Perfect Sequential Reciprocity and Dynamic Consistency

Penta, Antonio

Bocconi University, University of Pennsylvania

June 2004

Perfect Sequential Reciprocity and Dynamic Consistency.

Antonio Penta.*

Dept. of Economics,

University of Pennsylvania,

160 McNeil Building,

3718 Locust Walk Philadelphia, PA 19104

email: penta@sas.upenn.edu

tel.: (001).215.746.0363

September 10, 2007

Abstract

Dufwenberg and Kirchsteiger's (2004) extends Rabin's (1993) theory of *reciprocity* in a dynamic sense, introducing a rule of revision for player's beliefs. The *Sequential Reciprocity Equilibrium* [SRE] they define can be *dynamically inconsistent.* In this article it is argued that such dynamic inconsistency is not intrinsically related to issues of reciprocity, but rather to the particular way the beliefs' updating process is modeled. A refinement of the SRE, which is both dynamically consistent and, it is argued, more sound to assumptions usually made in the literature of information economics and philosophy, is proposed.

Keywords: Reciprocity – Dynamic Consistency.

JEL classification: A13; C70; D63

1. Introduction

Whithin the by now vast field of Behavioral Economics,[1] an interesting hint has been put forward by theories of *reciprocity*, pioneered by Rabin (1993). The key point in Rabin's theory of reciprocity is that a person's kindness depends on

*The present work was part of my undergraduate dissertation at Bocconi University, Milan (2004). I'm deeply grateful to my advisor, prof. Pierpaolo Battigalli. I also thank Giuseppe Cappelletti for helpful comments.

his *intentions*: a certain behavior is not necessarily kind nor unkind; its *kindness* depends on the consequences that it is supposed to produce, according to the beliefs held by who is displaying that behavior. This is the meaning of *kindness* as an *intention-driven* concept.[2] The explicit role for players' intentions requires the use of *psychological games*, as developed by Geanakoplos et al.(1989, GPS henceforth).

More recently, Dufwenberg and Kirchsteiger (2004, DK hereafter) have extended Rabin's theory of reciprocity in a dynamic sense: as the play unfolds, players update their beliefs, and their "reciprocational attitude" is affected by such a learning process. The dependence of players' preferences on the *updated beliefs* makes DK's go beyond GPS's theory of psychological games: in the latter, only *initial beliefes* are allowed to have a *direct* effect on payoffs. Indeed, besides extending Rabin's theory of reciprocity to a dynamic environment, DK's is also the first paper in which the logic of the psychological games is extended to encompass truly dynamic beliefs.[3]

The way DK pursue such an extension of the theory, with the definition of the *Sequential Reciprocity Equilibrium [SRE],* leads to at least one curious result, as emphasized by DK themselves: the *SRE* may be *dynamically inconsistent.* Dynamic inconsistency is another central topic in Behavioral Economics, and it has been extensively studied.[4] It seems natural then to investigate whether there is a direct relation between reciprocity and dynamic inconsistency, and to isolate the origin of the dynamic inconsistency which arises in DK's model.

In this article it is argued that the dynamic inconsistency of DK's model is not intrinsically related to issues of reciprocity, but rather to the particular way the beliefs' updating process is modeled. A refinement of DK's SRE, which is both dynamically consistent and, it is argued, more sound to assumptions usually made in the literature of information economics and philosophy, is proposed. Thus, the purpose of this note is twofold: first, to put in evidence that the source of the dynamic inconsistency in DK's model is not reciprocity per se, but rather specific implicit assumptions about the underlying epistemic model; second, to propose an equilibrium refinement that is more consistent with standard models in information economics.

The remaining part of the paper is organized as follows: section 2 introduces the basic notation and DK's theory of reciprocity; section 3 presents DK's SRE and the dynamic inconsistency is briefly discussed; in section 4 the *Perfect-SRE* is introduced, and its existence and dynamic consistency are proved; in section 5 we explain the intuition behind the new equilibrium concept, and its relations with contributions in literature of economics and philosophy.

2. DK's Theory of *Sequential Reciprocity* and basic notation.

Following Rabin's example, DK's theory of reciprocity consists of a "psychological extension" of a traditional dynamic game with perfect information, that we call *material game*. The *material game* is a dynamic game with perfect information, denoted by a tuple $\Gamma = \langle N, H, Z, (A_i, \pi_i)\rangle$,[5] where N is the set of players, H the set of *partial histories*, Z the set of *terminal histories*. For each $i \in N$, A_i represents the set of feasible actions of i and, for each $h \in H$, $A_i(h)$ represents the set of feasible actions at h.

For each $i \in N$, we denote by $S_i \subseteq [A_i]^H$ the set of i's pure strategies and by $B_i \subseteq [\Delta(A_i)]^H$ the set of his *behavioral strategies*. Let $B := \times_{i \in N} B_i$ and $B_{-i} := \times_{j \neq i} B_j$. For any $h \in H$, $s_i(h)$ represents the pure action s_i prescribes at

h, and by $\beta_{i,(h)}(a_i)$ the probability that the behavioral strategy β_i assigns to the action a_i at h. Therefore, $\forall h \in H, \beta_{i,(h)} \in \Delta[A_i(h)]$.

For each player $i \in N$, the *material payoff function* is defined as the *ex ante* payoff function: $\pi_i : B \to \mathbb{R}$ (payoffs depend on the profile of behavioral strategies).

Players' behavior will be represented by behavioral strategies. A behavioral strategy can be interpreted as an array of conditional beliefs about players' behavior at the partial histories. Thus, the set of *first order beliefs* of player i about player j are defined as $T^1_{i,j} := B_j$. Let be $T^1_i := \times_{j \neq i} T^1_{i,j}$ the set of player i's first order beliefs about the behavior of all his opponents. Furthermore, the set of *second order beliefs* of player i about what player j believes about player k's behavior is defined as $T^2_{i,j,k} := T^1_{j,k} = B_k$.

As anticipated, in order to overcome GPS's and Rabin's static frameworks, DK introduce a *beliefs' revision rule.* In order to cope with this, more notation is needed.

For each $h \in H$, let $S_i(h)$ be the set of i's (pure) strategies consistent with the history h. Also, letting $l(h)$ denote the *length* of h, i.e. the number of action profiles whose concatenation yields h, define, for each $i \in N$ and $h \in H$:

$$\begin{aligned} B_i^{(h)} \quad : \quad &= \Big\{ \beta_i \in B_i : \forall \left(a^1, ..., a^{l(h)-1}\right) \prec h, \\ &\forall k = 1, ..., (l(h) - 1), \beta_{i,\left(a^1,...,a^{l(h)-k-1}\right)} \left(a^{l(h)-1}\right) = 1 \Big\} \end{aligned}$$

In words, in $B_i^{(h)}$ there are only those *behavioral strategies* that prescribe the pure actions leading to h.at all the nodes on the path to h.

The basic intuition of DK's model of *beliefs' updating* is the following: as the play unfolds and history h is reached, players update their beliefs shifting to more pure beliefs. If at the beginning of the game they assigned a certain probability distribution to actions to be taken at histories on the path to h, once h has been reached they update their beliefs so to assign probability one to the actions *actually* taken by other players on the way through h, leaving their own beliefs over the players' behavior at the unreached histories unchanged. The same occurs for their beliefs about other players' beliefs as well. In terms of the notation above, for each $h \in H$, player i updates his first order beliefs about j shifting from elements $\beta_j^i \in B_j$ to elements $\beta_j^{i,(h)} \in B_j^{(h)}$.[6] Thus, the following definition of *the belief revision policy* is introduced (see DK):

Definition 1 *For each player i's initial behavioral strategy $\beta_i \in B_i \equiv B_i^{(\emptyset)}$, and for each $h \in H$, the updating rule yields a shift from β_i to $\beta_i^{(h)} \in B_i^{(h)}$, which is such that:*

1. $\forall h' \prec h, \beta_{i,(h')}^{(h)} \left(s_i\left(h'\right)\right) = 1$, *where* $s_i \in S_i\left(h\right)$.
2. $\forall h'' \nprec h, \beta_{i,(h'')}^{(h)} = \beta_{i,(h'')}$.

The following example may help to understand the definition above. Consider the game in figure 1: suppose that at the the beginning of the game (at history $\emptyset$) player 2's first order beliefs are represented by the following behavioral strategy:

$$\beta_1^{2,\emptyset} = \left\langle \left(\beta_{1,\emptyset}^{2,\emptyset} = pA + (1-p)\,B\right); \left(\beta_{1,(B,s)}^{2,\emptyset} = qC + (1-q)\,D\right)\right\rangle$$

According to this model, if history B occurs, player 2's beliefs about 1's strategy

C:/swp50/temp/graphics/J06I4801__1.pdf

must be:

$$\beta_1^{2,(B)} = \left\langle \left(\beta_{1,\emptyset}^{2,(B)} = B \right); \left(\beta_{1,(B,s)}^{2,(B)} = qC + (1-q)\, D \right) \right\rangle .$$

Note that such a belief revision rule is somewhat resembling to a form of Bayesian updating. Indeed, if it is $0 < p < 1$, the updating rule is actually a Bayesian one. But it is something more, since it determines a rule of beliefs revision also for the case in which $p = 1$ and B is reached, that is, conditional on zero probability events. In the language of modal logic and philosophy, DK's model of belief revision is consistent with the so called *conservativity principle*, which states that: <<when changing beliefs in response to new evidence, you should continue to believe as many of the old beliefs as possible>> (Harman (1986): p.46, cited in Battigalli and Bonanno (1999): p.156).[7]

The last bits of notation are directly related to the concept of *reciprocity*: as in Rabin (1993), DK define *kindness functions* as mappings from each player's "behaviors and intentions" to $\mathbb{R}$. This captures the idea of kindness as an *intention-driven* concept. The following concepts are needed:

- The set of player i's *Pareto efficient strategies*:

$$E_i \;:\; = \{\beta_i \in B_i : \nexists \beta_i' \in B_i \text{ s.t. } \forall h \in H, \forall \beta_{-i} \in B_{-i},$$
$$\pi_j\left(\beta_i', \beta_{-i}|h\right) \geq \pi_j\left(\beta_i, \beta_{-i}|h\right) \text{ for each } j \in N,$$
$$\text{with strict inequality for some } j, h, \beta_{-i}\}.$$[1]

When a player chooses a strategy $\beta_i \in B_i$, given his beliefs $\tilde{\beta}^i_{-i} \in B_{-i}$ about others' strategies, he actually chooses a profile of payoffs among those in the set $\Pi\left(\tilde{\beta}^i_{-i}\right) := \left\{\left(\pi_j\left(\beta_i, \tilde{\beta}^i_{-i}\right)\right)_{j\in N} : \beta_i \in B_i\right\}$. Thus, any strategy $\beta_i \in B_i$ should he choose, he would assign a payoff of $\pi_j\left(\beta_i, \tilde{\beta}^i_{-i}\right)$ to player j.

- The *equitable payoff* for player j, given i's belief β^i_{-i}, is:

$$\pi^e_j\left(\beta^i_{-i}\right) = \frac{\max\left\{\pi_j\left(\beta_i, \beta^i_{-i}\right) : \beta_i \in B_i\right\} + \min\left\{\pi_j\left(\beta_i, \beta^i_{-i}\right) : \beta_i \in E_i\right\}}{2}$$

The *equitable payoff* is used as a reference point to evaluate i's *kindness* to j. This is made through the definition of the *kindness function*: according to the same logic introduced by Rabin, player i's kindness to player j depends on his first order beliefs and measures the distance between the payoff player i "assigns" to j and the equitable payoff. Thus:

[1] For each $h \in H$, and for each $(\beta_i, \beta_{-i}) \in B_i \times B_{-i}$, it is, for each $j \in N$:

$$\pi_j\left(\beta_i, \beta_{-i}|h\right) \equiv \pi_j\left(\beta_i^{(h)}, \beta_{-i}^{(h)}\right).$$

Since $S_i \subset B_i$, the meaning of $\pi_j\left(s_i, \beta_{-i}|h\right)$ is clear too.

- the *kindness function* of i to j is a function $\kappa_{i,j} : B_i \times T^1_{i,j} \to \mathbb{R}$ s.t.:

$$\begin{aligned} \forall \beta_i &\in B_i, \forall \beta^i_{-i} \in T^1_{i,j}, \\ \kappa_{i,j}\left(\beta_i, \beta^i_{-i}\right) &= \pi_j\left(\beta_i, \beta^i_{-i}\right) - \pi^e_j\left(\beta^i_{-i}\right). \end{aligned}$$

- The *perceived kindness fuction* is defined shifting one step up in the hierarchy of beliefs, and represents how a player believes his opponents are being kind to him. Thus, it is represented by the function $\lambda_{i,j,i} : T^1_{i,j} \times T^2_{i,j,-j} \to \mathbb{R}$ s.t.:

$$\begin{aligned} \forall \beta^i_{-i} &\in T^1_{i,j}, \forall \beta^{i,j}_i \in T^2_{i,j,-j}, \\ \lambda_{i,j,i}\left(\beta^i_j, \beta^{i,j}_{-j}\right) &= \pi_i\left(\beta^i_j, \beta^{i,j}_{-j}\right) - \pi^e_i\left(\beta^{i,j}_{-j}\right). \end{aligned}$$

  In words, the kindness i perceives from j depends on what i believes j is doing ($\beta^i_j \in T^1_j$ is a first order belief of i about j), and on what i thinks j believes about the behavior of all the players other than j ($\beta^{i,j}_{-j} \in \times_{k \neq j} T^2_{i,j,k} \equiv T^2_{i,j,-j}$ are second order beliefs of player i).

- We are ready now to specify the *"psychological" utility function* that DK use to define their game of *sequential reciprocity*, that is:

$$\begin{aligned} \forall i &\in N, u_i \colon B_i \times \prod_{j \neq i} \left(T^1_{i,j} \times \prod_{k \neq j} T^2_{i,j,k} \right) \text{ s.t.:} \\ & u_i\left(\beta_i, \left(\beta^i_j, \left(\beta^{i,j}_k\right)_{k \neq j}\right)_{j \neq i}\right) \end{aligned}$$

$$= \pi_i \left(\beta_i, \left(\beta_j^i\right)_{j\neq i}\right) + \sum_{j\neq i} Y_{i,j} \cdot \left[\kappa_{i,j}\left(\beta_i, \beta_{-i}^i\right) \cdot \lambda_{i,j,i}\left(\beta_j^i, \beta_{-j}^{i,j}\right)\right] \qquad (1)$$

$\left(Y_{i,j}\right)_{j\neq i}$ are fixed and exogenous parameters, representing the sensitivity of player i to his reciprocation to j.

Thus, the Dynamic Reciprocity Game introduced by DK is defined as:

Definition 2 *The Dynamic Reciprocity Game obtained from the material game* $\Gamma = \langle N, H, Z, (A_i, \pi_i)\rangle$ *is a tuple* $\Gamma^R = \langle N, H, Z, (A_i, u_i)\rangle$ *in which the "psychological" utility functions* $(u_i)_{i\in N}$ *are defined as above.*

3. The *Sequential Reciprocity Equilibrium* and the *Dynamic Inconsistency.*

DK's idea of *sequential reciprocity* is that at each history, players take into account the kindness as evaluated according to the beliefs updated at that history. Thus, i's kindness to j at history h will be given by the function $\kappa_{i,j}\left(\beta_i, \beta_j^i | h\right) \equiv \kappa_{i,j}\left(\beta_i^{(h)}, \beta_j^{i,(h)}\right)$. Equivalently, the perceived kindness taken into account will be $\lambda_{i,j,i}\left(\beta_j^i, \beta_{-j}^{i,j} | h\right) \equiv \lambda_{i,j,i}\left(\beta_j^{i,(h)}, \beta_{-j}^{i,(h);j,(h)}\right)$.

Thus, DK introduce the following equilibrium concept, which is meant to capture such an idea of *sequential reciprocity*, in which updated beliefs play a direct effect:

Definition 3 (The Sequential Reciprocity Equilibrium) *The profile of behavioral strategies* $\left(\tilde{\beta}_i\right)_{i\in N}$ *is a "Sequential Reciprocity Equilibrium" [SRE] if* $\forall i \in N, \forall h \in H$:

1. $\tilde{\beta}_i(h) \in \arg\max_{\alpha_{i,h}\in\Delta[A_i(h)]} u_i \left(\tilde{\beta}_i \backslash \alpha_{i,h}, \left(\tilde{\beta}_j^{i,(h)}, \left(\beta_k^{i,(h);j,(h)}\right)_{k\neq j}\right)_{j\neq i}\right)$.[8]

2. $\tilde{\beta}_j^{i,\emptyset} = \tilde{\beta}_j$ *for all* $j \neq i$.

3. $\tilde{\beta}_k^{i,\emptyset;j,\emptyset} = \tilde{\beta}_k$ *for all* $j \neq i, k \neq j$.

Let R^{DK} *be the set of such SREs for* Γ^R.

Conditions 2 and 3 in the previous definition require nothing but the correctness of the *initial beliefs*, both of first (condition 2) and of second order (condition 3): this is a standard equilibrium requirement, also common to GPS's notions of equilibrium for psychological games. The key point of DK's model is condition 1: having fixed the equilibrium-consistent initial beliefs, the updated beliefs are obtained, at each history, through the *belief revision policy* discussed above. What condition 1 introduces is in fact a criterion of *sequential rationality* which is directly referred to the updated beliefs. Thus, DK's *SRE* is such that, for each history, players have no incentive to unilateral one-shot deviations, provided that reciprocity is evaluated according to the the beliefs held at *that* history.

The definition of the equilibrium as immune to one-shot deviations is crucial: as DK themselves point out, the *one-shot deviation principle* [OSDP] fails in their model, meaning that the SRE defined above, <<[...] does not exclude the possibility that a joint deviation at several [histories] might increase a player's utility as evaluated at the first [of the histories] where the player deviates>> (cit., DK). The failure of the OSDP prevents the applicability of backward induction arguments. Nonetheless, DK provide an existence result (see the Theorem in DK).

Intuitively, the source of the dynamic inconsistency of players' preferences is quite obvious: since in a psychological game players' preferences depend on their beliefs, and since players' beliefs change as the play unfolds, players' preferences may change as well. In fact, what condition 1 entails is that players have, at each history, a different objective function. The SRE indeed entails a sort of *multi-self* model.

An example might be useful to understand how DK's SRE works, and in particular how the dynamic inconsistency arises.

Example 4 *Consider the game in figure 1: it will be shown that if player* 1 *is enough "reciprocity-concerned" (that is if* Y_1 *is high enough), the strategy profile* $\langle (AD)\,;(s)\rangle$ *is an SRE of the game. If so it is, because of conditions 2 and 3 the candidate initial equilibrium beliefs must be:* $\beta_2^{1,\emptyset} = (s)$ *and* $\beta_1^{2,\emptyset} = (AD)$*, while the second order beliefs are* $\beta_1^{1,\emptyset;2,\emptyset} = (AD)$ *and* $\beta_2^{2,\emptyset;1,\emptyset} = (s)$*. Given these beliefs, it is easy to see that at the first node player* 1 *prefers* A *over* B*: because of his beliefs* $\beta_2^{1,\emptyset} = (s)$ *and* $\beta_1^{1,\emptyset;2,\emptyset} = (AD)$*,* 1 *interprets* 2*'s action* s *as "fair" (neither kind nor unkind: it is* $f'_{1,2,1}(AD, s) = 0$*), so that he is interested only in material payoffs, and so he prefers the terminal history* (A) *over the terminal history* (B, s, D) *(the one which would be reached, given* 2*'s strategy* (s) *and given the action* D *prescribed by* 1*'s candidate equilibrium strategy at the other node where he is active). At the second node, where* 2 *is active, the belief revision policy prescribes that* 2*'s beliefs are* $\beta_1^{2,(B)} = (BD)$ *and* $\beta_2^{2,(B);1,(B)} = (s)$*. Thus, player* 2 *interprets* 1*'s action* A *as unkind behavior, so that both for material and for "reciprocational" concerns he prefers* s *over* d*. At the leftmost node, player* 1*'s beliefs must be* $\beta_2^{1,(B,s)} = (s)$ *and* $\beta_1^{1,(B,s);2,(B,s)} = (BD)$*. With such beliefs,* 1 *interprets* 2*'s behavior as an unkind one, in that choosing* s *and accordig to what* 1 *believes* 2 *thought, player* 2 *assigned to* 1 *a payoff of* -1 *instead of one of* 5 *(which would have been reached had* 2 *chosen* d*). Thus, if player* 1 *is enough "reciprocity-concerned" (namely, if it is* $Y_1 > 1$*), at the last node he is willing to give up a payoff of* 2 *for one of* -1*, in order to reciprocate* 2*'s unkind behavior lowering* 2*'s payoff from* -2 *to*

-3. *That is, choosing D over C at the last node. Because of the definition, we have proved that the strategy profile $\langle (AD), (s) \rangle$ is indeed an SRE of the game in figure 1. Furthermore, such an SRE exhibits a problem of dynamic inconsistency: given the equilibrium initial beliefs, at the first node* 1 *thinks* 2*'s kindness to him is zero, so that he would like to deviate from* (AD) *to* (BC) *to increase his material payoff from* 0 *to* 2 *(thus, the equilibrium profile is not robust respect to "multistage" deviations). The dynamic inconsistency arises in that, if player* 1 *chooses B at the first node, and then (given* 2*'s strategy* (s)*) the rightmost node is reached, he doesn't perceive* 2*'s strategy* (s) *as fair anymore, but as an unkind one, and hence he will reciprocate choosing D at* (B, s)*, i.e. punishing* 2*'s unkind behavior.*

It should be clear from this example that the basic source of the OSDP failure is the fact that player 1 interprets his opponent's strategy (s) as *fair* at the beginning, and as *unkind* at the rightmost node. The unkindness of 2's (s) when 1 is at (B, s) is clear: the pattern of beliefs entails that 1 thinks 2 chooses for him a payoff of -1 instead of one of 5. But why does, at the initial history, 1 perceive the same action as *fair*? The reason is that at the beginning of the game 2's initial beliefs are $\beta_1^{2,\emptyset} = (AD)$: given these beliefs, player 2 simply cannot be neither kind nor unkind. In fact, the logic of the kindness functions is such that, for given beliefs of a player, we compare the payoff profile that a strategy of his own "chooses" with the others that are feasible according to his beliefs about the opponents. So, if 2 initially believes that his node will not be reached, he has actually no choice at all, and so, no opportunity to be neither kind nor unkind. Thus, in DK's model, if player 2 initially believes that 1 will play A, we are not able to distinguish, in terms of kindness, 2's strategies (s) and (d). Hence, a first drawback of the SRE is that it does not allow a complete kindness ranking of the strategies: in general it is not possible to make a kindness comparison among strategies that differ from each other only at histories reached with zero probability. Notice that such an incompleteness of the kindness ranking of strategies would not be an issue in a completely mixed strategy profile.

Another consequence of the SRE notion is that the kindness of a strategy is evaluated, at each history, according to the beliefs every player holds at *that* history, and not according to the beliefs he would hold at the histories in which he actually displays the actions purposed by his strategy (see condition 1 in the SRE's definition): despite the updating rule, DK treat the beliefs held at every history in isolation, without relating them to those held at the subsequent nodes. On the contrary, it seems natural to argue that since a player's strategy is a contingent plan of action, for all histories that might conceivably occur, its kindness should be considered as a contingent matter as well. Thus, the kindness of an action prescribed by a strategy at a given history, should depend on the beliefs that the player would have at *that* history. Under standard epistemic assumptions, a player that holds certain beliefs *knows* what he would believe at any node, whether his beliefs assign positive or zero probability to that node being reached. In the example above, 2's strategy (s) represents his *disposition to act at* B; but since at B he would never believe A, this means that the strategy (s) indeed represents 2's *disposition to act at* B*, believing that* B *has occured*. A notion of sequential reciprocity consistent with this kind of observations would not exhibit the kind of inconsistency discussed in the example above. Thus, the real source of the dynamic inconsistency in DK's equilibrium notion is in their model of treating beliefs, rather than in something peculiar to the psychological nature of the game.

4. Dynamic Consistency and the *Perfect-SRE.*

The previous example showed the non validity of the OSDP for the Dynamic Reciprocity Game *à la* DK. In the following it is introduced a refinement of DK's SRE that solves the problems of dynamic inconsistency, and an existence result is provided. The next section will discuss possible interpretations of such an equilibrium concept, and some related literature.

The following definition are necessary:

Definition 5 *Given a Dynamic Reciprocity Game à la DK*

$\Gamma^R = \left\langle N, H, Z, (A_i, u_i)_{i \in N} \right\rangle$, *we define the* ε*-perturbed game of* Γ^R *as the dynamic reciprocity game which is identical to* Γ^R *but in which, each player* $i \in N$ *can only choose such that* $\forall h \in H$, $\beta_i(h) \in \Delta^\varepsilon [A_i(h)]$ *where it is:*

$$\Delta^\varepsilon [A_i(h)] := \{\alpha_i \in \Delta [A_i(h)] : \forall a_i \in A_i(h), \alpha_i(a_i) \geq \varepsilon\}.$$

A test sequence for Γ^R *is a sequence of* ε*-perturbed games of* Γ^R *for* $\varepsilon \to 0$.

Definition 6 (The Perfect-SRE) *A profile of behavioral strategies* β^* *is a Perfect Sequential Reciprocity Equilibrium [P-SRE] of a Dynamic Reciprocity Game à la DK* Γ^R*, if it is the limit point of a sequence of SREs for a test sequence of* Γ^R*. Let* R^P *be the set of such P-SREs of* Γ^R.

The following theorem guarantees the existence of the P-SRE:

Theorem 7 (P-SRE's Existence) *Under the continuity assumptions of DK's Theorem, there exists a P-SRE for every Dynamic Reciprocity Game* Γ^R.

Proof. The existence of the sequence of SRE is guaranteed by DK's existence theorem, applied to each perturbed game Γ^{ε_n} of Γ^R in the test sequence $(\Gamma^{\varepsilon_n})_{n\geq 1}$. The compactness of the strategy space, together with the continuity assumptions, guarantees the existence of an accumulation point for the sequence of SREs of the games in the test sequence, hence of a P-SRE as we have defined it. ■

The following theorem proves that the P-SRE is indeed a refinement of DK's SRE:

Theorem 8 *Every P-SRE of a Dynamic reciprocity Game Γ^R is a SRE of it:* $R^P \subseteq R^{DK}$

Proof. All we have to prove is that a sequence of SRE of a test sequence $(\Gamma^{\varepsilon_n})_{n\geq 1}$ for Γ^R converges to a SRE of Γ^R: by definition, a sequence $(\beta^n)_{n\geq 1}$ is a sequence of SRE of a test sequence for Γ^R if and only if , $\forall n \geq 1, \forall i \in N, \forall h \in H$:

$\beta_i^n(h) \in \arg\max_{\alpha_{i,h} \in \Delta^{\varepsilon_n}[A_i(h)]} u_i \left(\beta_i^n \backslash \alpha_{i,h}, \left(\beta_j^n, (\beta_k^n)_{k\neq j}\right)_{j\neq i} |h\right)$ and for all $j \neq i$ $\beta_j^{i,\emptyset} = \beta_j$, and for all $j \neq i, k \neq j$ $\beta_k^{(i,\emptyset),(j,\emptyset)} = \beta_k$.

Note that $\forall n > m$ in a test sequence, $\Delta^{\varepsilon_m}[A_i(h)] \subseteq \Delta^{\varepsilon_n}[A_i(h)]$, therefore, $\forall n > m$, $\forall h \in H$,

$u_i(\beta^n|h) \geq u_i(\beta^n \backslash \alpha_{i,h}|h)$, $\forall \alpha_{i,h} \in \Delta^{\varepsilon_m}[A_i(h)]$.

If u_i is *continuous*, then the inequality holds also taking limits for $n \to \infty$:

$u_i(\beta^*|h) \geq u_i(\beta^* \backslash \alpha_{i,h}|h)$, $\forall \alpha_{i,h} \in \Delta^{\varepsilon_m}[A_i(h)]$, for each $h \in H$.

Such an inequality holds for every m. For all $h \in H$, the closure of the union of all $\Delta^{\varepsilon_m}[A_i(h)]$ is $\Delta[A_i(h)]$. This, together with the continuity of u_i yields:

$u_i(\beta^*|h) \geq u_i(\beta^* \backslash \alpha_{i,h}|h)$ $\forall \alpha_{i,h} \in \Delta[A_i(h)]$, for each $h \in H$.

For each $i \in N$, it yields nothing but the definition of the SRE for Γ^R. ■

Finally, the following theorem states the validity of the OSDP for the P-SRE:

Theorem 9 (OSPD for P-SRE) *If β^* is a P-SRE, then it holds (under the usual continuity assumptions):*

$$\begin{aligned} \forall i &\in N, \forall h \in H : \\ (\beta_i^*|h) &\in \arg\max_{\beta_i \in B_i(h)} u_i\left(\beta_i, \beta_{-i}^*|h\right) \end{aligned}$$

Proof. Ab absurdo, $\exists i \in N : \exists h' \in H, \exists \beta_i^{'} \in B_i\left(h\right) :$

$$u_i\left(\beta_i', \beta_{-i}^*|h'\right) > u_i\left(\beta^*|h'\right).$$

But, if so it is, by the continuity assumption (u_i is continuous in all its arguments) it must be that, for a *ε-perturbed game* Γ^ε in the test sequence $\left(\Gamma^{\varepsilon_n}\right)_{n\geq 1}$ for Γ (where for each $n \geq 1, \hat{\beta}^{\varepsilon_n}$ is the SRE of Γ^{ε_n} and such that $\lim_{k\to\infty}\left(\hat{\beta}^{\varepsilon_n}\right)_{k\geq 1} = \beta^*$) and "close enough" to Γ, we have

$$u_i\left(\beta_i', \hat{\beta}_{-i}^{\varepsilon}|h'\right) > u_i\left(\hat{\beta}^{\varepsilon}|h'\right),$$

which contradicts the fact that $\hat{\beta}^{\varepsilon}$ is a SRE in Γ^ε.[9] ■

Corollary 10 *The set R^P of the P-SREs of a dynamic reciprocity game à la DK is the subset of R^{DK} in which the equilibrium profile is dynamically consistent.*

With reference to the game in figure 1, an example of such an equilibrium is given by the strategy profile $\langle (BD), (s)\rangle$. It is immediate to prove that it is a SRE, in which the reciprocation occurs at each history as prescribed by DK's definition. But it is also easy to see that it is dynamically consistent: now, since player 1 thinks that 2 believes B, at both nodes where he is active 1 perceives 2's behavior as *unkind*, so that given 2's strategy s, player 1 wants to "punish" player 2 since the beginning of the game, that is to induce the terminal history (B, s, D). With such patterns of beliefs, and hence of perceived kindness, player 1 has no incentive to deviate anymore.

5. The P-SRE: interpretation and related concepts.

The *belief revision policy* introduced by DK can be cast into the family of the *KB-frame models* representing the paradigm in economics of information. Besides the *conservativity principle* already mentioned above, standard assumptions in that literature are the following: a) individuals always believe what they know; b) individuals always know their own beliefs. In a dynamic learning process, as a dynamic game is, these assumptions imply that an individual holding certain beliefs *knows* what he *would believe*, should a certain information be provided to him, whether such information is initially believed as possible or not. This is the *dynamic positive introspection property* of beliefs.[10]

In the section 3 it has been pointed out that the notion of reciprocity entailed by DK's SRE may not allow a complete ranking of a player's strategies, in terms of kindness. Such an incompleteness derived from the impossibility of comparing strategies differing one from the other only in actions taken at zero-probability histories. It has been argued that also off-the-equilibrium behavior should matter in the evaluation of a player's kindness, so that the kindness of a player's strategy should be evaluated according to the beliefs that player would hold when taking the actions prescribed by that strategy. The *dynamic positive introspection* property of beliefs discussed above provides a strong argument in favor of what counterfactual beliefs should be considered, and why they should matter. The P-SRE introduced in the previous section deals with all these three issues in a unified way.

As commented above, the logic of the kindness functions in DK entails a comparison of the outcomes that are feasible according to the players' beliefs, and the outcome actually realized by a given strategy, so that in general it is not possible to make a kindness comparison among actions taken at histories reached with zero probability. Indeed, in case of completely mixed behavioral strategy profiles, so that there are no unexpected histories, DK's model has none of the drawbacks discussed above. In particular, the problem of dynamic inconsistency disappears, together with the incompleteness in the ranking of the strategies' kindness.

The problem at hand is analogous to that studied by Blume et al. (1991a,b), who develop a decision theoretic model in which there exist no *Savage null* events, but in which the notion of zero probability is saved. The corresponding of the *Savage null* events here would be the strategies whose kindness players cannot distinguish because they differ only in actions taken at zero-probability histories. Blume et al.(1991a) solve the problem developing a *non-Archimedean decision theory* that applies to *lexicographic probability systems* to represent agents' beliefs, and their counterfactual beliefs (i.e. those that would be held after zero-probability events). In a similar fashion, one might consider a model of reciprocty in which ties in the kindness evaluation that are due to conditionalizations to zero-probability events are broken thanks to a lexicographic structure of the kindness functions, that directly reflects the lexicographic structure of beliefs both *on* and *off-the equilibrium path*, i.e. also with respect to the counterfactuals. As shown in the work of Stalnaker (1996, 1998), equilibrium concepts such as Selten's perfect equilibrium can be seen as a shortcut to represent a model of sequential rationality in a non-Archimedean decision theoretic framework, that dispenses with the use of lexicographic probability systems to represent counterfactual beliefs (see Blume et al., 1991b, on the latter point). Analogously, the notion of P-SRE can be interpreted as a shortcut for a model of sequential reciprocity that applies DK's notion of reciprocity to a lexicographic probability system, used to evaluate kindness at off-the-equilibrium histories using as counterfactual beliefs those consistent with the dynamic positive introspection property.

References

Battigalli, P. and Bonanno, G.,1999, Recent results on belief, knowledge and the epistemic foundations of game theory, Research in Economics 53, 149-225.

Battigalli, P. and Dufwenberg, M., 2005, Dynamic Psychological Games, IGIER Working Paper Series, n.287.

Battigalli, P. and Siniscalchi, M., 1999, Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games, Journal of Economic Theory 88, 188-230.

Blume, L., Brandenburger, A., and Dekel, E., 1991a, Lexicographic Probabilities and Choice Under Uncertainty, Econometrica 59, 61-79.

Blume, L., Brandenburger, A., and Dekel, E., 1991b, Lexicographic Probabilities and Equilibrium Refinements, Econometrica 59, 81-98.

Dufwenberg, M. and Kirchsteiger, G.,2004, A theory of sequential reciprocity, Games and Economic Behavior 47, 268-298.

Elster, J., 1998, Emotions and economic theory,. Journal of Economic Literature, 36, 47-74.

Frederick, S. Loewenstein, G. and O'Donoghue,T., 2002. Time Discounting and Time Preference: A Critical Review, Journal of Economic Literature, 40, 351-401.

Gärdenfors, P. 1988, Knowledge in Flux, MIT Press, Cambridge, MA.

Geanakoplos, J., Pearce, D., and Stacchetti, E.,1989, Psychological Games and Sequential Rationality, Games and Economic Behavior 1, 60-79

Gul, D. and Pesendorfer, W., 2006, The Canonical Type Space for Interdependent Preferences, mimeo, Princeton University.

Harman, G., 1986, Change in Views: Principles of Reasoning, MIT Press, Cambridge, MA.

Laibson, D., 1997, Golden Eggs and Hyperbolic Discounting, Quarterly Journal of Economics, 112, 443-477

O'Donoghue, T. and Rabin,M., 1999, Doing It Now or Later, American Economic Review, 89, 103-124

Penta, A., 2004, Psychological Games, Interactive Epistemology and Reciprocity, undergraduate dissertation, Bocconi University, Milan.

Rabin, M., 1993, Incorporating fairness into game theory and economics, American Economic Review 83, 1281-1302

Rabin, M., 1998, Psychology and Economics, Journal of Economic Literature, 36, 11-46

Reny, P., 1992, Backward Induction, Normal Form Perfection and Explicable Equilibria, Econometrica, 60, 627-649

Selten, R., 1975, Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games, International Journal of Game Theory, 4, 25-55.

Stalnaker, R., 1996, Knowledge, Belief and Counterfactual Reasoning in Games, Economics and Philosophy, 12, 133-163.

Stalnaker, R., 1998, Belief Revision in Games: Forward and Backward Induction, Mathematical Social Sciences, 36, 31-56.