Dynamic Psychological Games: Omitted Material

Pierpaolo Battigalli
Bocconi University

Martin Dufwenberg
University of Arizona

November 29, 2007

Abstract

This manuscript collects proofs and other material omitted from *Dynamic Psychological Games*, mimeo, November 2007 (DPG). To make it more self-contained, key definitions and results of the main paper are included. For the convenience of readers of DPG, the numbers of equations and statements coincide with DPG.

1 Extensive forms with observable actions

Here we provide a more complete definition of finite extensive forms with observable actions and related concepts used in results and proofs below.

Fix a finite player set N and finite action sets A_i $(i \in N)$. Let $A = \prod_{i\in N} A_i$. A history of length ℓ is a finite sequence of action profiles $h = (a^1, ..., a^\ell) \in A^\ell$. History $h = (a^1, ..., a^k)$ *precedes* $\overline{h} = (\overline{a}^1, ..., \overline{a}^\ell)$, written $h \prec \overline{h}$, if h is a prefix of $\overline{h}$, i.e. $k < \ell$ and $(a^1, ..., a^k) = (\overline{a}^1, ..., \overline{a}^k)$. In this case, we also write $\overline{h} = (h, \overline{a}^{k+1}, ..., \overline{a}^\ell)$. The empty sequence (the history with zero length) is denoted h^0. By convention h^0 precedes every proper history. A finite extensive form with observable actions is a structure $\langle N, H\rangle$ where $H \subseteq \{h^0\} \cup \left(\bigcup_{\ell=1}^{L} A^\ell\right)$ is a finite set of histories with the following properties:[1]

- $h^0 \in H$.
- $\forall \overline{h} \in H$, if $h \prec \overline{h}$ then $h \in H$.
- $\forall h \in H$, $\{a \in A : (h, a) \in H\} = \prod_{i\in N} A_i(h)$ where

$$A_i(h) = \left\{ a_i \in A_i : \exists a_{-i} \in \prod_{j\neq i} A_j, (h, (a_i, a_{-i})) \in H \right\}$$

  is the set of possible actions of player i at history h.

Note that $\langle H, \prec\rangle$ is a tree with distinguished root h^0; the symmetric closure of $\prec$ is denoted by $\preceq$.[2] We let $Z = \{h \in H : \prod_{i\in N} A_i(h) = \emptyset\}$ denote the set of terminal (or complete) histories.

We can now define the following derived elements:

- $S_i = \{s_i = (s_{i,h})_{h\in H} \in (A_i)^H : \forall h \in H \backslash Z, s_{i,h} \in A_i(h)\}$ is the set of strategies of player i, $S = \prod_{i=1}^n S_i$, $S_{-i} = \prod_{j\in N\backslash\{i\}} S_j$.
- $\zeta : S \to Z$ is the path function, that is, $z = (a^1, ..., a^K) = \zeta(s)$ iff $a^1 = (s_{i,h^0})_{i\in N}$,$\forall t \in \{1, ...K - 1\}$, $a^{t+1} = (s_{i,(a^1,...,a^t)})_{i\in N}$.
- For any $h \in H$, $S(h)$ is the set of strategy profiles consistent with h, i.e., $S(h) = \{s \in S : h \preceq \zeta(s)\}$. Since past actions are observed, it follows that $S(h) = \prod_{i=1}^n S_i(h)$, where $S_i(h)$ is the projection of $S(h)$ on S_i.

[1] Cf. Osborne & Rubinstein (1994, Chapter 6).

[2] Thus, $h \preceq h'$ iff either $h \prec h'$ or $h = h'$.

2 Infinite hierarchies of conditional beliefs

Here we just collect definitions and results about infinite hierarchies of conditional probability systems. This section is included to make the manuscript self-contained, but its content is also contained in DPG.

Fix a compact Polish space X, $\mathcal{B}$ is the Borel sigma algebra, $\mathcal{C}$ is a countable collection of clopen conditioning events.

Definition 1 *A* conditional probability system *(cps) on* $(X, \mathcal{B}, \mathcal{C})$ *is a function* $\mu(\cdot|\cdot) : \mathcal{B} \times \mathcal{C} \to [0,1]$ *such that for all* $E \in \mathcal{B}$, $F, F' \in \mathcal{C}$
(1) $\mu(\cdot|F) \in \Delta(X)$,
(2) $\mu(F|F) = 1$,
(3) $E \subseteq F' \subseteq F$ *implies* $\mu(E|F) = \mu(E|F')\mu(F'|F)$.

We regard the set $\Delta^{\mathcal{C}}(X)$ of cps' on $(X, \mathcal{B}, \mathcal{C})$ as a subset of the topological space $[\Delta(X)]^{\mathcal{C}}$, where $\Delta(X)$ is endowed with the topology of weak convergence of measures and $[\Delta(X)]^{\mathcal{C}}$ is endowed with the product topology.

From now on DM is a player i, and $(X, \mathcal{B}, \mathcal{C})$ is specified as follows: either $X = S_{-i}$ (a finite set), or $X = S_{-i} \times Y$, where Y is some compact Polish parameter space typically representing a set of opponents' beliefs; the Borel sigma-algebra $\mathcal{B}$ is implicitly understood, and conditioning events corresponds to histories, that is, $\mathcal{C} = \{F \subseteq S_{-i} \times Y : F = S_{-i}(h) \times Y, h \in H\}$ (or $\mathcal{C} = \{F \subseteq S_{-i} : F = S_{-i}(h), h \in H\}$ if $X = S_{-i}$). The set of cps' is denoted $\Delta^H(S_{-i} \times Y)$ a subset of $[\Delta(S_{-i} \times Y)]^H$. If conditioning event F corresponds to history h, then we abbreviate and write $\mu(\cdot|F) = \mu(\cdot|h)$.

Lemma 2 $\Delta^H(S_{-i})$ *is a compact Polish space. Furthermore, if* Y *is a compact Polish space, also* $\Delta^H(S_{-i} \times Y)$ *is a compact Polish space.*

Hierarchies of cps' are defined recursively as follows:

- $X_{-i}^0 = S_{-i}$ $(i \in N)$,
- $X_{-i}^k = X_{-i}^{k-1} \times \prod_{j \neq i} \Delta^H(X_{-j}^{k-1})$ $(i \in N;\ k = 1, 2, ...)$.

By repeated applications of Lemma 2, each X_{-i}^k is a cross-product of compact Polish spaces, hence compact Polish itself.[3] A cps $\mu_i^k \in \Delta^H(X_{-i}^{k-1})$ is called k-order cps. For $k > 1$, μ_i^k is a joint cps on the opponents' strategies and $(k-1)$-order cps'. A *hierarchy of cps'* is a countably infinite sequence of cps' $\boldsymbol{\mu}_i =$

[3]The cross-product of countably many compact Polish spaces is also compact Polish.

$(\mu_i^1, \mu_i^2, ...) \in \prod_{k>0} \Delta^H(X_{-i}^{k-1})$. $\boldsymbol{\mu}_i$ is *coherent* if the cps' of distinct orders assign the same conditional probabilities to lower-order events:

$$\mu_i^k(\cdot|h) = \text{marg}_{X_{-i}^{k-1}} \mu_i^{k+1}(\cdot|h) \ (k = 1, 2, ...;\ h \in H).$$

It can be shown that a coherent hierarchy $\boldsymbol{\mu}_i$ induces a cps ν_i on the cross-product of S_{-i} with the sets of hierarchies of cps' of i's opponents, a compact Polish space.

A coherent hierarchy $\boldsymbol{\mu}_i$ satisfies belief in coherency of order 1 if the induced cps ν_i is such that each $\nu_i(\cdot|h)$ $(h \in H)$ assigns probability one to the opponents' coherency; $\boldsymbol{\mu}_i$ satisfies belief in coherency of order k if it satisfies belief in coherency of order $k-1$ and the induced cps ν_i is such that each $\nu_i(\cdot|h)$ $(h \in H)$ assigns probability one the opponents' coherency of order $k-1$; $\boldsymbol{\mu}_i$ is *collectively coherent* if it satisfies belief in coherency of order k for each positive integer k. The set of collectively coherent hierarchies of player i is a compact Polish space, denoted by $\mathbf{M}_i$. We let M_i^k denote the set of k-order beliefs consistent with collective coherency, that is, the projection of $\mathbf{M}_i$ on $\Delta^H(X_{-i}^{k-1})$, and let $M_{-i}^k = \prod_{j \neq i} M_j^k$, $\mathbf{M}_{-i} = \prod_{j \neq i} \mathbf{M}_j$, $\mathbf{M} = \prod_{j \in N} \mathbf{M}_j$.

Lemma 3 *For each $i \in N$ there is a 1-to-1 and onto continuous function*

$$f_i = (f_{i,h})_{h \in H} : \mathbf{M}_i \to \Delta^H(S_{-i} \times \mathbf{M}_{-i})$$

whose inverse is also continuous. Furthermore, each coordinate function $f_{i,h}$ is such that for all $\boldsymbol{\mu}_i = (\mu_i^1, \mu_i^2 ...) \in \mathbf{M}_i$, $k \geq 1$

$$\mu_i^k(\cdot|h) = \text{marg}_{S_{-i} \times M_{-i}^1 \times ... \times M_{-i}^{k-1}} f_{i,h}(\boldsymbol{\mu}_i).$$

Definition 4 *A* psychological game *based on extensive form $\langle N, H \rangle$ is a structure $\Gamma = \langle N, H, (u_i)_{i \in N} \rangle$ where $u_i : Z \times \mathbf{M} \times S_{-i} \to \mathbb{R}$ is i's (measurable and bounded) psychological payoff function.*

3 Dynamic Programing on Beliefs-Induced Decision Trees

Fix a hierarchy of cps' $\boldsymbol{\mu}_i$ a non terminal history h and a strategy s_i consistent with h. The expectation of u_i conditional on h, given s_i and $\boldsymbol{\mu}_i$ is

$$\mathrm{E}_{s_i, \boldsymbol{\mu}_i}[u_i|h] := \int_{S_{-i} \times \mathbf{M}_{-i}} u_i(\zeta(s_i, s_{-i}), \boldsymbol{\mu}_i, \boldsymbol{\mu}_{-i}, s_{-i}) f_{i,h}(\boldsymbol{\mu}_i)(ds_{-i}, d\boldsymbol{\mu}_{-i}). \qquad \text{(EU)}$$

We want to relate the problem $\max_{s_i \in S_i(h)} \mathrm{E}_{s_i,\boldsymbol{\mu}_i}[u_i|h]$ to dynamic programming on a decision tree induced by $\boldsymbol{\mu}_i$. (Note that in this problem we allow for changing the strategy also at histories not weakly follow h. But by own-strategy independence of u_i this is irrelevant. We write the problem like this only to simplify the notation.)

For any fixed hierarchy of cps' μ_i, we obtain a well defined decision tree that can be solved by backward induction. Define value functions $V_{\boldsymbol{\mu}_i} : H \to \mathbb{R}$ and $\overline{V}_{\boldsymbol{\mu}_i} : (H \backslash Z) \times A_i \to \mathbb{R}$ as follows:

- For terminal histories $z \in Z$, let

$$V_{\boldsymbol{\mu}_i}(z) = \int_{S_{-i} \times \mathbf{M}_{-i}} u_i(z, \boldsymbol{\mu}_i, \boldsymbol{\mu}_{-i}, s_{-i}) f_{i,z}(\boldsymbol{\mu}_i)(ds_{-i}, d\boldsymbol{\mu}_{-i}).$$

- Assuming that $V_{\boldsymbol{\mu}_i}(h, a)$ has been defined for all the immediate successors (h, a) of history h, let

$$\overline{V}_{\boldsymbol{\mu}_i}(h, a_i) = \sum_{a_{-i} \in A_{-i}(h)} \mu_i^1(S_{-i}(h, a_{-i})|h) V_{\boldsymbol{\mu}_i}(h, (a_i, a_{-i}));$$

  for each $a_i \in A_i(h)$; then $V_{\boldsymbol{\mu}_i}(h)$ is defined as

$$V_{\boldsymbol{\mu}_i}(h) = \max_{a_i \in A_i(h)} \overline{V}_{\boldsymbol{\mu}_i}(h, a_i).$$

For any given strategy s_i and history $h \in H \backslash Z$, we use the following notation:

- For each k with $0 \le k < \ell(h)$ (recall that $\ell(h)$ denotes the length of history h), $a_i^k(h)$ is action taken by i in h at the prefix of h of length k. Thus, by definition $h = (a^0(h), a^1(h), ..., a^{\ell(h)-1}(h))$, where $a^k(h) = (a_1^k(h), ..., a_n^k(h))$.

- $(s_i|h)$ denotes the strategy that takes all the actions of player i in history h and behaves as s_i otherwise:

$$(s_i|h)_{h'} = \begin{cases} s_{i,h'} & \text{if } h' \not\prec h, \\ a_i^{\ell(h')}(h) & \text{if } h' \prec h. \end{cases}$$

- $(s_i|h, a_i)$ denote the strategy obtained from $(s_i|h)$ by replacing $s_{i,h}$ with $a_i \in A_i(h)$:

$$(s_i|h, a_i)_{h'} = \begin{cases} (s_i|h)_{h'} & \text{if } h' \neq h, \\ a_i & \text{if } h' = h. \end{cases}$$

- Finally, let $d(h) = \max_{h \preceq z}[\ell(z) - \ell(h)]$ denote the depth of the subtree with root h.

Lemma A (Dynamic Programming A). *Suppose that*

$$\forall h \in H \backslash Z,\ s^*_{i,h} \in \arg \max_{a_i \in A_i(h)} \overline{V}_{\boldsymbol{\mu}_i}(h, a_i).$$

Then

$$\forall h \in H \backslash Z, \mathrm{E}_{(s^*_i|h),\boldsymbol{\mu}_i}[u_i|h] = V_{\boldsymbol{\mu}_i}(h) = \max_{s_i \in S_i(h)} \mathrm{E}_{s_i,\boldsymbol{\mu}_i}[u_i|h]. \qquad \text{(DP)}$$

Proof. The proof is by induction on $d(h)$.

Basis step. Obviously (DP) holds for all h such that $d(h) = 1$.

Inductive step. Suppose (DP) holds for all h such that $1 \leq d(h) \leq k$. Let $d(h) = k + 1$. By the law of iterated expectations for all $a_i \in A_i(h)$

$$\mathrm{E}_{(s^*_i|h,a_i),\boldsymbol{\mu}_i}[u_i|h] = \sum_{a_{-i} \in A_{-i}(h)} \mu^1_i(S_{-i}(h, a_{-i})|h) \mathrm{E}_{(s^*_i|h,a_i),\boldsymbol{\mu}_i}[u_i|h, (a_i, a_{-i})].$$

By the inductive hypothesis, for all $a_i \in A_i(h)$, $a_{-i} \in A_{-i}(h)$

$$\mathrm{E}_{(s^*_i|h,a_i),\boldsymbol{\mu}_i}[u_i|h, (a_i, a_{-i})] = V_{\boldsymbol{\mu}_i}(h, (a_i, a_{-i})) = \max_{s_i \in S_i(h,(a_i,a_{-i}))} \mathrm{E}_{s_i,\boldsymbol{\mu}_i}[u_i|h, (a_i, a_{-i})].$$

Taking expectatons w.r.t. a_{-i}:

$$\mathrm{E}_{(s^*_i|h,a_i),\boldsymbol{\mu}_i}[u_i|h] = \overline{V}_{\boldsymbol{\mu}_i}(h, a_i).$$

Therefore

$$\begin{aligned}
\mathrm{E}_{(s^*_i|h),\boldsymbol{\mu}_i}[u_i|h] &= V_{\boldsymbol{\mu}_i}(h) = \max_{s_i \in S_i(h)} \mathrm{E}_{s_i,\boldsymbol{\mu}_i}[u_i|h] \\
& \quad \text{if and only if} \\
s^*_{i,h} &\in \arg \max_{a_i \in A_i(h)} \mathrm{E}_{(s^*_i|h,a_i),\boldsymbol{\mu}_i}[u_i|h] \\
& \quad \text{if and only if} \\
s^*_{i,h} &\in \arg \max_{a_i} \overline{V}_{\boldsymbol{\mu}_i}(h, a_i).
\end{aligned}$$

The latter condition holds by assumption; hence the inductive step is proved.■

4 Sequential Equilibrium

We focus on behavior strategies $\sigma_i = (\sigma_i(\cdot|h))_{h\in H\setminus Z} \in \prod_{h\in H\setminus Z} \Delta(A_i(h))$, interpreting σ_i as an array of common conditional first-order beliefs held by i's opponents. This interpretation is part of the notion of 'consistency' of profiles of strategies and hierarchical beliefs defined below.

In our framework an *assessment* is a profile $(\sigma, \boldsymbol{\mu}) = (\sigma_i, \boldsymbol{\mu}_i)_{i\in N}$ where σ is a behavioral strategy profile and $\boldsymbol{\mu} \in \mathbf{M}$. We extend the definition of consistency by adding a requirement concerning the higher-order beliefs that need to be specified in psychological games.

Let $\Pr_{\sigma_j}(\cdot|\hat{h}) \in \Delta(S_j(\hat{h}))$ denote the probability measure over j's strategies conditional on $\hat{h}$ derived from behavior strategy σ_j under the assumption of independence across histories:

$$\forall s_j \in S_j(\hat{h}),\ \Pr_{\sigma_j}(s_j|\hat{h}) := \prod_{h\in H\setminus Z: h \not\prec \hat{h}} \sigma_j(s_{j,h}|h)$$

($h \not\prec \hat{h}$ means that h is not a predecessor, or prefix, of $\hat{h}$).[4]

Definition 5 *A profile of first-order cps' $\mu^1 = (\mu_i^1)_{i\in N}$ is* derived *from a behavioral strategy profile $\sigma = (\sigma_i)_{i\in N}$ if for all $i \in N$, $s_{-i} \in S_{-i}$, $\hat{h} \in H$,*

$$\mu_i^1(s_{-i}|\hat{h}) = \prod_{j\neq i} \Pr_{\sigma_j}(s_j|\hat{h}). \qquad (1)$$

Clearly, if μ^1 is derived from σ then for any three players i, j, k, the beliefs of i and j about k coincide:

$$\forall \hat{h} \in H,\ \mathrm{marg}_{S_k}\mu_i^1(\cdot|\hat{h}) = \Pr_{\sigma_k}(\cdot|\hat{h}) = \mathrm{marg}_{S_k}\mu_j^1(\cdot|\hat{h}).$$

Definition 6 *Assessment $(\sigma, \boldsymbol{\mu})$ is* consistent *if*
(a) μ^1 is derived from σ,
(b) higher-order beliefs in $\boldsymbol{\mu}$ assign probability 1 to the lower-order beliefs:

$$\forall i \in N,\ \forall k > 1,\ \forall h \in H,\ \mu_i^k(\cdot|h) = \mu_i^{k-1}(\cdot|h) \times \delta_{\mu_{-i}^{k-1}}$$

where $\times$ denotes the product of measures and δ_x is the Dirac measure assigning probability 1 to singleton $\{x\}$.

[4] Cf. Kuhn (1953).

We now move to the main definition: a consistent assessment is a sequential equilibrium if it satisfies sequential rationality. Recall that

$$\mathrm{E}_{s_i,\boldsymbol{\mu}_i}[u_i|h] := \int_{S_{-i}\times\mathbf{M}_{-i}} u_i(\zeta(s_i,s_{-i}),\boldsymbol{\mu}_i,\boldsymbol{\mu}_{-i},s_{-i})f_{i,h}(\boldsymbol{\mu}_i)(ds_{-i},d\boldsymbol{\mu}_{-i}). \tag{2}$$

Definition 7 *An assessment* $(\sigma,\boldsymbol{\mu})$ *is a* sequential equilibrium (SE) *if it is consistent and for all* $i\in N$, $h\in H\backslash Z$, $s_i^*\in S_i(h)$,

$$\Pr_{\sigma_i}(s_i^*|h)>0 \ \Rightarrow\ s_i^*\in\arg\max_{s_i\in S_i(h)}\mathrm{E}_{s_i,\boldsymbol{\mu}_i}[u_i|h]. \tag{3}$$

Note that, by consistency, σ_i represents the first-order beliefs of i's opponents about i, and furthermore there is common certainty of the true belief profile $\boldsymbol{\mu}$ at every history; therefore the sequential rationality condition (3) can equivalently be written as

$$\forall j\neq i,\ \text{supp}\ \text{marg}_{S_i}\mu_j^1(\cdot|h)\subseteq\arg\max_{s_i\in S_i(h)}\sum_{s_{-i}\in S_{-i}(h)}\mu_i^1(s_{-i}|h)u_i(\zeta(s_i,s_{-i}),\boldsymbol{\mu},s_{-i}). \tag{4}$$

This clarifies that SE is a notion of equilibrium in beliefs. Indeed we could have given an equivalent definition of SE with no reference to behavioral strategies.

We can also take the point of view of an 'agent' (i,h) of player i, in charge of the move at history h, who seeks to maximize i's conditional expected utility given the consistent assessment $(\sigma,\boldsymbol{\mu})$. The expected utility of i conditional on h and $a_i\in A_i(h)$ given $(\sigma,\boldsymbol{\mu})$ can be expressed as

$$\mathrm{E}_{\sigma,\boldsymbol{\mu}}[u_i|h,a_i] := \sum_{s_{-i}\in S_{-i}(h)}\prod_{j\neq i}\Pr_{\sigma_j}(s_j|h)\sum_{s_i\in S_i(h,a_i)}\Pr_{\sigma_i}(s_i|h,a_i)u_i(\zeta(s),\boldsymbol{\mu},s_{-i}), \tag{5}$$

where $\Pr_{\sigma_i}(s_i|h,a_i):=\prod_{h'\in H\backslash Z:h'\not\preceq h}\sigma_i(s_{i,h'}|h')$ ($h'\not\preceq h$ means that h' is not h or a predecessor of h). This specification presumes that (i,h) assesses the probabilities of actions by other agents of player i in the same way as each player $j\neq i$; that is using the behavioral strategy σ_i.

Remark. *Suppose that* u_i *depends only on terminal histories and beliefs, not on* s_{-i}. *Then we obtain the more familiar formula*

$$\mathrm{E}_{\sigma,\boldsymbol{\mu}}[u_i|h,a_i]=\sum_z\Pr_\sigma(z|h,a_i)u_i(z,\boldsymbol{\mu}),$$

where $\Pr_\sigma(z|h,a_i)$ *is the probability of terminal history* z *conditional on* (h,a_i)

determined by σ.

Proposition 8 *A consistent assessment $(\sigma, \boldsymbol{\mu})$ satisfies (3) and hence is an SE if and only if for all $i \in N$, $h \in H \backslash Z$,*

$$\mathrm{supp}(\sigma_i(\cdot|h)) \subseteq \arg\max_{a_i \in A_i(h)} \mathrm{E}_{\sigma,\boldsymbol{\mu}}[u_i|h, a_i]. \tag{6}$$

Proof. Let $(\sigma, \boldsymbol{\mu})$ be consistent. Then for each $z \in Z$,

$$V_{\boldsymbol{\mu}_i}(z) = \sum_{s_{-i} \in S_{-i}(h)} \mu_i^1(s_{-i}|z) u_i(z, \boldsymbol{\mu}_i, \boldsymbol{\mu}_{-i}, s_{-i}),$$

and for all h with $d(h) = 1$ (recall that $d(h)$ is the depth of the tree wtih root h) we have

$$V_{\boldsymbol{\mu}_i}(h) = \max_{a_i \in A_i(h)} \mathrm{E}_{\sigma,\boldsymbol{\mu}}[u_i|h, a_i]. \tag{BI}$$

We must show that $(\sigma, \boldsymbol{\mu})$ satisfies the sequential rationality condition (3) if and only if the one-shot deviation condition (6) holds. The "if" part is obvious. Now suppose that $(\sigma, \boldsymbol{\mu})$ satisfies (6). Then a straightforward induction argument shows that (BI) holds for all $h \in H \backslash Z$. Therefore Lemma A implies that (3) holds.■

We obtain the following existence theorem:

Theorem 9 *If the psychological payoff functions are continuous, there exists at least one sequential equilibrium assessment.*

(The proof of the theorem is only sketched in DPG.)

Proof of Theorem 9. We first show how to associate a consistent assessment $(\sigma, \beta(\sigma))$ with each behavioral strategy profile σ. Let $\beta^1(\sigma) = (\beta_i^1(\sigma))_{i \in N} \in M^1$ denote the profile of first-order beliefs *derived* from σ according to Def. 5. The profile of belief hierarchies $\boldsymbol{\mu} = \beta(\sigma)$ is obtained by condition (b) in Definition 6:

$$\begin{aligned} \forall i &\in N,\ \mu_i^1 = \beta_i^1(\sigma), \\ \forall i &\in N,\ \forall k > 1,\ \forall h \in H,\ \mu_i^k(\cdot|h) = \mu_i^{k-1}(\cdot|h) \times \delta_{\mu_{-i}^{k-1}}. \end{aligned}$$

By construction, assessment $(\sigma, \beta(\sigma))$ is consistent. It is clear from the construction that $\beta(\cdot)$ is a continuous function.

Definition of ε-equilibrium. Fix a strictly positive vector $\varepsilon = (\varepsilon_{i,h}(a_i)_{a_i \in A_i(h)})_{i \in N, h \in H \backslash Z}$ s.t. $\forall h \in H \backslash Z$, $\sum_{a_i \in A_i(h)} \varepsilon(a_i) < 1$. An ε-*equilibrium* is a behavioral strategy profile σ s.t. $\forall i \in N$, $\forall h \in H$, $\forall a_i \in A_i(h)$,
(i) $\sigma_i(a_i|h) \geq \varepsilon_{i,h}(a_i)$,
(ii) $a_i \notin \arg\max_{a_i' \in A_i(h)} \mathrm{E}_{\sigma,\beta(\sigma)}[u_i|h, a_i']$ implies $\sigma_i(a_i|h) = \varepsilon_{i,h}(a_i)$.

Let Σ_ε denote the set of behavioral strategy profiles satisfying condition (i) of the definition above and let $r_\varepsilon : \Sigma_\varepsilon \twoheadrightarrow \Sigma_\varepsilon$ denote the "ε-best response correspondence" that assigns to each profile σ the subset of profiles in Σ_ε satisfying condition (ii) of the definition, that is,

$$\begin{aligned} & r_{\varepsilon,i}(\sigma) \\ = \; & \{\sigma_i' \in \Sigma_{\varepsilon,i} : \forall h, \forall a_i, a_i \notin \arg \max_{a_i' \in A_i(h)} \mathrm{E}_{\sigma,\beta(\sigma)}[u_i|h, a_i'] \Rightarrow \sigma_i(a_i|h) = \varepsilon_{i,h}(a_i)\}, \end{aligned}$$

$$r_\varepsilon(\sigma) = \prod_{i \in N} r_{\varepsilon,i}(\sigma).$$

$r_{\varepsilon,i}(\sigma)$ is a nonempty convex subset of $\Delta(A_i(h))$. Since $\mathrm{E}_{\sigma,\boldsymbol{\mu}}[u_i|h, a_i]$ is continuous in (σ, μ) and $\mu = \beta(\sigma)$ is a continuous function, $\mathrm{E}_{\sigma,\beta(\sigma)}[u_i|h, a_i]$ is continuous in σ. This implies that $r_{\varepsilon,i}(\sigma)$ has a closed graph. Thus, $r_\varepsilon(\cdot)$ is a nonempty convex valued correspondence with a closed graph from the compact and convex set $\prod_{h \in H \backslash Z} \prod_{i \in N} \Delta(A_i(h))$ to itself. By the Kakutani theorem $r_\varepsilon(\cdot)$ has a fixed point, which is an ε-psychological equilibrium.

Fix a sequence $\varepsilon^k \to 0$ and a corresponding sequence of ε^k-psychological equilibrium strategies σ^k. By compactness, the sequence (σ^k) has a limit point σ^*. We prove that $(\sigma^*, \beta(\sigma^*))$ is a sequential equilibrium. Assessment $(\sigma^*, \beta(\sigma^*))$ is consistent: to see this just note that, by continuity, $\beta(\sigma^*)$ is a limit point of $\beta(\sigma^k)$, and that the set of consistent assessments is closed. By continuity of $\mathrm{E}_{\sigma,\beta(\sigma)}[u_i|h, a_i]$ in σ (and finiteness of $A_i(h)$), for k sufficiently large

$$\arg \max_{a_i \in A_i(h)} \mathrm{E}_{\sigma^*,\beta(\sigma^*)}[u_i|h, a_i] = \arg \max_{a_i \in A_i(h)} \mathrm{E}_{\sigma^k,\beta(\sigma^k)}, [u_i|h, a_i].$$

This implies that

$$\mathrm{supp}(\sigma_i^*(\cdot|h)) \subseteq \arg \max_{a_i \in A_i(h)} \mathrm{E}_{\sigma^*,\beta(\sigma^*)}[u_i|h, a_i].$$

By Proposition 8, this implies that $(\sigma^*, \beta(\sigma^*))$ is a sequential equilibrium.■

Suppose that psychological utilities depend only on terminal nodes and beliefs: $u_i : Z \times \mathbf{M} \to \mathbb{R}$. For any such game $\Gamma = \langle N, H, (u_i)_{i \in N} \rangle$ and any profile of hierarchies of cps' $\boldsymbol{\mu} = (\boldsymbol{\mu}_i)_{i\in N}$, we can obtain a *standard* game $\Gamma^{\boldsymbol{\mu}} = \langle N, H, (v_i^{\boldsymbol{\mu}})_{i\in N}\rangle$ with payoff functions $v_i^{\boldsymbol{\mu}}(z) = u_i(z, \boldsymbol{\mu})$.

Remark 10 *Suppose that psychological payoff functions have the form $u_i : Z \times \mathbf{M} \to \mathbb{R}$. Then an assessment $(\sigma, \boldsymbol{\mu})$ is a sequential equilibrium if and only if it is consistent and σ is a subgame perfect (hence sequential) equilibrium of the standard game $\Gamma^{\boldsymbol{\mu}}$.*

Proof. Let $(\sigma, \boldsymbol{\mu})$ be a *consistent* assessment. Since i's utility in Γ has the form $u_i : Z \times \mathbf{M} \to \mathbb{R}$, the definition of $v_i^{\boldsymbol{\mu}}$ implies that

$$\mathrm{E}_{\sigma,\boldsymbol{\mu}}[u_i|h, a_i] = \sum_z \Pr_\sigma(z|h, a_i)u_i(z, \boldsymbol{\mu}) = \sum_z \Pr_\sigma(z|h, a_i)v_i^{\boldsymbol{\mu}}(z).$$

where $\Pr_\sigma(z|h, a_i)$ is the probability of terminal history z conditional on (h, a_i) determined by σ (see the previous remark). It easily follows that behavior strategy profile σ is sequentially rational in Γ given $\boldsymbol{\mu}$, if and only if it is a subgame perfect equilibrium a subgame perfect equilibrium of $\Gamma^{\boldsymbol{\mu}}$.■

5 Interactive Epistemology

States. The state of a player is therefore given by his strategy and his hierarchy of cps', $(s_i, \boldsymbol{\mu}_i)$. The set of states for player i is $\Omega_i = S_i \times \mathbf{M}_i$, and the set of *states of the world* is $\Omega = \prod_{i=1}^n \Omega_i$. We let $\Omega_{-i} = \prod_{j\neq i} \Omega_j$ and with a slight abuse of notation we also write $\omega = (\omega_i, \omega_{-i}) \in \Omega = \Omega_i \times \Omega_{-i}$.

Events. An *event* is a Borel subset $E \subseteq \Omega$; an event about i is any subset $E = E_i \times \Omega_{-i}$, where $E_i \subseteq \Omega_i$ is a Borel set. $\mathcal{E}_i$ is the family of events about i. $\mathcal{E}_{-i}$, the family of events about i's opponents, is the collections of events of the form $E = \Omega_i \times E_{-i}$ where $E_{-i} \subseteq \Omega_{-i}$ is a Borel set.

Finally we let $\mathcal{E}$ denote the collection of events of the form $E = \bigcap_{i\in N} E_i$ with $E_i \in \mathcal{E}_i$ for all $i \in N$. Equivalently $E \in \mathcal{E}$ iff E is a Borel set and $E = \prod_{i\in N} \mathrm{proj}_{\Omega_i} E$.

Conditional belief operators. $\forall h \in H$, $\mathrm{B}_{i,h} : \mathcal{E}_{-i} \to \mathcal{E}_i$ is defined as follows:

$$\forall E = \Omega_i \times E_{-i} \in \mathcal{E}_{-i},\ \mathrm{B}_{i,h}(E) = \{(s_i, \boldsymbol{\mu}_i, \omega_{-i}) : f_{i,h}(\boldsymbol{\mu}_i)(E_{-i}) = 1\}.$$

Conditional mutual belief. We define mutual belief operators only on the domain $\mathcal{E}$ of 'product events'. The mutual belief operator $\mathrm{B}_h : \mathcal{E} \to \mathcal{E}$ is defined as follows:

$$\forall E = \bigcap_{i\in N} E_i \in \mathcal{E},\, \mathrm{B}_h(E) = \bigcap_{i\in N} \mathrm{B}_{i,h}\left(\bigcap_{j\neq i} E_j\right).$$

Rationality. Recall the definition of $\mathrm{E}_{s_i,\boldsymbol{\mu}_i}[u_i|h]$ in eq. (EU). Let

$$H_i(s_i^*) = \{h \in H\backslash Z : s_i^* \in S_i(h)\}$$

denote the set of non-terminal histories allowed by s_i^*; the (weakly) sequential best response corrspondence $r_i(\cdot)$ is defined as follows:

$$\forall \boldsymbol{\mu}_i \in \mathbf{M}_i,\, r_i(\boldsymbol{\mu}_i) = \left\{ s_i^* : \forall h \in H(s_i^*),\, s_i^* \in \arg\max_{s_i\in S_i(h)} \mathrm{E}_{s_i,\boldsymbol{\mu}_i}[u_i|h] \right\} \tag{7}$$

The event "player i is rational" is

$$R_i = \{(s_i, \boldsymbol{\mu}_i, \omega_{-i}) : s_i \in r_i(\boldsymbol{\mu}_i)\}.$$

We note in passing a result that is omitted in DPG:

Remark. *If (σ^*, μ^*) is an SE assessment, then there exists a strategy profile s^* such that (a) for all $h \in H\backslash Z$, $i \in N$, $s_{i,h} \in$supp$\sigma_i^*(\cdot|h)$ and (b) for all $h \in H\backslash Z$, $\mathrm{Pr}_{\sigma^*}(h) > 0$ implies that there is rationality and common belief in rationality at state $((s^*|h), \mu^*)$, that is, $((s^*|h), \mu^*) \in R \cap \left(\bigcap_{k>0} (\mathrm{B}_h)^k(R)\right)$.*

This follow quite easily from the definition of sequential equilibrium: the sequential rationality condition implies that $\forall h \in H\backslash Z$, $\mathrm{Pr}_{\sigma^*}(h) > 0 \Rightarrow ((s^*|h), \mu^*) \in R$. By consistency there is common belief of μ^* conditional on h at each state $((s^*|h), \mu^*)$. It follows that $((s^*|h), \mu^*) \in R \cap \left(\bigcap_{k>0} (\mathrm{B}_h)^k(R)\right)$ at such histories.

The following lemma is stated without proof in the Appendix of DPG (see Lemma 17 of DPG):

Lemma B (Dynamic Programming B). *The sequential best reply correspondence $r_i : M_i \twoheadrightarrow S_i$ can be characterized as follows*

$$r_i(\boldsymbol{\mu}_i) = \left\{ s_i : \forall h \in H(s_i), s_{i,h} \in \arg\max_{a_i\in A_i(h)} \overline{V}_{\boldsymbol{\mu}_i}(h, a_i) \right\}.$$

Proof. Fix a strategy $\hat{s}_i$ and let $\hat{s}_i^*$ be the strategy obtained from $\hat{s}_i$ as follows: if $h \in H_i(\hat{s}_i)\backslash Z$ then $\hat{s}_{i,h}^* = \hat{s}_{i,h}$; if $h \in H\backslash(Z \cup H_i(\hat{s}_i))$ then $\hat{s}_{i,h}^* \in \arg\max_{a_i} \overline{V}_{\boldsymbol{\mu}_i}(h, a_i)$. By construction $\hat{s}_i$ and $\hat{s}_i^*$ are realization-equivalent: $H_i(\hat{s}_i) =$

$H_i(\hat{s}_i^*)$ and $\zeta(\hat{s}_i, s_{-i}) = \zeta(\hat{s}_i^*, s_{-i})$ for all s_{-i}. Therefore $u_i(\zeta(\hat{s}_i, s_{-i}), \boldsymbol{\mu}, s_{-i}) = u_i(\zeta(\hat{s}_i^*, s_{-i}), \boldsymbol{\mu}, s_{-i})$ for every $(\boldsymbol{\mu}, s_{-i})$. This implies that $\hat{s}_i$ belongs to $r_i(\boldsymbol{\mu}_i)$ if and only if $\hat{s}_i^*$ does, and similarly

$$\hat{s}_i \in \left\{ s_i : \forall h \in H(s_i)\backslash Z, s_{i,h} \in \arg\max_{a_i} \overline{V}_{\boldsymbol{\mu}_i}(h, a_i) \right\}$$

if and only if $\hat{s}_i^*$ does.

Suppose that $\hat{s}_i \notin \{s_i : \forall h \in H(s_i)\backslash Z, s_{i,h} \in \arg\max_{a_i} \overline{V}_{\boldsymbol{\mu}_i}(h, a_i)\}$. Since H is finite there is some history $\hat{h} \in H(s_i)\backslash Z$ such that $\hat{s}_{i,\hat{h}} \notin \arg\max_{a_i} \overline{V}_{\boldsymbol{\mu}_i}(\hat{h}, a_i)$ and $\hat{s}_{i,h} \in \arg\max_{a_i} \overline{V}_{\boldsymbol{\mu}_i}(h, a_i)$ for all $h \succ \hat{h}$ in $H_i(s_i)\backslash Z$. By the law of iterated expectations, for any s_i and any $h \in H_i(s_i)$

$$\mathrm{E}_{s_i,\boldsymbol{\mu}_i}[u_i|h] = \sum_{a_{-i} \in A_{-i}(h)} \mu_i^1(S_{-i}(h, a_{-i})|h)\mathrm{E}_{s_i,\boldsymbol{\mu}_i}[u_i|h, (s_{i,h}, a_{-i})].$$

By assumption,

$$\mathrm{E}_{\hat{s}_i,\boldsymbol{\mu}_i}[u_i|\hat{h}, (\hat{s}_{i,\hat{h}}, a_{-i})] = V_{\boldsymbol{\mu}_i}(\hat{h}, (\hat{s}_{i,\hat{h}}, a_{-i})).$$

Therefore

$$\mathrm{E}_{\hat{s}_i,\boldsymbol{\mu}_i}[u_i|\hat{h}] = \overline{V}_{\boldsymbol{\mu}_i}(\hat{h}, \hat{s}_{i,\hat{h}}).$$

Pick an action $\overline{a}_i \in \arg\max_{a_i} \overline{V}_{\boldsymbol{\mu}_i}(\hat{h}, a_i)$. Let $\overline{s}_i$ denote the strategy obtained from $\hat{s}_i^*$ by replacing $\hat{s}_{i,\hat{h}}^*$ with $\overline{a}_i$. By construction $\overline{s}_i \in S_i(\hat{h})$ and

$$\mathrm{E}_{\overline{s}_i,\boldsymbol{\mu}_i}[u_i|h] = V_{\boldsymbol{\mu}_i}(h) > \overline{V}_{\boldsymbol{\mu}_i}(h, \hat{s}_{i,\hat{h}}).$$

Therefore there is a history $\hat{h} \in H_i(\hat{s}_i)$ and a strategy $\overline{s}_i \in S_i(\hat{h})$ such that $\mathrm{E}_{\overline{s}_i,\boldsymbol{\mu}_i}[u_i|\hat{h}] > \mathrm{E}_{\hat{s}_i,\boldsymbol{\mu}_i}[u_i|\hat{h}]$, which implies that $\hat{s}_i \notin r_i(\boldsymbol{\mu}_i)$.

Now suppose that $\hat{s}_i \in \{s_i : \forall h \in H(s_i)\backslash Z, s_{i,h} \in \arg\max_{a_i} \overline{V}_{\boldsymbol{\mu}_i}(h, a_i)\}$. Then $\hat{s}_{i,h}^* \in \arg\max_{a_i} \overline{V}_{\boldsymbol{\mu}_i}(h, a_i)$ for all $h \in H\backslash Z$, and Lemma A implies that $\hat{s}_i^* \in r_i(\boldsymbol{\mu}_i)$. Hence $\hat{s}_i \in r_i(\boldsymbol{\mu}_i)$.■

Extensive form rationalizability. First define a '*strong belief* operator' SB_i as follows: $\mathrm{SB}_i(\emptyset) = \emptyset$ and

$$\forall E \in \mathcal{E}_{-i}\backslash\{\emptyset\}, \mathrm{SB}_i(E) = \bigcap_{[h] \cap E \neq \emptyset} \mathrm{B}_{i,h}(E).$$

For each $E = \bigcap_{i \in N} E_i \in \mathcal{E}$, the event "there is mutual strong belief in E" is defined by $\mathrm{SB}(E) = \bigcap_{i \in N} \mathrm{SB}_i \left(\bigcap_{j \neq i} E_j \right)$. Note that $\mathrm{SB}(E) \in \mathcal{E}$. Finally, we introduce an auxiliary 'correct strong belief' operator:

$$\forall E \in \mathcal{E},\ \mathrm{CSB}(E) = E \cap \mathrm{SB}(E).$$

Definition 11 *A state of the world ω is* rationalizable *if $\omega \in \bigcap_{k \geq 0} \mathrm{CSB}^k(R)$.*

To illustrate the full power of Definition 11, we analyze a *Generalized Trust Game with guilt aversion*, reminiscent of Ben-Porath & Dekel's (1992) money-burning game: $A_1 = \{D, Trust_1, ...Trust_L\}$ (Ann), $A_2 = \{Grab, Share\}$ (Bob). Ann moves first, action D terminates the game, action $Trust_\ell$ ($\ell = 1, ...L$) gives the move to Bob. Let $\alpha_\ell(\boldsymbol{\mu}_1) = \mu_1^1(\mathbf{Share}_\ell | h^0)$ and $\beta_\ell(\boldsymbol{\mu}_2) = \int \alpha_\ell(\mu_1^1)\mu_2^2(d\alpha_\ell(\mu_1^1)|Trust_\ell)$. As before we assume that Ann's utility is her material payoff, whereas Bob is averse to guilt. Applying the guilt formula of subsection 3.3 in DPG, the players' utilities are given by

$$\begin{aligned}
u_i(D) &= 1,\ i = 1, 2, \\
u_i(Trust_\ell, Share) &= \left(1 + \frac{\ell}{L}\right),\ i = 1, 2, \\
u_1(Trust_\ell, Grab) &= 0, \\
u_2(Trust_\ell, Grab) &= 2\left(1 + \frac{\ell}{L}\right) - \theta_2 \alpha_\ell \left(1 + \frac{\ell}{L}\right),
\end{aligned}$$

where θ_2 is Bob's sensitivity to guilt. Bob (strictly) prefers to share the yield of project ℓ if and only if $\theta_2 \beta_\ell > 1$.

For $L = 1$ and $\theta_2 = \frac{5}{2}$ we obtain Γ_3 (of DPG), and the forward induction argument used to solve Γ_3 (captured by 2 iterations of the CSB operator) works if and only if $\theta_2 > 2$. By contrast, when $L > 1$ rationalizability yields the efficient sharing outcome also for much lower values of θ_2, as shown in the following proposition, the proof of which is omitted in DPG:

Proposition 12 *In the Generalized Trust Game with guilt aversion, if $\theta_2 > 1 + \frac{1}{L}$ then, for every rationalizable state $(s_1, \boldsymbol{\mu}_1, s_2, \boldsymbol{\mu}_2)$, $s_1 = Trust_L$, $s_2 = (\mathbf{Share}_\ell)_{\ell=1}^L$, $\alpha_\ell(\boldsymbol{\mu}_1) = \beta_\ell(\boldsymbol{\mu}_2) = 1$ ($\ell = 1, ..., L$).*

Proof. For every even number $2k$ and event $\mathrm{CSB}^{2k}(R)$ we characterize the 'largest' project ℓ s.t., according to $\mathrm{CSB}^{2k}(R)$, action $Trust_\ell$ induces Bob to share.

Let $\hat{\ell}(\cdot)$ be defined by:

$$\hat{\ell}(0)=0, \text{ and } \hat{\ell}(2k)=\max\left\{\ell\in\{1,...,L\}:\theta_2\frac{L+\hat{\ell}(2k-2)}{L+\ell}>1\right\}, \ (k\geq 1).$$

Note that for each $k\geq 1$, $\theta_2\frac{L+\hat{\ell}(2k-2)}{L+1+\hat{\ell}(2k-2)}\geq\theta_2\frac{L}{L+1}>1$, where the latter inequality holds by assumption. Therefore function $\hat{\ell}(2k)$ is well-defined and strictly increasing in k until it attains its maximum, L. We claim that for each $k\geq 1$, event $\mathrm{CSB}^{2k}(R)$ implies that $\beta_\ell\geq\frac{L+\hat{\ell}(2k-2)}{L+\ell}$ for each $\ell>\hat{\ell}(2k-2)$, $s_{2,\ell}=\mathbf{Share}_\ell$ and $\alpha_\ell=1$ for each $\ell=1,...,\hat{\ell}(2k)$, and $s_1=Trust_\ell$ for some $\ell\geq\hat{\ell}(2k)$. This implies the thesis.

If Ann chooses project ℓ she signals that $\alpha_\ell\geq\frac{L}{L+\ell}$, because she can obtain \$1 by not investing. By forward induction and rationality [event $\mathrm{CSB}(R)$], $\beta_\ell\geq\frac{L}{L+\ell}$ and Bob shares if $\theta_2\frac{L+\hat{\ell}(0)}{L+\ell}:=\theta_2\frac{L}{L+\ell}>1$, that is $s_{2,\ell}=\mathbf{Share}_\ell$ for each $\ell=1,...,\hat{\ell}(2)$. Therefore event $\mathrm{CSB}^2(R)$ implies that $\alpha_\ell=1$ for each $\ell=1,...\hat{\ell}(2)$, hence (by rationality) $s_1=Trust_\ell$ for some $\ell\geq\hat{\ell}(2)$. This shows that the claim holds for $k=1$.

Suppose by way of induction that the claim holds for some k. If $\hat{\ell}(2k)=L$, we are done. Let $\hat{\ell}(2k)<L$. Event $\mathrm{CSB}^{2k+1}(R)$ implies that Bob interprets any project $\ell>\hat{\ell}(2k)$ as a signal that $\alpha_\ell\geq\frac{L+\hat{\ell}(2k)}{L+\ell}$, hence it implies that $\beta_\ell\geq\frac{L+\hat{\ell}(2k)}{L+\ell}$ and that Bob shares if $\theta_2\frac{L+\hat{\ell}(2k)}{L+\ell}>1$, that is $s_{2,\ell}=\mathbf{Share}_\ell$ for each $\ell=1,...,\hat{\ell}(2(k+1))$. Event $\mathrm{CSB}^{2k+2}(R)=\mathrm{CSB}^{2(k+1)}(R)$ implies that $\alpha_\ell=1$ for each $\ell=1,...,\hat{\ell}(2(k+1))$, hence $s_1=Trust_\ell$ for some $\ell\geq\hat{\ell}(2(k+1))$. Therefore the claim holds for $k+1$.■

6 Multi-self players and sequential reciprocity

We show how Dufwenberg & Kirchsteiger's reciprocity theory can be represented within an extended framework where players have 'local' utility functions $(u_{i,h}: Z\times\mathbf{M}\times S\to\mathbb{R})_{h\in H\setminus Z}$. We have already seen how our basic framework could reproduce their reciprocity theory in an example (Γ_6 of DPG), but to handle general games one needs a multi-selves approach.[5] We consider two-player games for simplicity. Recall that player i is inclined to be kind toward j if she believes j is kind toward her. Kindness depends on intentions. In particular, the *kindness of j toward i*, K_{ji}, given j's first-order belief $\nu\in\Delta(S_i)$ and strategy s_j is increasing

[5]This is not to suggest that one could not conceive of a different sort of reciprocity theory, which would not require a multi-selves approach.

in the difference between the expected material payoff of i and a belief-dependent 'equitable payoff' $\pi^e_{ji}(\nu)$ that j ascribes to i:

$$K_{ji}(s_j,\nu)=\sum_{s'_i}\nu(s'_i)\pi_i(\zeta(s'_i,s_j))-\pi^e_{ji}(\nu).$$

A player's kindness toward the co-player depends on his current first-order belief, which depends on the observed history. Therefore, for any fixed hierarchy of cps' $\boldsymbol{\mu}_j=(((\mu^1_j(\cdot|h))_{h\in H},(\mu^2_j(\cdot|h))_{h\in H},...)$, the kindness of j toward i at history h is $K_{ji}(s_j,\mu^1_j(\cdot|h))$, where $s_j\in S_i(h)$. Assume that at each history h player i maximizes the expected value of a linear combination of her material payoff and the product between her kindness at h toward the opponent and the opponent's kindness at h toward her, *i.e.*, i at h maximizes

$$\int_{S_j(h)\times\Delta(S_i(h))}\left[\pi_i(\zeta(s_i,s_j))+\theta_iK_{ij}(s_i,\mu^1_i(\cdot|h))K_{ji}(s_j,\mu^1_j(\cdot|h))\right]\mu^2_i(ds_j,d\mu^1_j(\cdot|h)|h).$$

At history h, i's preferences are represented by the payoff function

$$u_{i,h}(z,\boldsymbol{\mu},s)=\pi_i(z)+\theta_iK_{ij}(s_i,\mu^1_i(\cdot|h))K_{ji}(s_j,\mu^1_j(\cdot|h)),$$

or equivalently by

$$\pi_i(z)+\theta_iK_{ij}(s_i,\mu^1_i(\cdot|h))\hat{K}_{iji}(\mu^2_i(\cdot|h)),$$

where $\hat{K}_{iji}(\mu^2_i(\cdot|h))=\int_{S_j(h)\times\Delta(S_i(h))}K_{ji}(s_j,\mu^1_j(\cdot|h))\mu^2_i(ds_j,d\mu^1_j(\cdot|h)|h)$ is i's belief in j's kindness toward i. What we have here is, essentially, a reformulation of Dufwenberg & Kirchsteiger's model.

References

[1] Battigalli, P. and M. Dufwenberg (2007): "Dynamic Psychological Games", mimeo, (November 2007).

[2] Battigalli, P. and M. Siniscalchi (1999): "Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games," *Journal of Economic Theory*, **88**, 188-230.

[3] Ben-Porath, E. and E. Dekel (1992): "Signaling Future Actions and the Potential for Sacrifice, " *Journal of Economic Theory*, 57, 36-51.

[4] DUFWENBERG, M. AND G. KIRCHSTEIGER (2004): "A Theory of Sequential Reciprocity," *Games and Economic Behavior,* **47**, 268-298.

[5] KUHN, H.W. (1953): "Extensive Games and the Problem of Information," in *Contributions to the Theory of Games II*, ed. by H. W. Kuhn and A. W. Tucker. Princeton: Princeton University Press, pp. 193-216.

[6] OSBORNE, M. AND A. RUBINSTEIN (1994): *A Course in Game Theory.* Cambridge MA: MIT Press.