WARWICK

ECONOMICS

CRETA

Centre for Research in Economic Theory and its Applications

Discussion Paper Series

Affective interdependence and welfare

Aviad Heifetz, Enrico Minelli & Herakles Polemarchakis

February 2023 (Revised January 2026) No: 76

Affective interdependence and welfare [1]

A. Heifetz [2] E. Minelli [3] H. Polemarchakis [4]

January 17, 2026

[1]We thank Marco Ottaviani, Dirk Bergemann and three referees for their excellent observations and insights that helped us improve the paper.

[2]Open Un. Israel, aviadhe@openu.ac.il
[3]Un. Brescia, enrico.minelli@unibs.it
[4]Un. Warwick, h.polemarchakis@warwick.ac.uk

Abstract

Purely affective interaction allows the welfare of an individual to depend only the individual's own action and on the profile of welfare levels of others. Under an assumption on the structure of mutual affection that we interpret as "non-reinforcing mutual affection," we show that equilibria of affective interaction are Pareto optimal. Moreover, *if* purely affective interaction induces a standard game, then an equilibrium profile of actions is a Nash equilibrium of the induced game, and this Nash equilibrium and Pareto optimal profile of strategies is locally dominant.

Key words: purely affective interactions, Pareto optimality.

JEL classification: D62.

1 Introduction

Purely affective interaction allows the welfare of an individual to depend only on the individual's own action and on the profile of welfare levels of others. Importantly, actions of others do not affect directly the welfare of an individual. Affection can be positive or negative.

At an *affective equilibrium*, each individual chooses her action taking the welfare levels of other individuals as given, and these welfare levels are realized. Our main result is that, if the matrix of mutual affection satisfies a condition introduced by Gale and Nikaido (1965) that, in our setting, corresponds to *non-reinforcing mutual affection* (NRMA), then affective equilibria are Pareto efficient.

NRMA does not guarantee the existence of an *induced game*, in which utilities only depend on the profile of actions. Nevertheless, *if* an induced standard game exists, an equilibrium profile of actions of the affective interaction is a Nash equilibrium of the game, and, moreover, at this Nash equilibrium each player chooses a locally dominant strategy. This is an extension, under different assumptions, of the striking result of Ray and Vohra (2020) that Nash equilibria of games induced by purely affective interaction (that they call 'games of love and hate') are Pareto efficient.

In the last section of the paper we discuss *economies* with affective interaction. Indeed, Arrow (1981) in a gift-giving interaction with more than two individuals and Pearce (2008) in intergenerational cake-eating games, had concluded that affective interactions, even of pure love, are in general *not* conducive to optimality. We resolve this conundrum by arguing that when inefficiency arises, dependence of individual welfare on actions of other individuals lurks in the background either via the game form, as in Arrow's gift giving game[1] or via aggregate feasibility constraints as in Pearce (2008).

In an extension of the present work to the sequential setting, Heifetz (2023) defined backward induction much more simply and directly than in Pearce (2008), and he showed that backward induction paths of actions and utility levels are Pareto optimal, again under the assumptions of non-reinforcing *purely* affective interaction.[2]

[1] Bourles, Brammoullé, and Perez-Richet (2017)

[2] Optimality results indicate that games induced by purely affective interaction form a non-generic class within the class of games: Nash equilibria of generic games are suboptimal in Dubey (1986), and, likewise, backward-induction paths of generic sequential games are suboptimal in Heifetz, Minelli, and Polemarchakis (2021).

2 Purely affective interaction

Individuals are $i \in I = \{1, ..., n\}$, profiles of action are

$$x = (x_1, ..., x_n) \in X = \prod_{i=1}^{n} X_i,$$

and profiles of utility levels are

$$u = (u_1, ..., u_n) \in \mathbb{R}^n.$$

Utility functions are

$$V(x, u) = (V_i(x_i, u_{-i}) : i = 1, ...n), \tag{1}$$

and we also write

$$V_x(u) = V(x, u).$$

In a *purely affective interaction*, an individual controls her action and is aware of the direct effect of her choice on her well-being. Other individuals' choices do not affect her directly, but she cares about the well-being of others, and her perception of others' well-being may affect her preferences over her actions.

A profile of actions and utility levels (x, u) is *consistent* if, for every individual, u_i corresponds to the utility level at (x_i, u_{-i}) or

$$V_x(u) = u.$$

An *affective equilibrium* is profile of actions and utility levels, (x^*, u^*) such that

i) every individual maximizes V_i taking u^*_{-i} as given: [3]

$$V_x(u^*) \leq V_{x^*}(u^*), \quad and$$

ii) actions and utility levels are consistent:

$$V_{x^*}(u^*) = u^*.$$

[3] We employ the standard notation $u \leq \bar{u}$ for $u_i \leq \bar{u}_i$, $i = 1, .., n$ and $u < \bar{u}$ for $u_i \leq \bar{u}_i$, $i = 1, .., n$ with at least one strict inequality.

This definition captures the idea that, in a situation of purely affective interaction, an individual need not have a detailed knowledge of the others' network of affections, nor a complete understanding of the repercussions of her choice through the whole social network. Each individual takes an action based on her *perception* of the levels of well-being of the individuals she cares about, and, at equilibrium, these perceptions are confirmed. In other words, consistency between actions and utility levels is imposed *only* at equilibrium.

We make two assumptions on the structure of affective interaction.

Assumption 1. For every individual, X_i is an open subset of Euclidean space, and the utility function $V_i(\cdot, \cdot)$ is continuously differentiable.

The Jacobian of V_x at u is $J_x(u)$.

A square matrix is a *P-matrix* if all its principal minors are positive[4].

Assumption 2. For every $x \in X$ and $u \in \mathbb{R}^n$, the matrix $(I - J_x(u))$ is a P-matrix.

This is an assumption of *non-reinforcing mutual affection* (NRMA). We first illustrate in a simple example and then we discuss it for the general case.

Example 1. Consider the two person purely affective interaction where

$$V_1(x, u_2) = f(x) + a u_2,$$
$$V_2(y, u_1) = g(y) + b u_1,$$

For each (x, y) the Jacobian of the map $V_{(x,y)} : \mathbb{R}^2 \to \mathbb{R}^2$ is

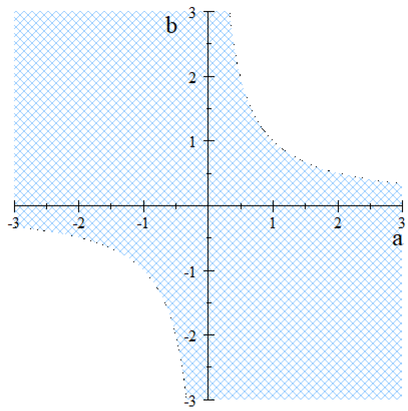
$$J_{(x,y)} = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}.$$

The matrix $(I - J_{(x,y)})$ has a unitary diagonal and determinant $\det(I - J_{(x,y)}) = 1 - ab$. It is a P-matrix if and only if $\det(I - J_{(x,y)})) > 0$ or $ab < 1$.

The two individuals can have positive or negative feelings towards each other, and these feelings may even be strong, but, if feelings go in the same direction, they both love or they both hate each other, they cannot be *too* strong – Figure 1.

[4]A principal minor is obtained by the elimination of rows and corresponding columns, but, importantly, without transpositions of rows or columns prior to elimination.

Figure 1



□

To interpret Assumption 2 in the general case, we recall a useful characterization.

Gale-Nikaido Lemma [Gale and Nikaido (1965), Theorem 2] *A matrix A is a P-matrix if and only if, for any non-zero $y \in \mathbb{R}^n$, there exists $i \in \{1, 2, \ldots n\}$, such that $y_i(Ay)_i > 0$.*

In words, P-matrices do not fully reverse the sign of any non-zero vector.

The Gale-Nikaido characterization allows us to write Assumption 2 as follows:

For every x and u and all $\Delta u \neq 0$, either there exists an i, such that

$$\Delta u_i > 0 \quad \text{and} \quad \Delta u_i > \sum_{j \neq i} \frac{\partial V_{x,i}(x_i, u)}{\partial u_j} \Delta u_j$$

or

$$\Delta u_i < 0 \quad \text{and} \quad \Delta u_i < \sum_{j \neq i} \frac{\partial V_{x,i}(x_i, u)}{\partial u_j} \Delta u_j.$$

This says that, if we start from a consistent pair $(\hat{x}, \hat{u})$,

$$\hat{u} - V_{\hat{x}}(\hat{u}) = 0,$$

then, for any exogenous change in the perceptions of utility levels, $\Delta u = u - \hat{u} \neq 0$, there is one individual i, for whom

$$u_i > \hat{u}_i, \quad \text{and} \quad V_i(\hat{x}, u) < u_i$$

or

$$u_i < \hat{u}_i \quad \text{and} \quad V_i(\hat{x}, u) > u_i.$$

That is, starting from a consistent pair of actions and utility levels, for any exogenous change in the perceptions of utility levels, there is always one individual whose *realized* utility does not reinforce the change in perceptions. Assumption 2 allows for a wide array of positive and negative individual feelings about changes in the well-being of (any subset) of other individuals, but it prevents mutually reinforcing loops of perceptions and realizations of well-being.

A benevolent and knowledgeable planner able to 'see through' the whole network of affective interactions may induce simultaneous changes in individual actions. Can this power and knowledge allow for the implementation of a consistent profile of actions and utility levels making everybody better-off with respect to the affective equilibrium?

A profile of actions and utility levels, $(\tilde{x}, \tilde{u})$ is a *Pareto improvement* over a profile (x, u) if

$$\tilde{u} > u.$$

A consistent profile of actions and utility levels, (x, u) is *Pareto optimal* if it does not permit a consistent Pareto improvement.

Theorem 1. *Under Assumptions* 1 *and* 2*, if* (x^*, u^*) *is an affective equilibrium, then it is Pareto optimal.*

Proof. Suppose, by way of contradiction, that $(\tilde{x}, \tilde{u})$ Pareto improves on (x^*, u^*) or

$$V_{\tilde{x}}(\tilde{u}) = \tilde{u} > u^*. \tag{2}$$

Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be defined by

$$F(u) = u - V_{\tilde{x}}(u).$$

Since (x^*, u^*) is an affective equilibrium,

$$F(u^*) = u^* - V_{\tilde{x}}(u^*) \geq 0,$$

and, since $(\tilde{x}, \tilde{u})$ is a consistent profile of actions and utility levels,

$$F(\tilde{u}) = \tilde{u} - V_{\tilde{x}}(\tilde{u}) = 0 \leq F(u^*). \tag{3}$$

By Assumptions 1 and 2, the matrix $(I - J_{\tilde{x}}(u))$, the Jacobian of F, is a P-matrix and, by Theorem 3 of Gale and Nikaido (1965), the inequalities (2) and (3) cannot obtain simultaneously for $\tilde{u} \neq u^*$. □

We end this section with a brief discussion of the issue of existence of an affective equilibrium. Our goal is not to prove a general theorem but rather to discuss one set of sufficient conditions for existence and illustrate in a simple example how they relate to the condition that we identified for efficency, NMRA.

Assume X_i is compact (rather than open) $\forall i \in I$. For every $u \in \mathbb{R}^n$ and every $i \in I$ choose a maximizer

$$\hat{x}_i(u_{-i}) \in \arg\max_{X_i} V_i\left(\cdot, u_{-i}\right)$$

which exists by Assumption 1 and the compactness of X_i. Define the function $H : \mathbb{R}^n \to \mathbb{R}^n$ by

$$H(u) = V_{\hat{x}(u)}(u)$$

where $\hat{x}(u) = (\hat{x}_i(u_{-i}))_{i\in I}$.[5]

Assume that for some integer $k \geq 1$ the function

$$H^k : \mathbb{R}^n \to \mathbb{R}^n$$

is a contraction. Then H has a unique fixed point u^* (see e.g. Pata (2019), Corollary 1.3), and $(\hat{x}(u^*), u^*)$ is an affective equilibrium, because (i) it is consistent,

$$u^* = H(u^*) = V_{\hat{x}(u^*)}(u^*)$$

and (ii) each individual i is best responding to u^*_{-i}.

For instance, consider again Example 1 above. Due to separability, we have

$$\begin{aligned} \hat{x}_1\left(u_2\right) &\in \arg\max_{X_1} f\left(\cdot\right) \\ \hat{x}_2\left(u_1\right) &\in \arg\max_{X_2} g\left(\cdot\right) \end{aligned}$$

and we can choose these maximizers $\hat{x}_1$, $\hat{x}_2$ to be independent of u_1, u_2. Then for $k = 2$ we have

$$\begin{aligned} H_1^2\left(u\right) &= V_1^2\left(\hat{x}_1, V_2(\hat{x}_2, u_1)\right) = f\left(\hat{x}_1\right) + a[g\left(\hat{x}_2\right) + bu_1] \\ H_2^2\left(u\right) &= V_2^2\left(\hat{x}_2, V_1(\hat{x}_1, u_2)\right) = g\left(\hat{x}_2\right) + b[f\left(\hat{x}_1\right) + au_2] \end{aligned}$$

[5] We note that the definition of H is independent of the choice of the particular maximizers $\hat{x}_i(u_{-i})$ of $i \in I$: If there are several such maximizers for individual i then they all yield i the same utility level, and others care only about i's utility level, not about i's particular choice.

The Jacobian of $H^2(u)$ is therefore

$$DH^2(u) = \begin{pmatrix} ab & 0 \\ 0 & ab \end{pmatrix}.$$

so H^2 is a contraction if $|ab| < 1$, a stronger condition than NMRA (our assumption 2), which is $ab < 1$ in this example.[6]

3 Induced games

The literature on *interdependent preferences*, that we discuss in Appendix 1, has identified conditions for consistency to hold at every profile of actions and utility levels, so that an *induced game* $U : X \to \mathbb{R}^n$ exists[7]:

$$U(x) = V_x(U(x)).$$

Our main interest lies in affective interactions as described by (1), independently of whether an induced standard game exists or not. A simple example illustrates why, even when it is well defined, the induced game may not be the most natural representation of a purely affective interaction.[8].

Example 2. Consider the situation of affective interaction described by:

$$\begin{aligned} V_1(x_1, u) &= f(x_1) + au_3, \\ V_2(x_2, u) &= g(x_2) + bu_1 \\ V_3(x_3, u) &= h(x_3) + cu_1. \end{aligned}$$

Individuals 1 and 3 are linked by mutual affection, while individual 2 cares about individual's 1 welfare but may not even be aware of the existence of individual 3.

[6] To prove equilibrium existence, Ray and Vohra (2020) relied on a boundedness condition implying in this example *both* $|a| < 1$ and $|b| < 1$ (see footnote 11 below) – a condition which is therefore even stronger than the assumption $|ab| < 1$ on which we rely here.

[7] This literature studies mostly the case of altruism. That is, positive affective interaction.

[8] A similar example was used by Pearce (2008) to argue that, in a sequential setting, the subgame perfect equilibria of the induced game are the wrong tool to analyze affective interactions.

In this example, an induced game is well defined and is easily seen to be:

$$\begin{aligned}
U_1(x_1, x_2, x_3) &= \frac{f(x_1) + ah(x_3)}{1 - ac}, \\
U_2(x_1, x_2, x_3) &= \frac{g(x_2)(1 - ac) + bf(x_1) + abh(x_3)}{1 - ac} \\
U_3(x_1, x_2, x_3) &= \frac{h(x_3) + cf(x_1)}{1 - ac}.
\end{aligned}$$

In the induced game, individual 2 is represented as 'seeing through' the utility and network of friendship of individual 1 and calculating the indirect effect of a change in the action of individual 3 on her own utility, even if, in the original affective interaction she may not even be aware of individual 3's existence. □

With the above caveat, we note that by our Assumption 2, Theorem 4 of Gale and Nikaido (1965) implies that for every $x \in X$ there is at most one solution to the system of equations:

$$F_x(u) = u - V_x(u) = 0 \tag{4}$$

That is, there is at most one profile of utility levels $U(x) = u_x$ consistent with the choice profile x. Pearce (2008), Bergstrom (1999) and Hori (2001) (propositions 3 and 4) proposed stronger assumptions that guarantee the existence of such a solution to (4) for *every* $x \in X$. Under such stronger assumptions[9], the solutions $U(x) = u_x$ of (4) for $x \in X$ therefore define a utility function $U_i : X \to \mathbb{R}$ for every individual i, that is a standard game induced by the fundamental affective interaction.

Example 3 below shows that without such stronger assumptions, a solution to (4) might not exist for some $x \in X$, so the induced game may be defined only on some subset of X.

Example 3. Consider the choice sets $X_1 = X_2 = (-1, 1)$ and the separable affective interaction on $X = X_1 \times X_2$

$$\begin{aligned}
V_1(x_1, u_2) &= \begin{cases} -x_1^2 + u_2 + e^{-u_2} & u_2 \geq 0 \\ -x_1^2 + u_2 + 2 - e^{u_2} & u_2 < 0 \end{cases} \\
V_2(x_2, u_1) &= \begin{cases} -x_2^2 + u_1 + e^{-u_1} & u_1 \geq 0 \\ -x_2^2 + u_1 + 2 - e^{u_1} & u_2 < 0 \end{cases}
\end{aligned}$$

[9] See Appendix 1

Then V_1, V_2 are continuously differentiable, with

$$\frac{\partial V_i}{\partial u_j} = \begin{cases} 1 - e^{-u_j} & u_j \geq 0 \\ 1 - e^{u_j} & u_j < 0 \end{cases}$$

and in particular for every $u_j \in \mathbb{R}$

$$0 \leq \frac{\partial V_i}{\partial u_j} < 1$$

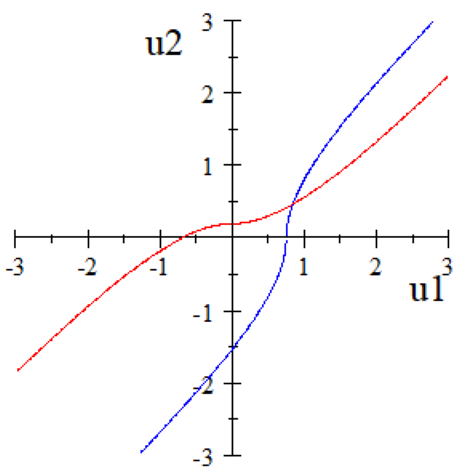
so in this example not only $I - J$ is a P-matrix, but it also has a dominant diagonal, and the Jacobian J has spectral radius smaller than 1 (see the Appendix for a discussion of these stronger requirements); and both individuals are altruistic.

For every $(x_1, x_2) \in X$ except for $(x_1, x_2) = (0, 0)$ there is a unique solution to the system of equations

$$u_1 = \begin{cases} -x_1^2 + u_2 + e^{-u_2} & u_2 \geq 0 \\ -x_1^2 + u_2 + 2 - e^{u_2} & u_2 < 0 \end{cases}$$
$$u_2 = \begin{cases} -x_2^2 + u_1 + e^{-u_1} & u_1 \geq 0 \\ -x_2^2 + u_1 + 2 - e^{u_1} & u_2 < 0 \end{cases}$$

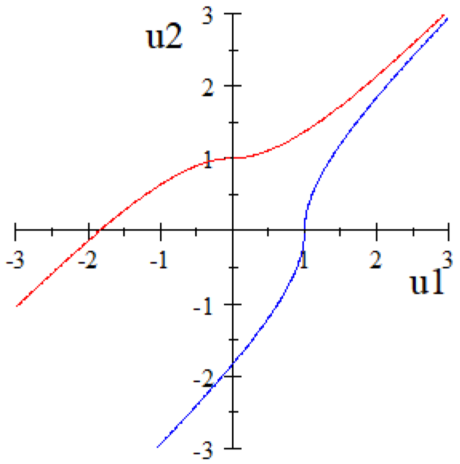
For example, for $x_1 = 0.9$, $x_2 = 0.5$, the solution is at the intersection of the two curves in Figure 2:

Figure 2



However, for $x_1 = x_2 = 0$ the "solution" is $u_1 = u_2 = \infty$, and there is no intersection:

Figure 3



□

In any affective interaction, we may then consider a subset of action profiles, $\bar{X} = \times_{i \in I} \bar{X}_i \subset X$ such that for every profile of actions $x \in \bar{X}$ a unique solution u_x to (4) exists, and define, for any $x \in \bar{X}$, $U(x) = u_x$. We call this game, $U : \bar{X} \to \mathbb{R}^n$, *an* induced game of the affective interaction, and ask what is the relationship between affective equilibria of the purely affective interaction and Nash equilibria of this induced game.

In this Section we explore this question (Theorems 2 and 3) and we compare our setting to the one discussed by Ray and Vohra (2020) (Theorem 4). We also prove an interesting property of Nash equilibria of the induced game, which highlights their special structure (Theorem 5).

Theorem 2. *Under Assumptions* 1 *and* 2*, if* (x^*, u^*) *is an affective equilibrium, and* $x^* \in \bar{X}$*, then* x^* *is a Nash equilibrium of the induced game* $U : \bar{X} \to \mathbb{R}^n$.

Proof. Suppose, by way of contradiction, that x^* is not a Nash equilibrium of the induced game $U : \bar{X} \to \mathbb{R}^n$. Then, for some individual i, there exists an alternative choice $\tilde{x}_i \in X_i$ for which the induced game is defined at

$$x = \left(\tilde{x}_i, x^*_{-i}\right),$$

and

$$\tilde{u}_i = U_i(x) > U_i(x^*) = u_i^*$$

Let

$$\tilde{u} = U(x) = (V_j(x_j, \tilde{u}_{-j}))_{j=1}^n,$$

and let $F : \mathbb{R}^n \to \mathbb{R}^n$ be defined by

$$F(u) = u - V_x(u).$$

By the definition of $\tilde{u}$,

$$F(\tilde{u}) = 0.$$

Also, for $j \neq i$, since $x_j = x_j^*$,

$$F_j(u^*) = 0.$$

At the same time, since (x^*, u^*) is an affective equilibrium and $x_i \neq x_i^*$,

$$F_i(u^*) \geq 0.$$

Altogether,

$$(u_k^* - \tilde{u}_k)(F_k(u^*) - F_k(\tilde{u})) \leq 0, \qquad k = 1, ..., n.$$

However, since by Assumption 2 the Jacobian $I - J_x(u)$ of $F(u)$ is a P-matrix for every $u \in \mathbb{R}^n$, by theorem 20.5 in Nikaido (1968), this set of inequalities cannot obtain for $\tilde{u} \neq u^*$.

□

A converse is also true:

Theorem 3. *Under Assumptions* 1 *and* 2*, if* $x^* \in \bar{X}$ *is a Nash equilibrium of the induced game* $U : \bar{X} \to \mathbb{R}^n$*, then* $(x^*, U(x^*)$ *it is an affective equilibrium of the affective interaction where choice sets are restricted to* $\bar{X}_i$ *for all* $i \in I$*.*

Proof. Denote $u^* = U(x^*)$ and suppose, by way of contradiction, that (x^*, u^*) is not an affective equilibrium. Then, for some individual i, there exists an alternative choice $\hat{x}_i \in \bar{X}$ with

$$\hat{u}_i = V_i\left(\hat{x}_i, u_{-i}^*\right) > V_i\left(x_i^*, u_{-i}^*\right) = u_i^*.$$

By assumption, the induced game is defined at

$$x = \left(\hat{x}_i, x_{-i}^*\right).$$

Let

$$\hat{u} = U(x) = \left(V_j(x_j, \hat{u}_{-j})\right)_{j=1}^{n}$$

and let $F : \mathbb{R}^n \to \mathbb{R}^n$ be defined by

$$F(u) = u - V_x(u).$$

By the definition of $\hat{u}$,

$$F(\hat{u}) = 0.$$

For $j \neq i$, since $x_j = x_j^*$,

$$F_j(u^*) = 0.$$

At the same time, since x^* is a Nash equilibrium, and $x_i \neq x_i^*$,

$$F_i(u^*) \geq 0.$$

Altogether,

$$\left(u_k^* - \hat{u}_k\right)\left(F_k(u^*) - F_k(\hat{u})\right) \leq 0, \qquad k = 1, ..., n.$$

However, since by Assumption 2 the Jacobian $I - J_x(u)$ of $F(u)$ is a P-matrix for every $u \in \mathbb{R}^n$, by theorem 20.5 in Nikaido (1968) this set of inequalities cannot obtain for $\hat{u} \neq u^*$. □

Of course the interest of Theorems 2 and 3 depends on the relationship between X and $\bar{X}$. If $\bar{X} = x^*$, they are true but trivial. More interestingly, from Theorem 3, we obtain:

Theorem 4. *Suppose an induced game U is uniquely defined everywhere on X (i.e. $\bar{X} = X$). Under Assumptions 1 and 2, if x^* is a Nash equilibrium of the induced game U, then it is Pareto optimal.*

Comparison with Ray and Vohra (2020)

Ray and Vohra (2020), in the paper that motivated our research, gave a different proof of Theorem 4. They restrict attention to 'coherent' purely affective interactions that satisfy the following two assumptions:

Assumption RV1 For every $x \in X$, the map $V(x, \cdot) : \mathbb{R}^n \to \mathbb{R}^n$ has a unique fixed point.

Assumption RV2 For every $x \in X$, there exists $B(x) < \infty$ such that, if $\| u \| > B(x)$ then $\| V_x(u) \| < \| u \|$.

Under these two assumptions[10] they prove the analog of Theorem 4 (their main Theorem: *Nash equilibria of games induced by coherent purely affective interactions are Pareto efficient*). In our proof of Theorem 4, we do make the assumption that the induced game is uniquely defined everywhere, i.e. Assumption RV1, but, differently from Ray and Vohra, we assume differentiability of $V(x, u)$ and we replace Assumption RV2 with Assumption 2 . In the smooth differentiable situations that we study, Assumption 2 *does not* imply Ray and Vohra's Assumption RV2. To see this, let us go back to the two person linearly separable affective interaction described in Example 1. In that example, Assumption 2 amounts to $ab < 1$, while Ray and Vohra's boundedness condition RV2 implies the stronger restrictions that $\mid a \mid < 1$ and $\mid b \mid < 1$.[11]

In our main result (Theorem 1), we show that a meaningful condition on mutual affection (Assumption 2, NMRA), is *sufficient for the Pareto efficiency of affective equilibrium, independently of whether an induced game is defined everywhere or not.*

Equilibria of induced games and local dominance

Our differentiable framework also allows for an interesting new result on the nature of Nash equilibria in games induced by purely affective interactions: at a Nash equilibrium of such a game, every player is choosing a strategy that is *locally* dominant, that is marginal changes in the other individuals' strategies do not induce any change in an individual's equilibrium choice.

[10]The assumptions are imposed also on every affective sub-interaction,with the payoff of a subset of players held fixed. See their paper for details.

[11]If $|a| > 1$ Ray and Vohra (2020) boundedness condition is not satisfied: for every (x, y) and every function $B(x, y) < \infty$, whenever $|u_2| > B(x, y) + |f(x)|$ we have

$$|u_2| - |f(x)| > B(x, y)$$

and since

$$|u_2| < |a| |u_2| = |au_2| = |(f(x) + au_2) - f(x)| \leq |f(x) + au_2| + |f(x)|$$

implies

$$|f(x) + au_2| > |u_2| - |f(x)|$$

with the sup norm $\|\cdot\|$

$$\|V((x, y), (u_1, u_2))\| = \sup(|f(x) + au_2|, |g(y + bu_1)|) \geq |f(x) + au_2| >$$

$$|u_2| - |f(x)| > B(x, y)$$

This property shed light on the very special nature of games induced by purely affective interactions.

Consider the function $F_x(u) = u - V(x,u)$ defined in equation 4. By Assumption 2, at every (x, u_x) such that $F_x(u_x) = 0$, it is the case that $\det(I - J_x(u_x)) \neq 0$, and we can apply the implicit function theorem at (x, u_x) to obtain the existence of smooth real-valued utility functions $U_x(\cdot) = (U_i(\cdot))_{i=1,\dots,n}$ defined on some rectangular neighborhood $\mathcal{O}_x = \prod_{i\in I} \bar{X}_i$ of x with

$$U_x(x) = u_x.$$

This defines an induced game $U : \bar{X} \to \mathbb{R}^n$ with $\bar{X} = \mathcal{O}_x$, and

$$\frac{\partial U_i(x)}{\partial x_j} = \frac{\partial V_j(x_j, (u_x)_{-j})}{\partial x_j} \left((I - J_x(u_x))^{-1}\right)_{ij}. \tag{5}$$

Proposition 1. *At a Nash equilibrium, x^* of the induced game U, the utility function of every individual is locally flat.*

Proof. From theorem 3, (x^*, u_{x^*}) is an affective equilibrium of the original affective interaction, and therefore, for all $j \in I$,

$$\frac{\partial V_j\left(x_j^*, (u_{x^*})_{-j}\right)}{\partial x_j} = 0. \tag{6}$$

Substituting (6) into (5) yields

$$\frac{\partial U_i(x^*)}{\partial x_j} = 0.$$

□

Theorem 5. *At a Nash equilibrium x^* of the induced game U, each individual's strategy is locally dominant.*

Proof. By Theorem 3, if x^* is a Nash equilibrium of an induced game, $(x^*, u^* = U(x^*))$ is an affective equilibrium, and, for each individual,

$$\frac{\partial V_i\left(x_i^*, u_{-i}^*\right)}{\partial x_i} = 0, \tag{7}$$

and therefore, by Proposition 1, the induced utility function U_i is flat as a function of x_i at x^*. Moreover, for $j \neq i$, it follows, again from (5), that

$$\frac{\partial U_i(x^*)}{\partial x_i \partial x_j} =$$

$$\frac{\partial\left(\frac{\partial U_i(x^*)}{\partial x_i}\right)}{\partial x_j} = \frac{\partial\left(\left(\frac{\partial V_i\left(x_i^*, u_{-i}^*\right)}{\partial x_i}\right)\left((I - J_{x^*}(u^*))^{-1}\right)_{ii}\right)}{\partial x_j} =$$

$$\frac{\partial V_i\left(x_i^*, u_{-i}^*\right)}{\partial x_i} \frac{\partial\left(\left((I - J_{x^*}(u^*))^{-1}\right)_{ii}\right)}{\partial x_j} + \frac{\partial V_i\left(x_i^*, u_{-i}^*\right)}{\partial x_i \partial x_j}\left((I - J_{x^*}(u^*))^{-1}\right)_{ii} = 0,$$

where the last equality is due to the first order condition (7) coupled with the fact that

$$\frac{\partial V_i\left(x_i^*, u_{-i}^*\right)}{\partial x_i \partial x_j} = 0$$

since V_i does not depend on x_j. □

Examples in Appendix 2 illustrate the results of this section.

4 Linearly separable affection

In *linearly separable purely affective interactions*, like those introduced in the Examples 1 and 2 above, the individuals' utility functions have the form

$$V_i(x_i, u) = f_i(x_i) + \sum_{j \neq i} a_{ij} u_j. \tag{8}$$

At every $x = (x_1, ..., x_n)$ the Jacobian J of V with respect to $u = (u_1, ..., u_n)$ has a zero diagonal $J_{ii} = 0$ and off-diagonal entries $J_{ij} = a_{ij}$.

Consistency in this special case takes the form

$$u = f(x) + Ju. \tag{9}$$

Under Assumption 2, we have that, in particular, $\det(I - J) \neq 0$, so we can uniquely solve the system of equations (9) at every x, thus obtaining the *induced game*:

$$U(x) = (I - J)^{-1} f(x) \equiv B f(x). \tag{10}$$

The utility functions $U = (U_1, ..., U_n)$ in the induced game are linear combinations of the 'base utilities' $f = (f_1, ... f_n)$. The matrix $B = (I - J)^{-1}$ summarizes the effect of changes in the base utilities $f(x)$ on the final well-being of individuals, taking into account the network of affective interactions between them. The simplicity of the structure of U in the case of linearly separable affective interaction allows us to i) illustrate the logic for the argument for Pareto efficiency of affective equilibria and ii) prove a useful further property of affective equilibria, again implied by Assumption 2 (see Proposition 2 below).

Indeed, in the case of linearly separable affection, a Pareto improvement generates a Δf such that $\Delta U = B\Delta f > 0$. If x^* is an affective equilibrium and $y = \Delta f$ a Pareto improvement on it, we have that $By > 0$. But then, by the Gale-Nikaido Lemma quoted above, there must exist i such that $y^i = \Delta f^i > 0$, a contradiction with x^* being an affective equilibrium.

To gain further understanding of the efficiency properties of affective equilibria, let us consider, for any given $\lambda \in \mathbb{R}^n_+ \setminus \{0\}$, the welfare function

$$W_\lambda(x) = \lambda U(x).$$

Given that $\lambda U(x) = \lambda B f(x)$, a social planner maximizing $W_\lambda(x)$ chooses a profile of actions x to maximize a linear combination of the 'base utilities' $f_i(x_i)$ with *social weights* λB reflecting the network of interactions. The maximization of each $f_i(x_i)$ at an affective equilibrium assures that the first order conditions for the maximization of $W_\lambda(x)$ are satisfied. To discuss second order conditions, let us assume concavity of base utilities.

Assumption 3. For every i, the action set X_i is convex and the function f_i is concave in x_i.

Even under Assumption 3, W_λ need not be concave in x, because Assumption 2 does not imply that the elements of B outside the diagonal are non-negative. Still, using Farkas's Lemma, Gale and Nikaido (1965) (Corollary 2) prove that

$$B \text{ is a P-matrix} \quad \Longrightarrow \exists\, \lambda \in \mathbb{R}^n_{++}, \;\; s.t. \;\; \lambda B >> 0.$$

Therefore, under Assumptions 2 and 3, *there exist* welfare weights λ such that, *for those weights*, W_λ is a sum of concave functions and therefore concave, leading to an alternative proof of Pareto optimality of affective equilibria:

Proposition 2. *Under Assumptions 1, 2 and 3, in a situation of linearly separable affection, if x^* is an affective equilibrium, there exists $\lambda \in \mathbb{R}^n_{++}$ such that x^* is a global maximum of W_λ.*

5 Affection and Wealth

Pure affection is a particular kind of externality: When individuals choose they do not take into account how a change in their choice, which would change their own utility, would also influence the wellbeing of others who care about that utility. Nevertheless, as we saw in Theorem 1, under the NRMA condition the affective equilibrium choice profile is efficient despite the presence of this externality.

An analogous phenomenon occurs in competitive exchange economies. When individuals choose they do not take into account how their own demand limits the feasible demand by others, and thus limits others' wellbeing. Nevertheless, by the first welfare theorem, the competitive equilibrium choice profile is efficient.

What happens when these two special externalities operate simultaneously? That is, what happens when each individual's choice influences both what the others can choose via an aggregate feasibility constraint, and also influences the others' wellbeing via their affection towards the individual's own wellbeing?

The following simple family of examples shows that in the presence of these two externalities together, both efficiency and inefficiency may be robust phenomena.

Example 4. Consider an economy with two individuals, $i = 1, 2$. There is a unique consumption good of total quantity 1, and no possibility of trade. Let $x_i \in (0, 1)$ be the share of individual i, where $x_1 + x_2 = 1$. The individuals' utility functions are

$$\begin{aligned} V_1(x_1, u_2) &= \log x_1 + a_{12}u_2 \\ V_2(x_2, u_1) &= \log x_2 + a_{21}u_1 \end{aligned}$$

with $a_{12}a_{21} < 1$ (i.e., Assumption 2 holds).

We note that unlike in the setting of Theorem 1, the aggregate feasibility constraint imposes a joint constraint on individual action sets. Under this aggregate feasibility constraint, can a social planner find an alternative division (x_1', x_2') such that $x_1' + x_2' = 1$ and which would make both individuals happier?

Substituting $u_1 = V_1(x_1, u_2)$, $u_2 = V_2(x_2, u_1)$ and then solving yields

$$u_i = \frac{\log x_i + a_{i,-i} \log(1 - x_i)}{1 - a_{12}a_{21}} \quad i = 1, 2$$

Each individual $i = 1, 2$ would have been happy to marginally increase her consumption at the expense of the other if

$$\frac{du_i}{dx_i} = \left(\frac{1}{1 - a_{12}a_{21}}\right)\left(\frac{1}{x_i} - \frac{a_{i,-i}}{1 - x_i}\right) > 0,$$

that is, if

$$a_{i,-i} < \frac{1 - x_i}{x_i} = \frac{x_{-i}}{1 - x_{-i}}$$

and would be happy to marginally reduce her consumption in favor of the other if the reverse inequality holds.

A social planner can therefore find a Pareto improving allocation (x_1', x_2') adhering to the aggregate feasibility constraint $x_1' + x_2' = 1$ when the parameters of the economy are in the open set

$$\left\{(x_1, a_{12}, a_{21}) \in (0,1) \times \mathbb{R}^2 : a_{12} > \frac{1 - x_1}{x_1}, a_{21} < \frac{x_1}{1 - x_1}, a_{12}a_{21} < 1\right\}$$

or

$$\left\{(x_1, a_{12}, a_{21}) \in (0,1) \times \mathbb{R}^2 : a_{12} < \frac{1 - x_1}{x_1}, a_{21} > \frac{x_1}{1 - x_1}, a_{12}a_{21} < 1\right\}$$

In contrast, such a Pareto improvement does not exist when

$$\left\{(x_1, a_{12}, a_{21}) \in (0,1) \times \mathbb{R}^2 : a_{12} < \frac{1 - x_1}{x_1}, a_{21} < \frac{x_1}{1 - x_1}, a_{12}a_{21} < 1\right\} \quad (11)$$

Both efficiency and inefficiency are robust phenomena in this family of examples, each obtaining for an open set of economies verifying Assumption 2. □

More generally, consider the case of *competitive economies with linearly separable affection.* As in section 4, we assume that the utility functions have the form

$$V_i(x_i, u) = f_i(x_i) + \sum_{j \neq i} a_{ij}u_j. \quad (12)$$

where now $x_i \in \mathbb{R}_+^L$ is the consumption bundle of individual i.

If the 'base utilities' $f_i(x_i)$ are concave in consumption, Assumption 3, and $e_i \in \mathbb{R}_+^L$ is the initial endowment of individual i, an *affective competitive equilibrium* is a triple (x^*, u^*, p^*) satisfying:

i) the first order conditions for utility maximization at prices p^*, for all l and all i

$$\frac{\partial f_i}{\partial x_{il}}(x_i^*) = \alpha_i^* p_l^*$$

where α_i^* is the marginal utility of revenue of individual i at equilibrium.

ii) feasibility

$$\sum_i x_i^* \leq \sum_i e_i$$

iii) consistency

$$V_{x^*}(u^*) = f(x^*) + Ju^* = u^*.$$

As in section 4, under Assumption 2, $\det(I - J) \neq 0$, so we can define the *induced utilities* :

$$U(x) = (I - J)^{-1} f(x) \equiv Bf(x). \tag{13}$$

The planner problem, is then, for any given $\lambda \in \mathbb{R}_+^n \setminus \{0\}$, to choose a consumption allocation x to maximize the welfare function

$$W_\lambda(x) = \lambda U(x) = \lambda B f(x)$$

under the feasibility constraint that, for all l:

$$\sum_i x_{il} \leq \sum_i e_{il}$$

If we let γ_l be the Lagrange multiplier of the l-th feasibility constraint, the first order conditions for the planner problem are, using (13), that, for all k and all l:

$$\frac{\partial \sum_i \lambda_i U_i(x)}{\partial x_{kl}} = [\sum_j \lambda_j b_{jk}] \frac{\partial f_k}{\partial x_{kl}}] = \gamma_l \tag{14}$$

We can then prove:

Proposition 3. *Under Assumptions 1, 2 and 3, in an economy with linearly separable affection, if i) x^* is an affective equilibrium, ii) α^* the equilibrium vector of marginal utilities of revenue, and, iii) for all i:*

$$\sum_{j \neq i} a_{ji} \frac{\alpha_i^*}{\alpha_j^*} < 1 \tag{15}$$

then x^* is Pareto efficient.

Proof. At the equilibrium allocation x^*, for all k and l:

$$\frac{\partial f_k}{\partial x_{kl}}(x_k^*) = \alpha_k^* p_l^*$$

Using (14), if, for all k:

$$\sum_j \lambda_j b_{jk} = \frac{1}{\alpha_k^*} \tag{16}$$

and, for all l:

$$\gamma_l^* = p_l^*$$

the equilibrium allocation x^* satisfies the first order conditions of the planner problem.

Write the equations in (16) in matrix form as:

$$\lambda B = [\alpha^*]^{-1}$$

where $[\alpha^*]^{-1}$ is a row vector with elements $(1/\alpha_k^*)$, $k = 1, \dots n$. Under Assumption 2, the matrix B is invertible and this system of equations has a unique solution:

$$\lambda^* = [\alpha^*]^{-1} B^{-1} = [\alpha^*]^{-1}[I - J] \tag{17}$$

If, for all i, we have $\lambda_i^* > 0$, then the welfare function W_λ is concave, as a sum of concave function (Assumption 3), and the second order conditions of the planner problem are also satisfied.

Thus, using (17), a sufficient condition condition for an affective equilibrium to be Pareto efficient is that, for all i:

$$\lambda_i^* = \frac{1}{\alpha_i^*} - \sum_{j \neq i} a_{ji} \frac{1}{\alpha_j^*} > 0 \tag{18}$$

which holds by (15). □

Condition (15) generalizes condition (11) of Example 4 above. Indeed, with a unique consumption good, logarithmic utilities for both individuals and $a_{12}a_{21} < 1$, the marginal utility of revenue of individual $i = 1, 2$ is

$$\alpha_i^* = \frac{1}{x_i} = \frac{1}{1 - x_{-i}}$$

so that equation (15), when applied to the example, is

$$a_{21}\frac{\frac{1}{x_1}}{\frac{1}{1-x_1}} < 1, \quad a_{12}\frac{\frac{1}{1-x_1}}{\frac{1}{x_1}} < 1$$

that is, condition (11).

With many goods, many individuals, and general consumption utilities, equation (15) is satisfied if, at equilibrium, the monetary value of the aggregate indirect effect, on all $j \neq i$, of a marginal increase in i's utility is less than the direct effect. Under this condition, the competitive equilbrium is Pareto efficient and it maximizes a welfare function W_{λ^*} where the implicit welfare weight of individual i is defined in (18).

References

K. J. Arrow. Optimal and voluntary income distribution. In S. Rosefielde, editor, *Economic Welfare and the Economics of Soviet Socialism:Essays in Honor of Abram Bergson*, pages 267–288. Cambridge University Press, 1981.

T. C. Bergstrom. Systems of benevolent utility functions. *Journal of Public Economic Theory*, 1(1):71–100, 1999.

R. Bourles, Y. Brammoullé, and A. Perez-Richet. Altruism in networks. *Econometrica*, 85:675–689, 2017.

P. Dubey. Inefficiency of nash equilibria. *Mathematics of Operations Research*, 11:363–376, 1986.

D. Gale and H. Nikaido. The jacobian matrix and global univalence of mappings. *Mathematische Anallen*, 159:81–93, 1965.

O. Galor. *Discrete Dynamical Systems*. Springer, 2007.

A. Heifetz. The non-dismal science of intergenerational affective interactions. *Games an Economic Behavior*, 140:575–584, 2023.

A. Heifetz, E. Minelli, and H. Polemarchakis. Liberal parentalism. *Journal of Public Economic Theory*, 23(6), 2021. URL http://www.polemarchakis.org/a89-lip.pdf.

H. Hori. Non-paternalistic altruism and utility interdepecence. *Japanese Economic Review*, 52:137–155, 2001.

P. J. Moylan. Matrices with positive principal minors. *Linear Algebra and Applications*, 17:53–98, 1977.

H. Nikaido. *Convex Sets and Economic Theory*. Academic Press, 1968.

V. Pata. *Fixed Point Theorems and Applications*. Springer, 2019.

D. G. Pearce. Nonpaternalistic sympathy and the inefficiency of consistent intertemporal plans. In M. O. Jackson and A. McLennan, editors, *Foundations in Microeconomic Theory: A volume in honor of Hugo Sonnenschein*, pages 213–231. Springer, 2008.

D. Ray and R. Vohra. Games of love and hate. *Journal of Political Economy*, 128:1789–1825, 2020.

Appendix 1: Stronger conditions

In this Appendix we discuss how Assumption 2 relates to the literature on interdependent preferences.

Towards this end, we start by recalling two conditions on the matrix $(I - J_x(u))$ that have been used (jointly with other assumptions) to prove the *existence and uniqueness of the induced game* for the case of altruism. We show that each of them implies our Assumption 2.

Spectral radius less than one

If the induced game U is defined at x, i.e. if $(x, U(x))$ is consistent, then, by definition, $U(x) = V_x(U(x))$, and therefore also $U(x) = V_x^k(U(x))$ for every $k \geq 1$[12]. Moreover, given Assumption 2, the first equality holds *only* for $U(x)$: by Gale and Nikaido (1965)Theorem 4, if $u \neq U(x)$ then $V_x(u) \neq u$.

Now, if u is some small perturbation of $U(x)$, representing a slight misassessment of the players regarding each other's utility levels with the action profile x, would the repeated re-assessments $V_x(u), V_x(V_x(u)), ..., V_x^k(u), ...$ converge back towards $U(x)$? This is a plausible requirement, because otherwise $U(x)$ is an *unstable* rest-point of V_x, and the definition of the induced game U is not robust to slight misperceptions.

The required convergence

$$V_x^k(u) \underset{k\to\infty}{\rightarrow} U(x)$$

is guaranteed in some small enough neighborhood of $U(x)$. That is, $U(x)$ is an asymptotically stable fixed point of V_x if the spectral radius of $J_x(U(x))$ (the largest of the absolute values of its eigenvalues), denoted $\rho(J_x(U(x)))$, satisfies[13]

$$\rho(J_x(U(x))) < 1,$$

whereas if, in contrast, $\rho(J_x(U(x))) > 1$ and no eigenvalue of $J_x(U(x))$ has absolute value equal to 1, then V_x is not asymptotically stable, and diverges away from arbitrarily small perturbations of $U(x)$.

In fact, the above re-assessments may take place among any subset $I_0 \subseteq I$ of the individuals, for fixed utility levels $\bar{u} = (\bar{u}_j)_{j\in I\setminus I_0}$ of the remaining

[12] V_x^k is defined inductively by

$$V_x^1(u) = V_x(u), \quad V_x^k(u) = V_x\left(V_x^{k-1}(u)\right)$$

for $k > 1$.

[13] Galor (2007), Theorem 4.8

individuals. The purely affective interaction V defines a *purely affective sub-interaction* $V^{\bar{u}}$ among the individuals in I_0,

$$V^{\bar{u}}(x,u) = (V_i(x,u,\bar{u}),)_{i\in I_0},$$

where $x = (x_i)_{i\in I_0}$ and $u = (u_i)_{i\in I_0}$. The set of *purely affective sub-interactions* of V is thus defined by ranging over all the non-empty subsets of individuals $I_0 \subseteq I$ and utility levels $\bar{u} = (\bar{u}_j)_{j\in I\setminus I_0}$ of the other individuals.

Assumption 4. For every $x \in X$ *and* $u \in R^n$,

$$\rho(J_x(u)) < 1,$$

and the same holds for all the sub-interactions of V.

This assumption implies our Assumption 2.

Proposition 4. *Under Assumption 4, for every* $x \in X$ *and* $u \in \mathbb{R}^n$, $(I - J_x(u))$ *is a P-matrix.*

Proof. $\rho(J_x(u)) < 1$ implies that all the eigenvalues $\lambda_1, ..., \lambda_n$ of $J_x(u)$ are within the open unit disk around the origin of the complex plane, and therefore that so are $-\lambda_1, ..., -\lambda_n$, which are the eigenvalues of $-J_x(u)$. Hence $1-\lambda_1, ..., 1-\lambda_n$, which are the eigenvalues of $I - J_x(u)$, all have positive real parts. These eigenvalues are the roots of the characteristic polynomial of $I - J_x(u)$. This characteristic polynomial has positive coefficients, and therefore its roots are all either real, and therefore positive by the above, and/or come in conjugate pairs of the form $c + di$ and $c - di$ whose product $c^2 + d^2$ is also positive. Hence the determinant of $I - J_x(u)$, which is the product of its eigenvalues, is positive.

All the above is true also for every sub-interaction involving only the subset I_0 of individuals, implying the positivity of the determinant of the principal submatrix of $I - J_x(u)$ with rows and columns in I_0, i.e. the positivity of the principal minor with rows and columns in I_0. We thus conclude that $I - J_x(u)$ is a P-matrix. □

Remark. The conclusion of Proposition 4, i.e. Assumption 2, is weaker than its premise, Assumption 4. For example, in the case of two individuals, denoting

$$J_x(U(x)) = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$$

$(I - J_x(U(x)))$ being a P-matrix means $ab < 1$, whereas $\rho(J_x(U(x))) < 1$ means the more stringent requirement $|ab| < 1$.

If $ab < -1$ then Assumption 2 holds, but Assumption 4 does not. In this case the eigenvalues of $J_x(U(x))$ are $\pm\sqrt{ab}$, whose absolute values are both larger than 1, and therefore V_x diverges away from $U(x)$ from arbitrarily small neighborhoods of $U(x)$.

Dominant Diagonal

Another property of the matrix $(I - J_x(u)$ that was considered in the literature, for example, Bergstrom (1999), is that the matrix has a *dominant diagonal*:

Assumption 5. For every x and u, the matrix $(I - J_x(u))$ is dominant diagonal: there exists $h(u) \in \mathbb{R}^n$, such that, for any $i = 1, \ldots n$,

$$h_i(u) > \sum_{j \neq i} h_j(u) \mid -\frac{\partial V_{x,i}}{\partial u_j} \mid .$$

That is, there is a way to rescale utilities at u, such that marginal changes in u_j, for $j \neq i$, have a total effect on $V_{x,i}$ less than 1.

Proposition 5. *Under Assumption 5, for every $x \in X$ and $u \in \mathbb{R}^n$, $(I - J_x(u))$ is a P-matrix*[14].

Remark. The conclusion of proposition 5, i.e. Assumption 2, is weaker than its premise, Assumption 5. For example, in the case of two individuals, denoting

$$J_x(U(x)) = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$$

$(I - J_x(U(x)))$ being a P-matrix means $ab < 1$, whereas dominant diagonal, Assumption 5, means the more stringent requirement $|a| < 1$ & $|b| < 1$.

The literature on Interdependent Preferences. As we mentioned in section 2, various authors have identified conditions for consistency to hold at *every* profile of actions, so that an induced game can be defined everywhere. For altruism, where the functions V_x are linear and separable in the others' utility levels, Pearce (2008) pointed out as sufficient the Hawkins-Simon condition on the matrix of cross derivatives of V_x, and Bergstrom (1999) pointed out the dominant diagonal condition as sufficient.

[14] Moylan (1977)

For altruism *without* linear separability, Hori (2001) (proposition 3) showed it is sufficient that the Hawkins-Simon condition hold uniformly or even non-uniformly if the functions V_x themselves are bounded (his proposition 4). Our example in section 4 shows in particular that in the absence of both requirements for V_x, indeed an induced game might not exist for x.

For the general case of altruism and/or spite, Ray and Vohra (2020) pointed out as a sufficient condition that V_x is a contraction, or alternatively that for each action profile, the associated directed network of payoff interdependencies is acyclic and the functions V_x are bounded.

Appendix 2: Two examples

In this Appendix we discuss two examples of *non-separable* purely affective interaction. The first illustrates the result in Theorem 5 (equilibrium strategies are locally dominant). The second one shows how non-separability allows for affective interactions in which the players attitudes towards one another may shift as they move in the action space.

Example 5 (locally dominant equilibrium strategies)

Let

$$V_1(x, u_2) = x(1-x) - 2xu_2,$$
$$V_2(y, u_1) = y(1-y) + \frac{1}{8}yu_1,$$

where $x, y \in (0,1)$. In this example, individual 1 is rather spiteful towards individual 2, and individual 2 is mildly sympathetic towards individual 1. For each $x, y \in (0,1)$ the Jacobian of V is

$$J_{(x,y)} = \begin{pmatrix} 0 & -2x \\ \frac{1}{8}y & 0 \end{pmatrix},$$

and $det[I - J_{(x,y)}] = (1 + \frac{1}{2}\sqrt{xy}) > 0$, so that Assumption 2 is satisfied.

The induced game is defined everywhere:

$$U_1(x, y) = \frac{8x((1-x) - 2y(1-y))}{8 + 2xy},$$
$$U_2(x, y) = \frac{y(8(1-y) + x(1-x))}{8 + 2xy},$$

with the best reply functions

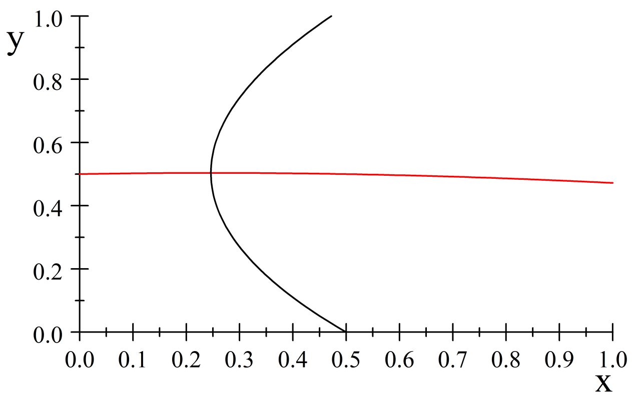
$$\beta_1(y) = \frac{2\sqrt{2y^3 - 2y^2 + y + 4} - 4}{y},$$
$$\beta_2(x) = \frac{\sqrt{-2x^3 + 2x^2 + 16x + 64} - 8}{2x},$$

whose intersection is the Nash equilibrium

$$x = 0.246\,20, \quad y = 0.503\,79.$$

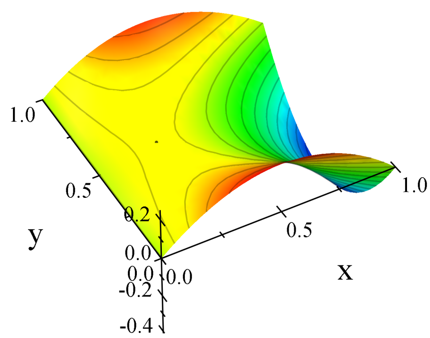
The reaction curves are locally flat (see Figure 4), an illustration of Theorem 5.

Figure 4



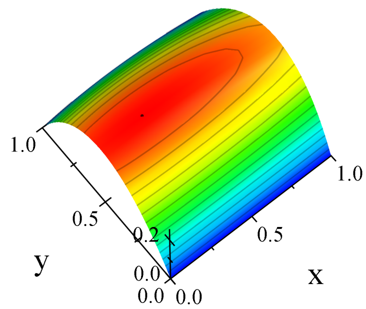
Also, at a Nash equilibrium utility functions are flat, an illustration of Proposition 1. For the spiteful individual 1, the Nash equilibrium is at a saddle point of his utility function (see Figure 5).

Figure 5



For the sympathetic individual 2, the Nash equilibrium is at a hilltop of her utility function (see Figure 6).

Figure 6



□

Example 6 (shifting attitudes)

Next,

$$V_1(x, u_2) = x^2(1 - x^2) + \frac{1}{2}xu_2,$$
$$V_2(y, u_1) = y^2(1 - y^2) + \frac{1}{2}yu_1,$$

for $x, y \in (-1, 1)$, so that each individual is sympathetic/spiteful with positive/negative actions respectively. The Jacobian of V at (x, y) is

$$J_{(x,y)} = \begin{pmatrix} 0 & \frac{1}{2}x \\ \frac{1}{2}y & 0 \end{pmatrix},$$

so that $det[I - J_{(x,y)}] = (1 - \frac{1}{4}xy) > 0$, and Assumption 2 is satisfied.

The induced game is defined everywhere:

$$U_1(x, y) = \frac{4x^2(1 - x^2) + 2xy^2(1 - y^2)}{4 - xy},$$
$$U_2(x, y) = \frac{4y^2(1 - y^2) + 2yx^2(1 - x^2)}{4 - xy}.$$

The graphs of U_1 and U_2 are in Figure 7 and Figure 8.

Figure 7

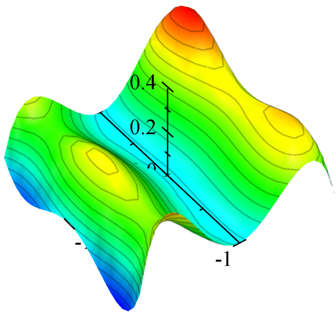
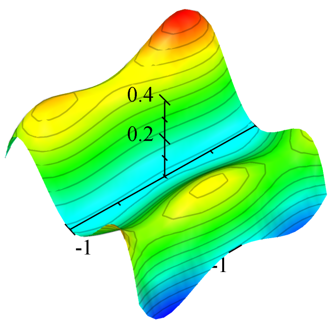


Figure 8



The unique Nash equilibrium is

$$x = y = 0.751\,97.$$

Both individuals are sympathetic and the Nash equilibrium is at the peak of their utility functions. The Nash equilibrium is Pareto optimal, and it maximizes the average of their utilities.

If, instead, the individuals were confined to negative actions $x, y \in (-1, 0)$, the unique Nash equilibrium would be

$$x = y = -0.68266$$

that is at a saddle point of U_1 and of U_2, and maximizes the individuals' average utility in that quadrant but not globally.

Similarly, if individual 1 were confined to positive actions (and thus sympathy) while individual 2 to negative actions (spite), there would be a unique Nash equilibrium within that quadrant

$$x = 0.724\,71, \quad y = -0.665\,76$$

with individual 1 at a hilltop and individual 2 at a saddle point, maximizing the average utility within that quadrant, but not globally. □