

Working Paper in Economics No. 747

Cooperation Between Emotional Players

Lina Andersson

Department of Economics, December 2018

ISSN 1403-2473 (Print)
ISSN 1403-2465 (Online)



UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Cooperation Between Emotional Players

Lina Andersson*

Abstract: This paper uses the framework of stochastic games to propose a model of emotions in repeated interactions. An emotional player, who transitions between different states of mind as a response to observed actions taken by the other player, can be in either a friendly, a neutral, or a hostile state of mind. The state of mind determines the player's psychological payoff that together with a material payoff constitutes his utility. In the friendly (hostile) state of mind the player has a positive (negative) concern for the other player's material payoffs. Emotions can both facilitate and obstruct cooperation in the repeated prisoners' dilemma game. If finitely repeated, then a traditional player (who cares only for own material payoffs) can have an incentive to manipulate an emotional player into a friendly state of mind for future gains. If infinitely repeated, then two emotional players may require less patience to sustain cooperation. However, emotions can also obstruct cooperation if the players are either unwilling to punish each other, or become revengeful when punished.

Keywords: emotions; cooperation; repeated prisoners' dilemma; stochastic games

JEL Classification: C73; D01; D91

*Department of Economics, University of Gothenburg; e-mail: lina.andersson@economics.gu.se. This paper started out as a term paper for Jörgen Weibull's Game Theory course at Stockholm School of Economics and I thank the course participants for helpful comments and suggestions. I thank the participants of the ESEM 2018 in Cologne, and I am also grateful for advice from Martin Dufwenberg, Olof Johansson-Stenman, and Jörgen Weibull.

1 Introduction

Repeated interactions are common: spouses sharing household work, a parent dressing their kid for school, two co-workers cooperating on a mutual project, or two countries negotiating a trade agreement. These interactions can be modeled as repeated games and, in some cases, as when the interaction takes the form of a finitely repeated prisoners' dilemma game, the game theoretical literature provides us with a unique prediction of the outcome. However, during the last decade, repeated game forms have attracted increasing interest from both theoretical and experimental economists (cf. Dreber, Fudenberg, and Rand, 2014; Rand, Fudenberg, and Dreber, 2015) since the outcomes from experimental studies do not line up with the theoretical predictions. One observation is that the subjects react emotionally to what other subjects are doing, and therefore perceive the stage game differently in different periods even when the material consequences are unaltered. Since a crucial assumption of repeated games is that one and the same stage game is played in each period, such interactions cannot be properly modeled as repeated games.

This paper proposes a model of players with emotionally driven social preferences in repeated interactions where the assumption of a unique stage game is relaxed through use of the stochastic games framework. Stochastic games, first introduced by Shapley (1953), are a generalization of repeated games, in which the stage game can change between periods. While stochastic games have been frequently used in models of oligopoly and competition, e.g. to model endogenous demand, this paper is, to my knowledge, the first to apply the stochastic games framework to a model of social preferences. In this model, the players can react emotionally to observed actions of other players, and, as a response, transition between different states of mind. A cooperative act by another player may induce a player to become friendly and value the other player's material payoff positively, whereas a hurtful act may induce the player to become hostile and value the other player's material payoff negatively. Each state of mind of a player defines a vector with stage game utilities. The transitions between the states of mind are determined by the player's type. The type defines a transition matrix which, given a current state and an action profile, defines a probability distribution over next period's states of mind.

Focus is on a set of relatively simple Markov types whose state transitions are

solely determined by the last observed action of the other player. The traditionally studied *Homo oeconomicus* corresponds to a type who cannot transition from the neutral state of mind, in which he cares only for his own material payoffs. The *Homo oeconomicus* is contrasted to two other types: one type that transitions between two states of mind, a neutral and a hostile, and one type that transitions between three states of mind, a friendly, a neutral, and a hostile. The model is applied to a repeated prisoners' dilemma game with fixed material payoffs, but where the players' states of mind affect their preferences over outcomes.

The paper shows how a history-sensitive concern for the other player's material payoffs can be beneficial for cooperation. A positive concern, for example after a good history of play, can encourage further cooperation. Likewise, a negative concern, after a bad history of play, can make threats of punishment more credible than in the standard model. Moreover, a *Homo oeconomicus* can have an incentive to manipulate the other player for future gains. The two main contributions of this paper are (i) the framework for modeling emotions in a repeated interaction using stochastic games and (ii) a better understanding of how emotions can affect cooperation.

The observation that altruistic preferences can be conditional has been made before. In the model by Levine (1998), altruistic preferences are conditional upon the degree of altruism of the other player. Altruistic individuals want to be altruistic toward other altruistic individuals, but not toward individuals who only care for their own well-being. Cox, Friedman, and Gjerstad (2007) and Cox, Friedman, and Sadiraj (2008) model altruistic preferences in extensive form games as conditional on the behavior of the other player. That is, the players can become more or less generous towards someone depending on how altruistically the other player has acted. In the model proposed in this paper, the altruistic preferences are conditional on the players' emotional state of mind, which is affected by the history of the interaction.

This paper is outlined as follows. Section 2 introduces the framework of stochastic games and discusses key results used in this paper. It can be skipped by the less technically interested reader. Section 3 presents the model of a stochastic prisoners' dilemma game with emotional players. Section 4 analyzes the finitely and infinitely repeated interaction between players with two states of mind. A third state of mind is then introduced in Section 5. Section 6 applies the model to

a repeated public goods provision game. It illustrates how the effect of emotions can depend on the structure of the interaction. Section 7 concludes by relating the model to earlier literature.

2 Stochastic Games

Stochastic games were first introduced by Shapley (1953). Stochastic games generalize repeated games by allowing the stage game to change between periods. A stochastic game can be defined as a tuple, $\Gamma = \langle I, S, A_i, u_i, P, \delta \rangle$. This paper is focused on a finite set of players, denoted $I = \{1, \dots, n\}$, and a finite state space, denoted S , with a stage game for each state $s \in S$. Each player $i \in I$ has a finite set A_i of actions $a_i \in A_i$, and $A = \times_{i \in I} A_i$ denotes the set of action profiles. The stage payoff function is denoted $u_i : S \times A \rightarrow \mathbb{R}$ for each player i . The transition probability to state s' from state s , when action profile a was taken in the previous period, is denoted $P(s'|s, a)$.

The game is finitely or infinitely repeated, and the players' common discount factor is denoted $\delta \in (0, 1)$. Since a stochastic game can have multiple stage games, there is a distinction between *ex ante* and *ex post* histories. The set of period- t ex ante histories, denoted $\tilde{\mathcal{H}}^t$, is the set $(S \times A)^t$ for each $t > 0$, which identifies the state and action profile in each period. The set of *ex post* histories also specifies the current state of the game, $(S \times A)^t \times S$, and is denoted by \mathcal{H}^t . Let $\mathcal{H} = \cup_{t=0}^{\infty} \mathcal{H}^t$, with $\tilde{\mathcal{H}}^0 = \{\emptyset\}$, so that $\mathcal{H}^0 = S$.

A pure behavior strategy profile, $\sigma : \mathcal{H} \rightarrow A$, defines an action for each player and each ex post history. Given a strategy profile $\sigma \in \Sigma$, the stochastic game defines a Markov chain on the finite state space S . The normalized expected present value for player i is¹

$$U_i(\sigma) = (1 - \delta)E \left[\sum_{t=1}^T \delta^{t-1} u_i(s_t, a_t) \right] \quad (1)$$

A strategy profile σ constitutes a *Nash equilibrium* of the stochastic game if $U_i(\sigma) \geq U_i(\sigma'_i, \sigma_{-i})$ for all $i \in I$ and all $\sigma'_i \in \Sigma_i$. A strategy profile is a *subgame perfect equilibrium* of the stochastic game if, for any ex post history $h \in \mathcal{H}$, the

1. Multiplying by $(1 - \delta)$ as in (1) is a non-essential normalization, made to simplify calculations.

continuation strategy profile $\sigma|h$ is a Nash equilibrium of the continuation game.

A strategy profile that prescribes the same action for any two histories of equal length and with the same last state, is a *Markov strategy profile*. While in general a strategy profile can depend on the period t , a *stationary Markov strategy profile* is a strategy profile which prescribed actions solely depends on the current state and not on the history leading up to the state. A strategy profile that is both Markov and a subgame perfect equilibrium is a *Markov perfect equilibrium*.

Let $s, s' \in S$ denote two states in a stochastic game. Then s' is *reachable* from s if the probability of visiting state s' after a finite number of periods is positive when starting in s . If s' is reachable from s and s is reachable from s' , the states are said to *communicate*. A stochastic game in which all states are communicating with each other, regardless of the actions of a single player, is called an *irreducible* game. Hoffman and Karp (1966) show that a Nash equilibrium in stationary Markov strategies always exists in irreducible zero-sum games.

Initially the literature on stochastic games was focused on zero-sum games, but the question soon arose whether some classes of nonzero-sum stochastic games have a Nash equilibrium. It was answered affirmatively by Parthasarathy and Sinha (1989): all nonzero-sum stochastic games with state independent transition probabilities have a Nash equilibrium in stationary Markov strategies. In addition, Mertens and Parthasarathy (1991) show that nonzero-sum stochastic games with bounded payoffs have a subgame perfect equilibrium in which each player's strategy only depends on the current and the previous state, and Horst (2005) shows that a Markov perfect equilibrium exists if the state space is compact and the action space is compact and convex.

While folk theorem results are more difficult to prove for stochastic games since they have multiple stage games, there exists folk theorems for subsets of stochastic games. Dutta (1995) proves a folk theorem for irreducible stochastic games, and Fudenberg and Yamamoto (2011) and Hörner et al. (2011) independently prove a folk theorem for irreducible games with imperfect public monitoring. The analysis in this paper builds on the results of Dutta (1995), presented next.

Since stochastic games may have more than one stage game, the definitions of feasible payoff vectors and minmax payoffs need to be generalized. Attention is on the set of pure Markov strategies, since any feasible payoff vector can be

realized by publicly mixing over pure Markov strategies (Dutta, 1995). Let Σ^M denote the set of pure Markov strategy profiles and $\sigma_1, \sigma_2 \in \Sigma^M$ denote two pure Markov strategy profiles. Then the set of feasible payoff vectors, given an initial state s_0 , and discount factor δ , is

$$V(\delta, s_0) = \{v \in \mathbb{R}^n : v = \lambda v_1 + (1-\lambda)v_2 \text{ for some } U(\sigma_1) = v_1, U(\sigma_2) = v_2, \lambda \in [0, 1]\}$$

The set of feasible payoff vectors in the infinitely repeated stochastic game is the convex hull of the range of normalized present values when only pure stationary Markov strategies are used. Given a finite state and action space, the payoffs to the pure Markov strategies can be shown to be continuous at $\delta = 1$. It then follows that the set of feasible payoffs is continuous at $\delta = 1$, so that $V(\delta, s_0) \rightarrow \hat{V}(s_0)$ as $\delta \rightarrow 1$, where the convergence is in the Hausdorff metric (Dutta, 1995).

The minmax payoffs are defined according to the one-period payoffs over the entire game and depend on the initial state and the discount factor. The minmax payoff for player i is

$$\underline{v}_i(\delta, s_0) = \inf_{\sigma_{-i} \in \Sigma_{-i}} \sup_{\sigma_i \in \Sigma_i} U_i(s_0, \sigma, \delta)$$

given an initial state s and discount factor δ . In other words, in a two player game the minmax payoff is the maximum payoff a player can guarantee himself given that the other player aims to minimize his payoff. Let $\hat{v}_i(s_0) = \limsup_{\delta \rightarrow 1} \underline{v}_i(\delta, s_0)$ denote the limit of the minmax payoff as $\delta \rightarrow 1$.

A stochastic game is said to have *asymptotic state independence* if three assumptions are fulfilled

(A1) The set of feasible payoff vectors, $\hat{V}(s) \subset \mathbb{R}^n$, is independent of the initial state: $\hat{V}(s) = \hat{V}(s') \forall s, s' \in S$.

(A2) The minmax payoff, $\hat{v}_i(s)$, is independent of the initial state for all players: $\hat{v}_i(s) = \hat{v}_i(s') \forall s, s' \in S$.

(A3) The dimension of the set of feasible payoff vectors is the same as the number of players: $\text{Dim}(\hat{V}) = n$.

Assume (A1) and let $V^* = \{v \in \hat{V} : v > \hat{v}\}$ denote the set of feasible and individually rational payoff vectors.

Theorem 1 (Dutta, (1995)). *Suppose that (A1)-(A3) hold, and that $\varepsilon > 0$ and $\hat{v} \in V^*$. Then there exist $\underline{\delta} < 1$ such that for every $\delta \geq \underline{\delta}$, there exists a subgame perfect equilibrium with payoff vector within ε of v after all histories with positive probability.*

In an irreducible stochastic game any state can be reached from any other state regardless the actions of a single player. As a consequence, the set of feasible payoff vectors and the minmax payoffs are independent of the initial state and all irreducible games fulfill both (A1) and (A2).

In many interactions at least one player can control some state transitions, a violation of (A2). In those cases, Theorem (1) still holds for any payoff vector that is strictly individually rational for all players and from all states.

Corollary 1 (Dutta, (1995)). *Suppose that (A1) and (A3) hold. Then the claim in Theorem 1 holds for any $v \in \hat{V}^*$ such that $v_i > \underline{v}_i(s) \forall i, s$.*

Theorem 1 and Corollary 1 are closely related to the Folk theorems for repeated games. All feasible and individually rational payoff vectors can be (approximately) received through play of a subgame perfect equilibrium strategy profile.

In standard repeated games the one-shot deviation principle is used to verify that a strategy profile is a subgame perfect equilibrium. In a stochastic game the players have two reasons to deviate from a strategy. As in the repeated game they may do so for the immediate payoff, but there is also an incentive to deviate in order to influence future state transitions. As the players become more patient the second incentive becomes more important. In a stochastic game with either a finite set of states or a deterministic transition function, a strategy $\hat{\sigma}_i$ is a *one-shot deviation* for player i , from strategy σ_i , if there is only one history at which the strategies disagree. A strategy profile σ in a stochastic game with finite set of states or with a deterministic transition function is subgame perfect if, and only if, there are no profitable one-shot deviations (Mailath and Samuelson, 2006).

3 Model

This paper studies repeated interactions between players with emotionally driven social preferences. Focus is on two player games, $n = 2$, and more precisely

on a repeated prisoners' dilemma game (form) with fixed *material payoffs*, as illustrated in Table 1. Each player has two actions, cooperate (C) and defect (D). If both players choose C , they receive a material payoff of c each; if both choose D , they receive a material payoff of d each. If one of them chooses C while the other chooses D , the second player receives a material payoff of b , and the first player receives a material payoff of 0. With $b > c > d > 0$ each player has a dominant action, D , but the payoff vector is Pareto dominated by the payoff vector received if both choose C .

Table 1: Material payoffs.

	C	D
C	c, c	$0, b$
D	$b, 0$	d, d

Let $A = \{C, D\}$ denote the set of available actions. An action profile, $a = (a_1, a_2) \in A^2$, is a pair of actions. Both players have the same finite set of potential states of mind, M . When the players' preferences over own and other's material payoffs depend on their current state of mind, the repeated interaction can be modeled as a stochastic game with state space $S = M^2$. Attention is in this paper restricted to Markov players. In other words, players whose state of mind is only affected by the last observed action of the other player, and not by any action further back in the history of play.² Each player i has transition probabilities $P_i(m|s, a)$ that, given an action profile, $a \in A^2$, and the current state, $s \in S$, gives the probability to transition to state of mind $m \in M$.

Given the interaction's material payoffs, each state of mind defines a vector of corresponding utilities. The one-period utility function $u_i(s, a)$ is

$$u_i(s, a) = \pi_i(a) + \phi(m_i)\pi_j(a) \tag{2}$$

The first term is utility derived from a material payoff, π_i . The second term is utility derived from a psychological payoff, $\phi(m_i)\pi_j$. In other words, utility derived from a positive or negative concern for the other player's material payoff.

2. Moreover, the transitions could also be assumed to be affected by exogenous factors such as the weather or a good night's sleep, or by one's own actions, such as feelings of guilt or a sense of generosity.

In this paper there are at most three states of mind: friendly (F), neutral (N), and hostile (H). In the friendly state of mind, the player cares positively for the other player's material payoff, $\phi(F) = \alpha \in \mathbb{R}_+$. In the neutral state of mind, he only cares for his own material payoff, $\phi(N) = 0$. In the hostile state of mind he cares negatively for the other player's material payoff, $\phi(H) = -\gamma \in \mathbb{R}_+$.

Focus is on two types of players, traditional and emotional. The player types differ in their transition probabilities. The traditional player, or *Homo oeconomicus*, starts in the neutral state of mind and cannot transition to any other state (such transitions have a zero probability). The *emotional player* also starts in the neutral state of mind, but can transition between all states of mind (such transitions have a positive probability).

In each period a simultaneous-move game of complete information and perfect monitoring is played, and the state is drawn and revealed to the players before they choose their action. The players have a common discount factor denoted $\delta \in (0, 1)$.³ Their total expected utility is presented in (1).

In Sections 4 and 5, focus is on play between two combinations of player types. The first consists of one *Homo oeconomicus* and one emotional player, and the second of two identical emotional players. Further, each combination is compared to the *benchmark case* of play between two *Homo oeconomicus*. In Section 4 the players have two states of mind, $M = \{N, H\}$, and hence in Section 5 the friendly state of mind is introduced, $M = \{F, N, H\}$.

Two subgame perfect strategy profiles, Grim Trigger and a mutual minmax, is contrasted for the infinitely repeated interaction. Grim Trigger is a well-known and commonly studied strategy profile in the repeated prisoners' dilemma game literature. In the Grim Trigger strategy profile, both players choose C as long as both have always chosen C , and play the Nash equilibrium otherwise. The mutual minmax strategy profiles have been shown to be sufficient for the folk theorem results (Fudenberg and Maskin, 1986). If both players use a mutual minmax strategy, and the players are sufficiently patient, then any feasible and individually rational payoff vector can be approximately realized (for more information see e.g. Mailath and Samuelson (2006), section 3). In the mutual minmax strategy profile, the players choose C as long as both have always chosen C ; they

3. While the players' discount factor may change between their states of mind, and different types of players can be assumed to discount the future differently, the model in this paper assumes a common discount factor, as is conventional for repeated interactions.

choose C also if the history contains L consecutive periods of mutual play of D after which only C has been played; they play the mutual minmax action, D , for L consecutive periods after all other histories. Let any history after which the players choose C be denoted *the cooperative phase*, and any history after which the players choose D be denoted *the punishment phase*.

4 Two states of mind

Assume the players have a neutral and a hostile state of mind, such that $M = \{N, H\}$. When the game begins, all player types are in the neutral state of mind. The emotional player transitions to the hostile state of mind if the other player chooses D . If the other player chooses C , the player transitions back to the neutral state of mind.⁴ Since the game is of complete information, the transition probabilities, displayed in Figure 1, are known to both players.

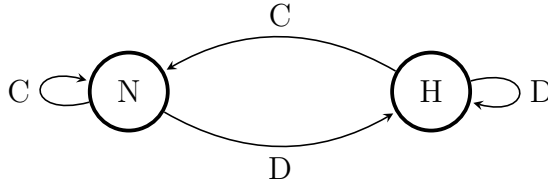


Figure 1: State transitions.

4.1 One *Homo oeconomicus* and one emotional player

Let player 1 be a *Homo oeconomicus*, and player 2 be an emotional player as defined above. The corresponding stochastic game has four states, M^2 , but only two are reachable from the initial state since by definition the *Homo oeconomicus* never leaves the neutral state of mind: $S = \{NN, NH\}$. The utility matrices are illustrated in Table 2. Player 1 is the row player and player 2 is the column player.

As seen in Table 2, player 2's utility from choosing D in the hostile state of mind is smaller compared to in his neutral state of mind. However, so is his utility from choosing C , and the decrease in utility from choosing C is larger than the decrease in utility from choosing D . Suppose that player 2 is in the

4. The state transitions are deterministic in this example, this is relaxed in Section 4.3.

illustrated in Figure 2, indicates the differences compared to the benchmark case. In this subset of utility vectors, player 2's utility is always lower while player 1's utility can be higher. Consider the case with $b = 4$, $c = 3$, $d = 2$, and $\gamma = 0.2$. The maximum utility for any player in the benchmark case is $3\frac{1}{3}$, whereas player 1 can receive a utility of 3.4 in this game.

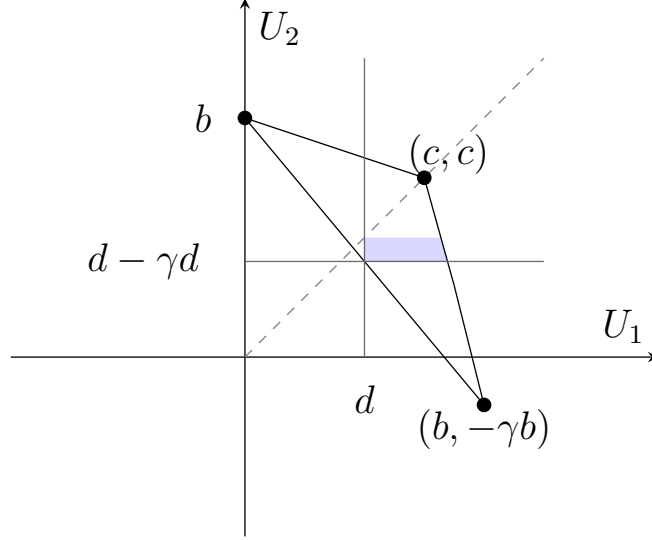


Figure 2: Feasible and individually rational utility vectors in the stochastic prisoners' dilemma game, for $\delta = 1$ and $\gamma = 0.2$.

The cooperative utility vector, (c, c) , belongs to the set of feasible and individually rational utility vectors. Moreover, as we will see, both the Grim Trigger and mutual minmax strategy profiles can sustain this outcome as a subgame perfect Nash equilibrium, given that the players are sufficiently patient.

4.1.1 Grim Trigger

The Grim Trigger strategy profile is subgame perfect if neither of the two players has an incentive to deviate from it. Player 1 has no incentive to deviate if his normalized discounted utility from using the strategy, c , is higher than the sum of his normalized discounted utility from his best one-shot deviation, $(1 - \delta)b$, and from continued play of the Nash equilibrium, δd .

Player 1 has no incentive to deviate iff

$$c \geq (1 - \delta)b + \delta d \quad (3)$$

This condition is identical to the corresponding one in the benchmark case.⁶

Now turn to the emotional player. If the players use the strategy, player 2 receives a normalized discounted utility of c . If player 2 deviates he receives a one-period normalized discounted utility of $(1 - \delta)b$. The players then use the strategy and choose D for the remaining interaction. However, in the first period after the deviation, player 2 still is in the neutral state of mind. He therefore receives a one-period normalized discounted utility of $(1 - \delta)\delta d$. Once player 1 has chosen D , player 2 transitions to the hostile state of mind. He remains there for the rest of the interaction, and receives a normalized discounted utility of $\delta^2(d - \gamma d)$.

Player 2 has no incentive to deviate iff

$$c \geq (1 - \delta)b + (1 - \delta)\delta d + \delta^2(d - \gamma d) \quad (4)$$

When $\gamma = 0$, (4) is identical to (3). An increase in γ decreases player 2's utility from deviating since choosing D forever is more costly for him due to his hostile state of mind. Hence, if (3) holds, so does (4). The *Homo oeconomicus* thus has the binding restriction on the minimal discount factor required for the Grim Trigger strategy profile to be subgame perfect.

Let $\underline{\delta}^*$ and $\underline{\delta}^{GT}$ be the minimal discount factors required for the Grim Trigger strategy profile to be subgame perfect in the case of play between one emotional player and one *Homo oeconomicus*, and in the benchmark case respectively.

Proposition 1. *Suppose that both players use the Grim Trigger strategy. Then $\underline{\delta}^* = \underline{\delta}^{GT}$.*

4.1.2 Mutual minmax

Consider next the mutual minmax strategy profile defined in Section 3, in which the players punish each other for L consecutive periods after a deviation. In the benchmark case, the players may have an incentive to deviate either during the cooperative phase or during the punishment phase. Since the mutual minmax action profile, (D, D) , coincides with the unique Nash equilibrium, the players have no incentive to deviate during the punishment phase. In the cooperative

6. Note that player 2 cannot hurt player 1 more than by choosing D , and since player 1 only cares for his own material payoff, he is unaffected by player 2's state of mind.

phase the players can have an incentive to deviate, and the mutual minmax strategy profile is subgame perfect in the benchmark case iff

$$c \geq (1 - \delta)b + (1 - \delta)\delta d + \dots + (1 - \delta)\delta^L d + \delta^{L+1}c \quad (5)$$

The normalized discounted utility from deviating is the sum of three terms: the one-shot deviation utility of $(1 - \delta)b$, the utility received during the punishment phase of L periods, $(1 - \delta)(\delta + \dots + \delta^L)d$, and the utility received from returning to choosing C after the punishment period, $\delta^{L+1}c$.

The above condition can also be presented as a cost-benefit calculation. The benefit of deviating is the utility from the one-period deviation, b , minus the one-period utility from using the strategy, c . The cost of deviating is the utility received if no one had deviated, c , minus the utility received during the punishment phase, d , for the L consecutive periods of punishment. Thus (5) can be rewritten as

$$b - c \leq (c - d)(\delta + \delta^2 + \dots + \delta^L) \quad (6)$$

In the case of one *Homo oeconomicus* and one emotional player, the condition required for the *Homo oeconomicus* not to deviate remains as in the benchmark case.

Now turn to the emotional player. As the *Homo oeconomicus*, the emotional player may have an incentive to deviate either during the cooperative phase or during the punishment phase. Further, his incentive to deviate also depends on his state of mind.

First consider the equilibrium path history, in which no player so far has chosen D . Start by assuming a punishment phase of length one, $L = 1$. Since no player so far has chosen D , player 2 is in the neutral state of mind. His benefit from his best one-shot deviation is $b - c$. His cost from deviating consists of two parts. In the first period of the punishment phase, player 2 is in the neutral state of mind and incurs a normalized discounted cost of $(c - d)\delta$. He then transitions to the hostile state of mind, and therefore receives a smaller utility from choosing C for one period before transitioning to the neutral state of mind. This implies a second cost of $\delta^2\gamma c$.

Suppose that $L = 1$, and $s = NN$. Then player 2 has no profitable one-shot

deviation in the cooperative phase iff

$$b - c \leq (c - d)\delta + \delta^2\gamma c \quad (7)$$

When $\gamma = 0$, (7) is identical to (6). When $\gamma > 0$, player 2 has a higher cost of deviating than do player 1.

Now assume $L = 2$. Player 2 is in the neutral state of mind, and his benefit from deviating is as before. So is his cost in the first punishment period. In the second punishment period, player 2 has transitioned to the hostile state of mind, and incurs a normalized discounted cost of $(c - d + \gamma d)\delta^2$. Finally, player 2 also incurs the cost associated with transitioning back to the neutral state of mind, $\delta^3\gamma c$, once the players return to cooperation.

Suppose that $L = 2$, and $s = NN$. Then player 2 has no profitable one-shot deviation in the cooperative phase iff

$$b - c \leq (c - d)\delta + (c - d + \gamma d)(\delta^2) + \delta^3\gamma c \quad (8)$$

For punishment phases of length $L > 1$, the above condition can be generalized to

$$b - c \leq (c - d)\delta + (c - d + \gamma d)(\delta^2 + \dots + \delta^L) + \delta^{L+1}\gamma c \quad (9)$$

When $\gamma = 0$, (9) is identical to (6). An increase in γ increases the cost of deviation and thus decreases the patience required to use the strategy.

Next consider the off the equilibrium path histories. If one of the players has deviated in the previous period, the players are supposed to play the mutual minmax action profile, (D, D) for L consecutive periods. If player 1 deviated, then player 2 is in the hostile state of mind and (D, D) is the unique Nash equilibrium of the stage game. Similarly, if player 2 deviated, then he is in the neutral state of mind and (D, D) is still the unique Nash equilibrium of the stage game. Consequently, regardless of who deviated, neither player has an incentive to deviate during the punishment phase.

After a punishment phase, the players are supposed to return to cooperation, but player 2 is then in the hostile state of mind. He is upset and revengeful, and receives a lower utility from cooperating with player 1. Consequently, player 2's cost-benefit calculation differ from when he is in the neutral state of mind. His benefit from deviating is higher, $b - c + \gamma c$, whereas his cost is unchanged.

Suppose that $L = 1$, and $s = NH$. Then player 2 has no profitable one-shot deviation in the cooperative phase iff

$$b - c + \gamma c \leq (c - d)\delta + \delta^2\gamma c \quad (10)$$

When $\gamma = 0$, (10) is identical to (6). An increase in γ increases both the benefit and the cost from deviating. When $L = 1$, the increased benefit dominates, and an increase in γ increases the patience required for player 2 to return to cooperation. Similarly for $L > 1$, player 2's benefit from deviating in the hostile state of mind is higher, whereas his cost is unchanged.

Suppose that $L > 1$, and $s = NH$. Then player 2 has no profitable one-shot deviation in the cooperative phase iff

$$b - c + \gamma c \leq (c - d)\delta + (c - d + \gamma d)(\delta^2 + \delta^3 + \dots + \delta^L) + \delta^{L+1}\gamma c \quad (11)$$

When $\gamma = 0$, (11) is identical to (6). An increase in γ increases both the benefit and the cost from deviating. When $L > 1$, γ also increases the cost of enduring the punishment phase in the hostile state of mind. As a consequence, an increase in γ can decrease the patience required for player 2 to return to cooperation if he is sufficiently patient. The cost increase dominates if

$$\frac{c}{d} < \frac{\delta^2}{1 - \delta} \quad (12)$$

If this inequality holds, then the *Homo oeconomicus* has the binding restriction on the minimal discount factor required for the mutual minmax strategy profile to be subgame perfect.

Let $\underline{\delta}^*$ and $\underline{\delta}^{MM}$ be the minimal discount factors required for the mutual minmax strategy profile to be subgame perfect in the case of play between one emotional player and one *Homo oeconomicus*, and in the benchmark case respectively.

Proposition 2. *Suppose that both players use the mutual minmax strategy with a punishment length $L > 1$. If (12) holds, then $\underline{\delta}^* = \underline{\delta}^{MM}$.*

To summarize, if the players use the Grim Trigger strategy, then player 2 prefers to choose C even when less patient than player 1 since he knows that the punishment of playing the Nash equilibrium forever after a deviation will be

more costly. On the other hand, if the players use the mutual minmax strategy, then player 2 might require more patience than player 1 due to his difficulty to commit to returning to cooperation after a punishment phase.

4.2 Two emotional players

Still assume two states of mind, $M = \{N, H\}$, but now both players can transition between them. The transition probabilities are as in Figure 1.⁷ The corresponding stochastic game has four states $S = \{NN, NH, HN, HH\}$, all reachable from the initial state. The corresponding utility matrices are displayed in Table 3.

Table 3: Utility matrices.

	C	D
C	c, c	$0, b$
D	$b, 0$	d, d

(a) $s = NN$

	C	D
C	$c, c - \gamma c$	$0, b$
D	$b, -\gamma b$	$d, d - \gamma d$

(b) $s = NH$

	C	D
C	$c - \gamma c, c$	$-\gamma b, b$
D	$b, 0$	$d - \gamma d, d$

(c) $s = HN$

	C	D
C	$c - \gamma c, c - \gamma c$	$-\gamma b, b$
D	$b, -\gamma b$	$d - \gamma d, d - \gamma d$

(d) $s = HH$

First consider the finitely repeated interaction. Each stage game has a unique Nash equilibrium in which each player choose D . Hence, the only subgame perfect Nash equilibrium of the finitely repeated interaction is for each player to choose D in each period.

4.2.1 Grim Trigger

Next consider the infinitely repeated interaction. In the hostile state of mind, the players care only for each other's material payoffs, and not each other's psychological payoff or utility. As a consequence, the condition for an emotional player

⁷ The players' transitions are independent of the other player's state of mind. As a consequence, one player can be in a neutral state of mind while the other is in a hostile state of mind. Thus, although emotions can be contagious, this is not reflected in the model.

not to deviate from a strategy profile does not depend on whether he interacts with another emotional player or a *Homo oeconomicus*.

In the infinitely repeated interaction between two identical emotional players, the Grim Trigger strategy profile is subgame perfect iff (4) holds. As previously noted, when $\gamma = 0$, (4) is identical to the condition for the benchmark case. An increase in γ increases the cost of deviating due to the negative concern for the other player's material payoff in the hostile state of mind. Since the players can anticipate their own future state of mind they prefer to choose C also for lower values of δ than in the benchmark case.

Two emotional players require less patience than two *Homo oeconomicus* for the Grim Trigger strategy profile to be subgame perfect, and an increase in γ decreases the required patience. Let $\underline{\delta}^*$ and $\underline{\delta}^{GT}$ be defined as before.

Proposition 3. *Suppose that both players use the Grim Trigger strategy. Then $\underline{\delta}^* < \underline{\delta}^{GT}$, and $\underline{\delta}^*$ is decreasing in γ .*

4.2.2 Mutual minmax

Suppose now that the players use the mutual minmax strategy profile, and let $L = 1$. On the equilibrium path, where no player so far has chosen D , the players have no profitable one-shot deviation if (7) holds. When $L > 1$, (9) needs to hold. Since an increase in γ increases the cost of deviation, the emotional players require less patience not to deviate on the equilibrium path than do a *Homo oeconomicus*. Further, the mutual minmax action profile (D, D) is the unique Nash equilibrium of all stage games, and regardless of who deviated neither player has an incentive to deviate during the punishment phase.

Once the players are in the hostile state of mind their utility from choosing C is decreased, and they might require more patience to return to cooperation. For $L = 1$ the condition is identical to (10), and an increase in γ increases the patience required to return to cooperation. For $L > 1$ the condition is identical to (11), and, given that (12) holds, an increase in γ decreases the patience required to return to cooperation. Hence, if the players are sufficiently patient, an increase in γ decreases the patience required for the mutual minmax strategy profile to be subgame perfect.

Let $\underline{\delta}^*$ and $\underline{\delta}^{MM}$ be defined as before.

Proposition 4. *Suppose that both players use the mutual minmax strategy with a punishment length $L > 1$. If (12) holds, then $\underline{\delta}^* < \underline{\delta}^{MM}$, and $\underline{\delta}^*$ is decreasing in γ .*

To summarize, two emotional players can sustain cooperation easier than two *Homo oeconomicus* if they use the Grim Trigger strategy. Emotionality can in this situation be interpreted as a substitute for patience.⁸ Sometimes the players need to punish each other during a finite period of time, e.g. if they have incomplete information. After a punishment phase, the players are in the hostile state of mind and find it more difficult to return to cooperation. Only if they are sufficiently patient can two emotional players find it easier to sustain cooperation than two *Homo oeconomicus* when using the mutual minmax strategy.

4.3 Two emotional players with stochastic transitions

The previous analysis is limited to the case of deterministic state transitions and although this assumption simplifies the analysis, this is not a restriction to the model. Let the two emotional players be of the same type with $M = \{N, H\}$ and stochastic state transitions; a generalization of the previously discussed emotional player.

The state transitions are illustrated in Figure 3, where $P|C$ denotes the probability, conditional on observing action C , of transitioning from the current state to the neutral state, and $P|D$ denotes the probability, conditional on observing action D , of transitioning from the current to the hostile state. The corresponding stochastic game has four states $S = \{NN, NH, HN, HH\}$, all reachable from the initial state.

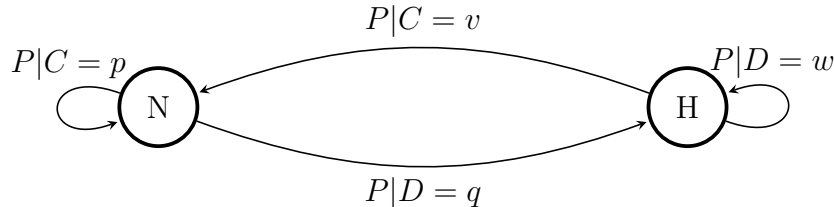


Figure 3: State transitions.

8. Emotional players are not by definition more patient than others. However, their states of mind can cause them to act as if they were more patient.

Focus is on the specific case where the players have $p = q = w = 1$ and $v < 1$, and remain longer in the hostile state of mind before transitioning to the neutral state of mind after observing C . This allows for more or less resentful players, where a player is less resentful the larger v is.

4.3.1 Grim Trigger

If the players use the Grim Trigger strategy, they cannot be distinguished from the emotional player type with deterministic state transitions. The players only transition to the hostile state of mind after a deviation, and since they then choose D forever they never get the chance to transition back to the neutral state of mind.

4.3.2 Mutual minmax

Consider instead the mutual minmax strategy profile defined in Section 3 with a punishment phase of length $L > 1$. If no player has chosen D , then $s = NN$ and neither player has a profitable one-shot deviation from cooperation iff

$$b - c \leq (c - d)\delta + (c - (1 - \gamma)d)(\delta^2 + \delta^3 + \dots + \delta^L) + \frac{\delta^{L+1}\gamma c}{1 - (1 - v)\delta} \quad (13)$$

When $v = 1$, (13) is identical to (9). A decrease in v increases the players' cost from deviating by increasing the number of periods they spend in the hostile state of mind.

Since v does not affect the stage games, (D, D) is still the unique Nash equilibrium of the one-shot stage games, and the players have no incentive to deviate during the punishment phase. After a punishment phase the players should return to cooperation, but once they are in the hostile state of mind they might be less inclined to do so.

Suppose that $L > 1$, and $s = HH$. Then the players have no profitable

one-shot deviation from cooperation iff

$$\begin{aligned}
b - (1 - \gamma)c &\leq [(1 - (1 - v)\gamma)(c - d)] \cdot \delta \\
&+ [(1 - (1 - v)^2\gamma)c - (1 - \gamma)d] \cdot \delta^2 \\
&+ [(1 - (1 - v)^3\gamma)c - (1 - \gamma)d] \cdot \delta^3 \\
&+ \dots \\
&+ [(1 - (1 - v)^L\gamma)c - (1 - \gamma)d] \cdot \delta^L \\
&+ \frac{(1 - (1 - v)^{L+1})\gamma c \cdot \delta^{L+1}}{1 - (1 - v)\delta}
\end{aligned} \tag{14}$$

When $v = 1$, (14) is identical to (11). The probability of remaining in the hostile state of mind even after cooperation has been resumed has two effects on the profitability of deviating for a player in the hostile state of mind. The cost of the punishment phase is lower since the player, with some probability, will remain in the hostile state of mind even if cooperation is resumed. Hence a decrease in v decreases the cost of the punishment and makes it more profitable for the player to deviate. The player also faces the cost of remaining in the hostile state of mind even after cooperation is resumed and a decrease in v increases the number of periods spent in the hostile state of mind, thereby making it less profitable for the player to deviate. The first effect dominates and a decrease in v increases the player's cost of deviating also in the hostile state of mind.

Let $\underline{\delta}^*$ be the minimal discount factor required for the mutual minmax strategy profile to be subgame perfect between the two emotional players with stochastic state transitions.

Proposition 5. *Suppose that both players use the mutual minmax strategy with a punishment length L . Then $\underline{\delta}^*$ is decreasing in v .*

In other words, more resentful players require more patience than less resentful players to sustain cooperation in the prisoners' dilemma game when using the mutual minmax strategy. This is because of their difficulty of returning to cooperation after a punishment phase.

5 Three states of mind

Now consider players with three states of mind, $M = \{F, N, H\}$. The emotional player has deterministic state transitions deterministic as illustrated in Figure 4. He starts in the neutral state of mind and transitions to the friendly state of mind if the other player chooses C , and to the hostile if the other player chooses D . In the friendly state of mind he transitions to the hostile if the other player chooses D , and remains if the other player chooses C . Finally, in the hostile state of mind he transitions to the neutral if the other player chooses C , and remains if the other player chooses D .

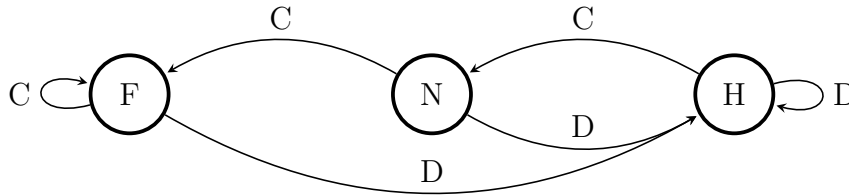


Figure 4: State transitions.

5.1 One *Homo oeconomicus* and one emotional player

Let player 1 be a *Homo oeconomicus*, and player 2 be an emotional player as defined above. The corresponding stochastic game has three states that are reachable from the initial state, $S = \{NF, NN, NH\}$. The utility matrix for each state is illustrated in Table 4. Player 1 is the row player and player 2 is the column player.

As seen in Table 4, player 2's utility from choosing D in the friendly state of mind is higher compared to in his neutral state of mind. However, so is his utility from choosing C , and the increase in utility from choosing C is larger than the increase in utility from choosing D . Suppose that player 2 is in the friendly state of mind and believes that player 1 will choose C with probability p . Then his expected utility from choosing C is: $p(c + \alpha c) + (1 - p)\alpha b$. Likewise, his expected utility from choosing D is: $pb + (1 - p)(d + \alpha d)$. Thus, given p , an increase in α increases player 2's utility from choosing C more than it increases his utility from choosing D .

Table 4: Utility matrices.

	C	D		C	D
C	$c, c + \alpha c$	$0, b$	C	c, c	$0, b$
D	$b, \alpha b$	$d, d + \alpha d$	D	$b, 0$	d, d
	(a) $s = NF$			(b) $s = NN$	

	C	D
C	$c, c - \gamma c$	$0, b$
D	$b, -\gamma b$	$d, d - \gamma d$
	(c) $s = NH$	

First consider the finitely repeated game. Still assume that $b > c > d > 0$. For (D, D) to be a Nash equilibrium in the stage game in the friendly state, α has to be sufficiently small. If $\alpha b > d + \alpha d$, then (D, C) is the unique Nash equilibrium in that stage game. Further, if $c + \alpha c > b$, then C is the emotional players best response to the other player choosing C . Moreover, as we will see, when the emotional player can care sufficiently for the other player, the other player has an incentive to exploit this to his own advantage.

Consider the finitely repeated game with a payoff structure that satisfies the above conditions, and let $T = 2$. In the last period, the players have to play the Nash equilibrium of that stage game: if they are in the friendly state, player 1 chooses D , and player 2 chooses C . In the first period, player 1 has to choose C for the players to transition to the friendly state. Player 2 best responds by choosing D . Player 1 uses the first period to manipulate player 2 into the friendly state of mind. He is willing to do this if he is sufficiently patient, $\delta > d/(b - d)$.

Let Σ^* be the set of subgame perfect equilibria of the finitely repeated interaction.

Proposition 6. *Suppose that $T = 2$. If $c + \alpha c > b$, $\alpha b > d + \alpha d$, and $\delta > d/(b - d)$ hold, then there exists $\sigma^* \in \Sigma^*$, such that $a_i(t, s|\sigma^*) = \{C\}$ for some $t \leq T$ and $s \in S$.*

Consider next the case of $T = 3$. Player 1 still has to be sufficiently patient to choose C in the first period. Further, in the second period, if the players are in the friendly state, then player 2 chooses C , his dominant action, whereas player 1 choose between C and D . By choosing C , he postpones the higher utility until

the last period, and ensures himself a utility of $\delta c + \delta^2 b$. By choosing D , he receives the higher utility in this period, but receives a lower utility in the last period, $\delta b + \delta^2 d$. If player 1 is sufficiently patient, $\delta > (b - c)/(b - d) \in (0, 1)$, he prefers to choose C in the second period. Finally, in the last period, the players play the Nash equilibrium of that stage game.

Following from the above, for $T > 2$, player 1 has more opportunities to benefit from player 2's friendly state of mind. Player 1 will take these opportunities if he is sufficiently patient. The patience required for player 1 to manipulate player 2 into the friendly state of mind in the first period is then decreasing in T .

For each finite time horizon T , there is a minimal discount factor, $\delta^\circ(T)$, required for the action profile (C, C) to be played on the equilibrium path of the finitely repeated game.

Proposition 7. *Suppose that $T > 2$. If $c + \alpha c > b$ and $\alpha b > d + \alpha d$ hold, then $\delta^\circ(T) > \delta^\circ(T')$ if $T < T'$.*

Now turn to the infinitely repeated interaction. The game is not irreducible since player 1 can control the state transitions, but, as in the previous section, Corollary 1 presented in Section 2 is applicable. The set of feasible utility vectors is spanned by the present values of the pure stationary Markov strategy profiles. Player 1's minmax utility is $\underline{v}_1(s) = d$. When $s_0 = NN$, player 2's minmax utility is $\underline{v}_2(\delta, s_0) = d - \gamma d$.

The shaded areas in Figure 5 mark the differences to the benchmark case of play between two *Homo oeconomicus*. Player 2 can achieve a higher utility than if he was a *Homo oeconomicus* if he is mostly in the friendly state of mind. He can also achieve a lower utility than if he was a *Homo oeconomicus* if he is mostly in the hostile state of mind. There is also a subset of feasible and individually rational utility vectors in which both players receive a higher utility than what is possible for two *Homo oeconomicus*. This is due to player 2's positive concern for the material payoff of player 1, which increases the maximum utility that can be received in each period. If $b = 4$, $c = 3$, $d = 2$, then the maximum utility for any player in the benchmark case is $3\frac{1}{3}$, and the cooperative utility vector is $(3, 3)$. If, in addition, $\gamma = \alpha = 0.2$, the maximum utility player 1 can receive from the interaction with the emotional player is 3.45, and the maximum utility the emotional player can receive is 3.74. The maximum symmetric utility vector is $(3.11, 3.11)$.

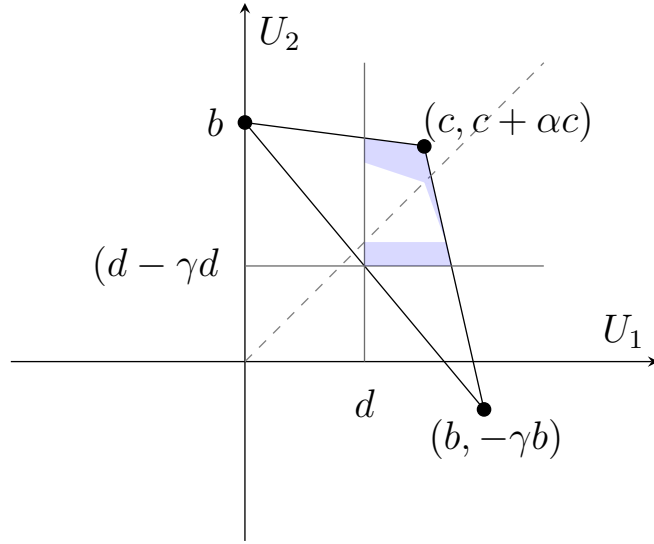


Figure 5: Feasible and individually rational utility vectors in the stochastic prisoners' dilemma game, for $\delta = 1$ and $\alpha = \gamma = 0.2$.

The cooperative utility vector, $(c, c + \alpha c)$, belongs to the set of feasible and individually rational utility vectors. Moreover, as we will see, both the Grim Trigger and the mutual minmax strategy profiles can sustain this outcome as a subgame perfect Nash equilibrium.

5.1.1 Grim Trigger

Consider first the Grim Trigger strategy profile. As before, player 1 is unaffected by player 2's states of mind, and his condition for Grim Trigger to be subgame perfect is identical to (3).

Now turn to the emotional player. If player 2 uses the strategy, he receives a one-period utility in the neutral state of mind before transitioning to the friendly state of mind and remaining there for the rest of the interaction. The normalized discounted utility from the first period is $(1 - \delta)c$, and the normalized discounted utility from the remaining interaction is $\delta(c + \alpha c)$. If player 2 instead deviates he chooses between deviating in the neutral or in the friendly state of mind. It is more profitable for him to deviate in the neutral state of mind, since the immediate benefit from cooperation is lower in that state. The normalized discounted utility in the first period is then $(1 - \delta)b$. In the second period the Nash equilibrium action profile (D, D) is played. Since player 1 chose C in the first period, player 2 is in the friendly state of mind. Player 2's normalized discounted

utility in the second period is $(1 - \delta)\delta(d + \alpha d)$. In the third period, player 2 has transitioned to the hostile state of mind and will stay there for the remaining interaction, receiving a normalized discounted utility of $\delta^2(d - \gamma d)$.

If $s = NN$, then player 2 has no profitable one-shot deviation from the Grim Trigger strategy profile iff

$$(1 - \delta)c + \delta(c + \alpha c) \geq (1 - \delta)b + (1 - \delta)\delta(d + \alpha d) + \delta^2(d - \gamma d) \quad (15)$$

When $\alpha = \gamma = 0$, (15) is identical to (3). An increase in either α or γ decreases player 2's profitability of deviating. His positive concern for the other player's material utility increases his benefit from cooperation, whereas his negative concern increases his cost of deviating. Player 2 requires less patience than player 1 not to deviate, and the common discount factor required for Grim Trigger to be subgame perfect is determined by the condition for player 1.

Let $\underline{\delta}^*$ and $\underline{\delta}^{GT}$ be defined as before.

Proposition 8. *Suppose that both players use the Grim Trigger strategy. Then $\underline{\delta}^* = \underline{\delta}^{GT}$.*

5.1.2 Mutual minmax

Consider next the mutual minmax strategy profile with a punishment phase of length L . Player 1's condition is identical to the benchmark case, (6).

Now turn to the emotional player, and assume that $L = 1$. If no player so far has deviated, player 2 finds it more profitable to deviate in the neutral than in the friendly state of mind. The one-period benefit from deviating is the utility from the best one-shot deviation, minus the utility from using the strategy: $b - c$. In the punishment period, player 2 has transitioned to the friendly state of mind. He receives a utility of $d + \alpha d$, instead of the utility he would have received had he not deviated, $c + \alpha c$. In addition, after the punishment period player 2 transitions to the hostile state of mind. As a consequence, he receives a smaller utility from choosing C when the players return to cooperation. Since it takes two periods for player 2 to transition to the friendly state of mind he incurs an additional cost of $\delta^2(\alpha + \gamma)c + \delta^3\alpha c$.

If $L = 1$ and $s = NN$, then player 2 has no profitable one-shot deviation in

the cooperative phase iff

$$b - c \geq (1 + \alpha)(c - d)\delta + \delta^2(\alpha + \gamma)c + \delta^3\alpha c \quad (16)$$

When $\alpha = \gamma = 0$, (16) is identical to (6).

Now assume $L = 2$. Player 2's one-period benefit from deviating is as before. So is the cost in the first period of the punishment phase. In the second period of the punishment phase, player 2 has transitioned to the hostile state of mind, and receives a lower utility, $d - \gamma d$, than in the friendly state of mind. Consequently, the cost of punishment is higher for player 2 when he is in the hostile state of mind. In addition, player 2 still faces the cost associated with transitioning back to the friendly state of mind.

For punishment phases of length $L > 1$, the condition can be generalized to

$$b - c \leq (1 + \alpha)(c - d)\delta + ((1 + \alpha)c - (1 - \gamma)d)(\delta^2 + \delta^3 + \dots + \delta^L) + \delta^{L+1}(\alpha + \gamma)c + \delta^{L+2}\alpha c \quad (17)$$

When $\alpha = \gamma = 0$, (17) is identical to (6). Player 2's states of mind have three effects on his profitability of deviating in the neutral state of mind. First, his utility from continued cooperation is larger since he, in the friendly state of mind, has a positive concern for the other player's material payoff. Second, his utility during the punishment phase is smaller if $L > 1$, due to his negative concern for the other players material payoff. Third, it takes time for player 2 to calm down and transition back to a friendly state of mind once the players have returned to cooperation.

After a deviation, the players are supposed to play the mutual minmax action profile for L periods. If the one who deviated was player 1, then player 2 is in the hostile state of mind and has no incentive to deviate from the punishment. However, if player 2 deviated, he has transitioned to the friendly state of mind. If, in addition, $ab > (1 + \alpha)d$, then he might have an incentive to deviate during the punishment phase.

If player 2 deviates from punishment, his immediate benefit is: $ab - d - \alpha d$. After this deviation, the players are supposed to restart the punishment phase, and now player 2 has transitioned to the hostile state of mind. The normalized discounted utility from the extra period of punishment is $\delta^{L+1}(d - \gamma d)$. If player

2 had used the strategy, then he would have received $\delta^{L+1}(c - \gamma c)$. In addition, player 2 also faces the delayed cost of transitioning back to the friendly state of mind.

If $s = NF$, then player 2 has no profitable one-shot deviation from punishment iff

$$\alpha b - (1 + \alpha)d \leq \delta^{L+1}(1 - \gamma)(c - d) + \delta^{L+2}\gamma c + \delta^{L+3}\alpha c \quad (18)$$

The friendly state of mind has two effects on the profitability of deviating during the punishment phase. It increases the cost of punishing the other player due to the positive concern for the other player's material payoff, but also increases the cost of deviating due to the cost of transitioning back to the friendly state of mind after the punishment phase. When $\delta^{L+3} > (b - d)/c$ the first effect dominates and an increase in α increases the profitability of deviating from the punishment. This is in line with Bernheim and Stark (1988) who find that a positive concern for the other players' payoff has two conflicting effects on cooperation. It facilitates cooperation by making it less profitable to deviate, but can also obstruct cooperation by making it more costly to punish the other player for deviating.

Further, the hostile state of mind also has two effects on the profitability of deviating during the punishment phase. It increases the cost of deviating from punishment by increasing the cost of remaining in the hostile state, but also decreases the cost of deviating by decreasing the difference in utility between cooperation and punishment. If player 2 is sufficiently patient, $\delta > (c - d)/c$, the first effect dominates, and an increase in γ increases the cost of deviating from the punishment. Thus, while player 2's positive concern for the other player's material payoff can make threats of punishment less credible, the possibility of transitioning to the hostile state of mind can offset this effect.

Finally, after a punishment phase the players are supposed to return to choosing C , but player 2 may require more patience to return to cooperation once he is in the hostile state of mind.

First suppose that $L = 1$, and player 2 is in the hostile state of mind. Then his immediate benefit from deviating is $b - c + \gamma c$. After the deviation he transitions to the neutral state of mind, and the cost of the punishment period is $(c - d)\delta$. After that punishment period, player 2 transitions back to the hostile state of mind. Once cooperation is resumed, he faces the cost associated with transitioning to the friendly state of mind.

If $L = 1$, and $s = NH$, then player 2 has no profitable one-shot deviation from cooperation iff

$$b - c + \gamma c \leq (c - d)\delta + \delta^{L+1}(\alpha + \gamma)c + \delta^{L+2}\alpha c \quad (19)$$

When $\alpha = \gamma = 0$, (19) is identical to (6). An increase in α increases player 2's cost of deviating, and an increase in γ increases player 2's benefit from deviating. Player 2 can thus require more patience than player 1 for the strategy profile to be subgame perfect.

Now suppose that $L = 2$. Then player 2's one-period benefit from deviating is as before, and so is the cost in the first period of the punishment phase. In the second period of the punishment phase, player 2 has transitioned to the hostile state of mind. In the hostile state of mind he receives a lower utility, $d - \gamma d$, than in the friendly state of mind. After the punishment phase has ended, player 2 still faces the cost associated with transitioning back to the friendly state of mind.

For punishment phases of length $L > 1$, the condition can be generalized to

$$b - (1 - \gamma)c \leq (c - d)\delta + ((1 + \alpha)c - (1 - \gamma)d)(\delta^2 + \delta^3 + \dots + \delta^L) + \delta^{L+1}(\alpha + \gamma)c + \delta^{L+2}\alpha c \quad (20)$$

When $\alpha = \gamma = 0$, (20) is identical to (6). An increase in α increases player 2's cost of deviating, whereas an increase in γ now increases both the benefit and the cost from deviating. As when $M = \{N, H\}$, the cost increase dominates if (12) holds.

Following from the above, given that α is sufficiently small, such that player 2 has no incentive to deviate during the punishment phase, and that (12) holds, player 2 requires less patience not to deviate from the mutual minmax strategy profile than do player 1. The minimal discount factor required is therefore determined by the condition for player 1.

Let $\underline{\delta}^*$ and $\underline{\delta}^{MM}$ be defined as before.

Proposition 9. *Suppose that both players use the mutual minmax strategy with a punishment length $L > 1$. If $\alpha b < (1 + \alpha)d$, and (12) hold, then $\underline{\delta}^* = \underline{\delta}^{MM}$.*

To summarize, if the players use the Grim Trigger strategy, player 2 prefers to choose C even when less patient than player 1 since his utility from cooperation

is higher while his utility from deviation is smaller. If the players use the mutual minmax strategy, player 2 might require more patience to return to cooperation after a punishment phase, and if α is sufficiently large he might also require more patience not to deviate during the punishment phase.

5.2 Two emotional players

Still assume three states of mind, $M = \{F, N, H\}$, but now both players can transition between them. The transition probabilities are as in Figure (4). The corresponding stochastic game has nine states that are reachable from the initial state, $S = M^2$.

Consider first the finitely repeated interaction. If α is sufficiently large, $c + \alpha c > b$, then the players can sustain cooperation by strategically manipulating each others states of mind. Notice that once both players are in the friendly state of mind the unique stage game Nash equilibrium is for both to choose C .

5.2.1 Grim Trigger

Consider next the infinitely repeated interaction. The condition for an emotional player not to deviate from a strategy profile does not depend on whether he interacts with another emotional player or a *Homo oeconomicus*. Thus, in the infinitely repeated interaction between two identical emotional players, the Grim Trigger strategy profile is subgame perfect iff (15) holds. When $\alpha = \gamma = 0$, (15) is identical to (3), and an increase in either α or γ decreases the patience required not to deviate; both states of mind facilitates cooperation when the Grim Trigger strategy is used.

Let $\underline{\delta}^*$ and $\underline{\delta}^{GT}$ be defined as before.

Proposition 10. *Suppose that both players use the Grim Trigger strategy. Then $\underline{\delta}^* < \underline{\delta}^{GT}$, and $\underline{\delta}^*$ is decreasing in α and γ .*

5.2.2 Mutual minmax

If the players use the mutual minmax strategy, the effect of the states of mind is less clear. Both a positive and a negative concern for the other player's material payoff can have conflicting effects on cooperation. The friendly state of mind makes it more costly to deviate from cooperation, but also to punish the other

player for deviating. The hostile state of mind can obstruct cooperation by making it more difficult to return to cooperation after a punishment phase, but makes threats of punishment more credible. When $L = 1$, the mutual minmax strategy profile is subgame perfect iff (16), (18), and (19) hold. When $L > 1$, it is subgame perfect iff (17), (18), and (20) hold.

Only if α is sufficiently small, such that the players have no incentive to deviate during the punishment phase, and the players are sufficiently patient, such that (12) holds, can two emotional players sustain cooperation easier than in the benchmark case when using the mutual minmax strategy.

Let $\underline{\delta}^*$ and $\underline{\delta}^{MM}$ be defined as before.

Proposition 11. *Suppose that both players use the mutual minmax strategy with a punishment length $L > 1$. If $\alpha b < (1 + \alpha)d$, and (12) hold, then $\underline{\delta}^* < \underline{\delta}^{MM}$, and δ^* is decreasing in γ .*

To summarize, two emotional players can sustain cooperation easier than two *Homo oeconomicus* if they use the Grim Trigger strategy. With that strategy, both the friendly and the hostile state of mind facilitates cooperation. If the players instead use the mutual minmax strategy, then the two states of mind have conflicting effects on cooperation. The friendly state of mind makes it more profitable to cooperate, but can also make it more difficult to punish the other player after a deviation. The hostile state of mind makes threats of punishment more credible, but can also make it more difficult to return to cooperation after the punishment phase has ended.

5.3 Two emotional players with stochastic transitions

Let the two emotional players be of the same type with $M = \{F, N, H\}$, but now with stochastic state transitions; a generalization of the previously discussed emotional player. The transition probabilities are presented in Figure 6, where $P|C$ denotes the probability, conditional on observing action C , of transitioning from the current state to the neutral or friendly state, and $P|D$ denotes the probability, conditional on observing action D , of transitioning from the current to the hostile state. The corresponding stochastic game has nine states, M^2 , all reachable from the initial state.

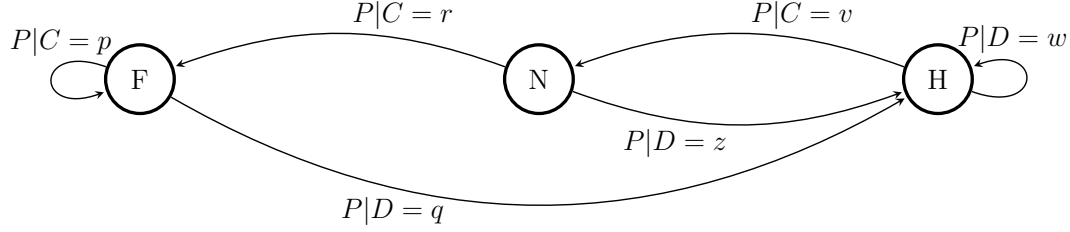


Figure 6: State transitions.

Focus is, as before, on the emotional player type who remains in the hostile state of mind with positive probability even after observing an act of cooperation, such that $p = q = r = z = w = 1$ and $v < 1$. This allows for more or less resentful players, where a player is less resentful the larger v is.

5.3.1 Grim Trigger

Consider the infinitely repeated interaction between the two emotional players. The patience required for the Grim Trigger strategy profile to be subgame perfect coincides with the patience required between two emotional players with three states of mind and deterministic state transitions since the players never return to choosing C .

5.3.2 Mutual minmax

Consider instead the mutual minmax strategy profile with a punishment phase of length $L > 1$. If no player so far has deviated, the players can either deviate in the first period, $s = NN$, or any subsequent period, $s = FF$. The players find it most profitable to deviate in the neutral state due to the lower utility from cooperation.

If $L > 1$, and $s = NN$, then the players have no profitable one-shot deviation from cooperation iff

$$\begin{aligned}
 b - c \leq & (1 + \alpha)(c - d)\delta + ((1 + \alpha)c - (1 - \gamma)d)(\delta^2 + \delta^3 + \dots + \delta^L) \\
 & + \frac{\delta^{L+1}(\alpha + \gamma)c}{1 - (1 - v)\delta} + \delta^{L+1+\frac{1}{v}}\alpha c
 \end{aligned} \tag{21}$$

When $v = 1$, (21) is identical to (17). A decrease in v increases the players' cost from deviating by increasing the number of periods they spend in the hostile

state of mind.

After a deviation, the players are supposed to choose their mutual minmax action for L consecutive periods. However, if they care sufficiently for the other player, they might have an incentive to deviate from punishment.

If the players are in the friendly state of mind, then they have no profitable one-shot deviation from punishment iff

$$\alpha b - (1 + \alpha)d \leq \delta^{L+1}(1 - \gamma)(c - d) + \frac{\delta^{L+1}(\alpha + \gamma)c}{1 - (1 - v)\delta} + \delta^{L+1+\frac{1}{v}}\alpha c \quad (22)$$

When $v = 1$, (22) is identical to (18), and an increase in v increases the cost of deviating from punishment by increasing the time spent in the hostile state of mind.

After a punishment phase the players should return to cooperation, but once they are in the hostile state of mind they might be less inclined to do so.

If $L > 1$, and $s = HH$, then the players have no profitable one-shot deviation from cooperation iff

$$\begin{aligned} b - (1 - \gamma)c &\leq [(1 - (1 - v)\gamma)(c - d)] \cdot \delta \\ &+ [((1 + \alpha) - v(1 - v)\alpha - (1 - v)^2(\alpha + \gamma))c - (1 - \gamma)d] \cdot \delta^2 \\ &+ [((1 + \alpha) - v(1 - v)^2\alpha - (1 - v)^3(\alpha + \gamma))c - (1 - \gamma)d] \cdot \delta^3 \\ &+ \dots \\ &+ [((1 + \alpha) - v(1 - v)^{L-1}\alpha - (1 - v)^L(\alpha + \gamma))c - (1 - \gamma)d] \cdot \delta^L \\ &+ \frac{(1 - (1 - v)^{L+1})(\alpha + \gamma)c \cdot \delta^{L+1}}{1 - (1 - v)\delta} + \frac{(\delta - (1 - v)^L)v\alpha \cdot \delta^{L+1}}{1 - (1 - v)\delta} \end{aligned} \quad (23)$$

When $v = 1$, (23) is identical to (20). In the hostile state of mind, the cost from the punishment phase is lower compared to the neutral state of mind. Hence a decrease in v decreases the cost of punishment and makes it more profitable for the players to deviate. The players also face the cost of remaining in the hostile state of mind even after cooperation is resumed. This probability is decreasing in v and makes it less profitable for the players to deviate. The first effect dominates, and a decrease in v increases the players' cost of deviating also in the hostile state of mind.

Let $\underline{\delta}^*$ be the minimal discount factor required for the mutual minmax strategy profile to be subgame perfect between the two emotional players with stochastic state transitions.

Proposition 12. *Suppose that both players use the mutual minmax strategy with a punishment length L . If $ab < (1 + \alpha)d$ holds, then $\underline{\delta}^*$ is decreasing in v .*

In other words, more resentful players require more patience to sustain cooperation in the prisoners' dilemma game when they use the mutual minmax strategy due to their difficulty of returning to cooperation after a punishment phase.

6 A repeated public goods production game

The prisoners' dilemma game is convenient to illustrate the tension between a Nash equilibrium and a Pareto dominant outcome. However, when studied as a repeated interaction, it is also special in the sense that the minmax action profile corresponds to the Nash equilibrium of the one-shot game. Further, in the prisoners' dilemma game the players are restricted to choose between two actions, potentially restricting their possibilities to act on emotions.

In this section the model is applied to a repeated interaction of public goods production. This type of interaction has also been studied by Bernheim and Stark (1988), who studies the interaction between two altruists, and by Alger and Weibull (2017) who studies differences between altruists and moralists. The interaction studied in this paper is between two players who can transition between different emotional states. The players have a continuum of actions and identical production costs. Their one-period utility is

$$u_i(s, x) = (1 + \phi(m_i))G(x) - c(x_i) - \phi(m_i)c(x_j), \quad (24)$$

where x is the vector of inputs $(x_1, x_2) \geq 0$, $G : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ is the production function which satisfies the INADA conditions, $c : \mathbb{R}_+ \rightarrow \mathbb{R}$ is the cost function which is convex and twice differentiable in x_i , and $\phi(m_i) : M_i \rightarrow \mathbb{R}$ the psychological payoff determined by i 's state of mind.

The first order condition for player i is

$$\frac{\partial u_i}{\partial x_i} = (1 + \phi(m_i)) \frac{\partial G(x)}{\partial x_i} - \frac{\partial c(x_i)}{\partial x_i} \quad (25)$$

The interaction is styled after Alger and Weibull (2017), with a Cobb-Douglas production function, $G(x) = x_i^\beta x_j^{(1-\beta)}$ with $\beta = 0.5$, and a quadratic cost function $c(x_i) = x_i^2$.

The aim of the analysis is to illustrate the differences with the prisoners' dilemma game. Therefore only the infinitely repeated interaction between a *Homo oeconomicus*, player 1, and an emotional player, player 2, who both use a Nash reversion strategy are analyzed.⁹ Further, as noted earlier, and in line with the results of Bernheim and Stark (1988), both a positive and a negative concern for the other player's material payoff can have conflicting effects on cooperation. Hence focus is on an interaction between two players with three states of mind, $M = \{F, N, H\}$.

The corresponding stochastic game has three states that are reachable from the initial state, $S = \{NF, NN, NH\}$. Each stage game has two Nash equilibria, and, similar to the prisoners' dilemma game, an outcome that Pareto dominates the equilibria. If one of the players contributes zero, then the other player's best response is to contribute zero. If both players are in the neutral state of mind, then their best response to the other player contributing 0.25 is to contribute an equal amount. However, the players could receive the highest utility if they could be forced to contribute 0.5 each. Assume for simplicity that player 2 has deterministic state transitions similar to the ones in the prisoner's dilemma game. If the other player contributes the social optimum (or more), then player 2 transitions to the friendly state of mind. If the other player contributes a smaller amount, then he transitions to the hostile state of mind.

Since there are two Nash equilibria, two possible Nash reversion exists. Focus is on the Nash reversion strategy profile in which the players contribute the socially optimal contributions and retreat to the positive contribution Nash equilibrium after a deviation since this equilibrium Pareto dominates the other. If

9. Nash reversion strategy profiles are strategy profiles in which the players retreat to a Nash equilibrium in the event of a deviation. The Grim Trigger strategy is a Nash reversion strategy for the prisoners' dilemma game. Nash reversion strategies are also commonly studied in oligopoly games (see e.g. Friedman, 1971).

the players would retreat to the Pareto dominated Nash equilibrium after a deviation, they would have an incentive to "renegotiate" this reversion since both players would prefer to retreat to the other Nash equilibrium (for a discussion of renegotiation in repeated games see e.g. Benoit and Krishna (1993)).

Remark. *The following analysis relies on the assumption of a utilitarian welfare function. This assumption is not uncontroversial when the players have other-regarding preferences. In particular, it implies that a spiteful individual should be compensated for being spiteful, whereas an altruist should contribute more (for a discussion on experienced utility see e.g. Kahneman, Wakker, and Sarin (1997); for an overview of behavioral welfare economics see e.g. Bernheim (2016) or Bernheim and Taubinsky (2018)).*

Under a utilitarian welfare function, the socially optimal contributions depend on the players' states of mind. When both players are in the neutral state of mind, the socially optimal contributions are for both to contribute 0.5, an equal amount. If player 2 is in the friendly state of mind the socially optimal contributions are for him to contribute more while player 1 contributes less. Likewise, if player 2 is in the hostile state of mind the socially optimal contributions are for him to contribute less while player 1 contributes more.

The socially optimal contributions in the friendly state are

$$x_1 = \frac{2 + \alpha}{4(1 + \alpha)^{\frac{3}{4}}} \quad (26)$$

and

$$x_2 = \frac{2 + \alpha}{4(1 + \alpha)^{\frac{1}{4}}} \quad (27)$$

The socially optimal contributions in the neutral state is found by setting $\alpha = 0$, and the socially optimal contributions in the hostile state is found by replacing α with $-\gamma$.

In this interaction, neither player has a dominant strategy, and the utility player 1 can gain by cooperating depends on player 2's state of mind. In addition, player 1's utility from other action profiles, for example the positive contribution Nash equilibrium, is also affected by player 2's state of mind.

Let $\alpha = \gamma = 0.2$. The players can choose to deviate either in the first period, $s = NN$, or in any subsequent period, $s = NF$. Player 1 finds it most profitable

to deviate in the first period, while player 2 finds it more profitable to deviate in any subsequent period.

If $s = NN$, then player 1 has no profitable one-shot deviation iff

$$0.25(1 - \delta) + 0.272 \cdot \delta \geq 0.291(1 - \delta) + 0.167 \cdot \delta \quad (28)$$

If $s = NF$, then player 2 has no profitable one-shot deviation iff

$$0.280 \geq 0.323(1 - \delta) + 0.233(1 - \delta) \cdot \delta + 0.145 \cdot \delta^2 \quad (29)$$

This strategy profile, in the interaction between a *Homo oeconomicus* and an emotional player, is subgame perfect for a common discount factor of $\delta \geq 0.355$. This is smaller than the one required in the benchmark case, $\delta \geq 0.396$. Player 2's positive concern for the other player's material payoffs in the friendly state of mind acts as a carrot by increasing both players' utility from cooperation. His negative concern in the hostile state of mind acts as a stick by lowering the utility from the (positive contribution) Nash equilibrium for both players. Accordingly, an increase in either α or γ decreases *both* players' profitability from deviating.

Let $\underline{\delta}^*$ and $\underline{\delta}^{NR}$ be the minimal discount factors required for the Nash reversion strategy profile to be subgame perfect in the case of play between one emotional player and one *Homo oeconomicus*, and in the benchmark case respectively.

Proposition 13. *Suppose that both players use the Nash reversion strategy. Then $\underline{\delta}^* < \underline{\delta}^{NR}$, and $\underline{\delta}^*$ is decreasing in α and γ .*

To summarize, cooperation is easier to sustain if at least one of the players is emotional, given that the players use the Nash reversion strategy. The friendly state of mind makes cooperation more profitable for both players since the emotional player contributes more in the friendly state of mind due to his positive concern for the other player's material payoff. The hostile state of mind makes deviation more costly for both players due to the emotional player's negative concern for the other player.

This is in contrast to the findings from the repeated prisoners' dilemma. In that interaction the *Homo oeconomicus* is unaffected by the emotional player's state of mind, and cooperation is only easier to sustain if both players are emotional.

The effect of emotions may vary depending on the structure of the interaction, just as emotions may have a different role in the interactions with one's boss as compared with one's spouse or child.

7 Concluding Remarks

This paper has proposed a model of repeated interactions between players who react emotionally to the history of play, and who can become friendly if the other player cooperates, or hostile if the other player defects. Focus is on a set of Markov players who transition between two or three states of mind. A player in a friendly (hostile) state of mind values the other player's material payoff positively (negatively), and a player in a neutral state of mind only cares for his own material payoffs. The traditionally studied *Homo oeconomicus* corresponds to a player who never leaves the neutral state of mind.

Cooperation in the prisoners' dilemma game is facilitated by the possibility that a player's concern for the other player's material payoff changes. A player in the friendly state of mind finds it less profitable to deviate from cooperation. If the other player deviates, the player transitions to a hostile state of mind in which he finds it less profitable to deviate from punishing the player. However, this comes with a caveat. While the possibility of transitioning to a hostile state of mind can make threats of punishment more credible, the players may also find it more difficult to return to cooperation after the punishment phase. If the interaction is of incomplete information, e.g. co-workers who cannot perfectly observe each others efforts, then individuals with strong emotions run the risk of a breakdown of cooperation due to a misunderstanding.

The strategic consequences of a player's state of mind depends on the structure of the game. In the infinitely repeated prisoners' dilemma game, the players need only concern themselves with their own state of mind, whereas in the public goods production game they also need to consider the other player's state of mind. One interpretation is that some interactions, as between two co-workers, have a more limited room for emotions than others, as with a spouse or a child.

An assumption of the proposed model is that the players only care about, or remember, the last observed action. This assumption can be relaxed and the effect of a longer memory depends on how the players reason about, and justify, their

own actions. If the players remember the last two periods, they may not only react to the last two observed actions of the other player, but also take into account how their own action in the second to last period affected the other player's choice of action in the last period. To what extent individuals take responsibility for how their own actions affect the outcome is an empirical question, but the growing literature on motivated beliefs in both economics and psychology suggests that individuals have a remarkable ability to justify their own behavior (Steele, 1988; Kunda, 1990; Epley and Gilovich, 2016; Gino, Norton, and Weber, 2016). An interesting avenue for further research is to study the difference in cooperative outcomes between individuals who justify their behavior, and individuals who acknowledges the consequences of their behavior.

Traditionally in game theory, and in this paper as well, utility depends solely on actions. Utility can also depend directly on the players' beliefs, and it is common for emotions to depend on the individual's beliefs or expectations (Elster, 1998). Emotions modeled using Psychological game theory (Geanakoplos, Pearce, and Stacchetti, 1989; Battigalli and Dufwenberg, 2009) can depend on the players' expectations and beliefs about intentions (cf. Charness and Dufwenberg, 2006; Battigalli and Dufwenberg, 2007; Battigalli, Dufwenberg, and Smith, 2018).

A related concern is that of social norms. When games are played in a historical and cultural context the players know how similar games have been played in the past, and as a consequence social norms specifying which actions the players are expected to consider usually develops (Myerson and Weibull, 2015). If the social norm is for everyone to defect in the prisoners' dilemma game, the emotional reaction is most likely not the same as if the social norm is for everyone to cooperate. The social norm affects the players' expectations which in turn affect the players' transition probabilities. This is left as an avenue for future work.

References

- Alger, Ingela, and Jörgen W Weibull. 2017. "Strategic behavior of moralists and altruists." *Games* 8 (3): 38.
- Battigalli, Pierpaolo, and Martin Dufwenberg. 2009. "Dynamic psychological games." *Journal of Economic Theory* 144 (1): 1–35.
- . 2007. "Guilt in games." *American Economic Review* 97 (2): 170–176.
- Battigalli, Pierpaolo, Martin Dufwenberg, and Alec Smith. 2018. "Frustration and Anger in Games." *Unpublished manuscript*.
- Benoit, Jean-Pierre, and Vijay Krishna. 1993. "Renegotiation in finitely repeated games." *Econometrica* 61 (2): 303–323.
- Bernheim, B Douglas. 2016. "The good, the bad, and the ugly: A unified approach to behavioral welfare economics." *Journal of Benefit-Cost Analysis* 7 (1): 12–68.
- Bernheim, B Douglas, and Oded Stark. 1988. "Altruism within the family reconsidered: Do nice guys finish last?" *The American Economic Review*: 1034–1045.
- Bernheim, B Douglas, and Dmitry Taubinsky. 2018. "Behavioral Public Economics," edited by B Douglas Bernheim, Stefano DellaVigna, and David Laibson. Elsevier.
- Charness, Gary, and Martin Dufwenberg. 2006. "Promises and partnership." *Econometrica* 74 (6): 1579–1601.
- Cox, James C, Daniel Friedman, and Steven Gjerstad. 2007. "A tractable model of reciprocity and fairness." *Games and Economic Behavior* 59 (1): 17–45.
- Cox, James C, Daniel Friedman, and Vjollca Sadiraj. 2008. "Revealed altruism." *Econometrica* 76 (1): 31–69.
- Dreber, Anna, Drew Fudenberg, and David G Rand. 2014. "Who cooperates in repeated games: The role of altruism, inequity aversion, and demographics." *Journal of Economic Behavior & Organization* 98:41–55.
- Dutta, Prajit K. 1995. "A folk theorem for stochastic games." *Journal of Economic Theory* 66 (1): 1–32.
- Elster, Jon. 1998. "Emotions and economic theory." *Journal of economic literature* 36 (1): 47–74.

- Epley, Nicholas, and Thomas Gilovich. 2016. "The mechanics of motivated reasoning." *Journal of Economic perspectives* 30 (3): 133–40.
- Friedman, James W. 1971. "A non-cooperative equilibrium for supergames." *The Review of Economic Studies* 38 (1): 1–12.
- Fudenberg, Drew, and Eric Maskin. 1986. "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information." *Econometrica* 54 (3): 533–554.
- Fudenberg, Drew, and Yuichi Yamamoto. 2011. "The folk theorem for irreducible stochastic games with imperfect public monitoring." *Journal of Economic Theory* 146 (4): 1664–1683.
- Geanakoplos, John, David Pearce, and Ennio Stacchetti. 1989. "Psychological games and sequential rationality." *Games and economic Behavior* 1 (1): 60–79.
- Gino, Francesca, Michael I Norton, and Roberto A Weber. 2016. "Motivated Bayesians: Feeling moral while acting egoistically." *Journal of Economic Perspectives* 30 (3): 189–212.
- Hoffman, Alan J, and Richard M Karp. 1966. "On nonterminating stochastic games." *Management Science* 12 (5): 359–370.
- Hörner, Johannes, Takuo Sugaya, Satoru Takahashi, and Nicolas Vieille. 2011. "Recursive methods in discounted stochastic games: An algorithm for $\delta \rightarrow 1$ and a folk theorem." *Econometrica* 79 (4): 1277–1318.
- Horst, Ulrich. 2005. "Stationary equilibria in discounted stochastic games with weakly interacting players." *Games and Economic Behavior* 51 (1): 83–108.
- Kahneman, Daniel, Peter P Wakker, and Rakesh Sarin. 1997. "Back to Bentham? Explorations of experienced utility." *The quarterly journal of economics* 112 (2): 375–406.
- Kunda, Ziva. 1990. "The case for motivated reasoning." *Psychological Bulletin* 108 (3): 480.
- Levine, David K. 1998. "Modeling altruism and spitefulness in experiments." *Review of economic dynamics* 1 (3): 593–622.
- Mailath, George J, and Larry Samuelson. 2006. *Repeated Games and Reputations: Long-Run Relationships*. Oxford university press.

- Mertens, Jean-Francois, and T. Parthasarathy. 1991. “Nonzero-Sum Stochastic Games.” In *Stochastic Games And Related Topics: In Honor of Professor L. S. Shapley*, edited by T. E. S. Raghavan, T. S. Ferguson, T. Parthasarathy, and O. J. Vrieze, 145–148. Springer.
- Myerson, Roger, and Jörgen Weibull. 2015. “Tenable strategy blocks and settled equilibria.” *Econometrica* 83 (3): 943–976.
- Parthasarathy, T, and S Sinha. 1989. “Existence of stationary equilibrium strategies in non-zero sum discounted stochastic games with uncountable state space and state-independent transitions.” *International Journal of Game Theory* 18 (2): 189–194.
- Rand, David G, Drew Fudenberg, and Anna Dreber. 2015. “It’s the thought that counts: The role of intentions in noisy repeated games.” *Journal of Economic Behavior & Organization* 116:481–499.
- Shapley, Lloyd S. 1953. “Stochastic games.” *Proceedings of the National Academy of Sciences* 39 (10): 1095–1100.
- Steele, Claude M. 1988. “The psychology of self-affirmation: Sustaining the integrity of the self.” In *Advances in experimental social psychology*, edited by Leonard Berkowitz, 21:261–302. Elsevier.