CRED

UNIVERSITÉ PARIS II
PANTHÉON-ASSAS

CENTRE DE RECHERCHES EN ÉCONOMIE ET DROIT
Paris Center for Law and Economics

CRED WORKING PAPER N^o 2019-19

A Model for Explaining Players' Decisions in Static Strategic Games

GILLES DUBAIL[*], DAMIEN GAUMONT[†]

[*]Department of Economic Research, Sherwood Global System, gilles.dubail@gmail.com, 128 rue de la Boétie, 75008 Paris, France.

[†]University of Paris 2, Panthéon-Assas, CRED EA 7321 and TEPP, damien.gaumont@u-paris2.fr, 12, Place du Panthéon.75231 Paris CEDEX 05, France.

Abstract:

The Bayes equivalent of the PD contains three essential information. 1) It provides the players with the basic probability distribution. 2) there always exists a support of these probabilities such that the difference between the expected payoff of cooperation and the one of defection is positive. Cooperation may be selected. 3) If this difference dominates the expected utility evaluated at the mixed strategies Nash equilibrium, then cooperation is trustable. Our modeling fits data with or without the use of subjective probabilities. Extensions to Zero-sum games, War and Peace, BoS, Coordination games, Head and Tail, Hawk Dove, Centipede game are studied.

JEL Classification C62, C72, D01.

Key Words: Mathematical Methods, Existence, Stability Conditions of Nash Equilibrium of One Shot PDs, Rational Choice of Cooperation, Defection and Trust.

1 Introduction

This paper reconciles theory (defection is the unique theoretical output in the Prisoner's Dilemma, hereafter PD, whatever the pure, mixed, sequential, trembling hand strategies, etc.) and empirics (cooperation is played at huge rates). Our objective is to explain with few assumptions this paradox, while keeping the rationality principle and the solution concept of Nash equilibrium. Prior to presenting them, let us shortly recall the history of useful concepts.

1.1 The history of the concepts

A first branch in game theory (game theoretic) started with Pascal (1654)[26], Fermat (1654)[13] then Huygens (1657)[19] and definitely with Bernouilli (1715)[8] who were the firsts to consider that an objective probability $p_i, i = 1, \ldots, k$ can be affected to a random event or a random gain g_i. They propose the concept of mathematical expectation $B(p_i, g_i) := \sum_{i=1}^{k} p_i g_i$. In this expression the gain g_i stands for a cardinal utility.
A second branch in game theory (bayesians) started with Bayes (1763) [7] considers that an action $a_j, j = 1, \ldots, n$ may have several consequences $\omega_{jk}, k = 1, \ldots, K$ each of them appearing with objective probabilities ${p_{ik}}$[1], $\mathcal{B}(a) := p(\omega_{|a})$. Laplace (1814)[22] proposes to affect the concept of moral utility $\overline{u}$ to an action, $L(p_j, \overline{u}(a_j)) := \sum_{j=1}^{n} p_j \overline{u}(a_j)$. He suggested that $L(p_j, \overline{u}(a_j)) = B(p_i, g_i)$. After Von Neumann and Morgenstern (1944)[31], economists associate an action to a random gain. They propose to replace g_i in $B(p_i, g_i)$ by its subjective valuation $\overline{s}(g_j(a_j))$, $EU(p_j, \overline{s}(g_j(a_j))) := \sum_{j=1}^{n} p_j \overline{s}(g_j(a_j))$. Savage (1939) suggests to replace the objective probability p_j by a subjective one $\overline{p}_j$, $SEU(\overline{p}_j, \overline{s}(g_j(a_j))) := \sum_{j=1}^{n} \overline{p}_j \overline{s}(g_j(a_j))$.

An important question has been debated between Kadane and Larkey (1982a[20], 1982b[21])) and Harsanyi (1982a[16],1982b[17]): how to choose the distribution of subjective probabilities $\overline{p}_j$? Harsanyi (1967[15] p159 and 1982, p. 120[17]) argued that subjective probability distributions entertained by different players are mutually "consistent" in the sense that they can be regarded as conditional probability distributions

[1]Bayes mentions that is theory is "applicable to events and appearance of Nature, [. . .], as the burning of wood on putting it into fire", p. 408.

derived from a certain "basic probability distribution", which is obtained as follows. Consider a sequential strategic game with incomplete information. Define the set of histories $H := \{\emptyset, 1, \ldots, h, \ldots, Z\}$, where z is the terminal history. Harsanyi suggested to associate to the previous game another game in complete information where Nature first conducts a lottery and randomly decides which particular sub-game will be played after each history, $h \in H \backslash Z$. Every player will know the "basic probability distribution" governing this lottery, which is objective, $p_j(h)(a_j(h))$. Harsanyi called it the Bayes equivalent of the original game.

Aumann (1987)[3] reconciles Game theoretic with Bayesians by assuming that some actions a_i are linked to state of the Nature ω_k. Nature plays first by choosing randomly the true state of the World. The exogenous probability p_k of a given state of the Nature is objective. This solves the problem of choosing the "right" probability distribution. The true state of the World is public information and revealed to players. The Aumann's criterion is the following $A(p_k, \overline{s}(a_i(\omega_k))$. In correlated equilibrium, players always choose a definite pure strategy and have no incentive in deviating.

1.2 Our concepts

1.2.1 Generalities

PRINCIPLE 1 *Actions of the opponents a_{-j} are considered the states of the world for player j, $a_{-j}(h) = \omega_j(h), \forall h \in H \backslash Z$.*

ASSUMPTION 1 *Players know and use all the Properties of the Bayes equivalent.*

1.2.2 Information

Players know the rule of a given PD, its primitive and all the properties of the Bayes equivalent of the game. In particular, prior to selecting a strategy, they use three following essential information directly extracted from Bayes equivalent of the game.

1.2.3 Information 1: the basic probabilities

Every player knows the "basic probability distribution" governing the lottery in the Bayes-equivalent of the PD. Due to the tree of decision of the sequential PD, it turns out that the basic probability distribution is unique, discrete and binomial.

1.2.4 Information 2: Emergence of cooperation

Let us denote $EP_{\mathcal{C}}$ the expected payoff of cooperation and $EP_{\mathcal{NC}}$ the one of defection.

PROPERTY 1 *In a sequential 2-action, 2-player PD, $H := \{\emptyset, 1, z\}$, the Bayes equivalent of the PD possesses the following property: there always exists a non empty support of basic probability such that the expected payoff of cooperation dominates the one of defection, $\Delta EP(p_j(\emptyset), p_j(1)) := EP_{\mathcal{C}}(p_j(\emptyset)), p_j(1)) - EP_{\mathcal{NC}}(p_j(\emptyset), p_j(1)) \geq 0$.*

An immediate consequence of Assumption 1 is that cooperation can be selected as a pure strategy by rational players. How do players trust their opponent to play it ?

1.2.5 Information 3. Emergence of trust

According to Assumption 1, the interval over which a pure strategy emerges as a possible choice is an endogenous, objective and public information.

ASSUMPTION 2 *Players evaluate the expected utility in terms of the mixed strategy Nash equilibrium $EU(\alpha^\star)$.*

Define $\Psi(p_{ji}(\emptyset), p_j(1), \alpha^\star) = \Delta EP(p_j(\emptyset), p_j(1)) - EU(\alpha^\star)$.

DEFINITION 1 *A strategy is called trustable if and only if the support of its basic probabilities is determined by using the criterion $C^\star := \max\{\Delta EP(p_j(\emptyset), p_j(1)), EU(\alpha^\star)\}$.*

Note that the choice of the criterion $C^\star$ is a Nash equilibrium: no player has an incentive in deviating. Given a basic probability p^0, cooperation is a trustable strategy if and only if $\Psi(p^0_{ji}(\emptyset), p^0_j(1), \alpha^\star) \geq 0$. As it will be shown, this is not always the case. Assumptions 1 and 2 are sufficient for explaining cooperation as a trustable strategy in PD.

From a theoretical point of view, the basic probability has the following economic interpretation. Using a first-order reasoning, the basic probability $p_j(1)$ is attributed by player i to his opponent j's choice of action $a_j(h = 1)$. Using a second-order reasoning, he simultaneously attributes to his opponent j a basic probability $p_{ji}(a_i(h = \emptyset))$ about his own strategy a_i.

1.3 Our modeling accommodates data

Prior to selecting a strategy, players can (or not) use subjective probabilities. According to Allais (1953)[1] (5. ELEMENT II p. 508):

DEFINITION 2 *A subjective probability is the consequence of a subjective distorsion of an objective probability,* $\exists\lambda \in \mathbb{R}_+ \exists\xi \leq 1 \mid \overline{p}_j(a_j(h)) := \lambda p_j(a_j(h)) + \xi$.

Subjective probabilities are private information.

ASSUMPTION 3 *If rational player use subjective probabilities, then they select their optimal strategies according to the subjective expected payoff of an action. Cooperation occurs iff* $\Delta SEP(\overline{p}_{ji}(\emptyset), \overline{p}_j(1)) \geq 0$ *(defection if* < 0*).*
Given observation $\hat{\mu}$, it always exists a couple (λ, ξ) that accommodates data $\hat{\mu} = \lambda p_j(a_j(h)) + \xi$.

1.4 Extensions

Extension to static strategic games, Zero Sum games, Aumann's War and Peace game, Battle of the Sexes, Coordination game, Head and Tail and Hawk Dove/Tragedy of Commons, as well as Centipede game are discussed in Section 4. It is shown that cooperation may emerge with trust, or without trust, non-cooperation is possible with trust or without trust. In the centipede game, McKelvey and Palfrey (1992) [24] we explain why a rational player cooperate by choosing to pass rather than to take the large pile. The end of the game is explained by the first period in the game for which the strategy "take it" provides the 2 players with more subjective expected payoff than the expected utility of the pure, mixed strategy Nash Equilibrium. Each player trusts the opponent to pass until this period.

Section 2 presents the available information. Section 3 is devoted to the formation of subjective probabilities. Section 4 presents the subjective expected payoff strategies in one-shot PD. Section 5 is devoted to the selection of equilibria. Section 6 faces our model to experiments, Section 7 discusses the results and proposes economic interpretations, including trust. Section 8 Proposes extensions to other famous strategic games, Section 9 concludes.

2 Information

The information available to the players consists in the rule of the game, the primitives of the game and the general properties of the Bayes-equivalent of the PD.

2.1 Rule of the PD

Consider a sequential 2-personal player / 2-action PD, $i = 1, 2$. If player i defects while player j cooperates, player i gets the payoff A_i and player j gets D_j. If i defects while j defects, both get the payoff C_i. If i cooperates while j cooperates, both get the payoff B_i. In any PD, $A_i > B_i > C_i > D_i, i = 1, 2$. Assume payoffs are symmetric such that $A_i = A_j = A, B_i = B_j = B, C_i = C_j = C, D_i = D_j = D$.

2.2 Primitives of the PD

Consider an extensive PD with imperfect information. Whoever leads the game, players know the primitives of the game according to the following definition.

DEFINITION 3 *An extensive game with imperfect information denoted* $\Gamma :=< N, H, P, f_0, I >$ *consists in*

1. $N = \{1, 2\}$ *is the set of players where players* $1, 2$ *are the personal players.*

2. H *is a set of (finite) sequences, each member* $h \in H$ *is a history, each component of an history is an action taken by a player* $i \in N$. Z *is the subset of terminal histories after which no action is to be taken.*

3. P *is the player function that assigns to each history* $h \in H \backslash Z$ *a player* $i \in N$.

4. f_0 *is the function that associates with every history* h *for which* $P(h) = i$ *a probability measure of* $f_0(. \mid h)$ *on* $A(h) \in \Gamma$ *the set of actions available* $\forall h \in H \backslash Z$.

5. $\forall i \in N$, $\mathcal{I}_i$ *is a partition of* $\{h \in H : P(h) = i\}$ *with the property that* $A(h) = A(h')$ *whenever* h *and* h' *are in the same member of the partition.* $\mathcal{I}_i$ *is the information partition of player* i, *and* $I_i \in \mathcal{I}_i$ *the information set of player* i.

2.3 General properties of the Bayes-equivalent of the PD

Section 2.3 presents the Bayes-equivalent of the PD and its useful properties. To define the Bayes-equivalent of the PD, replace N by $N_0 = \{0, 1, 2\}$ in Definition 3, where player 0 is the "random" player who is responsible for the random decisions in the game, Von Neuman and Morgenstern (1953)[32] p. 75 (10:A:$f^{\star}$, 10:A:$h^{\star}$) p. 159, Harsanyi (1967)[15] and Selten (1975)[29] p. 26. According to Harsaniy (1967)[15], p. 159, in the

"Bayes-equivalent of the PD", $\forall h \in H \backslash Z, P(h) = 0$, player 0 chooses each sub-game and assigns payoffs to players $i = 1, 2$. For this reason, we label it $\hat{\Gamma}_0$. In $\hat{\Gamma}_0$, $H := \{\emptyset, 1, Z\}$. The set of actions is $\forall i = 1, 2, \mathcal{A}_i(h) = \{\mathcal{C}_i : (h, \mathcal{C}_i) \in H \backslash Z, \mathcal{NC}_i : (h, \mathcal{NC}_i) \in H \backslash Z)\}$, where $\mathcal{C}_i$ represents cooperation and $\mathcal{NC}_i$ non-cooperation.

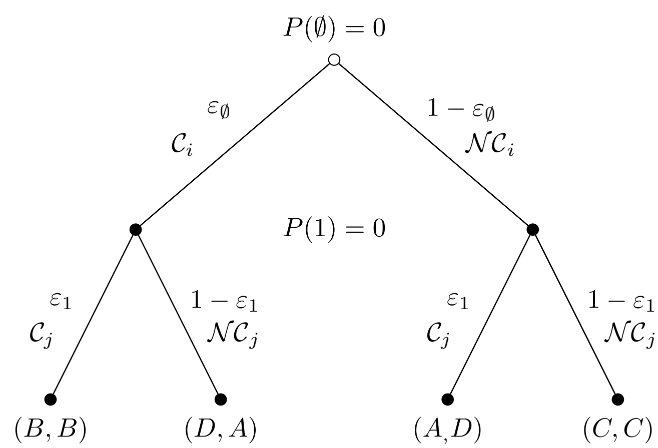
Players know the three following fundamental information: the basic probability distribution, the conditions of emergence of cooperation, and the conditions of the emergence of trust.

2.4 Information 1: the basic probability distribution

According to the "basic probability distribution", $\mathcal{C}_i|h$ occurs after history $h = \emptyset$ with the objective probability $\varepsilon_\emptyset$ and after history $h = 1$ with the objective probability ε_1. The action $\mathcal{NC}_i|h$ occurs after each history with the complementary probabilities, see Figure 1. There is no particular reasons to consider that the random mechanism is the same after each history. To illustrate that, consider a random draw after each history without replacement.

—————— insert Figure 1 here ——————

Figure 1: $\hat{\Gamma}_0$: the Bayes-equivalent of the PD



2.5 Information 2: Emergence of Cooperation

Consider Figure 1. Denote $EP_a(\varepsilon_\emptyset, \varepsilon_1)$ the expected payoff in terms of the " basic probability" of an action a which can either be $\mathcal{C}$ or $\mathcal{NC}$. Denote $\Delta EP_i(\varepsilon_\emptyset, \varepsilon_1) := EP_{\mathcal{C}_i}(\varepsilon_\emptyset, \varepsilon_1) -$

$EP_{\mathcal{NC}_i}(\varepsilon_\emptyset, \varepsilon_1)$ and $\Delta EP_j(\varepsilon_1, \varepsilon_\emptyset) := EP_{\mathcal{C}_j}(\varepsilon_1, \varepsilon_\emptyset) - EP_{\mathcal{NC}_j}(\varepsilon_1, \varepsilon_\emptyset)$ the variation of expected payoff between cooperation and non-cooperation for each player. The Bayes-equivalent of the PD leads to the following Theorem.

THEOREM 1 *In the Bayes-equivalent of any PD with symmetric payoffs, 2 players, 2 actions,* $\exists I_{\varepsilon_\emptyset} \neq \emptyset, I_{\varepsilon_\emptyset} := [\underline{\varepsilon}_\emptyset, \overline{\varepsilon}_\emptyset] \subset [0,1]$, $\exists I_{\varepsilon_1} \neq \emptyset, I_{\varepsilon_1} := [\underline{\varepsilon}_1, \overline{\varepsilon}_1] \subset [0,1]$ *such that given* $A, B, C, D,\ \forall (\tilde{\varepsilon}_\emptyset, \tilde{\varepsilon}_1) \in I_{\varepsilon_\emptyset} \times I_{\varepsilon_1} \mid S_0 : \left\{ \begin{array}{lcl} \Delta EP_i(\tilde{\varepsilon}_\emptyset, \tilde{\varepsilon}_1) & \geq & 0, \\ \Delta EP_j(\tilde{\varepsilon}_1, \tilde{\varepsilon}_\emptyset) & \geq & 0. \end{array} \right.$

Note that S_0 exhibits non-linear relations in term of the product of the basic probabilities $\varepsilon_\emptyset \varepsilon_1$. There is no methodology to solve S_0, one is proposed in Appendix A. The importance of this theorem (see Proposition 6 Appendix A) is that there always exists a non-empty support of probability for each basic probabilities over which the expected payoff of cooperation dominates the one of defection in the Bayes equivalent of the PD. Players knows this property prior to selecting their optimal strategy.

Theorem 2 proposes a typology of the Bayes-equivalent of any PD. Given the payoffs, Theorem 2 characterizes the endogenous non-empty support of probability over which system S_0 is satisfied

THEOREM 2 *Typology of PDs. Denote J the support of basic probabilities over which system S_0 is satisfied, where hereafter $\varepsilon, \underline{\varepsilon}, \overline{\varepsilon}, \hat{\varepsilon}$ are payoff-specific thresholds.*

1. *4 or 3 positive payoffs, the unique support of basic probabilities is $J_\varepsilon := [\varepsilon, 1]$.*

2. *2 positive payoffs, the supports of basic probabilities are $J_{\underline{\varepsilon}} := [0, \underline{\varepsilon}]$ and $J_{\overline{\varepsilon}} := [\overline{\varepsilon}, 1]$.*

3. *1 or 0 positive payoffs, the unique support of basic probabilities is $J_{\hat{\varepsilon}} := [0, \hat{\varepsilon}]$.*

See Appendix B for the proof of Theorem 2 and the details of each support of probabilities. When the support of probabilities is $J_\varepsilon := [\varepsilon, 1]$, the Bayes equivalent of the PD provides the players with the information that cooperation arises at huge rate from a random perspective, and reciprocally over $J_\varepsilon := [0, \varepsilon]$ at low rates. Such a typology indicates also that the length of each support of basic probabilities J is endogenous since payoff dependent. We now turn to study the emergence of trust.

2.6 Information 3: Emergence of Trust

Let us recall the literature relative to trust.

2.6.1 The literature relative to trust

The concept of trust is important for economics, politics, psychology and sociology. Prior to showing how it is related to our analysis, let us recall some important results in the literature. To be short, the sources of trust are direct or indirect. Direct sources of trust means that trust is founded on circumstances or facts, Sako 1992[28]. Indirect sources of trust have been identified to be credible commitments, Williamson (1983)[33], deterrence-based trust, Shapiro, Sheppard and Cheraskin (1992)[30] or trust as an encapsulated interest, Russell (2001). For Williamson (1993)[34], there is no tension between interest and trust. Glaeser, Laibon, Scheinkman, Soutter (2000)[14] and Coleman (1990) p. 99[10] underline that trust is compatible with the rational behavior principle.

As Adam Smith (1776) (Wealth of Nations, p. 22) noticed "trust" is empirically based and probabilistic. For Arrow (1969)[2] "Trust [...] is not a commodity which can be bought very easily. If you have to buy it, you already have some doubts about what you have bought". Departing Arrow, Dasgupta (1988)[11] considers trust a commodity, and Milgrom, North and Barry (1997)[25] an asset. For Coleman (1990) p.99[10] it is a bet: "An individual knows how much may be lost (the size of the bet), how much may be gained (the amount that might be won) and the chance of winning. These and only these are the relevant elements to define trust. If an individual has no risk aversion it is a simple matter for him to decide whether to place the bet". Buchanan (1991, p. 47) underlines "In that larger economic system, one is unlikely to have direct information on the trustworthiness of the other parties involved, so in a *laissez-faire* environment, one must resort to proxies: risk assessments by others who, in turn, must themselves be assessed for risk." For Kenneth Arrow (1969)[2] "In the absence of trust it would become very costly to arrange for alternative sanctions and guarantees, and many opportunities for mutually beneficial cooperation would have to be, foregone".

2.6.2 The link with our modeling

Whatever the typology (good, asset or bet), trust is intangible. More precisely, trust is an attitude that takes place *ex-ante* prior to making decision in $\hat{\Gamma}$ which is free of risk aversion considerations.

DEFINITION 4 *We call attitude opinions or feelings about something that affects individ-*

ual's behavior.

DEFINITION 5 *We call trust any rational and endogenous player's attitude such that player i leaves to player j the choice of an action that achieves i's objective (in whole or in part) and player i does not verify the j's action.*

REMARK 1 *In the PD, players cannot communicate. Decisions are made ex ante, simultaneously and for ever. Consequently, there is no verification of any action.*

REMARK 2 *In Definition 5, an objective is not necessary an optimum.*

Trust emerges under uncertainty (because of properties of the Bayes-equivalent of the PD $\hat{\Gamma}_0$) and is probabilistic as Adam Smith (1776) and Buchanan (1991) noticed. Indirect source of trust is compatible with the fact that players analyze the Bayes-equivalent of the PD $\hat{\Gamma}_0$. According to Coleman, the size of the bet is an important criterion for trust to emerge and in our Proposition 3 it depends on the sign of C, the payoff of bilateral cooperation.

An important question is: why should a rational player i leave to another player j the realization of his own objective ? Pure rational players find the best procedure to get the highest possible outcome. Denote $EU(\alpha^\star)$ the expected utility evaluated at the mixed (pure) strategy Nash equilibrium $\alpha^\star$.

Consider system S_0 with a unique random device after each history. In that particular case, $\varepsilon_\emptyset = \varepsilon_1 = \varepsilon$. Define $\Psi((\varepsilon, \alpha^\star) := \Delta EP(\varepsilon) - EU(\alpha^\star)$. If $\Psi(\varepsilon, \alpha^\star) \geq 0$ over some restriction $\hat{J}_\varepsilon$ of the support of basic probabilities of Theorem 2, then players get the information that a pure strategy can be selected according to the ΔEP criterion. If $\Psi(\varepsilon, \alpha^\star) < 0$, over some restriction of the support of basic probabilities of Theorem 2, then players get the information that the other pure strategy can be selected according to the $EU(\alpha^\star)$ criterion, see Figures 4, 5 and 6 below. To sum up, the set of supports of basic probabilities that is compatible with a given pure strategy is defined as $\cup \hat{J}_\varepsilon$.

2.6.3 Trust as the result of the properties of the Bayes equivalent of the PD

In the literature, trust is known to be the result of huge deviation from pure rationality. In our paper this is the opposite: it is an endogenous and rational attitude. Trust is the result of the properties of the Bayes equivalent of the PD. According to Definition 1, we have the following Theorem.

THEOREM 3 *In any PD, the support of trustable strategies is non empty.*

Proof. Consider the support of basic probabilities such that $\Delta EP(\varepsilon) \geq 0$. Over this support, $C \leq 0 \Rightarrow \Psi(\varepsilon) \geq 0$. Note that the discriminant of $\Psi(\varepsilon) = 0$ is $\Delta_\Psi = 8C(A+B-C-D)+(A-2C-D)^2$ which is negative for $C < \frac{1}{2}\Big(A+2B-D-\sqrt{2}\sqrt{(A-D)^2+2B(A+B-D)}\Big)$. If $\Delta_\Psi < 0$, then equation $\Psi(\varepsilon) = 0$ has no solutions in $\mathbb{R}$ and by the convexity of the function, $\Psi(\varepsilon) > 0$, see Figure 4. A negative discriminant is not a problem: it only means that the condition of positivity of the inequality is never satisfied in $\mathbb{R}$. It is easy to show that the discriminant cannot be nil, indeed $\varepsilon_1^{\star\star} = \varepsilon_2^{\star\star} \iff B = C$, which impossible in PDs. Consequently, $\Psi(\varepsilon) = 0$ admits two distinct solutions $\varepsilon_l^{\star\star}, l = 1, 2$, see Figure 5

$$\varepsilon_1^{\star\star} := \frac{A-2C-D-\sqrt{\Delta_\Psi}}{2(A+B-C-D)} \text{ and } \varepsilon_2^{\star\star} := \frac{A-2C-D+\sqrt{\Delta_\Psi}}{2(A+B-C-D)}.$$

Note that if $C > 0$ at $\varepsilon_1^{\star\star} = \varepsilon_2^{\star\star}$ we have $\partial\varepsilon_2^{\star\star}/\partial B < 0$ and $\lim\limits_{A\to\infty} = 1$, so that in the case of 4 and 3 positive payoffs, $\varepsilon_2^{\star\star} < 1$, see Figure 6.

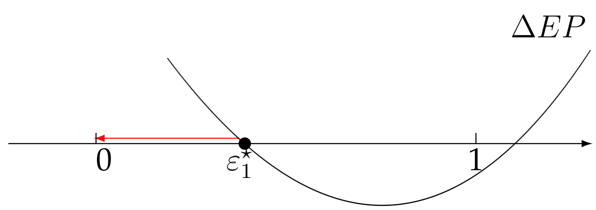


Figure 4

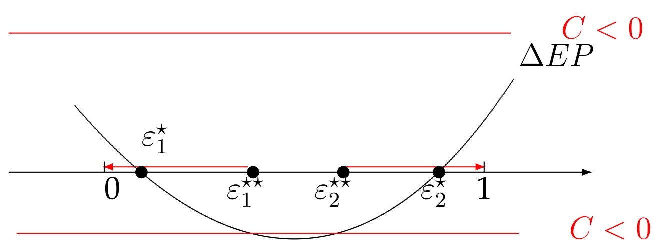


Figure 5

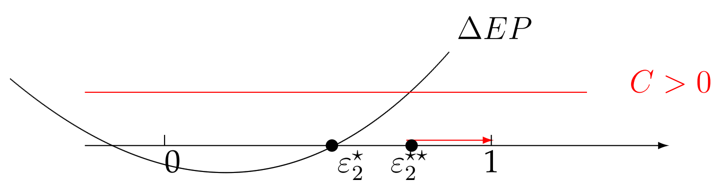


Figure 6

In any PDs, there always exists a non empty support of basic probabilities that induces a trustable strategy. This justifies our methodology. □

THEOREM 4 *In any PD, defection is always a trustable strategy.*

Proof. See Figure 4, 5 and 6. □

In Figure 4, 5 and 6, it is shown that if the basic probability of cooperation is very close to 0, then cooperation can be selected since it provides more objective expected payoff than defection for certainty (the mixed strategy Nash equilibrium $\alpha^\star = 1$). Each player can trust his opponent to be able to select cooperation.

COROLLARY 1 *Consider any PD, the support of cooperation with a unique random mechanism after each history is characterized as follows:*

1. *If 4 or 3 payoffs are* ≥ 0 *then cooperation emerges as a trustable strategy* $\forall \tilde{\varepsilon} \in \overline{J}_\varepsilon := [\varepsilon_2^{\star\star}, 1]$ *and defection as a trustable strategy over its complementary interval over* $[0, \varepsilon_2^{\star\star}]$*.*

2. *If 2 payoffs are* ≥ 0*, then cooperation emerges as a trustable strategy* $\forall \tilde{\varepsilon} \in J_\varepsilon := [0, \varepsilon_1^\star] \cup [\varepsilon_2^\star, 1]$*. Defection as a trustable strategy emerges over its complementary interval over* $[\varepsilon_1^\star, \varepsilon_2^\star]$*.*

3. *If 1 or 0 payoff are* ≥ 0*, then cooperation emerges as a trustable strategy either* $\forall \tilde{\varepsilon} \in \underline{J}_\varepsilon := [0, \varepsilon_1^\star]$*. Defection as a trustable strategy over its complementary interval over* $[\varepsilon_1^\star, 1]$*.*

Proof : it is a consequence of Theorem 3.

3 Our model matches experiments

Players use all the information the Bayes equivalent of the PD provides prior to playing.

3.1 Formation of subjective probabilities

All the results relative to the Bayes-equivalent of the PD, $\hat{\Gamma}_0$, are fundamental for explaining the formation of subjective probabilities. It helps solving the problem of selecting a prior probability distribution among a huge collection of possible priors. Indeed, in PDs the 'basic probability distribution' is unique. Each player uses all the information the Bayes-equivalent provides prior to playing a given PD, $\hat{\Gamma}$, see Definition 3. Since players do not communicate with each other and ignore the decision taken by the other, the PD is a game with imperfect information. Each personal player understands the game in the same way, and perfectly knows Theorem 1, Theorem 2, Theorem 3 and Theorem 4.

3.2 The Bayes-equivalent compatible subjective probabilities

3.2.1 Some general considerations

Denote μ the subjective probability of cooperation. By Definition 2: $h \in H \backslash Z := \{\emptyset, 1\}$, $\exists \lambda \in \mathbb{R}_+ \exists \xi \leq 1 \mid \mu(h) := \lambda \varepsilon(h) + \xi$. In this general notation, ξ allows to deform an objective probability equals to 0 into a subjective one equals to 1, with $\xi = 1$. Similarly, it also allows to deform an objective probability equals to 1 into a subjective probability equals to 0 with $\xi = -\lambda$, or $\lambda = \xi = 0$. These are two extreme cases and without loss of generality, for the remaining of the paper it will be assumed $\xi = 0$. Definition 2 reduces to $\exists \lambda \in \mathbb{R}_+^\star \mid \mu(h) := \lambda \varepsilon(h)$.

DEFINITION 6 *Define $\Delta SEP(\mu) := SEP_{\mathcal{C}}(\mu) - SEP_{\mathcal{NC}}(\mu)$ the variation of subjective expected payoff between cooperation and non-cooperation and S_μ the system*

$$S_\mu : \begin{cases} \Delta SEP_i(\mu) & \geq & 0 \\ \Delta SEP_j(\mu) & \geq & 0 \end{cases}$$

DEFINITION 7 *A subjective distortion λ of an objective probability ε is any subjective distortion such that $\Delta SEP(\mu) \geq 0$, where $\mu = \lambda \varepsilon$.*

DEFINITION 8 *A rational subjective distortion $\lambda^\star$ of an objective probability ε is any subjective distortion such that $\Psi(\mu^\star, \alpha^\star) \geq 0$, where $\mu^\star = \lambda^\star \varepsilon$.*

DEFINITION 9 *$\mu^\star$ is called a rational subjective probability.*

In PDs such a rational subjective distortion leads to select the support of rational subjective probabilities such that $\Psi(\mu^\star, \alpha^\star) \geq 0 \iff \Delta(\mu^\star) \geq C$, where C is the payoff of defection (in the PDs) reached for certainty in other alternative optimal strategies.

DEFINITION 10 *We call rational subjective expected payoffs strategy Nash equilibrium any support of rational subjective probabilities over which the choice of a given pure strategy a is made by using the criterion $C^\star := \max\{\Delta SEP_a(\mu^\star), EU(\alpha^\star)\}$.*

Note that according to Definition 1 any rational subjective expected payoffs strategy Nash equilibrium supports a trustable strategy which is always selected. Players extend the objective basic probability ε to rational subjective probability Nash Equilibrium $\mu^\star$. Note that if μ is the subjective probability of cooperation, players may be pessimistic for $\lambda^\star < 1$ or optimistic for $\lambda^\star \geq 1$. The same reasoning applies $\forall \tilde{\varepsilon} \in [\varepsilon_2^\star, 1]$

(Figure 5) or $[\varepsilon_2^{\star\star}, 1]$ (Figure 6). Similarly non-cooperation is possible as a trustable strategy explained by a rational subjective probability strategy Nash equilibrium over $[\varepsilon_1^{\star}, \varepsilon_1^{\star\star}] \cup [\varepsilon_2^{\star\star}, \varepsilon_2^{\star}]$, or by the mixed strategy Nash equilibrium (the expected utility criterion) over $[\varepsilon_1^{\star\star}, \varepsilon_2^{\star\star}]$ (Figure 5) and over $[0, \varepsilon_2^{\star\star}]$ (Figure 6). We have shown that the key concepts in presence are rationality and Nash equilibrium.

3.2.2 The methodology used to explain experiments

Let us detail the procedure by which our model fits data. According to Definition 2 (with $\xi = 0$) and S_μ, we hereafter restrict attention to the case of common rational subjective probabilities (i.e., a random device with replacement after each history). Assume that in experiments, cooperation rate is observed at $\hat{\mu} = \lambda \varepsilon_i^{\star\star}$, $i = 1, 2$ the two roots. We obtain $\lambda = \hat{\mu}/\varepsilon_l^{\star\star}$ which always fits data since $\lambda \in \mathbb{R}_+^{\star}$. Indeed, if $\hat{\mu} < \varepsilon_l^{\star\star}$ then $\hat{\lambda} < 1$ (and reciprocally). The following Figures 4 and 5 help understand our computations.

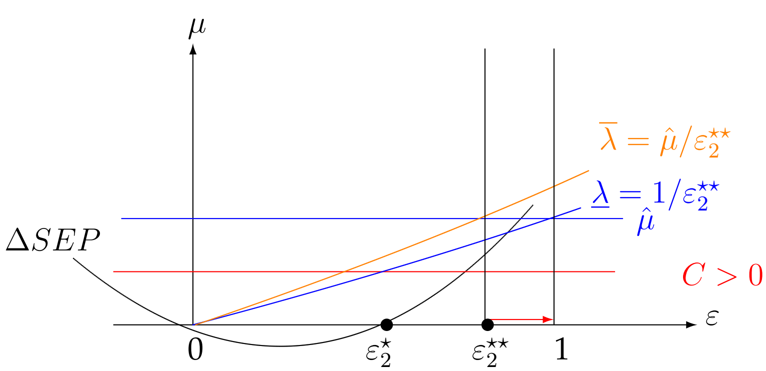


Figure 7

Figure 7 deals with the case of 4 or 3 positive payoffs. In those cases, $\forall \tilde{\varepsilon} \in [\varepsilon_2^{\star\star}, 1]$ corresponds a unique value of the rational subjective distortion λ of the objective probability $\tilde{\varepsilon}$ that belongs over the interval $[1/\varepsilon_2^{\star\star}, \hat{\mu}/\varepsilon_2^{\star\star}]$. Similarly for Figure 8.

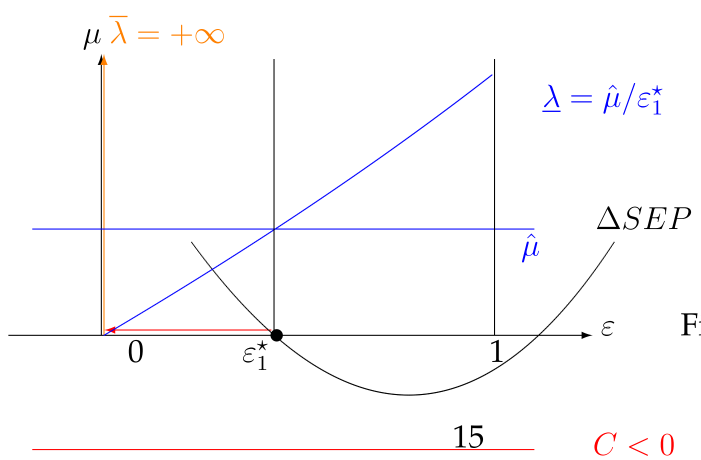


Figure 8

3.2.3 Application to our model

In the Bayes-equivalent of the PD, Nature chooses a sub-game according to a given random draw which provides the players with a basic probability distribution. From a theoretical point of view, a basic probability has the following economic interpretation. Consider $P(\emptyset) = i, P(1) = j$. To keep notation simple we don't use the upper-script $\star$s to characterize rational subjective probabilities or distortions. Denote μ the player i's rational subjective probability to cooperate and M the player j's rational subjective probability to cooperate. Denote λ the player i's rational subjective distortion of the basic probability of cooperation, and Λ the player j's one. Players formulate first-order reasoning and second-order reasoning.

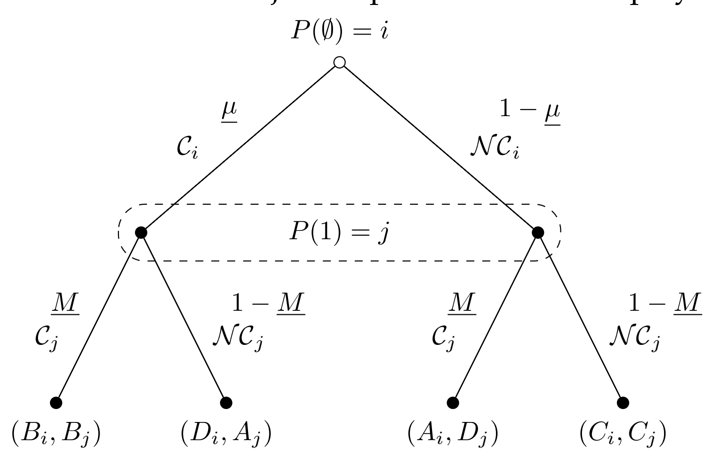
DEFINITION 11 *First-order reasoning: player i forms subjective probabilities about j's choice.*

DEFINITION 12 *Second-order reasoning: i forms subjective probabilities about j's subjective probabilities about i's choice.*

The PD from player i's point of view Player i forms the subjective probability $\underline{M} = \underline{\lambda}\varepsilon_1$ that player j chooses cooperation after $h = 1$. Player i takes into account that player j forms the subjective probability $\underline{\mu} = \underline{\Lambda}\varepsilon_\emptyset$ that Player i chooses cooperation after $h = \emptyset$.

——— insert Decision tree 2 here. ———

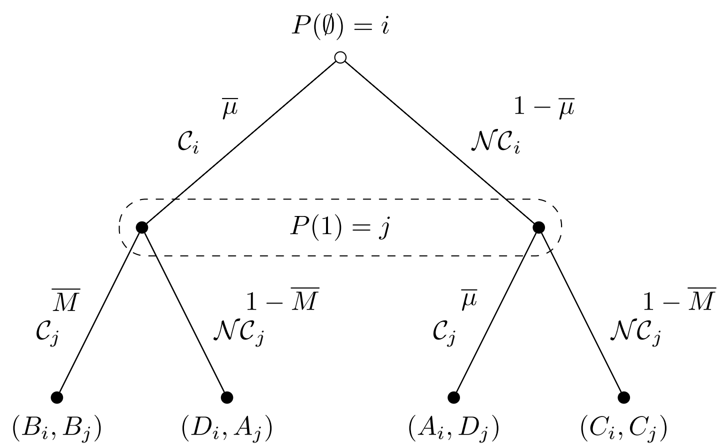
Figure 2: Decision Tree of $\hat{\Gamma}$ with subjective probabilities from player i's point of view



The PD from player j's point of view Player j forms the subjective probability $\overline{\mu} = \overline{\Lambda}\varepsilon_\emptyset$ that player i chooses cooperation after $h = \emptyset$. Player j takes into account the player i forms the subjective probability $\overline{M} = \overline{\lambda}\varepsilon_1$ that player j chooses cooperation after $h = 1$.

——— insert Figure 3 here. ———

Figure 3: Decision Tree of $\hat{\Gamma}$ with subjective probabilities from player j's point of view



3.3 Subjective expected payoffs strategies

We now analyze the PD with imperfect information in which players select strategies according to first-order reasoning and second-order reasoning in terms of their subjective probabilities. According to Definition 10, for cooperation to emerge in PDs, the following system $S_{\mu M}$ must be satisfied.

$$\mathcal{S}_{\mu M} : \begin{cases} \Delta SEP_i(\underline{\mu}, \underline{M}) & \geq & 0, \\ \Delta SEP_j(\overline{\mu}, \overline{M}) & \geq & 0. \end{cases}$$

3.3.1 First-order reasoning

LEMMA 1 *According to Property 3, given the properties of the Bayes-equivalent of the PD, $\hat{\Gamma}_0$, if players only formulate first-order reasoning in terms of their subjective probabilities in the PD, $\hat{\Gamma}$, then non cooperation is the only one subjective expected payoffs strategy Nash equilibrium.*

Proof. $\Delta SEP_i^{NC}(\underline{M}) := SEP_{\mathcal{NC}_i}(\underline{M}) - SEP_{\mathcal{C}_i}(\underline{M})$ and

$$\Delta SEP_j^{NC}(\overline{\mu}) := SEP_{\mathcal{NC}_j}(\overline{\mu}) - SEP_{\mathcal{C}_j}(\overline{\mu}), i = 1, 2, i \neq j.$$

$$\begin{cases} \Delta SEP_i^{NC}(\underline{M}) \geq 0 \\ \Delta SEP_j^{NC}(\overline{\mu}) \geq 0, \end{cases} \iff \begin{cases} (1-\underline{M})(A-B) + \underline{M}(C-D) \geq 0, \\ (1-\overline{\mu})(A-B) + \overline{\mu}(C-D) \geq 0. \end{cases}$$

According to Property 2, this system is always satisfied. □

PROPOSITION 1 *Given the properties of the Bayes-equivalent of the PD,* $\hat{\Gamma}_0$ *and depending on payoffs, if players only formulate first-order reasoning in terms of their subjective probabilities in the PD,* $\hat{\Gamma}$*, then non-cooperation is always selected, but may be a trustable or a non trustable strategy.*

Proof is given in Appendix C. If players only formulate first-order reasoning prior to formulating rational subjective probabilities then they play bilateral non-cooperation, but they cannot always trust each other to play it. For that reason, we investigate the first-order reasoning and the second-order reasoning.

3.3.2 1st-order reasoning and 2nd-order reasoning

Consider that players formulate first-order reasoning and second-order reasoning. Define $\Delta SEP_i(\underline{\mu}, \underline{M}) := SEP_{\mathcal{C}_i}(\underline{\mu}, \underline{M}) - SEP_{\mathcal{NC}_i}(\underline{\mu}, \underline{M})$ and $\Delta SEP_j(\overline{M}, \overline{\mu}) := SEP_{\mathcal{C}_j}(\overline{M}, \overline{\mu}) - SEP_{\mathcal{NC}_j}(\overline{M}, \overline{\mu})$. Replace into $S_{\mu M}$ the corresponding expected payoffs and obtain

$$S_{\mu M} : \begin{cases} (C - A + \underline{\mu}(A + B - C - D))\underline{M} + \underline{\mu}(C + D) - C \geq 0, & (1) \\ (C - A + \overline{M}(A + B - C - D))\overline{\mu} + \overline{M}(C + D) - C \geq 0. & (2) \end{cases}$$

System $S_{\mu M}$ is more complex than system S_0 since there are two non-linear inequalities with four unknowns. To solve it, we assume players formulate common subjective probabilities. There are two cases: the first one is a random device after each history, and the second one is a unique random device after each history.

3.3.3 Common subjective probabilities with a random device after each history

Assume that the random mechanism in the Bayes-equivalent of the PD consists in a different random draw[2] after each history, i.e., $\varepsilon_\emptyset \neq \varepsilon_1$. In that particular case the subjective probabilities according to the first-order reasoning and the second-order reasoning

[2]For example: throw a dice after $h = \emptyset$ and choose a card in a 52 cards game after $h = 1$, or simply a random draw after $h = \emptyset$ without replacement after $h = 1$.

are the following $\underline{\mu} = \underline{\Lambda}\varepsilon_\emptyset, \underline{M} = \underline{\lambda}\varepsilon_1, \overline{\mu} = \overline{\Lambda}\varepsilon_\emptyset, \overline{M} = \overline{\lambda}\varepsilon_1$. By the assumption of common subjective probabilities, we have $\underline{\Lambda} = \underline{\lambda} = \overline{\Lambda} = \overline{\lambda} = \lambda$. Consequently,

$$\begin{cases} \underline{\mu} &= \overline{\mu} &= \lambda\varepsilon_\emptyset &= \mu_\emptyset, \\ \overline{M} &= \underline{M} &= \lambda\varepsilon_1 &= \mu_1. \end{cases}$$

According with these new notations, system $S_{\mu M}$ becomes $S_{\mu_\emptyset \mu_1}$

$$S_{\mu_\emptyset\mu_1}: \begin{cases} (C - A + \mu_\emptyset(A + B - C - D))\mu_1 + (C + D)\mu_\emptyset - C &\geq 0, \\ (C - A + \mu_1(A + B - C - D))\mu_\emptyset + (C + D)\mu_1 - C &\geq 0. \end{cases}$$

REMARK 3 *Note that system $S_{\mu_\emptyset\mu_1}$ is exactly the same as S_0.*

THEOREM 5 *In any PD with symmetric payoffs, 2 players, 2 actions, there always exists a subjective expected payoff strategy Nash equilibrium that supports cooperation. More precisely $\exists I_{\mu_\emptyset} \neq \emptyset, I_{\mu_\emptyset} := \left[\underline{\mu}_\emptyset, \overline{\mu}_\emptyset\right] \subset [0,1]\,, \exists I_{\mu_1} \neq \emptyset, I_{\mu_1} := \left[\underline{\mu}_1, \overline{\mu}_1\right] \subset [0,1]\,, \mid\ \forall \tilde{\mu}_\emptyset \in I_{\mu_\emptyset}, \forall \tilde{\mu}_1 \in I_1$, we have $\Delta SEP_i(\tilde{\mu}_\emptyset, \tilde{\mu}_1) \geq 0$, and $\Delta SEP_j(\tilde{\mu}_1, \tilde{\mu}_\emptyset) \geq 0$, where $\underline{\mu}_\emptyset, \overline{\mu}_\emptyset, \underline{\mu}_1, \overline{\mu}_1$ are interval-specific thresholds.*

Proof. See Theorem 1. □

THEOREM 6 *Typology of PDs. Denote J_μ the support of subjective probabilities over which system $S_{\mu_\emptyset\mu_1}$ is satisfied, where $m, \underline{m}, \overline{m}, \hat{m}$ are interval-specific thresholds.*

1. *4 or 3 positive payoffs, the unique support of basic probabilities is $J_\mu := [\ m, 1\,[$.*

2. *2 positive payoffs, the supports of basic probabilities are $J_{\underline{\mu}} :=]\,0, \underline{m}\ ]$ and $J_{\overline{\mu}} := [\ \overline{m}, 1\,[$.*

3. *1 or 0 positive payoffs, the unique support of basic probabilities is $J_{\hat{\mu}} :=]\,0, \hat{m}\ ]$.*

Proof. See Theorem 2. □

Moreover, by Theorems 5 and 6, if players simultaneously formulate first-order reasoning and second-order reasoning prior to formulating rational subjective probabilities, then players can trust each other and play bilateral cooperation. Depending on payoffs, if the PD has 4 or 3 positive payoffs, then the rational subjective distortion λ of the basic probability ε is less that 1, i.e. players underestimate the possibility of cooperation. On the opposite, if there are 4 or 3 negative payoffs, player may overestimate ε. Consequently, optimism or pessimism are trust compatible endogenous attitudes.

3.3.4 Common subjective probabilities with a unique random device

From Definition 10, reversing some or all inequality signs into system $S_{\mu_\emptyset \mu_1}$ would generate 4 possible Nash equilibria. Even if this is empirically observed that all combinations of strategies are played, from a theoretical point of view it is interesting to restrict the set of subjective payoff strategy Nash equilibria. Assume the same random draw after each history, we have $\varepsilon_\emptyset = \varepsilon_1 = \varepsilon$, and consequently, $\mu_\emptyset = \mu_1 = \mu$. System $S_{\mu_\emptyset \mu_1}$ reduces to S_μ as a single inequality $(C-A+\mu(A+B-C-D))\mu+(C+D)\mu-C \geq 0$. Consider it as an equality and denote $\mu_1^\star$ and $\mu_2^\star$ the two distinct solutions.

THEOREM 7 *Consider any PD, the support of cooperation with common subjective probabilities and a unique random mechanism after each history is characterized as follows:*

1. *If 4 or 3 payoffs are* ≤ 0 *then cooperation emerges* $\forall \tilde{\mu} \in \overline{J}_\mu := [\mu_2^\star, 1]$ *and defection over its complementary interval over* $[0, \mu_2^\star]$.
2. *If 2 payoffs are* ≤ 0*, then cooperation emerges* $\forall \tilde{\mu} \in J_\mu := [0, \mu_1^\star] \cup [\mu_2^\star, 1]$*. Defection emerges over its complementary interval over* $[\mu_1^\star, \mu_2^\star]$.
3. *If 1 or 0 payoff are* ≤ 0*, then cooperation emerges either* $\forall \tilde{\mu} \in \underline{J}_\mu := [0, \mu_1^\star]$*. Defection over its complementary interval over* $[\mu_1^\star, 1]$.

Proof. See Theorem 2. □

According to Theorem 7, there are only two subjective expected payoff strategy Nash equilibria: defection or cooperation. It is unimportant wether or not players agree on the nature of the random device. Indeed, if one belief in two different random devices, while the other one in a single one after each history, then the first player acts according to the set-valued functions defined by (17) and (18) (see Appendix A) while the other one acts on the restriction of these zone over the 45 degree line.

3.4 Applications to experiments

All the following simulations have been computed with Mathematica. Clark and Sefton (2001)[9] p. 53 study sequential games with 4 positive payoffs. According to Theorem 6 , Table 1 shows how our results match experiments.

——————— insert Table 1 here ———————

Table 1: % of cooperation in 1-shot PD: Clark and Sefton

PD Game #	# of < 0 payoffs	A	B	C	D	$\hat{\mu}$	$\tilde{\varepsilon}_2 \in [\varepsilon_2^{\star\star}, 1]$	$\lambda \in \left[\hat{\mu}, \frac{\hat{\mu}}{\varepsilon_2^{\star\star}}\right]$
1	0	500	400	100	0	0.425	$[0.7215, 1]$	$[0.425, 0.58]$

Marcus (2009)[23] experiments one-shot PD with 0, 1, 2, 3 or 4 negative payoffs. According to Theorem 6 Table 2 shows that our results match experiments.

———————— insert Table 2 here ————————

Table 2: % of cooperation in 1-shot PD: Marcus

PD Game #	# of < 0 payoffs	A	B	C	D	$\hat{\mu}$	$\tilde{\varepsilon}_2 \in [\varepsilon_2^{\star\star}, 1]$	$\lambda \in \left[\hat{\mu}, \frac{\hat{\mu}}{\varepsilon_2^{\star\star}}\right]$
1	0	150	110	50	30	0.32	$[0.80, 1]$	$[0.32, 0.40]$
2	1	110	70	10	-10	0.22	$[0.71, 1]$	$[0.22, 0.31]$
							$\tilde{\varepsilon} \in [0, \varepsilon_1^{\star}] \cup [\varepsilon_2^{\star}, 1]$	$\lambda \in [\ \frac{\hat{\mu}}{\varepsilon_1^{\star}}, \infty\ [\ \cup \left[\hat{\mu}, \frac{\hat{\mu}}{\varepsilon_2^{\star}}\right]$
3	2	70	30	-30	-50	0.27	$[0, 0.21] \cup [0.79, 1]$	$[\ 1.28, +\infty\ [\ \cup [0.27, 0.34]$
							$\tilde{\varepsilon} \in [0, \varepsilon_1^{\star}]$	$\lambda \in [\ \frac{\hat{\mu}}{\varepsilon_1^{\star}}, \infty\ [$
4	3	30	-10	-70	-90	0.24	$[0, 0.36]$	$[\ 0.67, +\infty\ [$
5	4	-10	-50	-110	-130	0.32	$[0, 0.41]$	$[\ 0.77, +\infty\ [$

Rapoport and Chammah (1970)[27] study the behavior in PD in relation to the payoffs. They build an index derived from the payoffs and underline the related behavior to this index through the experiments of 7 different PD (among 12), played by 70 pairs of males student 300 times. Each series of game is built in order to study the impact on cooperation of the variation of one payoff at the time. All the games chosen by Rapoport and Chammah have the property that the pure strategy $\mathcal{N}$ash Equilibrium corresponds to a negative payoff $C < 0$. For that reason, these experiments are very interesting for our purpose. Results are summed up in Tables 3, 4 and 5. According to Theorem 7, two negative payoffs involve two supports of subjective probabilities $[0, \mu_1^{\star}] \cup [\mu_2^{\star}, 1]$. In games 1, 11 and 3, A, C, D are constants and B changes $1, 5, 9$

———————— insert Table 3 here ————————

In games 4, 3, 5 B, C are constant but A and D evolve symmetrically.

———————— insert Table 4 here ————————

Table 3: Rapoport and Chammah (1970) P-D Experiments versus our theoretical results

PD Game #	# of < 0 payoffs	A	B	C	D	$\hat{\mu}$	$\tilde{\varepsilon} \in [0, \varepsilon_1^\star] \cup [\varepsilon_2^\star, 1]$	$\lambda \in [\frac{\hat{\mu}}{\varepsilon_1^\star}, \infty[\cup \left[\hat{\mu}, \frac{\hat{\mu}}{\varepsilon_2^\star}\right]$
1	2	10	9	-1	-10	0.73	$[0, 0.05] \cup [0.68, 1]$	$[14.99, +\infty[\cup [0.73, 1.07]$
11	2	10	5	-1	-10	0.63	$[0, 0.05] \cup [0.80, 1]$	$[13.07, +\infty[\cup [0.63, 0.79]$
3	2	10	1	-1	-10	0.46	$[0, 0.04] \cup [0.95, 1]$	$[9.63, +\infty[\cup [0.46, 0.48]$

Table 4: Rapoport and Chammah (1970) P-D Experiments versus our theoretical results

PD Game #	# of < 0 payoffs	A	B	C	D	$\hat{\mu}$	$\tilde{\varepsilon} \in [0, \varepsilon_1^\star] \cup [\varepsilon_2^\star, 1]$	$\lambda \in [\frac{\hat{\mu}}{\varepsilon_1^\star}, \infty[\cup \left[\hat{\mu}, \frac{\hat{\mu}}{\varepsilon_2^\star}\right]$
4	2	2	1	-1	-2	0.66	$[0, 0.21] \cup [0.79, 1]$	$[3.12, +\infty[\cup [0.66, 0.84]$
3	2	10	1	-1	-10	0.46	$[0, 0.04] \cup [0.95, 1]$	$[9.63, +\infty[\cup [0.46, 0.48]$
5	2	50	1	-1	-50	0.27	$[0, 0.01] \cup [0.99, 1]$	$[27.27, +\infty[\cup [0.27, 0.2727]$

In the last series of games 2, 12, 3 A, B, D are constant and the pure strategy Nash equilibrium $C < 0$ evolves $-9, -5, -1$.

———— insert Table 5 here ————

Table 5: Rapoport and Chammah (1970) P-D Experiments versus our theoretical results

PD Game #	# of < 0 payoffs	A	B	C	D	$\hat{\mu}$	$\tilde{\varepsilon} \in [0, \varepsilon_1^\star] \cup [\varepsilon_2^\star, 1]$	$\lambda \in [\frac{\hat{\mu}}{\varepsilon_1^\star}, \infty[\cup \left[\hat{\mu}, \frac{\hat{\mu}}{\varepsilon_2^\star}\right]$
2	2	10	1	-9	-10	0.77	$[0, 0.31] \cup [0.95, 1]$	$[2.44, +\infty[\cup [0.77, 0.80]$
12	2	10	1	-5	-10	0.59	$[0, 0.20] \cup [0.95, 1]$	$[2.92, +\infty[\cup [0.59, 0.62]$
3	2	10	1	-1	-10	0.46	$[0, 0.04] \cup [0.95, 1]$	$[9.63, +\infty[\cup [0.46, 0.48]$

Fudenberg, Rand and Dreber (2012) manipulate A and B.

———— insert Table 6 here ————

———— insert Table 7 here ————

4 Extensions to other famous games

How to start with other games? This section extends our methodology in the direction of Zero-sum games, War and Peace, BoS, Coordination games, Head and Tail, Hawk Dove and Centipede Game. We compute the rational subjective expected payoff strategy Nash equilibrium for these famous static strategic games. We do not present all the

Table 6: Fudenberg, Rand and Dreber AER 2012 P-D Perfect adjusted results

PD Game #	# of < 0 payoffs	A	B	C	D	$\hat{\mu}$	$\tilde{\varepsilon}_2 \in [\varepsilon_2^{\star\star}, 1]$	$\lambda \in \left[\hat{\mu}, \frac{\hat{\mu}}{\varepsilon_2^{\star\star}}\right]$
1	1	3	1	0	-2	0.54	$[0.83, 1]$	$[0.54, 0.65]$
2	1	4	2	0	-2	0.75	$[0.75, 1]$	$[0.75, 1]$

Table 7: Fudenberg, Rand and Dreber AER 2012 P-D Partial adjusted results

PD Game	# of < 0	A	B	C	D	$\hat{\mu}$	$\tilde{\varepsilon}_2 \in [\varepsilon_2^{\star\star}, 1]$	$\lambda \in \left[\hat{\mu}, \frac{\hat{\mu}}{\varepsilon_2^{\star\star}}\right]$
3	1	5	3	0	-2	0.79	$[0.70, 1]$	$[0.79, 1.12]$
4	1	8	6	0	-2	0.76	$[0.64, 1]$	$[0.76, 1.17]$

details but only significant examples in order to exhibit the potential of our technique for other famous static strategic games.

In all the following games, we restrict attention to common subjective probabilities, i.e., system S_μ, which solution are $\mu_1^\star$ and $\mu_2^\star$. Zones of trustable strategies are the result of $\Psi(\mu, \alpha) \geq 0$.

4.1 Zero-sum games

The methodology used in the context of sequential PD can be extended to zero-sum games in normal form. Table 8 presents a 2-players, 2-actions zero-sum game which has no pure strategy Nash equilibrium. Payoffs $0 \geq A > B > C > D$ and K constant, have been chosen since this game exhibits a subjective expected payoff strategy Nash equilibrium without trust.

——————— insert Table 8 here ———————

Emergence of cooperation We show that there exists a unique subjective expected payoff strategy Nash equilibrium. The condition $\Delta SEP(\mu) := SEP_{b_i}(\mu) - SEP_{a_i}(\mu) \geq 0, i = 1, 2$ leads to the following system

$$S_\mu^{0\Sigma G} \begin{cases} -(A-B-C+D)\mu^2 - (2B+C-D)\mu + B \geq 0, & (3) \\ (A-B-C+D)\mu^2 + (2B+C-D)\mu - B + (1-2\mu)K \geq 0. & (4) \end{cases}$$

For $K = 0$, $S_\mu^{0\Sigma G}$ leads to one unique 2nd degree polynomial which roots $r = 1, 2$ are

$$\mu_r^\star := \frac{D - 2B - C \pm \sqrt{(C-D)^2 + 4AB}}{2(A-B-C+D)}.$$

Table 8: Zero-sum game

Subj. Proba. by 1	by 2	SORSUP_2 / FORSUP_1	$\underline{\mu}$ / $\underline{M}$	$1-\underline{\mu}$ / $1-\underline{M}$
SORSUP_1	FORSUP_2	Actions	a_2	b_2
$\overline{M}$	$\overline{\mu}$	a_1	A, $K-A$	C, $K-C$
$1-\overline{M}$	$1-\overline{\mu}$	b_1	D, $K-D$	B, $K-B$

FORSUP_i: First-order reasoning / subjective probability of player i,
SORSUP_i: Second-order reasoning / subjective probability of player i.

For $A = 0$, $\mu^\star_{i2} = 1$ and $\mu^\star_{i1} = 1/(1 + (C - D)/B) > 1$ for $|(C - D)/B| < 1$. In that case, the unique rational subjective expected payoff strategy Nash equilibrium is $\forall \mu \in [0, 1], N_{RSEPS} := \{(b_1, b_2)\}$. If $|(C - D)/B| > 1$ then the root is negative and $\forall \mu \in [0, 1], N_{RSEPS} := \{(a_1, a_2)\}$.

Emergence of trust As above $A = K = 0$. The mixed strategy Nash equilibrium is $0 < \alpha^\star := 1/(1 + (A - D)(B - C)) < 1$, $0 < \beta^\star := 1/(1 + (A - C)(B - D)) < 1$ and $EU1(\alpha^\star, \beta^\star) = \frac{-CD}{B-C-D}$, $EU2(\alpha^\star, \beta^\star) = \frac{CD}{B-C-D}$, so that $EU1(\alpha^\star, \beta^\star) = -EU2(\alpha^\star, \beta^\star)$. Consider $B = -5, C = -10, D = -15$ and find $\forall \varepsilon \in [0, 1]], \Psi(\varepsilon) < 0$ for player 1 as well as for player 2. There is no trust.

4.2 Aumann's game: War and Peace

Aumann (2006) [6] presents a static strategic game for which the cooperative outcome cannot be reached, whatever the pure or mixed Nash strategies. Player 1 must decide whether both she and player 2 will receive the same amount A or whether she will receive $n > 1$ times more, and player 2 will receive n times less. Simultaneously, player 2 must decide whether or not to take a punitive action, which will harm both players; if she does so, the division is cancelled, and instead, each player gets nothing.

——————————— insert Table 9 here ———————————-

Table 9: Aumann' game War and Peace

Common subj. proba by 2 / by 1		$SOSUP_2$ / $FOSUP_1$	μ / μ	$1-\mu$ / $1-\mu$
$SOSUP_1$	$FOSUP_2$	Actions	a_2	b_2
μ	μ	a_1	A / A	0 / 0
$1-\mu$	$1-\mu$	b_1	nA / $\frac{A}{n}$	0 / 0

$FOSUP_i$: First-order subjective probabilities of player i,
$SOSUP_i$: Second-order subjective probabilities of player i.

Emergence of cooperative equilibrium System S_0^μ defined as the difference between the subjective expected payoff of the cooperative action $a_i, i = 1, 2$ and the subjective expected payoff of action b_i, leads to

$$\begin{cases} \mu^2 A - \mu(1-\mu)nA & \geq & 0, \\ \mu^2 A + \mu(1-\mu)\frac{A}{n} & \geq & 0, \end{cases} \iff \begin{cases} (n+1)\mu - n & \geq & 0, \\ (n-1)\mu + 1 & \geq & 0. \end{cases} \quad \begin{cases} \mu_1^\star & = & \frac{n}{1+n}, \\ \mu_2^\star & = & \frac{1}{1-n}. \end{cases}$$

Note that payoffs are irrelevant. Define $0 < \mu^\star = 1/(1+1/n) < 1$. Since $1/(1-n) < 0$, the cooperative action is played $\forall \tilde{\mu} \in [\mu^\star, 1]$. The smaller n the higher the support of the subjective distortion of the objective basic probability that supports the Aumann's cooperative outcome.

Emergence of trust The pure or mixed strategy Nash equilibrium provides player 1 with the expected utility $EU_1 = nA$ and player 2 with $EU_1 = A/n$. The new system to solve is

$$\begin{cases} \mu^2 A - \mu(1-\mu)nA - nA & \geq & 0, \\ \mu^2 A + \mu(1-\mu)\frac{A}{n} - \frac{A}{n} & \geq & 0, \end{cases} \iff \begin{cases} (1+n)\mu^2 - \mu - n & \geq & 0, \\ (n-1)\mu^2 + \mu - 1 & \geq & 0. \end{cases}$$

$$\begin{cases} \mu_{11}^{\star\star} & = & \frac{n-\sqrt{n}\sqrt{4+5n}}{2(1+n)} & \mu_{12}^{\star\star} & = & \frac{n+\sqrt{n}\sqrt{4+5n}}{2(1+n)} \\ \mu_{12}^{\star\star} & = & \frac{1+\sqrt{-3+4n}}{2(1-n)} & \mu_{22}^{\star\star} & = & \frac{1-\sqrt{-3+4n}}{2(1-n)} \end{cases}$$

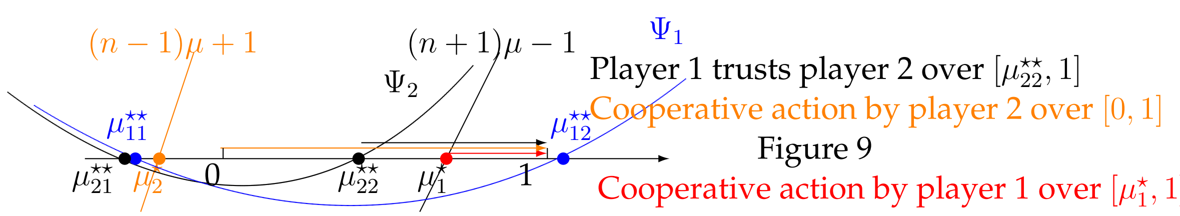


Figure 9

4.3 Battle of the Sexes

Consider the Battle of the Sexes. If player $i = 1, 2$ goes to his favorite concert (action a_i for player i while player $j \neq i$ joins him, or action b_j for the player j) then player i gets A and player j gets B and reciprocally. If they go to a separate concert, each player choose action a_i or b_i, both get C, where $A > B > C$ and $A > B \geq 0$.

Emergence of cooperation Assume that players evaluate $\Delta SEP_i(\mu) = SEP_{a_i}(\mu) - SEP_{b_i}(\mu) \geq 0$, which leads to the following system S_0^μ

$$\begin{cases} (A-B)\mu^2 + 2B\mu - B & \geq 0 \\ -(A-B)\mu^2 + 2A\mu - A & \geq 0 \end{cases}$$

which has four solutions $\mu_1^\star = -\frac{\sqrt{B}}{\sqrt{A}-\sqrt{B}} \leq 0, \mu_{11}^\star = \frac{\sqrt{B}}{\sqrt{A}+\sqrt{B}} < 1, \mu_2^\star = \frac{\sqrt{A}}{\sqrt{A}+\sqrt{B}} < 1$ and $\mu_{22}^\star = \frac{\sqrt{A}}{\sqrt{A}-\sqrt{B}} \geq 0$. Consider the particular class of games where $B = 0, \forall A, C$: $\mu_1^\star = \mu_{11}^\star = 0$. By the convexity of the first relation in system S_0^μ, the solution is exactly $\mu^\star = 0$ for player 1. By the concavity of the second relation, the solution is $\mu^\star = 1$. Consequently, there is a unique subjective expected payoff strategy Nash equilibrium in $\mu^\star = 1$: Player 2 reneges to go to his first best concert. The set of optimal actions is $N_{RSEPS} = \{(a_1, b_2)\}$. Note that if both players evaluate $\Delta SEP_i(\mu) = SEP_{a_i}(\mu) - SEP_{b_i}(\mu) \geq 0$ then player 1 reneges. One can also see that for $A = 0$ results are reversed.

Emergence of trust The mixed strategy Nash equilibrium leads to $\alpha^\star = (B-C)/(A-C+B-C)$ for player 1 and $\beta^\star = (A-C)/(A-C+B-C)$ for player 2. The expected utility evaluated in terms of the mixed strategy is $EU_i = (AB - C^2)/(A-C+B-C)$. If $B = 0$ then $EU_i < 0$ and by definition all subjective probabilities are rational. For $A = 0$ it remains the same. Each player can trust each other.

4.3.1 Coordination games

Emergence of cooperation Similarly, consider the coordination game where payoffs are A for each player if both go to player's 1 concert, $B = 0$ if they both go to player's 2 concert and $C < B$ if they go alone, where $A > B > C$. Assume that the players evaluate $\Delta SEP_i(\mu) = SEP_{a_i}(\mu) - SEP_{b_i}(\mu) \geq 0$. System S_0^μ reduces to only one relation, which is the same convex relation of BoS. If $B = 0$ then $\mu_1^\star = \mu_{11}^\star = 0$. The subjective expected payoff strategy Nash equilibrium is unique and both players go to player's 1 concert. Again, reversing the criterium leads to the opposite conclusion. Both players go to player's 1 concert.

Emergence of trust The mixed strategy Nash equilibrium leads to $\alpha^\star = \beta^\star = (A - C)/(A - C + B - C)$. The expected utility evaluated in terms of the mixed strategy is $EU_i = (AB - C^2)/(A - C + B - C)$. If $B = 0$ then $EU_i < 0$ and by definition all subjective probabilities are rational. Each player can trust each other.

This game can be extended to Stag and Hare or Assurance Game or Trust Dilemma. Stag and Hare describes the conflict between safety and social cooperation. JJ Rousseau was the first to point out this game in which 2 hunters cannot communicate and individually have to choose to hunt a stag or a hare. Assume that payoffs are ranked as follows $A > B \geq C > D$. If both players hunt a stag, they have to cooperate and each gets A. If one hunts a stag while the other hunts a hare then the first player gets D while the other gets B. If both hunt a hare, then each gets C. The set of pure strategy Nash equilibrium is $\mathcal{N}_p = \{(S_1, S_2); (H_1, H_2)\}$. Assume that the players evaluate $\Delta SEP_i(\mu) = SEP_{s_i}(\mu) - SEP_{h_i}(\mu) \geq 0$. System S_0^μ reduces to only one relation $(A + B - C - D)\mu^2 + (-B + 2C + D)\mu - C \geq 0$. If $A = -(B - D)^2/4C$ the previous inequality admits a single root and both players hunt a stag.

4.4 Head and Tail

The game Head and Tail is a particular case of the zero-sum game of subsection 4.1, where $k = 0, A = B = -1$ and $C = D = 1$. System S_0^μ has only one point solution $\mu^\star = 1/2$, which is exactly the mixed strategy Nash equilibrium. There is neither cooperation nor trust in that game.

4.5 Hawk Doves or Tragedy of commons

Two players are fighting to share some benefit. Each can act as a hawk or a dove. The best outcome for each is reached when she acts like a hawk A when the other acts like a dove C. If both act like a Dove each gets B. If both act like a hawk each gets D, where $A > B > C > D$.

Emergence of cooperation Assume that the players evaluate $\Delta SEP_i(\mu) = SEP_{\mathcal{D}_i}(\mu) - SEP_{\mathcal{H}_i}(\mu) \geq 0$. In the case of common subjective probabilities with the same random mechanism for each player, system S_0^{μ} leads to a unique convex relation which has two distinct roots.

$$\mu_k^{\star} := \frac{A - C + 2B \pm \sqrt{(A-C)^2 + 4BD}}{2(A + B - C - D)}, k = 1, 2.$$

If $D = 0$, then $0 < \mu_1^{\star} < 1$ and $\mu_2^{\star} = 1$, by the convexity of the relation, $\forall \mu \in [0, \mu_1^{\star}]$, players choose to act as a Dove. Moreover $\forall \mu \in [\mu_1^{\star}, 1]$, players choose to act as a Hawk. Note that Hawk Doves and the Tragedy of commons are the same games, see Hardin's in Diekert (2012)[12] p. 1779.

Emergence of trust The mixed strategy Nash equilibrium is $\alpha^{\star} = \beta^{\star} = (A - B)/(A - B + C - D)$. The expected utility evaluated in terms of the mixed strategy is $EU_i = (AC - CD)/(A - B + C - D)$. If $D = 0$ then $EU_i > 0$. $\Delta SEP_i(\mu) - EU_i \geq 0$ leads to

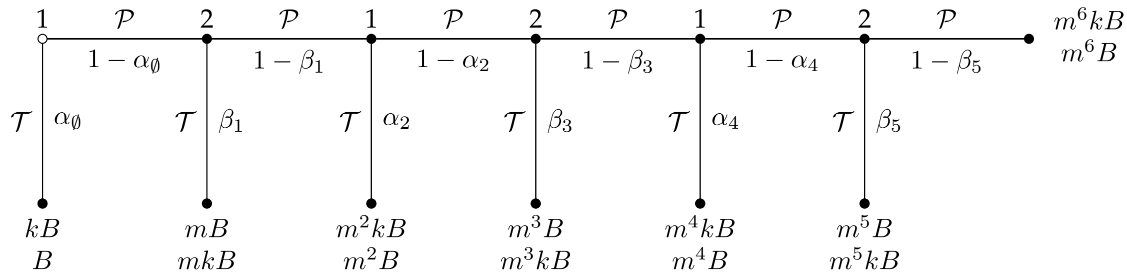
$$\mu(1-\mu)(A + C) + (1-\mu)^2 B - (AC)/(A - B + C) \geq 0$$

Solutions are: $\mu_1^{\star\star} := \dfrac{C - B}{A - B + C} < 0,\ 0 < \mu_2^{\star\star} := \dfrac{A - B}{A - B + C} = \dfrac{1}{1 + \frac{C}{A-B}} < 1.$

and by definition all subjective probabilities belonging to $[\mu_2^{\star\star}, 1]$ are rational. The condition for players to trust each other to act as a Dove is $\mu_2^{\star\star} < \mu_1^{\star}$ which is possible if $A < B + C$. Consequently, each player can trust each other over $[\mu_2^{\star\star}, \mu_1^{\star}]$. If $A > B + C$ there are no trusts.

4.6 The centipede game

As in McKelvey and Palfrey (1992) [24], consider two piles of money are on the table. One pile $k \times B$ is larger by a factor $k > 1$ than the other, denoted B . There are two players, $N := \{1, 2\}$, each of whom alternately gets a turn in which he can choose either to take T the larger of the two piles of money or to pass, P. The set of actions is $\mathcal{A}_i := \{T, P\}$. When one player takes, the game ends, with the player whose turn it is getting the large pile and the other player getting the small pile. On the other hand, whenever a player passes, both piles are multiplied by some fixed amount $m > 1$, and the play proceeds to the next player. The finite number of moves to the game is known in advance to both players. The next Figure presents the Bayes equivalent of the game, where α_h or $\beta_{h'}$ stand for the basic probability to take the pile.



Emergence of cooperation In the Bayes equivalent of the centipede game with common subjective probabilities (the same random mechanism after each period), $\forall h \in H, \varepsilon_h = \varepsilon_{h'} = \varepsilon$. The player function $P(h) = i \in N$ indicates that player i chooses an action $a(h)$, $h \in H$. Assume that the players evaluate $\Delta SEP_1(\varepsilon) = SEP_{\mathcal{P}_{P(\emptyset)=1}}(\varepsilon) - SEP_{\mathcal{T}_{P(\emptyset)=1}}(\varepsilon = 1) \geq 0$. It is important to note that $SEP_{\mathcal{T}_i}(\varepsilon = 1)$ corresponds to the certain event "$P(h) = i$ *takes it*". System S_0^ε for player 1 leads to a unique concave relation $-B(k+m(-1+\varepsilon))\varepsilon$ which has two distinct roots $\varepsilon_0 = 0$ and $\varepsilon_1 = 1-k/m$. Consequently, Pass is possible. Apply this to McKelvey and Palfrey (1992) [24] where $B = 0.1$, $k = 4$ and $m = 2$ and find $\varepsilon_1 = -1$. The the unique admissible solution is to Pass: $\varepsilon_0 = 0$. Similarly in Aumann (1988) [4] centipede game $B = 0.5$, $k = 20$ and $m = 10$ and find $\varepsilon_0 = 0$ and $\varepsilon_1 = -1$.

System S_0^ε for player 2 leads to a unique cubic relation $Bm(k + m(-1+\varepsilon))(-1+\varepsilon)\varepsilon$ which has three solutions $\varepsilon_0 = 0$, $\varepsilon_1 = 1$ and $\varepsilon_2 = 1 - k/m$. Again, Pass is possible. Apply this to McKelvey and Palfrey (1992) there are two admissible solutions Pass for ε_0 or take it for $\varepsilon_1 = 1$. Similarly in Aumann centipede game.

Both player have an incentive to pass after each history. The reason why players continue to pass and not to take it (since ε_1 is a solution) is due to the emergence of trust in the other player as we turn now to explain.

Emergence of trust In the Bayes-equivalent of the centipede game, it has been shown that $\varepsilon = 0$ is a solution. Our concept of trust enlightens players's behavior who cooperate during few periods. Indeed, as in previous subsections, the mixed strategy Nash equilibrium is $\forall h \in H, \alpha_\emptyset^\star = 1$. The game ends at the first history and payoffs are kB for player 1 and B for player 2. For trust to emerge, each player solves the following problem $\exists ? h^\star \in H \mid \Delta SEP_{P(h^\star)=i}(\varepsilon) - EU_i(\alpha_\emptyset^\star) \geq 0$. We show that during some few periods the previous inequality is not satisfied, but generally, as in most experiments, players start to choose action $\mathcal{T}(h^\star)$ for $h^\star = 4$ or 5. Let us compute $\Delta SEP_{P(2)=1}(\varepsilon) - EU_1(\alpha^\star_\emptyset) = Bk(-1+m^2(-1+\varepsilon)^2\varepsilon) \geq 0$ which is a cubic relation that admits only one real root in $\mathbb{R}$, $\varepsilon = 1.14 > 1$. We have $\Delta SEP_{P(2)=1}(\varepsilon) - EU_1(\alpha^\star_\emptyset) < 0$ over $[0,1]$, consequently $h = 2$ is not solution for player 1. Compute $\Delta SEP_{P(4)=1}(\varepsilon) - EU_1(\alpha^\star_\emptyset) = Bk(-1+m^4(-1+\varepsilon)^4\varepsilon)$. In McKelvey and Palfrey (1992), the criterion is positive over $[0.092, 0.35]$. There exists a support of bayes equivalent probabilities over which player 1 may choose to take it all. In Aumanns, the criterion is positive over the entire interval $[0,1]$. Still pass is possible since 0 belongs to the interval.

For player 2 in $h = 1$ $\Delta SEP_{P(1)=2}(\varepsilon) - EU_2(\alpha^\star_\emptyset) = B(-1-km(-1+\varepsilon)\varepsilon)$ which admits two solutions $\varepsilon_0 = \frac{1}{2} - \frac{\sqrt{-4+km}}{2\sqrt{km}}$ or $\varepsilon_1 = \frac{1}{2} + \frac{\sqrt{-4+km}}{2\sqrt{km}}$. In McKelvey and Palfrey (1992), the criterion is positive over $[0.1465, 0.8535]$ and in Aumann over $[0.005, 0.0995]$. Whatever the history, one can adjust $\hat{\mu} = \lambda\varepsilon$ so that the theory matches data.

4.7 Asymmetric payoffs in PDs

In this paper we assumed from beginning symmetric payoffs. Considering asymmetric payoffs, one can easily see that as long as the sign of each player's payoffs are the same, all the previous results in terms of cooperation hold. Considering common subjective probabilities, the only modification is that system S_0^μ leads to two different inequalities (thus two roots per each) and cooperation will emerge as a subjective expected payoff strategy Nash equilibrium. As long as ΔSEP dominates EU in terms of the mixed strategy Nash equilibrium, trust emerges in the same way, but not necessarily for all players.

5 Conclusion

The Bayes equivalent of the PD provides players with three information: Information 1. The objective basic probability distribution of the game. This is a useful information since players do not have to select the same priors among a huge collection of plausible priors. Indeed, it turns out that the basic probability distribution is unique in PDs. Information 2. The length of the support of these probabilities over which the expected payoff of cooperation dominates the one of defection. This explains the emergence of cooperation. Information 3. The length of the support of these probabilities over which the difference between the expected payoff of cooperation and defection is greater (or equal) to the expected utility evaluated at the mixed strategy Nash equilibrium. This explains the emergence of trust.

Prior to playing the PD, all players use these three information. Using information 1. players evaluate the expected payoff of an action in terms of the basic probabilities. Using information 2. They compare the expected payoff of cooperation with the one of defection. This helps players select cooperation as an expected payoff strategy Nash equilibrium (or defection depending on the support of the basic probability distribution). For cooperation to be another Nash equilibrium, players use first-order reasoning and second-order reasoning to evaluate the expected payoff of each action. Cooperation is endogenous and consistent with the full rationality principle. Using information 3. They evaluate the length of the support over which the chosen strategy is trustable. According to this support, they formulate subjective probabilities which are the result of a rational subjective distortion of the objective basic probabilities.

Each Nash equilibrium remains locally stable even if players formulate different (rational) subjective probabilities. No assumption about common beliefs in rationality are required for cooperation or defection to be a Nash equilibrium. The methodology reconciles *ex-ante* and *ex-post* rationality.

Our methodology matches observed cooperation rates. It can be applied to other static strategic games, zero sum games, the War and Peace Aumann's game, the Battle of the Sexes, the coordination game, Head and Tail and Hawk Dove / the Tragedy of Commons, the centipede game as discussed in Section 4. In Head and Tail the unique equilibrium coincides with the mixed strategy Nash equilibrium. In War and Peace by Aumann (2007)[3] the cooperative outcome can be played. Finally, there is no more paradox in PDs.

5.1 Acknowledgements

Gilles Dubail thanks the university of Paris 2, Panthéon Assas, and CRED for their support. This paper has been written while visiting the university. We particularly thank Rabah Amir, Francis Bloch, Claudine Desrieux, Roger Guesnerie, Meglena Jeleva, Daniel Leonard, Christina Pawkowitsch, Julian Revalski, Larry Samuelson, Randall Wright, George Zaccour and Yu Zhu for valuable discussions on earlier versions of this paper. We thank all participants in the CRED seminar and those of the PhD program at Paris 2, for their fruitful comments. Remaining errors are ours.

A Appendix

Proof. In S_0, replace the expected payoffs of cooperation and non cooperation by their respective expression and factorize according to the priors, see Figure 1: ε_1 for player i and $\varepsilon_\emptyset$ for player j.

$$S_0 : \begin{cases} [C - A + \varepsilon_\emptyset(A + B - C - D)]\,\varepsilon_1 + \varepsilon_\emptyset(C + D) - C \geq 0, & (5) \\ [C - A + \varepsilon_1(A + B - C - D)]\,\varepsilon_\emptyset + \varepsilon_1(C + D) - C \geq 0. & (6) \end{cases}$$

Relation (5) is relative to player $i = 1, 2$ and relation (6) to player $j \neq i$.

PROPERTY 2 *In any PD* $A + B - C - D > 0$*.*

By definition of a PD, we have $A > B > C > D$, in particular $A - C > 0$ and $B - D > 0$ thus $A + B - C - D > 0$. By Property 2, relations 5 and 6 of system S_0 are non-linear and continuous in priors, $\varepsilon_\emptyset$ and ε_1. Note that S_0 exhibits non-linear relations in term of the product $\varepsilon_\emptyset \varepsilon_1$. To our knowledge, there is no methodology to solve S_0. We propose one. Consider relation (5) as a linear inequality in ε_1 and respectively relation (6) in $\varepsilon_\emptyset$. For the rest of the paper, $s_i(\varepsilon_\emptyset) := C - A + \varepsilon_\emptyset(A + B - C - D)$ and $s_j(\varepsilon_1) := C - A + \varepsilon_1(A + B - C - D)$ denote the positive slopes of inequalities (5) and (6) in S_0 or $-\sigma_i(\varepsilon_\emptyset)$ and $-\sigma_j(\varepsilon_1)$ the negative ones. Similarly, $g_i(\varepsilon_\emptyset) := \varepsilon_\emptyset(C + D) - C$ and $g_j(\varepsilon_1) := \varepsilon_1(C + D) - C$ denote the positive intercepts of inequalities (5) and (6) or $-\gamma_i(\varepsilon_\emptyset)$ and $-\gamma_j(\varepsilon_1)$ the negative ones.

PROPERTY 3 $\forall A > B > C > D$, $\forall h = \emptyset, 1, ds_i(\varepsilon_h)/d\varepsilon_h > 0.$

Proof. From Property 2, inequality (5) reveals a slope in ε_1 which has a positive slope in $\varepsilon_\emptyset$ and conversely for inequality (6). □

For any PD, Theorem 1 proves the existence of two non empty sets of solutions $I_{\varepsilon_\emptyset}$ and I_{ε_1} such that $\forall \tilde{\varepsilon}_\emptyset \in I_{\varepsilon_\emptyset} \subset [0, 1]$ and $\forall \tilde{\varepsilon}_1 \in I_{\varepsilon_1} \subset [0, 1]$ system S_0 is satisfied. Theorem 2 characterizes these intervals over $[0, 1]$.

Proof. Theorem 1 is proved with Propositions 2, 3, 4, 5, 6 and 7 which help prove the existence of at least one (at most two) support(s) of probability over $[0, 1]$ for which system S_0 is satisfied.

PROPOSITION 2

1. *Whatever the payoffs of a given PD* $\exists \varepsilon^\star \in \;]0,1[\;|\; s_i(\varepsilon^\star) = 0$ *and* $\forall h = \emptyset, 1, \forall \tilde{\varepsilon}_h \geq \varepsilon^\star, s_i(\tilde{\varepsilon}_h) > 0$*, otherwise the slope is negative,* $-\sigma_i(\tilde{\varepsilon}_h)$*.*

2. *Whatever the payoffs of the PD,* $\exists \varepsilon^{\star\star} \in \mathbb{R} \mid g_i(\varepsilon^{\star\star}) = 0$*.*

3. $\varepsilon^\star > \varepsilon^{\star\star} \iff \frac{AD-BC}{C+D} > 0$*.*

4. $\varepsilon^{\star\star} \neq \varepsilon^\star$*.*

5.a $\exists A_{-\gamma} > B_{-\gamma} > C_{-\gamma} > D_{-\gamma} \mid \varepsilon^{\star\star} \leq 0$*. In that case* $\forall h = \emptyset, 1, d(-\gamma_i(\tilde{\varepsilon}_h))/d\tilde{\varepsilon}_h \leq 0 \Rightarrow -\gamma_i(0) < 0$*. The case* $dg_i(\tilde{\varepsilon}_h)/d\tilde{\varepsilon}_h > 0$ *and* $g_i(0) > 0$ *is impossible.*

5.b $\exists \underline{A}_g > \underline{B}_g > \underline{C}_g > \underline{D}_g, \mid 0 \leq \varepsilon^{\star\star} \leq 1$*, the slope of the intercept is negative over* $[0,1]$*.*

5.c $\exists \overline{A}_g > \overline{B}_g > \overline{C}_g > \overline{D}_g \mid 0 \leq \varepsilon^{\star\star} \leq 1$*, the slope of the intercept is positive over* $[0,1]$*.*

5.d $\exists A_g > B_g > C_g > D_g \mid 1 \leq \varepsilon^{\star\star}, \forall h = \emptyset, 1, \forall \tilde{\varepsilon}_h \in [0,1], d(-\gamma_i(\tilde{\varepsilon}_h)/d\tilde{\varepsilon}_h) \geq 0 \Rightarrow -\gamma_i(0) < 0$*. The case* $dg_i(\tilde{\varepsilon}_h)/d\tilde{\varepsilon}_h < 0 \Rightarrow g_i(0) > 0$ *is impossible.*

Proof. Proof of item 1. $\forall h = \emptyset, 1 \; s(\varepsilon_h) = 0$, is obtained for $0 < \varepsilon^\star := \frac{1}{1+\frac{B-D}{A-C}} < 1$ By Property 2 and Property 3, $s(\varepsilon_h)$ is increasing in ε_h. Consequently, $\forall h = \emptyset, 1, \forall \tilde{\varepsilon}_h \in [0, \varepsilon^\star], d(-\sigma_i(\tilde{\varepsilon}_h))/d\tilde{\varepsilon}_h > 0 \Rightarrow -\sigma_i(0) < 0$ and $\forall \tilde{\varepsilon}_h \in [\varepsilon^\star, 1], d(s_i(\tilde{\varepsilon}_h))/d\tilde{\varepsilon}_h > 0 \Rightarrow s_i(1) > 0$. For small values of $\tilde{\varepsilon}_h$ the slope is negative and positive above.

Proof of item 2 $\forall h = \emptyset, 1 \; g(\varepsilon_h) = 0$ leads to $\varepsilon^{\star\star} := \frac{C}{C+D}$.

Proof of item 3. Define $z_\varepsilon := \varepsilon^\star - \varepsilon^{\star\star}$. Compute $z_\varepsilon = \frac{AD-BC}{(A+B-C-D)(C+D)} = \frac{N_\varepsilon}{D_\varepsilon}$, where N_ε is the numerator and D_ε the denominator. By Property 2, the condition reduces to $\frac{AD-BC}{C+D} > 0$.

Proof of item 4. The condition for $\varepsilon^\star = \varepsilon^{\star\star}$ is $AD = BC$ which by Property 2 is impossible.

Proof of item 5a.

- If $C \geq 0 > D$ and $\mid C \mid < \mid D \mid$ then $C + D < 0 \Rightarrow \varepsilon^{\star\star} \leq 0$ and $\forall h = \emptyset, 1, \forall \tilde{\varepsilon}_h \in [0,1], d(-\gamma_i(\tilde{\varepsilon}_h))/d\tilde{\varepsilon}_h \leq 0 \Rightarrow -\gamma_i(0) < 0$. The slope of the intercept is negative, and the intercept never crosses the horizontal axis over $[0,1]$.

- If $C \geq 0 > D$ and $\mid C \mid > \mid D \mid$ then $C + D > 0 \Rightarrow \varepsilon^{\star\star} \geq 0$, a contradiction.

- If $0 > C > D$ or if $C > D > 0$ then $\varepsilon^{\star\star} \geq 0$, a contradiction.

Consequently, the case $\varepsilon^{\star\star} \leq 0,\ dg_i(\tilde{\varepsilon}_h)/d\tilde{\varepsilon}_h > 0$ and $g_i(0) > 0$ is impossible.

Proof of item 5b. Assume $0 \geq C \geq D \iff 0 \leq \varepsilon^{\star\star} \leq 1$ and $C + D < 0, \forall h = \emptyset, 1, \forall \tilde{\varepsilon}_h \in [0, \varepsilon^{\star\star}], dg_i(\tilde{\varepsilon}_h)/d\tilde{\varepsilon}_h \leq 0, \Rightarrow g_i(0) > 0$ and $\forall \tilde{\varepsilon}_h \in [\varepsilon^{\star\star}, 1], d(-\gamma_i(\tilde{\varepsilon}_h))/d\tilde{\varepsilon}_h \leq 0 \Rightarrow -\gamma_i(1) < 0$. This case splits into two subcases depending on $z_\varepsilon > 0$ or $z_\varepsilon < 0$.

Proof of item 5c. Assume $C \geq D \geq 0 \iff 0 \leq \varepsilon^{\star\star} \leq 1$ and $C + D > 0, \forall h = \emptyset, 1, \forall \tilde{\varepsilon}_h \in [0, \varepsilon^{\star\star}], d(-\gamma_i(\tilde{\varepsilon}_h))/d\tilde{\varepsilon}_h \geq 0 \Rightarrow -\gamma_i(0) < 0$ and $\forall \tilde{\varepsilon}_h \in [\varepsilon^{\star\star}, 1], dg_i(\tilde{\varepsilon}_h)/d\tilde{\varepsilon}_h \geq 0, \Rightarrow g_i(1) > 0$. This case splits into two subcases depending on $z_\varepsilon > 0$ or $z_\varepsilon < 0$.

Proof of item 5d. If $C \geq 0 > D$ and $\mid C \mid > \mid D \mid$ then $C + D > 0 \Rightarrow \varepsilon^{\star\star} \geq 1$ and $\forall h = \emptyset, 1, \forall \tilde{\varepsilon}_h \in [0, 1], d(-\gamma_i(\tilde{\varepsilon}_h))/d\tilde{\varepsilon}_h \geq 0 \Rightarrow -\gamma_i(0) < 0$. The slope of the intercept is positive, and the intercept never crosses the horizontal axis over $[0, 1]$. □

PROPOSITION 3 *If the value of the payoff of the pure strategy Nash equilibrium of the PD is negative, $C < 0$, and if in the Bayes-equivalent of the PD the priors $(\varepsilon_\emptyset, \varepsilon_1)$ tends toward 0^+, then System S_0 is always satisfied.*

Proof. If $C < 0$, then $-C > 0$, and if $\varepsilon_\emptyset \to 0$ and if $\varepsilon_1 \to 0^+$ then S_0 reduces to $-C > 0$. □

We now study the conditions for which $\Delta EP(\varepsilon_\emptyset, \varepsilon_1) \geq 0$. This generates a typology of PD, since system S_0 turns into 4 different expressions S_1, S_2, S_3, S_4.

PROPOSITION 4 *Recall $\forall h = \emptyset, 1, h \neq k, I_{\varepsilon_h} := [\underline{\varepsilon}_h, \overline{\varepsilon}_h] \subset [0, 1]$.*

1. *$\Delta EP(\varepsilon_k, \varepsilon_h) = s(\varepsilon_k)\varepsilon_h + g(\varepsilon_k)$ for $z_\varepsilon < 0$ and either $\tilde{\varepsilon}_h \in I^{11}_{\varepsilon_h} := [\varepsilon^\star, \varepsilon^{\star\star}]$ (1 and 5b of Proposition 2) or $\tilde{\varepsilon}_h \in I^{12}_{\varepsilon_h} := [\varepsilon^{\star\star}, 1]$ (1 and 5c of Proposition 2). For $z_\varepsilon > 0$ $\tilde{\varepsilon}_h \in I^{13}_{\varepsilon_h} := [\varepsilon^\star, 1]$ (1 and 5c of Proposition 2).*

2. *$\Delta EP(\varepsilon_k, \varepsilon_h) = s(\varepsilon_k)\varepsilon_h - \gamma(\varepsilon_k)$: for either $z_\varepsilon < 0$ and either $\tilde{\varepsilon}_h \in I^{21}_{\varepsilon_h} := [\varepsilon^\star, \varepsilon^{\star\star}]$ (1 and 5c of Proposition 2) or $\tilde{\varepsilon}_h \in I^{22}_{\varepsilon_h} := [\varepsilon^{\star\star}, 1]$ (1 and 5b of Proposition 2), or $\tilde{\varepsilon}_h \in I^{23}_{\varepsilon_h} := [\varepsilon^\star, 1]$ (1 and 5d of Proposition 2). For $z_\varepsilon > 0$ and $\tilde{\varepsilon}_h \in I^{23}_{\varepsilon_h} := [\varepsilon^\star, 1]$ (either 1 and 5a, or 1 and 5b of Proposition 2).*

3. *$\Delta EP(\varepsilon_k, \varepsilon_h) = -\sigma(\varepsilon_k)\varepsilon_h + g(\varepsilon_k)$: for either $z_\varepsilon < 0$ and $\tilde{\varepsilon}_h \in I^{31}_{\varepsilon_h} := [0, \varepsilon^\star]$ (1 and 5b of Proposition 2) or for $z_\varepsilon > 0$ and either $\tilde{\varepsilon}_h \in I^{32}_{\varepsilon_h} := [\varepsilon^{\star\star}, \varepsilon^\star]$ (1 and 5c of Proposition 2) or $\tilde{\varepsilon}_h \in I^{33}_{\varepsilon_h} := [0, \varepsilon^{\star\star}]$ (1 and 5b of Proposition 2).*

4. *$\Delta EP(\varepsilon_k, \varepsilon_h) = -\sigma(\varepsilon_k)\varepsilon_h - \gamma(\varepsilon_k)$: for $z_\varepsilon < 0$ and $\forall \tilde{\varepsilon}_h \in I^{41}_{\varepsilon_h} := [0, \varepsilon^\star]$ (either 1 and 5c or 1 and 5d of Proposition 2). For $z_\varepsilon > 0$ and either $\forall \tilde{\varepsilon}_h \in I^{41}_{\varepsilon_h} := [0, \varepsilon^\star]$ (1 and 5a of*

Proposition 2, or $\forall \tilde{\varepsilon}_h \in I^{42}_{\varepsilon_h} := [\varepsilon^{\star\star}, \varepsilon^{\star}]$ *(1 and 5b of Proposition 2) or* $\tilde{\varepsilon}_h \in I^{43}_{\varepsilon_h} := [0, \varepsilon^{\star\star}]$ *(1 and 5c of Proposition 2).*

Proof. Proposition 4 is an immediate consequence of Proposition 2. One can easily draw each corresponding case of Proposition 2 and check the existence of the corresponding interval $I^{kl}_{\varepsilon_h}, k = 1, 2, 3, 4$ and $l = 1, 2, 3$ for each system. Indications of proof are given in parenthesis for each item. □

Using Propositions 2 and 4, S_0 can take four different forms labeled S_1, S_2, S_3 or S_4.

$$\forall \tilde{\varepsilon}_h \in I^{11}_{\varepsilon_h} \text{ or } I^{12}_{\varepsilon_h} \text{ or } I^{13}_{\varepsilon_h}, \quad S_1 : \begin{cases} s_i(\varepsilon_\emptyset)\varepsilon_1 + g_i(\varepsilon_\emptyset) \geq 0, & (7) \\ s_j(\varepsilon_1)\varepsilon_\emptyset + g_j(\varepsilon_1) \geq 0. & (8) \end{cases}$$

$$\forall \tilde{\varepsilon}_h \in I^{21}_{\varepsilon_h} \text{ or } I^{22}_{\varepsilon_h} \text{ or } I^{23}_{\varepsilon_h}, \quad S_2 : \begin{cases} s_i(\varepsilon_\emptyset)\varepsilon_1 - \gamma_i(\varepsilon_\emptyset) \geq 0, & (9) \\ s_j(\varepsilon_1)\varepsilon_\emptyset - \gamma_j(\varepsilon_1) \geq 0. & (10) \end{cases}$$

$$\forall \tilde{\varepsilon}_h \in I^{31}_{\varepsilon_h} \text{ or } I^{32}_{\varepsilon_h} \text{ or } I^{33}_{\varepsilon_h}, \quad S_3 : \begin{cases} -\sigma_i(\varepsilon_\emptyset)\varepsilon_1 + g_i(\varepsilon_\emptyset) \geq 0, & (11) \\ -\sigma_j(\varepsilon_1)\varepsilon_\emptyset + g_j(\varepsilon_1) \geq 0. & (12) \end{cases}$$

$$\forall \tilde{\varepsilon}_h \in I^{41}_{\varepsilon_h} \text{ or } I^{42}_{\varepsilon_h} \text{ or } I^{43}_{\varepsilon_h}, \quad S_4 : \begin{cases} -\sigma_i(\varepsilon_\emptyset)\varepsilon_1 - \gamma_i(\varepsilon_\emptyset) \leq 0, & (13) \\ -\sigma_j(\varepsilon_1)\varepsilon_\emptyset - \gamma_j(\varepsilon_1) \leq 0. & (14) \end{cases}$$

Note that S_4 always supports defection. We now prove that the domains over which the previous systems are satisfied coincide with the intervals of Proposition 4 for system S_1 and S_4, and are restrictions of them for systems S_2 and S_3. We label them $I^{S_1}_{\varepsilon_h}, I^{S_2}_{\varepsilon_h}, I^{S_3}_{\varepsilon_h}$ and $I^{S_4}_{\varepsilon_h}$. In Proposition 5, $\varepsilon^\star_1$ and $\varepsilon^\star_2$ characterize the restrictions of the domain of validity for S_2 and S_3 and are defined in the proof.

PROPOSITION 5 *Depending on the various possible values of payoffs:*

1. *Either* $\forall \tilde{\varepsilon}_h \in \underline{I}^{S_1}_{\varepsilon_h} := I^{13}_{\varepsilon_h}, \forall \tilde{\varepsilon}_h \in \hat{I}^{S_1}_{\varepsilon_h} := I^{11}_{\varepsilon_h}$*, or* $\forall \tilde{\varepsilon}_h \in \overline{I}^{S_1}_{\varepsilon_h} := I^{12}_{\varepsilon_h}, S_1 \geq 0.$
2. *Either* $\forall \tilde{\varepsilon}_h \in \underline{I}^{S_2}_{\varepsilon_h} := [\varepsilon^\star_2, 1], \forall \tilde{\varepsilon}_h \in \hat{I}^{S_2}_{\varepsilon_h} := [\varepsilon^\star_2, \varepsilon^{\star\star}]$ *or* $\forall \tilde{\varepsilon}_h \in \overline{I}^{S_2}_{\varepsilon_h} := [\varepsilon^{\star\star}, \varepsilon^\star_1], S_2 \geq 0.$
3. *Either* $\forall \tilde{\varepsilon}_h \in \underline{I}^{S_3}_{\varepsilon_h} := [0, \varepsilon^\star_1], \forall \tilde{\varepsilon}_h \in \hat{I}^{S_3}_{\varepsilon_h} := [\varepsilon^\star_2, \varepsilon^\star]$ *or* $\tilde{\varepsilon}_h \in \overline{I}^{S_3}_{\varepsilon_h} := [0, \varepsilon^\star], S_3 \geq 0.$
4. *Either* $\forall \tilde{\varepsilon}_h \in \underline{I}^{S_4}_{\varepsilon_h} := I^{41}_{\varepsilon_h}, \forall \tilde{\varepsilon}_h \in \hat{I}^{S_4}_{\varepsilon_h} := I^{42}_{\varepsilon_h}$ *or* $\tilde{\varepsilon}_h \in \overline{I}^{S_4}_{\varepsilon_h} := I^{43}_{\varepsilon_h}, S_4 \leq 0.$

Proof. Item 1 of Proposition 5: For system S_1 to be satisfied, the 2 following conditions must hold together

$$\Delta EP_i(\varepsilon_\emptyset, \varepsilon_1) \geq 0 \iff \varepsilon_1 \ \geq \ 0 \geq \hat{\varepsilon}_1(\varepsilon_\emptyset) := \frac{-g_i(\varepsilon_\emptyset)}{s_i(\varepsilon_\emptyset)}, \tag{15}$$

$$\Delta EP_j(\tilde{\varepsilon}_1, \tilde{\varepsilon}_\emptyset) \geq 0 \iff \varepsilon_\emptyset \ \geq \ 0 \geq \hat{\varepsilon}_\emptyset(\varepsilon_1) := \frac{-g_j(\varepsilon_1)}{s_j(\varepsilon_1)}. \tag{16}$$

Consequently, according to Proposition 4 item 1, $\Delta EP(\varepsilon_k, \varepsilon_h) \geq 0$ over either $I^{11}_{\varepsilon_h}, I^{12}_{\varepsilon_h}$ or $I^{13}_{\varepsilon_h}$. Item 2 of Proposition 5: For system S_2 to be satisfied, (17) and (18) must hold together

$$\Delta EP_i(\varepsilon_\emptyset, \varepsilon_1) \geq 0 \iff \varepsilon_1 \ \geq \ \hat{\varepsilon}_1(\varepsilon_\emptyset) := \frac{\gamma_i(\varepsilon_\emptyset)}{s_i(\varepsilon_\emptyset)} \geq 0, \tag{17}$$

$$\Delta EP_j(\varepsilon_1, \varepsilon_\emptyset) \geq 0 \iff \varepsilon_\emptyset \ \geq \ \hat{\varepsilon}_\emptyset(\varepsilon_1) := \frac{\gamma_j(\varepsilon_1)}{s_j(\varepsilon_1)} \geq 0. \tag{18}$$

Consequently, the domain of definition of system S_2 is a restriction of the interval $I^{2l}_{\varepsilon_h}$ defined in Proposition 4 item 2. To characterize it, we need to prove that the set-valued functions defined by $\varepsilon_1 \geq \hat{\varepsilon}_1(\varepsilon_\emptyset)$ (17) and $\varepsilon_\emptyset \geq \hat{\varepsilon}_\emptyset(\varepsilon_1)$ (18) intersect the interval $[0,1]^2$ over a non-empty subset. The two thresholds $\hat{\varepsilon}_1(\varepsilon_\emptyset)$ and $\hat{\varepsilon}_\emptyset(\varepsilon_1)$ represents the probabilistic frontier under/above which players choose cooperation. Such a frontier is defined as follows. Consider the particular points $\varepsilon_\emptyset = \varepsilon_1 = \varepsilon$ for which the two previous set-valued functions intersect the 45 degree line over $[0,1]^2$. Over the 45 degree line, (17) and (18) becomes a unique relation $\varepsilon = \gamma(\varepsilon)/s(\varepsilon)$. Consequently, $\exists \varepsilon_1^\star \in \mathbb{R}, \exists \varepsilon_2^\star \in \mathbb{R} \mid \Delta EP_i(\varepsilon_1^\star, \varepsilon_2^\star) = 0$ and $\Delta EP_j(\varepsilon_2^\star, \varepsilon_1^\star) = 0$. Replace γ and s by their respective expressions and obtain the following condition:

$$(A + B - C - D)\varepsilon^2 - (A - 2C - D)\varepsilon - C \ \geq \ 0. \tag{19}$$

Let us denote $\varepsilon_1^\star$ and $\varepsilon_2^\star$ the two solutions of (19) as an equality.

$$\varepsilon_1^\star = \frac{A - 2C - D - \sqrt{(A-D)^2 + 4BC}}{2(A + B - C - D)}, \ \ \varepsilon_2^\star = \frac{A - 2C - D + \sqrt{(A-D)^2 + 4BC}}{2(A + B - C - D)}.$$

System S_2 is satisfied out of the set of solutions, i.e., $\forall \tilde{\varepsilon}_h \leq \varepsilon_1^\star$ and $\forall \tilde{\varepsilon}_h \geq \varepsilon_2^\star$. Proposition 7 proves that at mean one of these two solutions belongs to $[0, 1]$.

For system S_3 to be satisfied, the 2 following conditions must hold together

$$\Delta EP_i(\varepsilon_\emptyset, \varepsilon_1) \geq 0 \iff \varepsilon_1 \ \leq \ \hat{\varepsilon}_1(\varepsilon_\emptyset) := \frac{g_i(\varepsilon_\emptyset)}{\sigma_i(\varepsilon_\emptyset)}, \tag{20}$$

$$\Delta EP_j(\varepsilon_1, \varepsilon_\emptyset) \geq 0 \iff \varepsilon_\emptyset \ \leq \ \hat{\varepsilon}_\emptyset(\varepsilon_1) := \frac{g_j(\varepsilon_1)}{\sigma_j(\varepsilon_1)}. \tag{21}$$

Note that $\hat{\varepsilon}_1(\varepsilon_\emptyset) > 0$ and $\hat{\varepsilon}_\emptyset(\varepsilon_1) > 0$ and that cooperation emerges below these two thresholds, contrary to the previous cases. Using the same methodology as for System S_2, system S_3 leads to the following condition:

$$-(A+B-C-D)\varepsilon^2 + (A-2C-D)\varepsilon + C \leq 0. \tag{22}$$

Note that due to the symmetry of payoffs, inequality 19 is the same as inequality 22. Recall that S_4 is never satisfied.

PROPOSITION 6 *$\varepsilon_1^\star = \varepsilon_2^\star$ is impossible.*

Proof. If $\varepsilon_1^\star = \varepsilon_2^\star$ then the discriminant is zero, which is impossible by Property 2.

$$\begin{aligned} A-D > B-C \iff & (A-D)^2 > (B-C)^2 \\ & \iff (A-D)^2 + 4BC > (B-C)^2 + 4BC \\ & \qquad \iff (A-D)^2 + 4BC > (B+C)^2 > 0. \end{aligned}$$

From what we learn that there are two distinct solutions. □

PROPOSITION 7 *$\varepsilon_1^\star < 0$ and $\varepsilon_2^\star < 0$ is impossible. $\varepsilon_1^\star > 1$ and $\varepsilon_2^\star > 1$ is impossible. If $\varepsilon_1^\star < 0$ then $\varepsilon_2^\star \in [0,1]$. If $\varepsilon_1^\star \in [0,1]$ then $\varepsilon_2^\star \in [0,1]$ or $\varepsilon_2^\star > 1$. There is at least one interior solution.*

Proof. There are two main cases:

1 If $A-2C-D < 0$ then $\varepsilon_1^\star < 0$ and $\varepsilon_2^\star > 0$. Indeed, $\varepsilon_2^\star > 0 \iff \sqrt{(A-D)^2+4BC} > -(A-2C-D) \iff (A-D)^2+4BC > (A-D)^2-4C(A-D)+4C^2$ which simplifies to be $A+B-C-D > 0$, always true. Moreover, $\varepsilon_2^\star < 1 \iff \sqrt{(A-D)^2+4BC} < 2(A+B-C-D)-(A-2C-D)$ which simplifies to become $(A-D)^2+4BC > (A-D+2B)^2 \iff (A-D)^2+4BC > (A-D)^2-4B(A-D)+4B^2 \iff A+B-C-D > 0$. Conclusion if $\varepsilon_1^\star < 0$ then $0 < \varepsilon_2^\star < 1$.

2.a If $A-2C-D > 0$ and $A-2C-D < \sqrt{(A-D)^2+4BC}$, then $\varepsilon_1^\star < 0$ and $\varepsilon_2^\star > 0$. As above $\varepsilon_2^\star < 1$.

2.b If $A-2C-D > 0$ and $A-2C-D > \sqrt{(A-D)^2+4BC}$ then $\varepsilon_1^\star > 0$ and $\varepsilon_1^\star < 1$ leads to $-\sqrt{\Delta} < A+2B-D$ which again is $A+B-C-D > 0$. Since the numerator and the denominator are positive, $\varepsilon_2^\star > 0$. The condition for $\varepsilon_2^\star < 1$ is $B > 0$. If $B < 0$ then $\varepsilon_2^\star > 1$ and $0 < \varepsilon_1^\star < 1$ (see Lemma 2 below).

Table 18 illustrates those cases. □ This ends Proposition 7. □

LEMMA 2 *In any PD, the condition for $\varepsilon_1^\star > 0$ is $C < 0$, $\varepsilon_2^\star < 1$ is $B > 0$.*

Proof. $\varepsilon_1^\star > 0$ if the numerator is positive, since the denominator is always positive. Compute $(A-2C-D)^2-((A-D)^2+4BC)$ and get $-4C(A+B-C-D)$. Consequently, using Property 2 $C < 0 \Rightarrow \varepsilon_1^\star > 0, \varepsilon_2^\star < 1$. Note $B = 0 \Rightarrow \varepsilon_2^\star = 1$ and $\partial\varepsilon_2^\star/\partial B_{|B=0} = \frac{C}{\sqrt{4BC+(A-D)^2}(A+B-C-D)} - \frac{A-2C-D+\sqrt{4BC+(A-D)^2}}{2(A+B-C-D)^2}$ and $B = 0, \partial\varepsilon_2^\star/\partial B_{|B=0} < 0$. □

This ends the proof of Theorem 1. □

□

B Appendix

Proof. Theorem 2 is proved using Lemma 3. We study the properties of $\hat\varepsilon_h$.

$$\frac{\partial\hat\varepsilon_h(\varepsilon_k)}{\partial\varepsilon_k} = \frac{AD-BC}{(C-A+\varepsilon_k(A+B-C-D))^2},\quad \frac{\partial^2\hat\varepsilon_h(\varepsilon_k)}{\partial\varepsilon_k^2} = \frac{2(A+B-C-D)(BC-AD)}{(C-A+\varepsilon_k(A+B-C-D))^3}.$$

If $AD > BC$ and $\varepsilon_k \geq (A-C)/(A+B-C-D)$ then $\hat\varepsilon_h$ is increasing in ε_k and concave. From what all the other cases are deduced. From Property 2 and Proposition 2 items 1 and 3, depending on the value of the slope of the slope (s or $-\sigma$), the first derivative and second derivative may be positive or negative. All the following various cases are ranked according to the number of positive payoffs.

LEMMA 3 *For systems S_1 and S_2, sign $(\partial\hat\varepsilon_h(\varepsilon_k)/\partial\varepsilon_k) = -$sign $(\partial^2\hat\varepsilon_h(\varepsilon_k)/\partial\varepsilon_k^2)$, and for S_3 and S_4 sign $(\partial\hat\varepsilon_h(\varepsilon_k)/\partial\varepsilon_k) =$ sign $(\partial^2\hat\varepsilon_h(\varepsilon_k)/\partial\varepsilon_k^2)$.*

Proof. For systems S_1 and S_2 rewrite

$$\frac{\partial\hat\varepsilon_h(\varepsilon_k)}{\partial\varepsilon_k} = \frac{(A+B-C-D)(C+D)z_\varepsilon}{s(\varepsilon_k)^2},\quad \frac{\partial^2\hat\varepsilon_h(\varepsilon_k)}{\partial\varepsilon_k^2} = \frac{2(A+B-C-D)^2(C+D)(-z_\varepsilon)}{s(\varepsilon_k)^3}.$$

For systems S_3 and S_4 rewrite

$$\frac{\partial\hat\varepsilon_h(\varepsilon_k)}{\partial\varepsilon_k} = \frac{(A+B-C-D)(C+D)z_\varepsilon}{(-\sigma(\varepsilon_k))^2},\quad \frac{\partial^2\hat\varepsilon_h(\varepsilon_k)}{\partial\varepsilon_k^2} = \frac{2(A+B-C-D)^2(C+D)(-z_\varepsilon)}{(-\sigma(\varepsilon_k))^3}.$$

Either $\hat\varepsilon_h$ is increasing and concave, or decreasing and convex. □

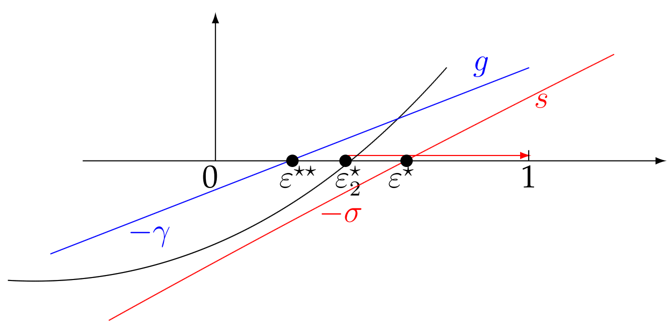
Let us turn to the typology of PD.

1. Suppose $A > B > C > D > 0$. From Lemma 2, $\varepsilon_1^\star < 0$ and $\varepsilon_2^\star < 1$.

(a) $z_\varepsilon > 0$, $\varepsilon^\star > \varepsilon^{\star\star} \iff N_\varepsilon > 0, D_\varepsilon > 0$ and $\partial\hat{\varepsilon}/\partial\varepsilon > 0$. From Proposition 2 item 2 $D_\varepsilon > 0 \Rightarrow C > -D$ and $\varepsilon^{\star\star} > 0$. Consequently, if $D_\varepsilon > 0$ the slope of the intercept is necessarily positive. In that case, $\varepsilon_2^\star < \varepsilon^\star$. Indeed,

$$\frac{A - 2C - D + \sqrt{(A-D)^2 + 4BC}}{2(A+B-C-D)} < \frac{A-C}{A+B-C-D} \iff AD > BC,$$

which is exactly the assumption $N_\varepsilon > 0$. From Lemma 3, the representative curve of $\hat{\varepsilon}(\varepsilon)$ is increasing convex over S_3 and S_4, and increasing concave over S_1. Consider the surface $[0,1]^2$. The function $\hat{\varepsilon}(\varepsilon)$ crosses horizontal line (defined by coordinates $(0,1)(1,1)$) or vertical line (defined by coordinates $(1,0)(1,1)$) for $\hat{\varepsilon}(\varepsilon) = 1$ that is $\varepsilon_\emptyset^1 = A/(A+B)$. Results are summed up in Table 10. In all the following Graphics, the red line corresponds to the slope of the slope and the blue one to the slope of the intercept. The black curve is the second-order equation.



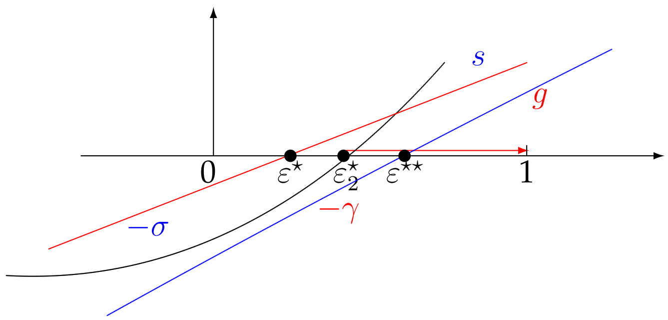
————— insert Table 11 here —————-

Table 10: $EP_{\mathcal{C}} \geq EP_{\mathcal{NC}}$ over a high support of the basic probabilities

ε	0	$\varepsilon^{\star\star}$	$\varepsilon^\star$ 1
value of the slope	$-\sigma$	$-\sigma$	s
value of the intercept	$-\gamma$	g	g
system	S_4	S_3	S_1
$EP_C(\tilde{\varepsilon}) - EP_{NC}(\tilde{\varepsilon}) \geq 0$		$\varepsilon_2^\star$ →	
$\Delta EP_i(\tilde{\varepsilon}) \geq 0$		$\forall\tilde{\varepsilon} \in J_\varepsilon := [\varepsilon_2^\star, 1]$	

(b) $z_\varepsilon < 0$, $\varepsilon^\star < \varepsilon^{\star\star}$ and $\partial\hat{\varepsilon}(\varepsilon)/\partial\varepsilon < 0$. By Proposition 2 item 1: $0 < \varepsilon^\star < 1$. $N_\varepsilon < 0$ and $D_\varepsilon > 0 \Rightarrow C > -D$ and $1 > \varepsilon^{\star\star} > 0$. Consequently, if $D_\varepsilon > 0$ the slope of the intercept is necessarily positive. $1 > \varepsilon_2^\star > \varepsilon^\star$, which involves $D > 0$. From Lemma 3, the representative curve of $\hat{\varepsilon}(\varepsilon)$ is decreasing concave over

S_4 and decreasing convex over S_2, S_1. On $[0,1]^2$, it crosses horizontal line (defined by coordinates $(0,1)(1,1)$) or vertical line (defined by coordinates $(1,0)(1,1)$) for $\hat{\varepsilon}(\varepsilon) = 1$ that is $\varepsilon^1_\emptyset = A/(A+B)$. SeeTable 11 and Table 18 for an example.



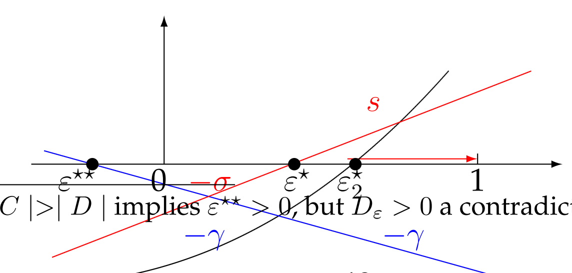
——————— insert Table 12 here ———————

Table 11: $EP_{\mathcal{C}} \geq EP_{\mathcal{NC}}$ over a high support of the basic probabilities

ε	0	$\varepsilon^\star$	$\varepsilon^{\star\star}$ 1
value of the slope	$-\sigma$	s	s
value of the intercept	$-\gamma$	$-\gamma$	g
system	S_4	S_2	S_1
$EP_C(\tilde{\varepsilon}) - EP_{NC}(\tilde{\varepsilon}) \geq 0$		$\varepsilon^\star_2$ →	
$\Delta EP_i(\tilde{\varepsilon}) \geq 0$		$\forall \tilde{\varepsilon} \in J_\varepsilon := [\varepsilon^\star_2, 1]$	

2. Case $A > B > C > 0 > D$. From Lemma 2, $\varepsilon^\star_1 < 0$ and $\varepsilon^\star_2 < 1$.

   (a) $z_\varepsilon > 0$, $\varepsilon^\star > \varepsilon^{\star\star} \iff N_\varepsilon < 0, D_\varepsilon < 0$ and $\partial\hat{\varepsilon}(\varepsilon)/\partial\varepsilon < 0$. By Proposition 2 item 1: $0 < \varepsilon^\star < 1$. From Proposition 2 item 2, if $C > 0 > D$ and $\mid C \mid < \mid D \mid$ then $\varepsilon^{\star\star} < 0$[3]. From Lemma 3, the representative curve of $\hat{\varepsilon}(\varepsilon)$ is decreasing concave over S_4 and decreasing convex over S_2. On $[0,1]^2$, it crosses horizontal line (defined by coordinates $(0,1)(1,1)$) or vertical line (defined by coordinates $(1,0)(1,1)$) for $\hat{\varepsilon}(\varepsilon) = 1$ that is $\varepsilon^1_\emptyset = A/(A+B)$. See Table 12 and Table 18 for an example.



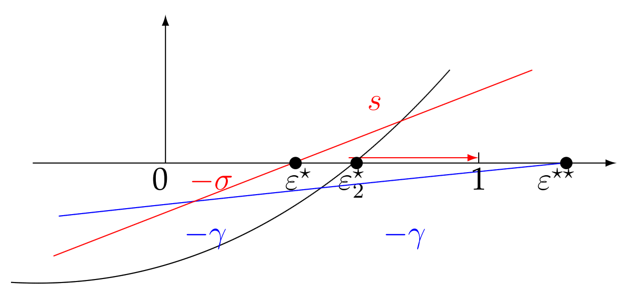
[3]The case where $\mid C \mid > \mid D \mid$ implies $\varepsilon^{\star\star} > 0$, but $D_\varepsilon > 0$ a contradiction. This case cannot exist.

——————— insert Table 13 here ———————

Table 12: $EP_{\mathcal{C}} \geq EP_{\mathcal{NC}}$ with high support of the basic probabilities

ε	0	$\varepsilon^\star$ 1
value of the slope	$-\sigma$	s
value of the intercept	$-\gamma$	$-\gamma$
system	S_4	S_2
$EP_C(\tilde{\varepsilon}) \geq EP_{NC}(\tilde{\varepsilon})$		$\varepsilon_2^\star \rightarrow$
$\Delta EP_i(\tilde{\varepsilon}) \geq 0$		$\forall \tilde{\varepsilon} \in J_\varepsilon := [\varepsilon_2^\star, 1]$

(b) $z_\varepsilon < 0$, $\varepsilon^\star < \varepsilon^{\star\star}$ and for $\mid C \mid > \mid D \mid$ we have $\partial\hat{\varepsilon}(\varepsilon)/\partial\varepsilon < 0$. By Proposition 2 item 1: $0 < \varepsilon^\star < 1$. From Lemma 3, the representative curve of $\hat{\varepsilon}(\varepsilon)$ is decreasing concave over S_4 and decreasing convex over S_2. On $[0,1]^2$, it crosses horizontal line (defined by coordinates $(0,1)(1,1)$) or vertical line (defined by coordinates $(1,0)(1,1)$) for $\hat{\varepsilon}(\varepsilon) = 1$ that is $\varepsilon_\emptyset^1 = A/(A+B)$. Results are summed up in Table 13 and an example is provided in Table 18.



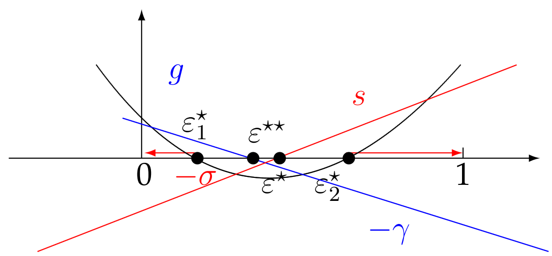
——————— insert Table 14 here ———————

Table 13: $EP_{\mathcal{C}} \geq EP_{\mathcal{NC}}$ with high support of the basic probabilities

ε	0	$\varepsilon^\star$ 1
value of the slope	$-\sigma$	s
value of the intercept	$-\gamma$	$-\gamma$
system	S_4	S_2
$EP_C(\tilde{\varepsilon}) \geq EP_{NC}(\tilde{\varepsilon})$		$\varepsilon_2^\star \rightarrow$
$\Delta EP_i(\tilde{\varepsilon}) \geq 0$		$\forall \tilde{\varepsilon} \in J_\varepsilon := [\varepsilon_2^\star, 1]$

3. $A > B > 0 > C > D$. From Lemma 2, $0 < \varepsilon_1^\star < 1$ and $0 < \varepsilon_2^\star < 1$.

(a) $z_\varepsilon > 0$, $\varepsilon^\star > \varepsilon^{\star\star} \iff N_\varepsilon < 0, D_\varepsilon < 0$ and $\partial\hat{\varepsilon}(\varepsilon)/\partial\varepsilon < 0$. By Proposition 2 item 1: $0 < \varepsilon^\star < 1$. From Lemma 3, the representative curve of $\hat{\varepsilon}(\varepsilon)$ is decreasing concave over S_3, S_4 and decreasing convex over S_2. This case is more complex than the previous one. Indeed, on $[0,1]^2$ it crosses horizontal axis or vertical axis for $\hat{\varepsilon}(\varepsilon) = 0$ that is $\varepsilon^0_\emptyset = \varepsilon^{\star\star} = C/(C+D)$, or it crosses horizontal line (defined by coordinates $(0,1)(1,1)$) or vertical line (defined by coordinates $(1,0)(1,1)$) for $\hat{\varepsilon}(\varepsilon) = 1$ that is $\varepsilon^1_\emptyset = A/(A+B)$. See Table 14 and Table 18 for an example.



———— insert Table 15 here ————

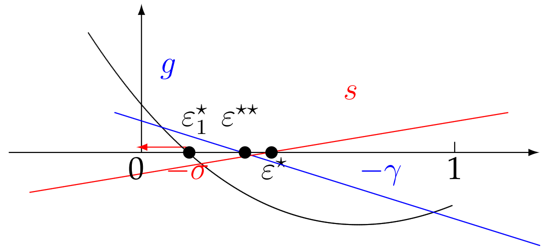
Table 14: $EP_{\mathcal{C}} \geq EP_{\mathcal{NC}}$ over a high support of the basic probabilities

ε	0	$\varepsilon^{\star\star}$	$\varepsilon^\star$ 1
value of the slope	$-\sigma$	$-\sigma$	s
value of the intercept	g	$-\gamma$	$-\gamma$
system	S_3	S_4	S_2
$EP_C(\tilde{\varepsilon}) \geq EP_{NC}(\tilde{\varepsilon})$	$\leftarrow$ $\varepsilon^\star_1$		$\varepsilon^\star_2$ $\rightarrow$
$\Delta EP_i(\tilde{\varepsilon}) \geq 0$	$\forall\tilde{\varepsilon} \in J_{\underline{\varepsilon}} := [0, \varepsilon^\star_1] \cup J_{\overline{\varepsilon}} := [\varepsilon^\star_2, 1]$		

(b) $z_\varepsilon < 0$ is impossible. Indeed, $0 > C > D \Rightarrow -BC > 0, AD < 0$ and in any PD, $\mid AD \mid > \mid BC \mid$. Consequently, $N_\varepsilon < 0$ and $D_\varepsilon = C + D < 0$ too: $\Rightarrow z_\varepsilon > 0$.

4. $A > 0 > B > C > D$. From Lemma 2, $\varepsilon^\star_1 > 0$ and $\varepsilon^\star_2 > 1$.

(a) $z_\varepsilon > 0$, $\varepsilon^\star > \varepsilon^{\star\star} \iff N_\varepsilon < 0, D_\varepsilon < 0$ and $\partial\hat{\varepsilon}(\varepsilon)/\partial\varepsilon < 0$. By Proposition 2 item 1: $0 < \varepsilon^\star < 1$. From Lemma 3, the representative curve of $z_\varepsilon(\varepsilon)$ is decreasing concave over S_3, S_4 and decreasing convex over S_2. It crosses horizontal axis or vertical axis for $\hat{\varepsilon}(\varepsilon) = 0$ that is $\varepsilon^0_\emptyset = \varepsilon^{\star\star} = C/(C+D)$. In that game, the support of the prior probability of cooperation ε for $EP_{\mathcal{C}} > EP_{\mathcal{NC}}$ is low. See Table 15 for details and an example is provided in Table 18.



——————— insert Table 16 here ———————

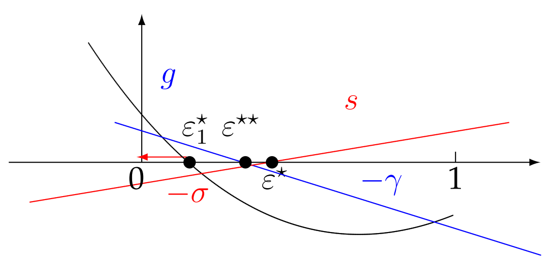
Table 15: $EP_{\mathcal{C}} \geq EP_{\mathcal{NC}}$ over a high support of the basic probabilities

ε	0	$\varepsilon^{\star\star}$	$\varepsilon^{\star}$ 1
value of the slope	$-\sigma$	$-\sigma$	s
value of the intercept	g	$-\gamma$	$-\gamma$
system	S_3	S_4	S_2
$EP_C(\tilde{\varepsilon}) \geq EP_{NC}(\tilde{\varepsilon})$	$\leftarrow$ $\varepsilon_1^\star$		
$\Delta EP_i(\tilde{\varepsilon}) \geq 0$	$\forall \tilde{\varepsilon} \in J_{\underline{\varepsilon}} := [0, \varepsilon_1^\star]$		

(b) $z_\varepsilon < 0$ is impossible. Indeed, the numerator is negative since $AD < 0, -BC < 0$ and the denominator is negative too, since $C + D < 0$.

5. $0 \geq A > B > C > D$. From Lemma 2, $\varepsilon_1^\star > 0$ and $\varepsilon_2^\star > 1$.

(a) $z_\varepsilon > 0$, $\varepsilon^\star > \varepsilon^{\star\star} \iff N_\varepsilon < 0, D_\varepsilon < 0$ and $\partial\hat{\varepsilon}(\varepsilon)/\partial\varepsilon < 0$. By Proposition 2 item 1: $0 < \varepsilon^\star < 1$. From Lemma 3, the representative curve of $z_\varepsilon(\varepsilon)$ is decreasing concave over S_3, S_4 and decreasing convex over S_2. It crosses horizontal axis or vertical axis for $\hat{\varepsilon}(\varepsilon) = 0$ that is $\varepsilon_\emptyset^0 = \varepsilon^{\star\star} = C/(C + D)$. See Table 16 for details and an example is provided in Table 18.

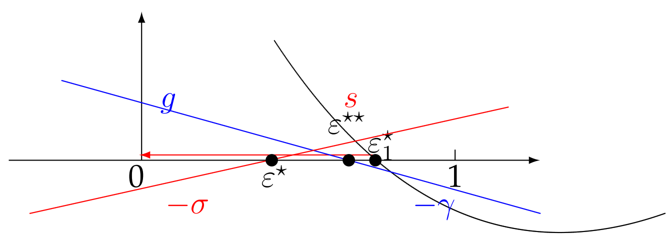


——————— insert Table 17 here ———————

Table 16: $EP_{\mathcal{C}} \geq EP_{\mathcal{NC}}$ over a high support of the basic probabilities

ε	0	$\varepsilon^{\star\star}$	$\varepsilon^{\star}$ 1
value of the slope	$-\sigma$	$-\sigma$	s
value of the intercept	g	$-\gamma$	$-\gamma$
system	S_3	S_4	S_2
$EP_C(\tilde{\varepsilon}) \geq EP_{NC}(\tilde{\varepsilon})$	← $\varepsilon_1^{\star}$		
$\Delta EP_i(\tilde{\varepsilon}) \geq 0$	$\forall \tilde{\varepsilon} \in J_{\underline{\varepsilon}} := [0, \varepsilon_1^{\star}]$		

(b) $z < 0$, $\varepsilon^{\star} < \varepsilon^{\star\star}$ and $\partial \hat{\varepsilon}(\varepsilon)/\partial \varepsilon > 0$. By Proposition 2 item 1: $0 < \varepsilon^{\star} < 1. N_{\varepsilon} > 0$ and $D_{\varepsilon} < 0 \Rightarrow 0 < \varepsilon^{\star\star} < 1$. Indeed, from Proposition 2 item 2, we have $C + D < 0$ and $\varepsilon^{\star\star} < 0$. This implies $\varepsilon^{\star\star} < 1$. From Lemma 3, the representative curve of $z_{\varepsilon}(\varepsilon)$ is increasing convex over S_3 and increasing concave over S_2. It crosses horizontal axis or vertical axis for $\hat{\varepsilon}(\varepsilon) = 0$ that is $\varepsilon_{\emptyset}^{0} = \varepsilon^{\star\star} = C/(C + D)$. See Table 17 for details and an example is provided in Table 18.



——————— insert Table 18 here ——————

Table 17: $EP_{\mathcal{C}} \geq EP_{\mathcal{NC}}$ over a high support of the basic probabilities

ε	0	$\varepsilon^{\star}$	$\varepsilon^{\star\star}$ 1
value of the slope	$-\sigma$	s	s
value of the intercept	g	g	$-\gamma$
system	S_3	S_1	S_2
$EP_C(\tilde{\varepsilon}) \geq EP_{NC}(\tilde{\varepsilon})$	←	←	$\varepsilon_1^{\star}$
$\Delta EP_i(\tilde{\varepsilon}) \geq 0$	$\forall \tilde{\varepsilon} \in J_{\underline{\varepsilon}} := [0, \varepsilon_1^{\star}]$		

□

——— insert Table 19 here ———-

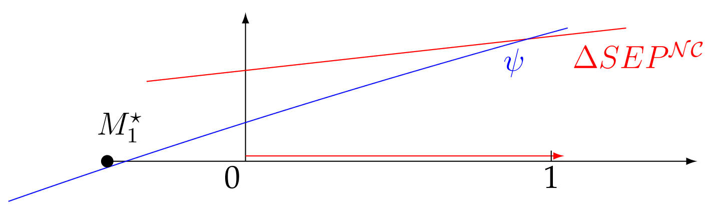
C Appendix

Proof. Compute $\Delta SEP_j^{\mathcal{NC}}(\overline{\mu}) := SEP_{\mathcal{NC}_j}(\overline{\mu}) - SEP_{\mathcal{C}_j}(\overline{\mu}), i = 1, 2, i \neq j$, the difference between non-cooperation and cooperation. Consider the following system of inequalities

$$\begin{cases} \Psi_i(\underline{M}) \leq 0 \\ \Psi_j(\overline{\mu}) \leq 0, \end{cases} \iff \begin{cases} (C - D - A + B)\underline{M} + A - B - C \leq 0, \\ (C - D - A + B)\overline{\mu} + A - B - C \leq 0. \end{cases}$$

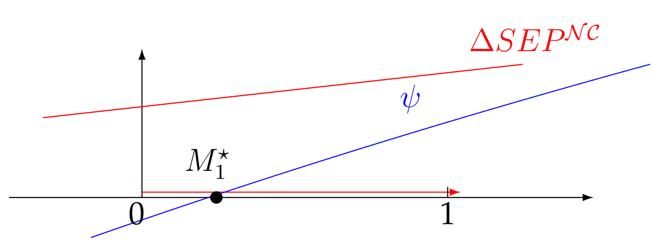
From a mathematical point of view, the two inequalities are the same. We only develop the proof for one inequality which can also be interpreted as player formulate common subjective probabilities. Denote the slope t if $(C - D - A + B) \geq 0$ and $-T$ if $(C - D - A + B) < 0$. Denote the intercept r if $A - B - C \geq 0$ and $-R$ if $A - B - C < 0$.

1. First case: $t\underline{M} + r \leq 0 \iff \underline{M} \leq M_1^\star := \frac{-r}{T} < 0$. In that case non-cooperation is a trustable strategy, and always selected.

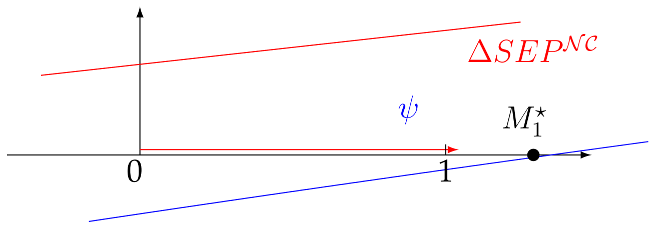


Graphic 1

2. Second case: $t\underline{M} - R \leq 0 \iff \underline{M} \leq M_1^\star := \frac{R}{T} > 0$. Since the slope is positive we have the two possible sub-cases: $0 \leq M_1^\star \leq 1$ or $M_1^\star > 1$. The non cooperation is trustable over $[M_1^\star, 1]$ on Graphic 2, but is never trustable on Graphic 3, but always selected.

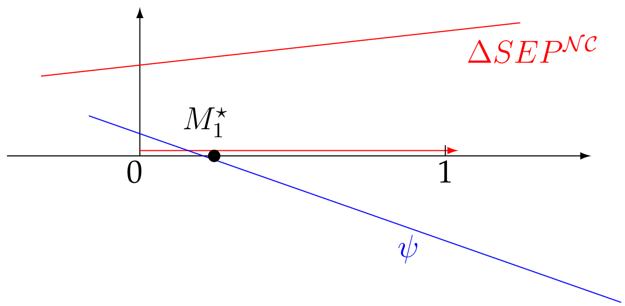


Graphic 2

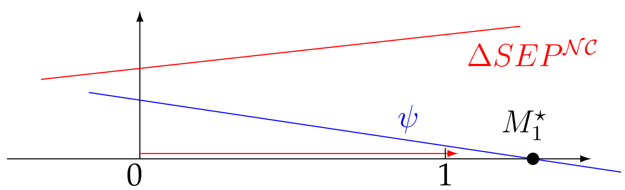


Graphic 3

3. Third case $-T\underline{M}+r \leq 0 \iff \underline{M} \leq M_1^\star := \frac{r}{T} > 0$. On Graphic 4 non-cooperation is trustable over $[0, M_1^\star]$ while it is always trustable on Graphic 5, but always selected.

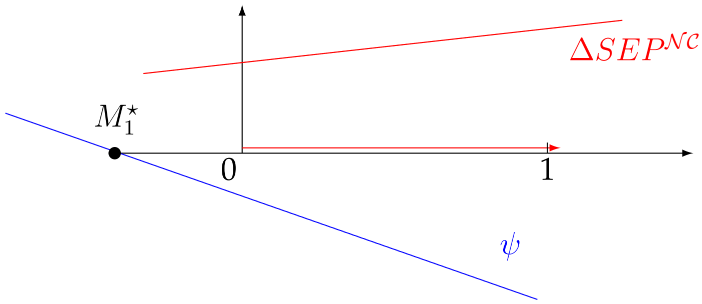


Graphic 4



Graphic 5

4. Fourth case $-T\underline{M} - R \leq 0 \iff \underline{M} \leq M_1^\star := \frac{-R}{T} < 0$. Non cooperation is never trustable, but always selected.



Graphic 6

□

References

[1] Allais, Maurice (1953) "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine", *Econometrica*, 21(4):503-546.

[2] Arrow, Kenneth (1969) "*The Organization of Economic Activity: Issues Pertinent to the Choice of Market versus Non-market Allocation*" Chapter 1, Rationality: Individual And Social, p. 23.

[3] Aumann Robert J. (1987) "Correlated Equilibrium as an Expression of Bayesian Rationality" *Econometrica: the Journal of the Econometric Society* 1-18.

[4] Aumann Robert J. (1988) "Preliminary Notes on Integrating Irrationality into Game Theory," Mimeo, International Conference on Economic Theories of Politics, Haifa.

[5] Aumann Robert J. (1995) "Backward Induction and Common Knowledge of Rationality" *Games and Economic Behavior*, 8, 6-19.

[6] Aumann Robert (2006) "War and Peace" PNAS November 14, 2006 103 (46) 17075-17078; https://doi.org/10.1073/pnas.0608329103

[7] Bayes, Thomas (1763) "An Essay towards solving a Problem in the Doctrine of Chances" By the late Rev. Mr. Bayes, F.R.S. communicated by Mr Price, in a Letter to John Canton, A.M.F.R.S.

[8] Bernouilli, Jacques (1715) "Art Conjectandis"

[9] Clark, Kenneth and Martin Sefton (2001) "The Sequential PD: Evidence from Reciprocation" *The Economic Journal* 111(468):51-68.

[10] Coleman, James S., (1990) "Relations of Trust," in: *Foundations of Social Theory*, James S. Coleman (dir.), Cambridge, Harvard University Press:91-116, p. 99.

[11] Dasgupta, Partha, (1988) "Trust as a Commodity, in: *Trust: Making and Breaking Cooperative Relationships*," Diego Gambetta (dir.), New York, Blackwell, 1988, pp. 49-72, p. 59.

[12] Diekert Florian K. (2012) "The Tragedy of the Commons from a Game-Theoretic Perspective" *Sustainability*, 4:1776-1786

[13] Fermat, Pierre de (1654) "Correspondance avec Blaise Pascal"

[14] Glaeser, Edward L., David I. Laiton, Jose A. Scheinkman and Christine L. Soutter (2000) "Measuring Trust" *The Quarterly Journal of Economics* 115(3):811-846.

[15] Harsanyi, John (1967) "Games With Incomplete Information Played By " Bayesian Players" , *Management Science* 14(3):159-182.

[16] Harsanyi John C. (1982a) "Comment — Subjective Probability and the Theory of Games: Comments on Kadane and Larkey's Paper" *Management Science*, 28(2):120-124.

[17] Harsanyi John C. (1982b) "Rejoinder to Professors Kadane and Larkey" *Management Science*, 28(2):124-125.

[18] Harsanyi John and Selten Rainard (1988) "*A General Theory of Equilibrium Selection in Games*, MITPress Books.

[19] Huygens Christiaan (1657) "De ratiocinais in ludo aleae" Tome XIV, p.1.

[20] Kadane J.B and Larkey PD (1982a) "Subjective Probability and the Theory of Games" *Management Science*, 28:113-20.

[21] Kadane J.B and Larkey PD (1982b) "Reply to Professor Harsanyi" *Management Science*, 28:124.

[22] Laplace (1814) "Essai philosophique sur les probabilitï£¡s", Mme Veuve Courcier, Imprimeur-Libraire, p21, p72-73 and p 92.

[23] Marcus, James Charles (2009) "Losses, Gains, and Asymmetry in the 1-Shot PD Game," Thesis, University of Illinois at IUrbana-Champaign, 2009.

[24] McKelvey, R.D., and T.R. Palfrey (1992) "An experimental study of the Centipede Game" *Econometrica*, 60(4),:804-836.

[25] Milgrom, Paul R., Douglass C. North et Barry R. Weingast, (1997) "The Role of Institutions in the Revival of Trade: Law Merchant, Private Judges, and the Champagne Fairs, " in: *Reputation: Studies in the Voluntary Elicitation of Good Conduct*, Daniel B. Klein (dir.), Ann Arbor, The University of Michigan Press, 1997, pp. 243-266, p. 243.

[26] Pascal, Blaise (1654) "Lettre ï£¡ l'Acadï£¡mie des Sciences".

[27] Rapoport, Anatol and Chammah (1970) "*PD: A Study in Conflict and Cooperation*" with the collaboration of Carol J. Orwant, Ann Arbor Paperbacks, The University of Michigan Press.

[28] Sako, Mari (1992) "*Prices, Quality and Trust: Inter-firm Relations in Britain and Japan*," Cambridge, Cambridge University Press, 1992, pp. 37-38.

[29] Selten, Reinhard (1975) "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, 4(1):25-55.

[30] Shapiro, Debra L., Blair H. Sheppard et Lisa Cheraskin (1992) "Business on a Handshake," *Negotiation Journal* 8:365-377.

[31] Von Neumann, John and Oskar Morgenstern (1944) "*Theory of Games and Economic Behavior*" Princeton University Press.

[32] Von Neumann, John and Oskar Morgenstern (1953) "*Theory of Games and Economic Behavior*" Princeton University Press.

[33] Williamson, Oliver E., (1983) "Credible Commitments: Using Hostages to Support Exchange" , *The American Economic Review* 73(4):519-540.

[34] Williamson, Oliver (1993) "Calcultiveness, Trust and Economic Organization" *Journal of Law and Economics* 36(1):453-486.

Table 18: Examples of all possible characterizations of solutions of $\hat{\Gamma}_0$

# of > 0 payoffs	A	B	C	D	$z_\varepsilon = \frac{N_\varepsilon}{D_\varepsilon}$	$\varepsilon^\star$	$\varepsilon^{\star\star}$	$\varepsilon_1^\star$	$\varepsilon_2^\star$	$sslope_\varepsilon^\circ$	$islope_\varepsilon^{\circ\circ}$
Table 1	7	3	2	1	$\frac{1}{3} > 0$	0.71	0.66	-0.41	0.69	$7\varepsilon - 5$	$3\varepsilon - 2$
4											
Table 2	5	3	2	1	$\frac{-1}{3} < 0$	0.60	0.66	-0.63	0.63	$5\varepsilon - 3$	$3\varepsilon - 2$
Table 3	5	3	2	-3	$\frac{-21}{-1} > 0$	0.33	-2	-0.29	0.74	$9\varepsilon - 3$	$-\varepsilon - 2$
3											
Table 4	5	3	2	-1	$\frac{-11}{1} < 0$	0.42	2	-0.41	0.69	$7\varepsilon - 3$	$\varepsilon - 2$
Table 5	5	3	-2	-3	$\frac{-9}{-5} > 0$	0.53	0.40	0.21	0.70	$13\varepsilon - 7$	$-5\varepsilon + 2$
2	$z < 0$ is an impossible case										
Table 6	0	-1	-3	-4	$\frac{-3}{-7} > 0$	0.5	0.42	0.39	1.27	$6\varepsilon - 3$	$-7\varepsilon + 3$
1	$z < 0$ is an impossible case										
Table 7	-1	-2	-3	-4	$\frac{-2}{-7} > 0$	0.5	0.42	0.40	1.84	$4\varepsilon - 2$	$-7\varepsilon + 3$
0											
Table 8	-1	-2	-3	-7	$\frac{1}{-10} < 0$	0.28	0.30	0.304	1.41	$7\varepsilon - 2$	$-10\varepsilon + 3$

$^\circ$ Where $sslope_\varepsilon$ is the slope of the slope,
$^{\circ\circ}$ $islope_\varepsilon$ is the slope of the intercept.