Affective interdependence and welfare

A. Heifetz [1] E. Minelli [2] H. Polemarchakis [3]

August 28, 2023

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

Abstract

Purely affective interaction allows the welfare of an individual to depend on her own actions and on the profile of welfare levels of others. Under an assumption on the structure of mutual affection that we interpret as "non-explosive mutual affection," we show that equilibria of simultaneous-move affective interaction are Pareto optimal independently of whether or not an induced standard game exists. Moreover, *if* purely affective interaction induces a standard game, then an equilibrium profile of actions is a Nash equilibrium of the 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 on her own actions 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.

Ray and Vohra (2020) demonstrated a striking result: *if* a purely affective interaction induces a standard game, Nash equilibria of the induced game are Pareto optimal.

Here we argue that the efficiency properties of equilibria of purely affective interactions are better understood *without* reference to the existence of an induced standard game. Once a natural equilibrium notion is defined, the Pareto efficiency of equilibrium follows immediately from the assumption that the matrix of mutual affection satisfies a condition introduced by Gale and Nikaido (1965) which, in our setting, we interpret as *non-explosive mutual affection*. We then show that, *if* an induced standard game exists, then an equilibrium profile of actions is a Nash equilibrium of the game, and, moreover, at this Nash equilibrium each player chooses a locally dominant strategy.

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 (2022) 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-explosive *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

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

and we also write

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

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

$$V_x(u) = u.$$

In a purely affective interaction, an individual controls her action x_i, and is aware of the direct effect of her choice on her well-being. Other individuals' choices do not affect her directly[3], but she cares about the well being of others, and her perception of others' well-being may affect her preferences over her actions. As in a competitive economy, an individual need not have a detailed knowledge of others' action sets and preferences. At a consistent profile of actions and utility levels, perceptions are confirmed.

A *parametric equilibrium* is a consistent profile of actions and utility levels, (x^*, u^*) that satisfies individual optimization: every individual maximizes V_i taking u^*_{-i} as given or[4]

$$V_x(u^*) \leq V_{x^*}(u^*).$$

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.

---

[3] Maybe because every externality has been priced – see the last section of the paper.

[4] 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.

Assumption 1. For every individual, X_i is an open subset of Euclidean space, and the utility function $V_i(\cdot,\cdot)$ is twice 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[5].

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

To interpret the assumption, we recall a useful characterization of P-matrices introduced by Gale and Nikaido:

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.

In our context, the Gale-Nikaido characterization allows us to interpret Assumption 2 as an assumption of *non-explosive mutual affection.*

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

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

or

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

Suppose now that we start from a consistent pair $(\hat{x}, \hat{u})$,

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

Then, for any exogenous change in utility levels while holding the profile of actions fixed, $\Delta u = u - \hat{u} \neq 0$, there is one individual i, for whom

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

or

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

---

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

That is, under Assumption 2, starting from a consistent pair of actions and utility levels, for any exogenous change in the perception of utility levels there is always one individual whose resulting utility, after the change, does not reinforce the exogenous change. Thus, 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 explosive affective interaction.

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

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

$$V_{\tilde{x}}(\tilde{u}) = \tilde{u} > u^*. \tag{1}$$

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

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

Since (x^*, u^*) is a parametric equilibrium,

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

and

$$F(\tilde{u}) = \tilde{u} - V_{\tilde{x}}(\tilde{u}) = 0 \leq F(u^*). \tag{2}$$

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 (1) and (2) cannot obtain simultaneously for $\tilde{u} \neq u^*$. □

3 The case of linearly separable affection

In a linearly separable purely affective interaction, the individuals' utility functions have the form

$$V_i(x_i, u) = f_i(x_i) + \sum_{j \neq i} a_{ij} u_j. \tag{3}$$

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

$$u = f(x) + Ju. \tag{4}$$

Under Assumption 2, $\det(I-J) \neq 0$, and we can uniquely solve the system of equations (4) at every x, thus obtaining the *induced game* corresponding to (3),

$$U(x) = (I-J)^{-1} f(x) = Bf(x). \tag{5}$$

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.

Under Assumption 2, $B = (I-J)^{-1}$ is also a P-matrix[6]. In particular, the diagonal of B is positive:

$$b_{ii} = \frac{|\,(I-J)_{ii}\,|}{Det(I-J)} > 0, \tag{6}$$

and, using (5):

Claim 1. *Under Assumption 2, in a situation of linearly separable affection, an action profile x^* is a Nash equilibrium of the induced game U if and only if $(x^*, U(x^*))$ is a parametric equilibrium of V.*

A planner may try to obtain a Pareto improvement over a Nash equilibrium in the induced game U by *jointly* changing the actions of everybody. In the case of linearly separable affection, a Pareto improvement is a Δf_x such that $\Delta U = B\Delta f_x > 0$, while a Nash deviation in the induced game is a $\Delta f_x^i = (0, \ldots, df^i, \ldots, 0)$ such that $\Delta U^i = B\Delta f_x^i > 0$. If x^* is a Nash equilibrium of U and $y = \Delta f_{x^*}$ a Pareto improvement, then $By > 0$ and $y \neq 0$. But then, by the Gale-Nikaido Lemma, there must exist i such that $y^i = \Delta f_{x^*}^i > 0$ (also, given (6), $B\Delta f_{x^*}^i > 0$), a contradiction with x^* being a Parametric equilibrium (also, with x^* being Nash equilibrium):

Claim 2. *Under Assumption 2, in a situation of linearly separable affection, if x^* is a Nash equilibrium of the induced game U, then x^* is Pareto optimal.*

For given $\lambda \in \mathbb{R}^n_+ \setminus \{0\}$, consider the welfare function

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

a linear combination of the 'base utilities' $f_i(x_i)$. The maximization of each $f_i(x_i)$ thus assures that the first order conditions for the maximization of $W_\lambda(x)$ are satisfied. Let us assume concavity of base utilities.

---

[6] Horn and Johnson (2013), Theorem 4.3.2

Assumption 3. For every i, 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 Nash equilibrium:

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

4 Locally induced game

In the general, non additively separable case, Assumption 2 does not guarantee that at any given x the system of equations

$$F_x(u) = u - V_x(u) = 0 \tag{7}$$

has a solution, so the induced game $U(x)$ need not be well defined everywhere on X.

Still, under Assumption 2, another result of Gale and Nikaido (1965), Theorem 4, implies that *if* a solution u_x of (7) exists, then it is unique.

Also, again by Assumption 2, $\det(I - J_x(u_x)) \neq 0$, and we can apply the implicit function theorem to $F : X \times \mathbb{R}^n \to \mathbb{R}^n$ 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 neighborhood $\mathcal{O}_x$ of x with

$$U_x(x) = u_x,$$

and

$$\frac{\partial U_i(x)}{\partial x_j} = \frac{\partial V_j(x_i, u_{x-i})}{\partial x_j} \left((I - J_x(u_x))^{-1}\right)_{ij}. \tag{8}$$

We call $U_x : \mathcal{O}_x \to \mathbb{R}^n$ the *locally induced game* of V at x.

In the locally induced game, parametric equilibrium strategies are *locally dominant*:

Theorem 2. *At a parametric equilibrium* (x^*, u^*), *each individual's action is locally dominant in the locally induced game* U_{x^*}.

Proof. At a parametric equilibrium, for each individual,

$$\frac{\partial V_i\left(x_i^*, u_{-i}^*\right)}{\partial x_i} = 0, \tag{9}$$

and therefore, by (8) with $j = i$, 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 (8), that

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

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

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

where the last equality is due to the first order condition (9) 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. With a marginal change in x_j from x^*, the function U_i remains constant as a function of x_i, and x_i^* therefore remains a local maximizer of U_i. □

One can also derive analogs of Claims 1 and 2:[7]

Claim 4. *Under Assumptions 1 and 2, if* (x^*, u^*) *is a parametric equilibrium of* V, *then* x^* *is a Nash equilibrium of the locally induced game* U_{x^*}.

Claim 5. *Suppose the induced game* U *is defined on the entirety of* X. *Under Assumptions 1 and 2, if* x^* *is a Nash equilibrium of the induced game, then* $(x^*, U(x^*))$ *is a parametric equilibrium.*

Corollary 1. *Suppose the induced game* U *is defined on the entirety of* X. *Under Assumptions 1 and 2, if* x^* *is a Nash equilibrium of the induced game* U, *then it is Pareto optimal.*

---

[7] See our working paper for the proofs.

Claims 4 and 5 and the Corollary are the analog in our setting of Ray and Vohra (2020)'s Lemma 1 and Theorem 1. Their results do not imply, nor are implied by ours; they do not assume differentiability (our Assumption 1) but their coherence condition (bundedness and uniqueness, see their paper p.1793) is stronger than our Assumption 2.[8]

5 Stronger conditions

Stronger conditions imply that $(I - J_x(u))$ is a P-matrix.

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 $k \geq 1$[9]. 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

---

[8] Consider the two person purely affective interaction where

$$V_1(x, u_2) = f(x) + au_2,$$
$$V_2(y, u_1) = g(y) + bu_1,$$

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

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

and the matrix $(I - J_{(x,y)})$ has a unitary diagonal and determinant $\det(I - J_{(x,y)}) = 1 - ab > 0$, and therefore it is a P-matrix if $ab < 1$. When $|a| > 1$ (and $|b| < \frac{1}{|a|}$) Ray and Vohra (2020)' s boundedness condition is not satisfied – for every (x, y) and every function $B(x, y) < \infty$, whenever $|u_2| > B(x, y) + |f(x)|$, in the sup norm $\|\cdot\|$,

$$\|U((x, y), (u_1, u_2))\| \geq |f(x) + au_2| \geq |u_2| - |f(x)| > B(x, y).$$

[9] 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 $k > 1$.

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

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

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[10]

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

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

---

[10] Galor (2007), Theorem 4.8

Proof. $\rho\left(J_x\left(u\right)\right) < 1$ implies that all the eigenvalues $\lambda_1, ..., \lambda_n$ of $J_x\left(u\right)$ 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 $-J_x\left(u\right)$. Hence $1-\lambda_1, ..., 1-\lambda_n$, which are the eigenvalues of $I - J_x\left(u\right)$, all have positive real parts. These eigenvalues are the roots of the characteristic polynomial of $I - J_x\left(u\right)$. 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 $I - J_x\left(u\right)$, 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 $I - J_x\left(u\right)$ 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 $I - J_x\left(u\right)$ is a P-matrix. □

Remark. The conclusion of proposition 1, i.e. Assumption 2, is weaker than its premise, Assumption 4. For example, in the case of two individuals, denoting

$$J_x\left(U\left(x\right)\right) = \left(\begin{array}{cc} 0 & a \\ b & 0 \end{array}\right)$$

$\left(I - J_x\left(U\left(x\right)\right)\right)$ being a P-matrix means $ab < 1$, whereas $\rho\left(J_x\left(U\left(x\right)\right)\right) < 1$ 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 $J_x\left(U\left(x\right)\right)$ are $\pm\sqrt{ab}$, whose absolute values are both larger than 1, and therefore V_x diverges away from $U\left(x\right)$ from arbitrarily small neighborhoods of $U\left(x\right)$.

Dominant Diagonal

Another property of the matrix $(I - J_x(u)$ that we may consider is that the matrix is *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 2. *Under Assumption 5, for every* $x \in X$ *and* $u \in \mathbb{R}^n$*,* $(I - J_x(u))$ *is a P-matrix*[11]*.*

6 Examples

Examples illustrate the results and their implications.

Example 1 (Two person linearly separable affection)*.* Consider the two person purely affective interaction where

$$V_1(x, u_2) = f(x) + au_2,$$
$$V_2(y, u_1) = g(y) + bu_1,$$

with

$$ab < 1.$$

That is, the two individuals can have positive or negative feelings towards each other, and these feelings may even be strong, but they satisfy the assumption of moderate *reciprocal* affection: if feelings go in the same direction (both love or both spite) they cannot be too strong.

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

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

and the matrix $(I - J_{(x,y)})$ has a unitary diagonal and determinant $\det(I - J_{(x,y)}) = 1 - ab > 0$, and therefore it is a P-matrix[12].

The induced game is

$$U_1(x, y) = \frac{f(x) + ag(y)}{1 - ab},$$
$$U_2(x, y) = \frac{g(y) + bf(x)}{1 - ab},$$

---

[11] Moylan (1977)

[12] Such separable interactions in which $|a| > 1$ (and $|b| < \frac{1}{|a|}$) do not induce "a game of love and hate" in the sense of Ray and Vohra (2020), and therefore are not covered by their analysis, because their boundedness condition (i) is not satisfied – for every (x, y) and every function $B(x, y) < \infty$, whenever $|u_2| > B(x, y) + |f(x)|$, in the sup norm $\|\cdot\|$,

$$\|U((x, y), (u_1, u_2))\| \geq |f(x) + au_2| \geq |u_2| - |f(x)| > B(x, y).$$

whose Nash equilibria (if there are any) are (x^*, y^*) where

$$x^* \in \arg\max f(x), \quad y^* \in \arg\max g(y).$$

That is, every Nash equilibrium is in dominant strategies.

In particular, if f and g are concave then there exists at most one Nash equilibrium (x^*, y^*). A social welfare function of the form

$$W(x, y) = \lambda_1 U_1(x, y) + \lambda_2 U_2(x, y),$$

where $(\lambda_1, \lambda_2) > 0$, is then concave and maximized at the unique Nash equilibrium (x^*, y^*) if and only if

$$\frac{1}{1 - ab} \begin{pmatrix} \lambda_1 & \lambda_2 \end{pmatrix} \begin{pmatrix} 1 & a \\ b & 1 \end{pmatrix} > 0.$$

With $ab < 1$ and therefore $\frac{1}{1-ab} > 0$, such $(\lambda_1, \lambda_2) > 0$ indeed exists since

(i) if both $a \geq 0$ and $b \geq 0$ (mutual sympathy) then U_1 and U_2 are already concave themselves, the Nash equilibrium (x^*, y^*) is their global maximum, and any $(\lambda_1, \lambda_2) > 0$ will do,

(ii) if $a \geq 0$ but $b < 0$ (individual 1 is sympathetic and individual 2 is spiteful) then U_1 is concave and the Nash equilibrium (x^*, y^*) is its global maximum, while U_2 is not concave, and (x^*, y^*) is a saddle point of it. One can then choose $\lambda_1 > 0$ and $\lambda_2 = 0$ to get a concave welfare function W which is maximized at (x^*, y^*),

(iii) similarly, if $b \geq 0$ but $a < 0$ one can choose $\lambda_2 > 0$ and $\lambda_1 = 0$ to the desired effect, and

(iv) finally, if both $a < 0$ and $b < 0$ (mutual spite) then both U_1 and U_2 are not concave, and the Nash equilibrium (x^*, y^*) is a saddle point of each of them. Nevertheless, since $ab < 1$ (the reciprocal extent of spite is moderate), the vectors $(1, b)$ and $(a, 1)$ are within a half-plane containing the positive orthant, and therefore both $(1, b)$ and $(a, 1)$ form an acute angle with vectors $(\lambda_1, \lambda_2) > 0$ in some positive cone[13].

We therefore conclude that *in all cases*, the Nash equilibrium (x^*, y^*) is Pareto optimal.

---

[13] This cone becomes narrower as $ab \nearrow 1$. The weights $(\lambda_1, \lambda_2) > 0$ in this cone 'strike the right balance' between the curvatures of U_1 and U_2 at (x^*, y^*) – between the concavity of U_1 in x and the convexity of U_2 in x so that the linear combination W – the social welfare function – is concave in x, and *at the same time* also between the convexity of U_1 in y and the concavity of U_2 in y, so that the linear combination W is concave also in y.

Example 2 (Non-separable affection). Now let

$$\begin{aligned} V_1\left(x,u_2\right) &= x\left(1-x\right)-2xu_2, \\ V_2\left(y,u_1\right) &= y\left(1-y\right)+\frac{1}{8}yu_1, \end{aligned}$$

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

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

whose eigenvalues are $\pm\frac{1}{2}i\sqrt{xy}$, and its spectral radius is therefore $\frac{1}{2}\sqrt{xy}<1$.

The induced game is

$$\begin{aligned} U_1\left(x,y\right) &= \frac{8x\left(\left(1-x\right)-2y\left(1-y\right)\right)}{8+2xy}, \\ U_2\left(x,y\right) &= \frac{y\left(8(1-y)+x\left(1-x\right)\right)}{8+2xy}, \end{aligned}$$

with the best reply functions

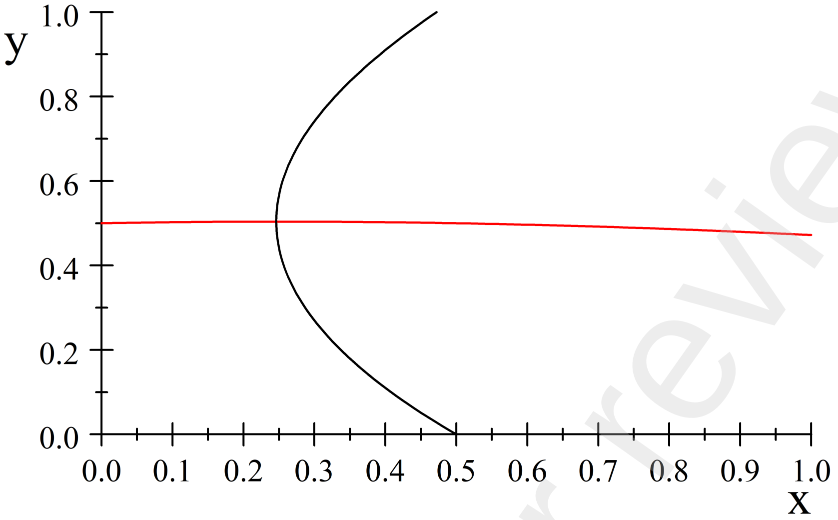
$$\begin{aligned} \beta_1\left(y\right) &= \frac{2\sqrt{2y^3-2y^2+y+4}-4}{y}, \\ \beta_2\left(x\right) &= \frac{\sqrt{-2x^3+2x^2+16x+64}-8}{2x}, \end{aligned}$$

whose intersection is the Nash equilibrium

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

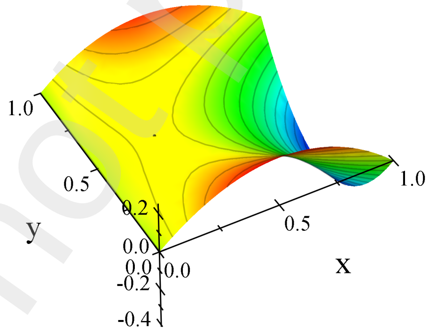
where the reaction curves are locally flat (see Figure 1).

[Figure 1]



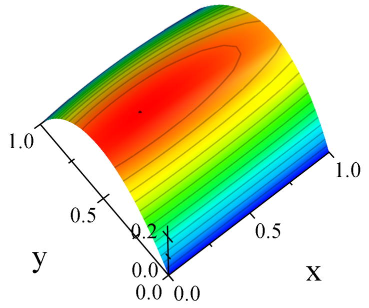
For the spiteful individual 1, the Nash equilibrium is at a saddle point of his utility function (see Figure 2).

[Figure 2]



For the sympathetic individual 2, the Nash equilibrium is at a hilltop of her utility function (see Figure 3).

[Figure 3]



Example 3 (Shifting attitudes). Next,

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

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

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

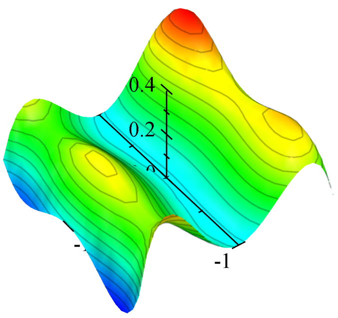
whose eigenvalues are $\pm\frac{1}{2}\sqrt{xy}$. This implies that the spectral radius of the Jacobian is smaller than $\frac{1}{2}$.

The induced game is

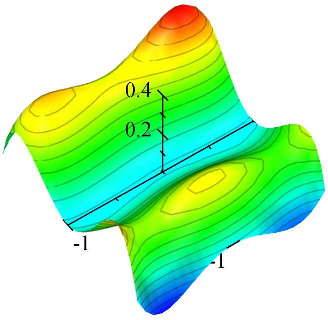
$$U_1(x,y) = \frac{4x^2(1-x^2) + 2xy^2(1-y^2)}{4-xy},$$
$$U_2(x,y) = \frac{4y^2(1-y^2) + 2yx^2(1-x^2)}{4-xy}.$$

The graphs of U_1 and U_2 are in Figure 4 and Figure 5.

[Figure 4]



[Figure 5]



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.

7 Economies with affective interaction

Consider a slightly modified version of Bergstrom (1989) 'Love and Spaghetti' example. Romeo and Juliet both care about spaghetti (x), and money (m), and they also care about the other's happiness:

$$\begin{aligned} V_1(x_1,u_2) &= \sqrt{x_1}+m_1+au_2,\\ V_2(x_2,u_1) &= \sqrt{x_2}+m_2+bu_1. \end{aligned}$$

with initial endowments $e_i=(1,M)$, $i=1,2$.

Differently from Bergstrom, we allow for negative a and b, but, by Assumption 2, we assume that $ab<1$ (so here Romeo and Julliet do not necessarily love each other, but their affective interdependence, positive or negative, is moderate).

We can solve for the induced utilities $U_1(x,m)$ and $U_2(x,m)$:

$$\begin{aligned} U_1(x,m) &= \tfrac{1}{1-ab}\left(\sqrt{x_1}+m_1\right)+\tfrac{a}{1-ab}\left(\sqrt{x_2}+m_2\right),\\ U_2(x,m) &= \tfrac{b}{1-ab}\left(\sqrt{x_1}+m_1\right)+\tfrac{1}{1-ab}\left(\sqrt{x_2}+m_2\right). \end{aligned}$$

If the action set of each player could be defined independently from the action of the other, this would just be a special case of a linearly separable purely affective interaction satisfying our assumption of non-explosive mutual affection, and the parametric equilibrium would be Pareto efficient. Here, though, feasibility imposes a joint restriction on individual actions.

Of course, we know from the first theorem of welfare economics that, in a perfectly competitive economy *without* affective interaction, aggregate feasibility constraints do not prevent the Pareto efficiency of equilibria. A natural question to ask is thus wether competitive equilibria are Pareto efficient in a perfectly competitive economy with *non-explosive* mutual affection.

At a *competitive equilibrium of an economy with affective interaction*, each individual chooses (x_i,m_i) to maximize U_i taking (x_j,m_j) as given, under the budget constraint: $p_x x_i+p_m m_i=p_x+p_m M$, and prices adjust to guarantee feasibility.

If we fix $p_m=1$ (and M is large enough), at the unique competitive equilibrium,

$$\hat{p}_x=1,$$

$$\hat{x}_i=e_i=1,\quad i=1,2.$$

A *benevolent non-myopic social planner* chooses the allocation (x_1,x_2) of spaghetti to solve

$$\max_{x_1,x_2} W_\lambda(x_1,x_2,m)=\sum_i \lambda_i U_i(x_1,x_2,m),$$

and the feasibility constraint on spaghetti is:

$$x_1 + x_2 = 2,$$

leading to the first order condition

$$\frac{\sqrt{x_1}}{\sqrt{2 - x_1}} = \frac{\lambda_1 + \lambda_2 b}{\lambda_1 a + \lambda_2}. \tag{10}$$

We see that, in an economy, non-explosive mutual affection, $ab < 1$ *does not* guarantee that there exist $(\lambda_1, \lambda_2) > 0$ such that the equilibrium allocation $\hat{x}_1 = 1$ solves the planner problem (i.e. Claim 3 does not hold). For example, if $a = 2$ and $b = 1/4$, at $\hat{x}_1 = 1$, (10) becomes $\lambda_1 = -\frac{3}{4}\lambda_2$. Indeed, the equilibrium allocation of good, $\hat{x}_1 = \hat{x}_2 = 1$, generates utilities $\hat{u}_1 = 2\sqrt{\hat{x}_1} + 4\sqrt{\hat{x}_2} = 6$, $\hat{u}_2 = \frac{1}{2}\sqrt{\hat{x}_1} + 2\sqrt{\hat{x}_2} = 2.5$, while a planner solving (10) with $\lambda_1 = \lambda_2 = 1$ would rather choose $\tilde{x}_1 = 0.29$, $\tilde{x}_2 = 2 - \tilde{x}_1 = 1.71$, generating utilities $\tilde{u}_1 = 2\sqrt{\tilde{x}_1} + 4\sqrt{\tilde{x}_2} = 6.28$ and $\tilde{u}_2 = \frac{1}{2}\sqrt{\tilde{x}_1} + 2\sqrt{\tilde{x}_2} = 2.87$, a Pareto improvement: the planner reallocates the good taking into account the strong love of individual 1 towards individual 2.

For the specific preferences and endowments of the example, positive welfare weights such that the competitive equilibrium maximizes social welfare, and is therefore Pareto efficient, do exist under the stronger condition: $ab < 1$ *and* $a < 1$, $b < 1$.

A characterization of economies with affective interaction in which competitive equilibrium allocations are Pareto efficient is an open problem.[14]

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. Love and spaghetti: the opportunity cost of virtue. *Journal of Economic Perspectives*, 3:165–173, 1989.

---

[14] Winter (1969) and Dufwenberg, Heidhues, Kirchsteiger, Riedel, and Sobel (2011) identified a condition, *social monotonicity* that, under assumptions of monotonicity and convexity in own consumption and, in particular, separability of utilities between own utility and the profile of utilities of others, implies that Pareto optimal allocations can be supported as competitive allocations and can be attained with redistributions of revenue. They also showed by examples that social monotonicity *does not* guarantee the efficiency of competitive equilibria.

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.

M. Dufwenberg, P. Heidhues, P. Kirchsteiger, F. Riedel, and J. Sobel. Other-regarding preferences in general equilibrium. *Review of Economic Studies*, 78:613–639, 2011.

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. Unpublished manuscript, 2022.

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.

R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 2nd edition, 2013.

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

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.

S. G. Winter. A simple remark on the second optimality theorem of welfare economics. *Journal of Economic Theory*, 1:99–103, 1969.