quad (C2a)$$

$$\left(\frac{1 - r_{DC}}{r_{CD}} \right) (P - S) + (T + R) + \left(1 - \frac{1}{r_{CD} / w_{CD}} \right) (R - P) - \left(1 + \frac{1}{r_{CD} / w_{CD}} \right) (P - S) + \frac{r_{DC}}{r_{CD} + r_{CD} / w_{CD}} [(P - S) - (T - R)] > 0 \quad (C2b)$$

Since $r_{CD}/w_{CD} \rightarrow \infty$ as $(r_{CD}, w_{CD}) \rightarrow (0, 0)$, expression (C2b) converges to,

$$\left(\lim_{r_{CD} \rightarrow 0} \left(\frac{1 - r_{DC}}{r_{CD}} \right) \right) (P - S) + (T - R) + (R - P) - (P - S) \quad (C2c)$$

The left-most term just above implies (C2c) equals ∞ if the DC-players' signal probability r_{DC} is less than 1; otherwise, $r_{DC} = 1$ implies (C2c) equals the same difference as before, $(T - R) + (R - P) - (P - S) = (T - P) - (P - S)$. So either case ($r_{DC} < 1$ or $r_{DC} = 1$) implies (C2b) must become positive as $(r_{CD}, w_{CD}) \rightarrow (0, 0)$; which in turn implies $\Delta(CD, DC | CD, p, Z) > 0$ must also hold as $(r_{CD}, w_{CD}) \rightarrow (0, 0)$, because $(T - P) > (P - S)$ was already assumed.

Next consider the last comparison of CD against CλD. Algebraic manipulation of expected payoffs in the CD

column of Table 1 yields the following inequality for strictly positive $(r_{CD}, w_{CD}) > (0, 0)$:

$$\Delta(CD, C\lambda D | CD, p, Z)/r_{CD} > 0 \quad \text{if and only if} \quad (C3a)$$

$$\frac{1}{r_{CD}} (P - S) + (T - P) - (P - S) - \frac{(T - P)}{r_{CD}/w_{CD}} + (r_{CD} - \frac{1}{r_{CD}/w_{CD}})[(P - S) - (T - R)] > 0 \quad (C3b)$$

The left-most term of (C3b) will rise to ∞ as $(r_{CD}, w_{CD}) \rightarrow (0, 0)$ if $\lambda > 0$; otherwise, $\lambda = 0$ implies (C3b) converges to the same difference as before, $(T - P) - (P - S)$. So either way ($\lambda > 0$ or $\lambda = 0$), (C3b) must become positive as $(r_{CD}, w_{CD}) \rightarrow (0, 0)$; which in turn implies $\Delta(CD, C\lambda D | CD, p, Z) > 0$ must also hold as $(r_{CD}, w_{CD}) \rightarrow (0, 0)$, since $(T - P) > (P - S)$ was already assumed. This completes the four cases: showing CD against another CD always does better than (CC, DC, DD, C λ D) against CD, whenever (r_{CD}, w_{CD}) shifts close enough to $(0, 0)$ along a convex NN ROC curve; thus proving Theorem 2.

REFERENCES

- Ahn, T. K., Elinor Ostrom, David Schmidt, Robert Shupp, and James Walker. (2001) "Cooperation in PD Games: Fear, Greed, and History of Play", *Public Choice*, 106, pp. 137 - 155.
- Albert, Max, and Heiner, Ronald A. (2002) "An Indirect-Evolution Approach To Newcomb's Problem," to appear in, *Homo Economicus*.
- Axelrod, Robert, and Hamilton, William D. (1981) "The Evolution of Cooperation," *Science*, Vol. 211, pp. 1390 - 1396.
- Axelrod, Robert, (1984) *The Evolution Of Cooperation*, Basic Books, New York.
- Axelrod, Robert and Dion, Douglas (1988) "The Further Evolution of Cooperation," *Science*, Vol. 242, pp. 1385 - 1390.
- Banerjee, Abhijit, and Weibull, Jörgen (2000) "Neutrally Stable Outcomes In Cheap-Talk Games," *Games and Economic Behavior*, Vol. 32, pp. 1 - 24.
- Ben-Shakhar, G. Lieblich, I. and Bar-Hillel, Y. (1986) *Journal Of Applied Psychology*, Vol. 67, pp. 701 - 718.
- Bhaskar, V. (1998) "Noisy Communication and the Evolution of Cooperation," *Journal of Economic Theory*, 92, pp. 110 - 131.
- Binmore, Ken, (1987) "Modeling Rational Players I", *Economics and Philosophy*, 3, pp. 179 - 214.
- Binmore, Ken, (1988) "Modeling Rational Players II", *Economics and Philosophy*, 4, pp. 9 - 55.
- Binmore, Ken (1994) *Game Theory and the Social Contract I: Playing Fair*, MIT Press, Cambridge.
- Bolton, Gary E., and Ockenfels, Axel (2000) "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review*, 90 (No. 1), pp. 166 - 93.
- Coleman, Andrew (1983) *Game Theory and Experimental Work*, Pergamon Press, London.
- Cooper, R., Dejong, D., Forsythe, R., and Ross, T. (1996) "Cooperation Without Reputation: Experimental Evidence From Prisoners' Dilemma Games," *Games and Economic Behavior*, Vol. 12, pp. 187 - 218.
- Egan, James P. (1975) *Signal Detection Theory and ROC Analysis*, Academic Press, New York.
- Frank, Robert (1988) *Passions Within Reason: The Strategic Role Of Emotions*, Norton, New York.
- Frey, Bruno & Bohnet, Iris (1995) "Institutions Affect Fairness: Experimental Investigations," *Journal of Institutional & Theoretical Economics*, 286, pp. 286 - 303.
- Fudenberg, Drew, and Tirole, Jean (1991) *Game Theory*, MIT Press, Cambridge.

- Gauthier, David (1986) *Morals By Agreement*, Clarendon Press, Oxford.
- Geanakoplos, J. Pearce, D. and Schacchetti, E. (1989) "Psychological Games and Sequential Rationality," *Games and Economic Behavior*, Vol. 1, pp. 60 - 79.
- Getty, D. J., Pickett, R., D'Orsi, C. J. and Swets, John (1988) *Investigations In Radiology*, Vol. 23, pp. 240 - 251.
- Gintis, Herbert (2000) *Game Theory Evolving*, Princeton University Press.
- Green, David M., Swets, John (1966) *Signal Detection Theory and Psychophysics*, Wiley, New York; reprinted with updated topical bibliographies by Krieger, New York, 1974, and by Peninsula Publishing, Los Angeles, 1989.
- Kim, Y.-G., and Sobel, J. (1995) "An Evolutionary Approach to Pre-Play Communication," *Econometrica*, Vol. 63, pp. 1181 - 1193.
- Lave, Lester B. (1962) "An Empirical Approach to the Prisoner's Dilemma," *Quarterly Journal of Economics*, 76, pp. 424 - 436.
- Lave, Lester B. (1965) "Factors Affecting Cooperation in the Prisoner's Dilemma," *Behavioral Science*, 10, pp. 26 - 38.
- Ledyard, John (1995) "Public Goods: A Survey of Experimental Research," in John H. Kagel & Alvin Roth, eds., *Handbook of Experimental Economics*, Princeton University, pp. 111 - 194.
- Liebrand, W., David Messick, & Henk Wilke (1992) *Social Dilemmas: Theoretical Issues and Research Findings*, Pergamon Press, New York.
- Lusted, L. B. (1968) *Introduction To Medical Decision Making*, Thomas, Springfield, Il.
- Minas, J. S., Scodel, A., Marlowe, D., & Rowson, H. (1960) "Some Descriptive Aspects of Two Person Non-zero Sum Games," *Journal of Conflict Resolution*, 4, pp. 193 - 197.
- Maynard-Smith, John (1982) *Evolution and the Theory of Games*, Cambridge University, London.
- Orwell, John, Schwartz-Shea, Peregrina, & Simons, R. (1984) "Do Cooperators Exit More Readily Than Defectors ?" *American Political Science Review*, Vol. 78, pp. 147 - 162.
- Ostrom, Elinor (2000) "Collective Action and the Evolution of Social Norms", *Journal of Economic Perspectives*, 14, Summer, pp. 137 - 158.
- Palfrey, T. R. & Prisbrey, J. E. (1997) Anomalous Behavior In Public Goods Experiments: How Much and Why?" *American Economic Review*, 87, pp. 829 - 846.
- Rapoport, Anotal (1967) A Note On The 'Index of Cooperation' For Prisoners' Dilemma," *Journal of Conflict Resolution*, 11, pp. 101 - 103.
- Rapoport, Anotal, and Chammah, Albert (1965) *Prisoner's Dilemma: A Study in Conflict and Cooperation*, University of Michigan, Ann Arbor.
- Rapoport, Anotal, Guyer, M. & Gordon, D. (1976) *The 2 x 2 Game*, University of Michigan, Ann Arbor.
- Rasmussen, Eric (2001) *Readings In Games And Information*, Blackwell, Oxford.
- Sally, David (1995) "Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments From 1958 - 1992". *Rationality and Society*, vol. 7(1), pp. 58 - 92.
- Saxe, L., Dougherty, D., and Scott, B. (1985) *American Psychological Review*, Vol. 40, pp. 794 - 811.
- Schuessler, Rudolf (1989) "Exit Threats and Cooperation Under Anonymity," *Journal of Conflict Resolution*, Vol. 33, pp. 728 - 749.

Swets, John (1961) *Science*, Vol. 134, pp. 168 - 172.

Swets, John, Ed. (1964) *Signal Detection and Recognition By Human Observers*, Wiley, New York; reprinted by Peninsula Publishing, Los Angeles, 1989.

Swets, John (1973) *Science*, Vol. 182, pp. 990 - 995.

Swets, John, and Pickett, R. (1982) *Evaluation of Diagnostic Systems: Methods From Signal Detection Theory*, Academic Press, New York.

Swets, John (1988) "Measuring The Accuracy Of Diagnostic Systems," *Science*, Vol. 240, pp. 1285 - 1293.

Trivers, Robert L. (1971) "The Evolution of Reciprocal Altruism," *Quarterly Review of Biology*, Vol. 46, pp. 35 - 57.

Vanberg, Victor, and Congleton, Roger (1992) "Rationality, Morality, Exit," *American Political Science Review*, Vol. 86(2), pp. 418 -431.

Vanberg, Victor, and Congleton, Roger (2001) "Help, Harm, or Avoid ? On The Personal Advantage of Dispositions To Cooperate and Punish In Multilateral PD Games With Exit," *Journal of Economic Behavior & Organization*, Vol. 44, pp. 145 - 167.

Wärneryd, Karl (1993) "Cheap Talk, Coordination, and Evolutionary Stability," *Games and Economic Behavior*, Vol. 5, pp. 532 - 546.

Weibull, Jörgen, W. (1997) *Evolutionary Game Theory*, MIT Press, Cambridge.