Selfishness Level of Strategic Games

Krzysztof R. Apt* Guido Schäfer†

Abstract

We introduce a new measure of the discrepancy in strategic games between the social welfare in a Nash equilibrium and in a social optimum, that we call ***selfishness level***. It is the smallest fraction of the social welfare that needs to be added to the players' payoffs to ensure that a Nash equilibrium of the resulting game is also its social optimum. This notion is unrelated to that of price of stability. We compute the selfishness level for some selected games. In particular, the selfishness level of finite ordinal potential games is finite, while that of a Cournot competition oligopoly game and Tragedy of the Commons game is infinite. We also provide an estimate on the selfishness level of linear congestion games and fair cost sharing games.

The intelligent way to be selfish is
to work for the welfare of others
Dalai-Lama[1]

*Center for Mathematics and Computer Science (CWI) and University of Amsterdam, The Netherlands. Email: k.r.apt@cwi.nl.

†Center for Mathematics and Computer Science (CWI) and VU University Amsterdam, The Netherlands. Email: g.schaefer@cwi.nl.

[1] from [4, page 109]

1 Introduction

1.1 Motivation

The discrepancy in strategic games between the social welfare in a Nash equilibrium and in a social optimum has been long recognized by the economists. One of the flagship examples is Cournot competition, a strategic game involving firms that simultaneously choose the production levels of a homogeneous product. The payoff functions in this game describe the firms' profit in the presence of some production costs, under the assumption that the price of the product depends negatively on the total output. It is well-known, see, e.g., [3, Section 10.2], that the price in the social optimum is strictly higher than in the Nash equilibrium, which shows that the competition between the producers of a product drives its price down.

In computer science the above discrepancy led to the introduction of the notions of the *price of anarchy*, see [10], and the *price of stability*, see [17], that measure the ratio between the social welfare in a worst and, respectively, a best Nash equilibrium and a social optimum. This originated a huge research effort aiming at determining both ratios for specific strategic games that possess (pure) Nash equilibria.

These two notions are *descriptive* in the sense that they refer to an existing situation. In this paper we propose another notion that measures the discrepancy between the social welfare in a Nash equilibrium and a social optimum, which is *normative*, in the sense that it refers to a modified situation. It draws on the concept of *altruistic games* (see, e.g., [11] and more recent [12]). In these games each player's payoff is modified so that it also depends positively on a fraction of the social welfare in the considered joint strategy.

In our approach the minimal fraction for which such a modification of the original game yields the price of stability 1 is viewed as the ***selfishness level*** of the game. That is, the selfishness level of a game is the smallest fraction of the social welfare that needs to be offered to each player to achieve that a social optimum is realized in a Nash equilibrium.

So in a finite game, if some Nash equilibrium of the game is also a social optimum, then its selfishness level is 0. Otherwise if $\alpha > 0$ is the smallest fraction of the social welfare that needs to be added to the players' payoffs to ensure that a Nash equilibrium of the resulting game is also its social optimum, then the selfishness level is α. If such an α does not exist, then the selfishness level of the game is ∞. For infinite games one needs additionally to consider the possibility that a minimum of a set of αs may not exist.

1.2 Related work

On an abstract level, the proposed approach is discussed in [1], in chapter "How to Promote Cooperation", from where we cite (see page 134): "An excellent way to promote cooperation in a society is to teach people to care about the welfare of others."

There are only few articles in the algorithmic game theory literature that study the influence of altruism in strategic games [7, 9, 8, 5, 6]. In these works, altruistic player behavior is modeled by altering each player's perceived payoff in order to account also for the welfare of others. The models differ in the way they combine the player's individual payoff with the payoffs of the other players. All these studies are descriptive in the sense that they aim at understanding the impact of altruistic behavior on specific strategic games.

Closest to our work are the articles [8] and [6]. Elias et al. [8] study the inefficiency of equilibria in network design games with altruistic (or, as they call it, socially-aware) players. As we do here, they define each player's cost function as his individual cost plus α times the social cost. They derive lower and upper bounds on the price of anarchy and the price of stability, respectively, of the modified game. In particular, they show that the price of stability is at most $(H_n+\alpha)/(1+\alpha)$, where n is the number of players.

In an independent work, Chen et al. [6] introduce a framework to study the *robust price of anarchy*, which refers to the worst-case inefficiency of more general solution concepts such as coarse correlated equilibria (see [16]), of altruistic extensions of strategic games. In their model, player i's perceived cost is a convex combination of $(1-\beta_i)$ times his individual cost plus β_i times the social cost, where $\beta_i \in [0,1]$ is the altruism level of i. If all players have a uniform altruism level $\beta_i = \beta$, this model relates to the one we consider here by setting $\alpha = \beta/(1-\beta)$. Although not being the main focus of the paper, the authors also provide upper bounds of $2/(1+\beta)$ and $(1-\beta)H_n+\beta$ on the price of stability for linear congestion games and fair cost sharing games, respectively.

Note that in all three cases the price of stability approaches 1 as α goes to ∞. This seems to suggest that that the selfishness level of these games is ∞. However, this is not the case: We derive a characterization result that allows us to determine the selfishness level of a strategic game. Using this characterization, we can show that the selfishness level of finite potential games is finite, thereby showing that the selfishness level of the games mentioned above is finite. We also derive explicit bounds on the selfishness level of linear congestion games and fair cost sharing games (which include network design games as a special case) that do not depend on the number of players.

1.3 Outline of the paper

In what follows we provide in Section 3 the definition of the selfishness level and show that the selfishness level of a finite game can be an arbitrary real number that is unrelated to the price of stability. Then in Section 4 we provide a characterization that allows us to determine when the selfishness level of a game is finite. In the case of finite games this boils down to a simple test, namely the existence of a specific social optimum that we call *stable*. In particular, the selfishness level of a finite game with a unique social optimum is finite.

Finally, in Section 5 we compute the selfishness level for some selected games. In particular, the selfishness level of the n-players Prisoner's Dilemma game is $1/(2n-3)$, and that of the Traveler's Dilemma game is $\frac{1}{2}$. We also show that the selfishness level of finite ordinal potential games is finite, while those of Cournot competition for n firms (an example of an infinite ordinal potential game) and of a Tragedy of the Commons game are infinite. Finally, we provide an estimate on the selfishness level of linear congestion games and fair cost sharing games.

2 Preliminaries

A ***strategic game*** (in short, a game) $G = (N, \{S_i\}_{i\in N}, \{p_i\}_{i\in N})$ is given by a set $N = \{1,\ldots,n\}$ of players, a non-empty set of ***strategies*** S_i for every player $i \in N$, and a ***payoff function*** p_i for every player $i \in N$ with $p_i : S_1 \times \ldots \times S_n \rightarrow \mathbb{R}$. The interpretation is that every player $i \in N$ aims at choosing a strategy $s_i \in S_i$ so as to maximize his individual payoff $p_i(s)$, where $s = (s_1,\ldots,s_n)$.

We call $s \in S_1 \times \ldots \times S_n$ a ***joint strategy***, denote its ith element by s_i, denote $(s_1,\ldots,s_{i-1},s_{i+1},\ldots,s_n)$ by s_{-i} and similarly with S_{-i}. Further, we write (s'_i, s_{-i}) for $(s_1,\ldots,s_{i-1},s'_i,s_{i+1},\ldots,s_n)$, where we assume that $s'_i \in S_i$. Sometimes, when focussing on player i we write (s_i, s_{-i}) instead of s.

A joint strategy s a ***Nash equilibrium*** if for all $i \in \{1,\ldots,n\}$ and $s'_i \in S_i$

$$p_i(s_i, s_{-i}) \geq p_i(s'_i, s_{-i}).$$

Further, given a joint strategy s we call the sum $SW(s) := \sum_{j=1}^{n} p_j(s)$ the ***social welfare*** of s. When the social welfare of s is maximal we call s a ***social optimum***.

3 Selfishness level

Given a strategic game $G := (N, \{S_i\}_{i \in N}, \{p_i\}_{i \in N})$ and $\alpha \geq 0$ we define the game $G(\alpha) := (N, \{S_i\}_{i \in N}, \{r_i\}_{i \in N})$ by putting

$$r_i(s) := p_i(s) + \alpha SW(s).$$

So when $\alpha > 0$ the payoff of each player in the $G(\alpha)$ game depends on the social welfare of the players. $G(\alpha)$ is then an altruistic version of the game G.

Suppose now that for some $\alpha \geq 0$ a Nash equilibrium of $G(\alpha)$ is a social optimum of $G(\alpha)$. Then we say that G is α-***selfish***. We define now the ***selfishness level*** of a game by a case analysis.

If for no $\alpha \geq 0$, G is α-selfish, then we say that its selfishness level is ∞. If for some $\alpha \geq 0$, G is α-selfish and

$$\min_{\alpha \in \mathbb{R}_+} (G \text{ is } \alpha\text{-selfish})$$

exists, then we call this minimum the ***selfishness level*** of G, and otherwise we stipulate that the selfishness level of G is undefined.

Of course, when the game G is finite and for some $\alpha \geq 0$, G is α-selfish, the above minimum does exist. We show below (Theorem 2) that this does not need to be the case when G is infinite, that is, for some games their selfishness level is undefined.

Note that the social welfare of a joint strategy s in $G(\alpha)$ equals $(1 + \alpha n)SW(s)$, so the social optima of G and $G(\alpha)$ coincide. Hence we can replace in the above definition the reference to a social optimum of $G(\alpha)$ by one to a social optimum of G. This is what we shall do in the proofs below.

The above definitions refer to strategic games in which each player i maximizes his payoff function p_i and the social welfare of a joint strategy s is given by $SW(s)$. These definitions obviously apply to strategic games in which every player i minimizes his cost function c_i and the social cost of a joint strategy s is defined as $SC(s) := \sum_{j=1}^{n} c_j(s)$.

Intuitively, a low selfishness level means that the share of the social welfare needed to induce the players to choose a social optimum is small. This share can be viewed as an 'incentive' needed to realize a social optimum. Let us illustrate this definition on three simple examples.

Example 1. Prisoner's Dilemma

	C	D
C	2,2	0,3
D	3,0	1,1

	C	D
C	6,6	3,6
D	6,3	3,3

Consider the Prisoner's Dilemma game G (on the left) and the resulting game $G(\alpha)$ for $\alpha = 1$ (on the right). In the latter game the social optimum, (C,C), is also a Nash equilibrium. One can easily check that for $\alpha < 1$, (C,C) is also a social optimum of $G(\alpha)$ but not a Nash equilibrium. So the selfishness level of this game is 1.

Example 2. Battle of the Sexes

	F	B
F	2,1	0,0
B	0,0	1,2

Here each Nash equilibrium is also a social optimum, so the selfishness level of this game is 0.

Example 3. Matching Pennies

	H	T
H	$1,-1$	$-1,\ 1$
T	$-1,\ 1$	$1,-1$

Since the social welfare of each joint strategy is 0, for each α the game $G(\alpha)$ is identical to the original game in which no Nash equilibrium exists. So the selfishness level of this game is ∞. More generally, the selfishness level of a constant sum game is 0 if it has a Nash equilibrium and otherwise it is ∞.

Recall that, given a finite game G that has a Nash equilibrium, its ***price of stability*** is the ratio $SW(s)/SW(s')$ where s is a social optimum and s' is a Nash equilibrium with the highest social welfare in G. So the price of stability is 1 iff the selfishness level of G is 0. However, in general there is no relation between these two notions. The following observation also shows that the selfishness level of a finite game can be an arbitrary real number.

Theorem 1. *For every finite $\alpha > 0$ and $\beta > 1$ there is a finite game whose selfishness level is α and whose price of stability is β.*

Proof. Consider the following generalized form of the Prisoner's Dilemma game G to which we refer by $PD(\alpha,\beta)$:

	C	D
C	$1,1$	$0,x+1$
D	$x+1,0$	$\frac{1}{\beta},\frac{1}{\beta}$

where $x = \frac{\alpha}{\alpha+1}$.

In this game and in each game $G(\gamma)$ with $\gamma \geq 0$, (C,C) is the unique social optimum. To compute the selfishness level we need to consider a game $G(\gamma)$ and stipulate that (C,C) is its Nash equilibrium. This leads to the inequality $1+2\gamma \geq (\gamma+1)(x+1)$, from which it follows that $\gamma \geq \frac{x}{1-x}$, i.e., $\gamma \geq \alpha$. So the selfishness level of G is α. Moreover, its price of stability is β. □

We now use the above games $PD(\alpha,\beta)$ to establish the following result showing that for some games the selfishness level is undefined.

Theorem 2. *There exists a game that is α-selfish for every $\alpha > 0$, but is not 0-selfish.*

Proof. We construct a strategic game $G = (N, \{S_i\}_{i\in N}, \{p_i\}_{i\in N})$ with two players $N = \{1,2\}$ by combining, for an arbitrary but fixed $\beta > 1$, infinitely many $PD(\alpha,\beta)$ games with $\alpha > 0$ as follows: For each $\alpha > 0$ we rename the strategies of the $PD(\alpha,\beta)$ game to, respectively, $C(\alpha)$ and $D(\alpha)$ and denote the corresponding payoff functions by p_i^{α}. The set of strategies of each player $i \in N$ is $S_i = \{C(\alpha) \mid \alpha > 0\} \cup \{D(\alpha) \mid \alpha > 0\}$ and the payoff of i is defined as

$$p_i(s_i,s_{-i}) := \begin{cases} p_i^{\alpha}(s_i,s_{-i}) & \text{if } \{s_i,s_{-i}\} \subseteq \{C(\alpha),\ D(\alpha)\} \text{ for some } \alpha > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Every social optimum of G is of the form $(C(\alpha),C(\alpha))$, where $\alpha > 0$. (Note that we exploit that $\beta > 1$ here.) By the argument given in the proof of Theorem 1, $(C(\alpha),C(\alpha))$ with $\alpha > 0$ is a Nash equilibrium in the game $G(\alpha)$ because the deviations from $C(\alpha)$ to a strategy $C(\gamma)$ or $D(\gamma)$ with $\gamma \neq \alpha$ yield a payoff of 0. Thus, G is α-selfish for every $\alpha > 0$. Finally, observe that G is not 0-selfish because every Nash equilibrium of G is of the form $(D(\alpha),D(\alpha))$, where $\alpha > 0$. □

4 A characterization result

We now characterize the games with a finite selfishness level. To this end we shall need the following notion. We call a social optimum s ***stable*** if for all $i \in N$ and $s'_i \in S_i$ the following holds:

$$\text{if } (s'_i,s_{-i}) \text{ is a social optimum, then } p_i(s_i,s_{-i}) \geq p_i(s'_i,s_{-i}).$$

In other words, a social optimum is stable if no player is better off by unilaterally deviating to another social optimum.

Lemma 1. *Consider a strategic game* $G := (N, \{S_i\}_{i\in N}, \{p_i\}_{i\in N})$ *and* $\alpha \geq 0$.

(i) If s is both a Nash equilibrium of $G(\alpha)$ *and a social optimum of G, then s is a stable social optimum of G.*

(ii) If s is a stable social optimum of G, then s is a Nash equilibrium of $G(\alpha)$ *iff for all* $i \in N$ *and* $s'_i \in R(i,s)$

$$\alpha \geq \frac{p_i(s'_i, s_{-i}) - p_i(s_i, s_{-i})}{SW(s_i, s_{-i}) - SW(s'_i, s_{-i})}$$

where

$$R(i,s) := \{s'_i \in S_i \mid p_i(s'_i, s_{-i}) > p_i(s_i, s_{-i}) \text{ and } SW(s_i, s_{-i}) > SW(s'_i, s_{-i})\}.$$

Proof.
(i) Suppose that s is both a Nash equilibrium of $G(\alpha)$ and a social optimum of G. Consider some joint strategy (s'_i, s_{-i}) that is a social optimum. By the definition of a Nash equilibrium

$$p_i(s_i, s_{-i}) + \alpha SW(s_i, s_{-i}) \geq p_i(s'_i, s_{-i}) + \alpha SW(s'_i, s_{-i}),$$

so $p_i(s_i, s_{-i}) \geq p_i(s'_i, s_{-i})$, as desired.

(ii) Suppose that s is a stable social optimum of G. Then s is a Nash equilibrium of $G(\alpha)$ iff for all $i \in N$ and $s'_i \in S_i$

$$p_i(s_i, s_{-i}) + \alpha SW(s_i, s_{-i}) \geq p_i(s'_i, s_{-i}) + \alpha SW(s'_i, s_{-i}). \tag{1}$$

If $p_i(s_i, s_{-i}) \geq p_i(s'_i, s_{-i})$, then (1) holds for all $\alpha \geq 0$ since s is a social optimum. If $p_i(s'_i, s_{-i}) > p_i(s_i, s_{-i})$, then, since s is a stable social optimum of G, we have $SW(s_i, s_{-i}) > SW(s'_i, s_{-i})$.

So (1) holds for all $i \in N$ and $s'_i \in S_i$ iff

$$\alpha \geq \frac{p_i(s'_i, s_{-i}) - p_i(s_i, s_{-i})}{SW(s_i, s_{-i}) - SW(s'_i, s_{-i})}$$

holds for all $i \in N$ and $s'_i \in R(i,s)$. □

This leads us to the following result.

Theorem 3. *Consider a strategic game* $G := (N, \{S_i\}_{i\in N}, \{p_i\}_{i\in N})$.

(i) The selfishness level of G is finite iff a stable social optimum s exists for which

$$\alpha(s) := \max_{i\in N,\ s'_i \in R(i,s)} \frac{p_i(s'_i, s_{-i}) - p_i(s_i, s_{-i})}{SW(s_i, s_{-i}) - SW(s'_i, s_{-i})}$$

is finite.

(ii) If G is finite, then its selfishness level is finite iff it has a stable social optimum. In particular, if G has a unique social optimum, then its selfishness level is finite.

(iii) If G is finite and has a stable social optimum, then its selfishness level equals $\min_{s\in SSO} \alpha(s)$, *where* SSO *is the set of stable social optima.*

(iv) If $\beta > \alpha \geq 0$ *and G is* α*-selfish, then G is* β*-selfish.*

Proof. (i) and (iv) follow by Lemma 1, (ii) by (i) and (iii) by (ii) and Lemma 1. $\square$

Using the above theorem we now exhibit a class of games for n players for which the selfishness level is unbounded. In fact, the following more general result holds.

Theorem 4. *For each function $f : \mathbb{N} \rightarrow \mathbb{R}_+$ there exists a class of games for n players, where $n > 1$, such that the selfishness level of a game for n players equals $f(n)$.*

Proof. Assume $n > 1$ players and that each player has two strategies, 1 and 0. Denote by $\mathbf{1}$ the joint strategy in which each strategy equals 1 and by $\mathbf{1}_{-i}$ the joint strategy of the opponents of player i in which each entry equals 1. The payoff for each player i is defined as follows:

$$p_i(s) := \begin{cases} 0 & \text{if } s = \mathbf{1} \\ f(n) & \text{if } s_i = 0 \text{ and } \forall j < i,\ s_j = 1 \\ -\frac{f(n)+1}{n-1} & \text{otherwise.} \end{cases}$$

So when $s \neq \mathbf{1}$, $p_i(s) = f(n)$ if i is the smallest index of a player with $s_i = 0$ and otherwise $p_i(s) = -\frac{f(n)+1}{n-1}$. Note that $SW(\mathbf{1}) = 0$ and $SW(s) = -1$ if $s \neq \mathbf{1}$. So $\mathbf{1}$ is a unique social optimum.

We have $p_i(0, \mathbf{1}_{-i}) - p_i(\mathbf{1}) = f(n)$ and $SW(\mathbf{1}) - SW(0, \mathbf{1}_{-i}) = 1$. So by Theorem 3$(iii)$ the selfishness level equals $f(n)$. $\square$

5 Examples

We now use the above characterization result to determine or compute an upper bound on the selfishness level of some selected games. First, we exhibit a well-known class of games (see [13]) for which the selfishness level is finite.

5.1 Potential games

Given a game $G := (N, \{S_i\}_{i \in N}, \{p_i\}_{i \in N})$, a function $P : S_1 \times \ldots S_n \rightarrow \mathbb{R}$ is called an ***ordinal potential function*** for G if for all $i \in N$, $s_{-i} \in S_{-i}$ and $s_i, s_i' \in S_i$

$$p_i(s_i, s_{-i}) > p_i(s_i', s_{-i}) \text{ iff } P(s_i, s_{-i}) > P(s_i', s_{-i}).$$

A game that possesses an ordinal potential function is called an ***ordinal potential game***.

Theorem 5. *Every finite ordinal potential game has a finite selfishness level.*

Proof. Each social optimum with the largest potential is a stable social optimum. So the claim follows by Theorem 3(ii). $\square$

In particular, every finite congestion game (see [15]) has a finite selfishness level. We shall derive explicit bounds for two special cases of these games in Sections 5.7 and 5.8.

5.2 Prisoner's dilemma for n players

We assume that each player $i \in N = \{1, \ldots, n\}$ has two strategies, 1 (cooperate) and 0 (defect). We put

$$p_i(s) := 1 - s_i + 2 \sum_{j \neq i} s_j.$$

Proposition 1. *The selfishness level of the n-players Prisoner's Dilemma game is* $\frac{1}{2n-3}$.

Proof. Denote by $\mathbf{1}$ the joint strategy in which each strategy equals 1. In this game $\mathbf{1}$ is the unique social optimum, with for each $i \in N$, $p_i(\mathbf{1}) = 2(n-1)$ and $SW(\mathbf{1}) = 2n(n-1)$.

Consider now the joint strategy s in which player i deviates from $\mathbf{1}$ to the strategy 0, while the other players remain at 1. We have then $p_i(s) = 2(n-1)+1$ and $SW(s) = 2(n-1)+1+2(n-1)(n-2)$. Hence

$$\frac{p_i(s) - p_i(\mathbf{1})}{SW(\mathbf{1}) - SW(s)} = \frac{1}{2n-3}$$

The claim now follows by Theorem 1(iii). In particular, for $n = 2$ we get, as already argued in Example 1, that the selfishness level of the original Prisoner's Dilemma game is 1. □

5.3 Traveler's dilemma

This is a strategic game discussed in [2] with two players $N = \{1,2\}$, strategy set $S_i = \{2,\dots,100\}$ for every player i, and payoff function p_i for every i defined as

$$p_i(s) := \begin{cases} s_i & \text{if } s_i = s_{-i} \\ s_i + 2 & \text{if } s_i < s_{-i} \\ s_{-i} - 2 & \text{otherwise.} \end{cases}$$

Proposition 2. *The selfishness level of the Traveler's Dilemma game is* $\frac{1}{2}$.

Proof. The unique social optimum of this game is $(100,100)$, while $(2,2)$ is its unique Nash equilibrium.

If player i deviates from the social optimum to a strategy $s'_i \leq 99$, while the other player remains at 100, the respective payoffs become $s'_i + 2$ and $s'_i - 2$, so the social welfare becomes $2s'_i$. So

$$\frac{p_i(s'_i,100) - p_i(100,100)}{SW(100,100) - SW(s'_i,100)} = \frac{s'_i - 98}{200 - 2s'_i}$$

The maximum, $\frac{1}{2}$, is reached when $s'_i = 99$. So the claim follows by Theorem 1(iii). □

5.4 War of attrition

This is a strategic game, see, e.g., [14, Section 3.4], with two players $N = \{1,2\}$, strategy set $S_i = \mathbb{R}_+$ for every player i, and payoff function p_i for every i defined as follows, where $v > 0$

$$p_i(s) := \begin{cases} -s_i & \text{if } s_i < s_{-i} \\ \frac{1}{2}v - s_i & \text{if } s_i = s_{-i} \\ v - s_{-i} & \text{otherwise.} \end{cases}$$

Proposition 3. *The selfishness level of the war of attrition game is* 0.

Proof. A joint strategy s is a Nash equilibrium iff either $s_1 = 0$ and $s_2 \geq v$ or $s_2 = 0$ and $s_1 \geq v$. So for each Nash equilibrium s we have $SW(s) = v$, i.e., each Nash equilibrium is a social optimum. □

5.5 Cournot competition

We consider a symmetric oligopoly Cournot competition with the same linear cost function for all players. We assume that each player $i \in N = \{1,\ldots,n\}$ has a strategy set $S_i = \mathbb{R}_+$ and payoff function

$$p_i(s) := s_i\Big(a - b\sum_{j=1}^{n} s_j\Big) - cs_i$$

for some given a,b,c, where $a > c$ and $b > 0$.

The price of the product is represented by the expression $a - b\sum_{j=1}^{n} s_j$ and the production cost corresponding to the production level s_i by cs_i. In what follows we rewrite the payoff function as $p_i(s) := s_i(d - b\sum_{j=1}^{n} s_j)$, where $d := a - c$.

Proposition 4. *The selfishness level of the n-players Cournot competition game is* ∞*.*

Intuitively, this result means that in this game no matter how much we 'involve' the players in sharing the social welfare we cannot achieve that they will select a social optimum.

Proof. We first determine the stable social optima of this game. Fix a joint strategy s and let $t := \sum_{j=1}^{n} s_j$. Then $SW(s) = t(d - bt)$. This expression becomes maximal precisely when $t = \frac{d}{2b}$. So this game has infinitely many social optima and each of them is stable.

Take now a stable social optimum s. So $\sum_{j=1}^{n} s_j = \frac{d}{2b}$. Fix $i \in N$. Let $u := \sum_{j\neq i} s_j$. For every strategy z of player i

$$p_i(z,s_i) = -bz^2 + (d - bu)z$$

and

$$SW(z,s_i) = -bz^2 + (d - 2bu)z + u(d - bu).$$

Denote now s_i by y and consider a strategy x of player i such that $p_i(x,s_{-i}) > p_i(y,s_{-i})$. Then $u + x \neq \frac{d}{2b}$, so $SW(y,s_{-i}) > SW(x,s_{-i})$.

We have

$$\begin{aligned} p_i(x,s_{-i}) - p_i(y,s_{-i}) &= -b(x^2 - y^2) + (d - bu)(x - y) \\ = -b(x-y)(x+y+u-\tfrac{d}{b}) &= -b(x-y)(x - \tfrac{d}{2b}), \end{aligned}$$

where the last equality holds since $u - \frac{d}{b} = -(y + \frac{d}{2b})$ on the account of the equality $u + y = \frac{d}{2b}$.

Further,

$$\begin{aligned} SW(y,s_{-i}) - SW(x,s_{-i}) &= b(x-y)^2 - (d - 2bu)(x - y) \\ = b(x-y)(x+y+2u-\tfrac{d}{b}) &= b(x-y)^2, \end{aligned}$$

where the last equality holds since $2u - \frac{d}{b} = -2y$ on the account of the equality $u + y = \frac{d}{2b}$.

We have $x \neq y$. Hence

$$f(x) := \frac{p_i(x,s_{-i}) - p_i(y,s_{-i})}{SW(y,s_{-i}) - SW(x,s_{-i})} = -\frac{x - \frac{d}{2b}}{x - y} = -1 + \frac{y - \frac{d}{2b}}{y - x}.$$

Since

$$p_i(x,s_{-i}) - p_i(y,s_{-i}) = b(y - x)\Big(x - \frac{d}{2b}\Big)$$

we have $p_i(x,s_{-i}) - p_i(y,s_{-i}) > 0$ iff $y < x < \frac{d}{2b}$ or $y > x > \frac{d}{2b}$. But $y \leq \frac{d}{2b}$, since $u + y = \frac{d}{2b}$. So the conjunction of $p_i(x,s_{-i}) > p_i(y,s_{-i})$ and $SW(x,s_{-i}) > SW(y,s_{-i})$ holds iff $y < x < \frac{d}{2b}$.

Now

$$\max_{y<x<\frac{d}{2b}} f(x) = \infty.$$

But s was an arbitrary stable social optimum, so the claim follows by Theorem 3(i). □

5.6 Tragedy of the commons

Assume that each player $i \in N = \{1,\dots,n\}$ has the real interval $[0,1]$ as its set of strategies. Each player's strategy is his chosen fraction of a common resource. Let (see also [14, Exercise 63.1] and [18, pages 6–7]):

$$p_i(s) := \max\Big(0, s_i\Big(1 - \sum_{j=1}^{n} s_j\Big)\Big).$$

This payoff function reflects the fact that player's enjoyment of the common resource depends positively from his chosen fraction of the resource and negatively from the total fraction of the common resource used by all players. Additionally, if the total fraction of the common resource by all players exceeds a feasible level, here 1, then player's enjoyment of the resource becomes zero.

Proposition 5. *The selfishness level of the n-players Tragedy of the Commons game is* ∞*.*

Proof. We first determine the stable social optima of this game. Fix a joint strategy s and let $t := \sum_{j=1}^{n} s_j$. If $t > 1$, then the social welfare is 0. So assume that $t \leq 1$. Then $SW(s) = t(1-t)$. This expression becomes maximal precisely when $t = \frac{1}{2}$ and then it equals $\frac{1}{4}$. So this game has infinitely many social optima and each of them is stable.

Take now a stable social optimum s. So $\sum_{j=1}^{n} s_j = \frac{1}{2}$. Fix $i \in \{1,\dots,n\}$. Denote s_i by a and consider a strategy x of player i such that $p_i(x,s_{-i}) > p_i(a,s_{-i})$. Then $\sum_{j \neq i} s_j + x \neq \frac{1}{2}$, so $SW(a,s_{-i}) > SW(x,s_{-i})$.

We have $p_i(a,s_{-i}) = \frac{a}{2}$ and $SW(a,s_{-i}) = \frac{1}{4}$. Further, $p_i(x,s_{-i}) > p_i(a,s_{-i})$ implies $\sum_{j \neq i} s_j + x < 1$ and hence

$$p_i(x,s_{-i}) = x\Big(a + \frac{1}{2} - x\Big)$$

and

$$SW(x,s_{-i}) = \Big(\frac{1}{2} - a + x\Big)\Big(1 - \frac{1}{2} + a - x\Big) = \frac{1}{4} - (a-x)^2.$$

Also $x \neq a$. Hence

$$f(x) := \frac{p_i(x,s_{-i}) - p_i(a,s_{-i})}{SW(a,s_{-i}) - SW(x,s_{-i})} = \frac{(a-x)(x-\frac{1}{2})}{(a-x)^2} = \frac{x-\frac{1}{2}}{a-x} = -1 + \frac{a-\frac{1}{2}}{a-x}$$

Since

$$p_i(x,s_{-i}) - p_i(a,s_{-i}) = (a-x)\Big(x - \frac{1}{2}\Big)$$

we have $p_i(x,s_{-i}) > p_i(a,s_{-i})$ iff $a < x < \frac{1}{2}$ or $a > x > \frac{1}{2}$. But $a \leq \frac{1}{2}$, since $\sum_{j \neq i} s_j + a = \frac{1}{2}$. So the conjunction of $p_i(x,s_{-i}) > p_i(a,s_{-i})$ and $SW(x,s_{-i}) < SW(a,s_{-i})$ holds iff $a < x < \frac{1}{2}$.

Now

$$\max_{a < x < \frac{1}{2}} f(x) = \infty.$$

But s was an arbitrary stable social optimum, so the claim follows by Theorem 3(i). □

5.7 Linear congestion games

In a congestion game (see [15]) $G = (N, E, \{S_i\}_{i \in N}, \{d_e\}_{e \in E})$ we are given a set of players $N = \{1,\dots,n\}$, a set of facilities E with a delay function $d_e : \mathbb{N} \to \mathbb{Q}$ for every facility $e \in E$, and a strategy set $S_i \subseteq 2^E$ for every player $i \in N$. For a joint strategy $s \in S_1 \times \cdots \times S_n$, define $x_e(s)$ as the number of players using facility $e \in E$, i.e., $x_e(s) = |\{i \in N : e \in s_i\}|$. The goal of a player is to minimize his individual cost $c_i(s) = \sum_{e \in s_i} d_e(x_e(s))$. The social cost function is given by $SC(s) = \sum_{i=1}^{n} c_i(s)$. In a linear congestion game, the delay function of every facility $e \in E$ is of the form $d_e(x) = a_e x + b_e$, where $a_e, b_e \in \mathbb{Q}_+$ are non-negative rational numbers. Using standard scaling arguments, we can assume without loss of generality that $a_e, b_e \in \mathbb{N}$.

Given a linear congestion game, we define L as the maximum number of facilities that any player can choose, i.e., $L := \max_{i \in N,\ s_i \in S_i} |s_i|$. Moreover, let $\Delta_{\max} := \max_{e \in E}(a_e + b_e)$ and $\Delta_{\min} := \min_{e \in E}(a_e + b_e)$.

Proposition 6. *The selfishness level of a linear congestion game is at most* $\frac{1}{2}(L \cdot \Delta_{\max} - \Delta_{\min} - 1)$.

Note that $|L| \leq |E|$, so the exhibited bound does not depend on the number of players.

Proof. Let s be a stable social optimum. Note that s exists by Theorems 3(*ii*) and 5. Because we consider a cost minimization game here the condition in Theorem 3(i) reads

$$\alpha(s) := \max_{i \in \{1,\dots,n\},\ s'_i \in R(i,s)} \frac{c_i(s_i, s_{-i}) - c_i(s'_i, s_{-i})}{SC(s'_i, s_{-i}) - SC(s_i, s_{-i})}, \tag{2}$$

where

$$R(i,s) := \{s'_i \in S_i \mid c_i(s'_i, s_{-i}) < c_i(s_i, s_{-i}) \text{ and } SC(s'_i, s_{-i}) > SC(s_i, s_{-i})\}.$$

Fix some player i and let $s' = (s'_i, s_{-i})$ for some $s'_i \in R(i,s)$. We use x_e and x'_e to refer to $x_e(s)$ and $x_e(s')$, respectively. Note that

$$x'_e = \begin{cases} x_e + 1 & \text{if } e \in s'_i \setminus s_i, \\ x_e - 1 & \text{if } e \in s_i \setminus s'_i, \\ x_e & \text{otherwise.} \end{cases} \tag{3}$$

Exploiting (3), we obtain

$$\begin{aligned} c_i(s_i, s_{-i}) - c_i(s'_i, s_{-i}) &= \sum_{e \in s_i} (a_e x_e + b_e) - \sum_{e \in s'_i} (a_e x'_e + b_e) \\ &= \sum_{e \in s_i \setminus s'_i} (a_e x_e + b_e) - \sum_{e \in s'_i \setminus s_i} (a_e(x_e + 1) + b_e). \end{aligned}$$

Similarly,

$$\begin{aligned} SC(s'_i, s_{-i}) - SC(s_i, s_{-i}) &= \sum_{e \in E} x'_e(a_e x'_e + b_e) - \sum_{e \in E} x_e(a_e x_e + b_e) \\ &= \sum_{e \in s'_i \setminus s_i} (x_e + 1)(a_e(x_e + 1) + b_e) - x_e(a_e x_e + b_e) \\ &\quad + \sum_{e \in s_i \setminus s'_i} (x_e - 1)(a_e(x_e - 1) + b_e) - x_e(a_e x_e + b_e) \\ &= \sum_{e \in s'_i \setminus s_i} (a_e(2x_e + 1) + b_e) - \sum_{e \in s_i \setminus s'_i} (a_e(2x_e - 1) + b_e). \end{aligned}$$

Given a congestion vector $\mathbf{x} = (x_e)_{e \in E}$, define $P(\mathbf{x}) := \sum_{e \in s_i \setminus s'_i} (a_e x_e + b_e)$ and $Q(\mathbf{x}) := \sum_{e \in s'_i \setminus s_i} (a_e(x_e + 1) + b_e)$. Note that $P(\mathbf{x})$ and $Q(\mathbf{x})$ are integers because $a_e, b_e \in \mathbb{N}$ for every facility $e \in E$. Note that with these definitions, $P(\mathbf{1}) = \sum_{e \in s_i \setminus s'_i} (a_e + b_e)$ and $Q(\mathbf{0}) = \sum_{e \in s'_i \setminus s_i} (a_e + b_e)$. We have

$$\frac{c_i(s_i, s_{-i}) - c_i(s'_i, s_{-i})}{SC(s'_i, s_{-i}) - SC(s_i, s_{-i})} = \frac{P(\mathbf{x}) - Q(\mathbf{x})}{2Q(\mathbf{x}) - Q(\mathbf{0}) - 2P(\mathbf{x}) + P(\mathbf{1})}.$$

Because $s_i' \in R(i,s)$, we know that $P(\mathbf{x}) > Q(\mathbf{x})$ and $2Q(\mathbf{x}) - Q(\mathbf{0}) > 2P(\mathbf{x}) - P(\mathbf{1})$. So we obtain

$$Q(\mathbf{x}) + 1 \leq P(\mathbf{x}) \leq Q(\mathbf{x}) + \frac{1}{2}(P(\mathbf{1}) - Q(\mathbf{0}) - 1).$$

Exploting these inequalities, we obtain

$$\begin{aligned}\frac{P(\mathbf{x}) - Q(\mathbf{x})}{2Q(\mathbf{x}) - Q(\mathbf{0}) - 2P(\mathbf{x}) + P(\mathbf{1})} &\leq Q(\mathbf{x}) + \frac{1}{2}(P(\mathbf{1}) - Q(\mathbf{0}) - 1) - Q(\mathbf{x}) \\ &= \frac{1}{2}(P(\mathbf{1}) - Q(\mathbf{0}) - 1) \\ &= \frac{1}{2}\left(\sum_{e \in s_i \setminus s_i'} (a_e + b_e) - \sum_{e \in s_i' \setminus s_i} (a_e + b_e) - 1 \right) \\ &\leq \frac{1}{2}(|s_i \setminus s_i'| \cdot \Delta_{\max} - |s_i' \setminus s_i| \cdot \Delta_{\min} - 1).\end{aligned}$$

Note that $|s_i' \setminus s_i| \geq 1$; otherwise, $s_i' \subseteq s_i$ and thus $SC(s_i', s_{-i}) \leq SC(s)$ which contradicts $s_i' \in R(i,s)$. The above expression is thus at most

$$\frac{1}{2}(L \cdot \Delta_{\max} - \Delta_{\min} - 1).$$

Because this bound holds for every player i and $s_i' \in R(i,s)$, we conclude by Theorem 3(*iii*) that the selfishness level α is at most $\frac{1}{2}(L \cdot \Delta_{\max} - \Delta_{\min} - 1)$. □

Proposition 6 is tight for certain values of L, $\Delta_{\max}$ and $\Delta_{\min}$. As an example, it yields an upper bound of $n-1$ for $L = 1$, $\Delta_{\max} = 2n$ and $\Delta_{\min} = 1$, which is tight as the following example shows. Consider a symmetric congestion game with $2n$ players and two facilities e_1 and e_2 with delay functions x and $2n$, respectively. Clearly, a socially optimal strategy profile s splits the $2n$ players evenly among the facilities and has cost $SC(s) = n^2 + 2n^2 = 3n^2$. Consider a player i that uses facility e_2. We have $c_i(s) = 2n$. If i switches to facility e_1, we obtain $SC(s_i', s_{-i}) = (n+1)^2 + 2n(n-1) = 3n^2 + 1$ and $c_i(s_i', s_{-i}) = n + 1$. Thus

$$\alpha \geq \frac{c_i(s) - c_i(s_i', s_{-i})}{SC(s_i', s_{-i}) - SC(s)} = n - 1.$$

5.8 Fair cost sharing games

In a fair cost sharing game players allocate facilities and share the cost of the used facilities in a fair manner. Formally, a fair cost sharing game is given by $G = (N, E, \{S_i\}_{i \in N}, \{c_e\}_{e \in E})$, where $N = \{1, \ldots, n\}$ is the set of players, E is the set of facilities, $S_i \subseteq 2^E$ is the set of facility subsets available to player i, and $c_e \in \mathbb{Q}_+$ is the cost of facility $e \in E$. As for congestion games, we let $x_e(s)$ be the number of players using facility $e \in E$ in a joint strategy $s \in S_1 \times \cdots \times S_n$. The cost of a facility $e \in E$ is evenly shared among the players using it. That is, the cost of player is defined as $c_i(s) = \sum_{e \in s_i} c_e / x_e(s)$. The social cost function is given by $SC(s) = \sum_{i=1}^n c_i(s)$. Using standard scaling arguments, we can assume without loss of generality that $c_e \in \mathbb{N}$.

Given a cost sharing game, we define L as the maximum number of facilities that any player can choose, i.e., $L := \max_{i \in N,\ s_i \in S_i} |s_i|$. Moreover, let $c_{\max} := \max_{e \in E} c_e$.

Proposition 7. *The selfishness level of a fair cost sharing game is at most* $\frac{1}{2}L \cdot c_{\max} - 1$.

Note that $|L| \leq |E|$, so the exhibited bound does not depend on the number of players.

Proof. Let s be a stable social optimum. Note that s exists by Theorems 3(ii) and 5. Fix some player i and let $s' = (s'_i, s_{-i})$ for some $s'_i \in R(i,s)$. We use x_e and x'_e to refer to $x_e(s)$ and $x_e(s')$, respectively. It is not difficult to verify that

$$SC(s'_i, s_{-i}) - SC(s_i, s_{-i}) = \sum_{e \in s'_i : x'_e = 1} \frac{c_e}{x'_e} - \sum_{e \in s_i : x_e = 1} \frac{c_e}{x_e}.$$

By definition, we have

$$c_i(s_i, s_{-i}) - c_i(s'_i, s_{-i}) = \sum_{e \in s_i} \frac{c_e}{x_e} - \sum_{e \in s'_i} \frac{c_e}{x'_e}.$$

Thus

$$\frac{c_i(s_i, s_{-i}) - c_i(s'_i, s_{-i})}{SC(s'_i, s_{-i}) - SC(s_i, s_{-i})} = \frac{\sum_{e \in s_i : x_e \geq 2} \frac{c_e}{x_e} - \sum_{e \in s'_i : x'_e \geq 2} \frac{c_e}{x'_e}}{SC(s'_i, s_{-i}) - SC(s_i, s_{-i})} - 1.$$

Note that the denominator is at least 1 because $s'_i \in R(i,s)$ and each c_e belongs to $\mathbb{N}$. We conclude

$$\frac{c_i(s_i, s_{-i}) - c_i(s'_i, s_{-i})}{SC(s'_i, s_{-i}) - SC(s_i, s_{-i})} \leq \sum_{e \in s_i : x_e \geq 2} \frac{c_e}{x_e} - 1 \leq \frac{1}{2} L \cdot c_{\max} - 1.$$

The claim follows by Theorem 3(iii). □

6 Conclusions

We presented in this paper a new discrepancy measure between the social welfare in a Nash equilibrium and in a social optimum, that we call the selfishness level. In contrast to the concepts of price of anarchy and price of stability this measure is normative in that it indicates by what fraction of altruism the original game needs to be modified to achieve a desired situation.

The proposed measure can be also used for other games and for other solution concepts, for instance extensive games and subgame perfect equilibria. As an example consider the six-period version of the centipede game (see, e.g., [14]) depicted in Figure 1.

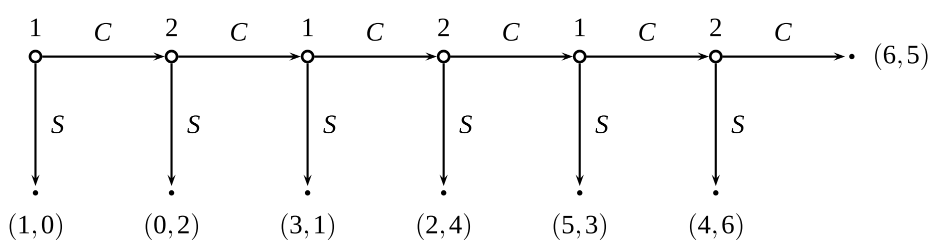

Figure 1: A centipede game.

In its unique subgame perfect equilibrium each player chooses S in every period and the resulting payoffs are $(1,0)$. Since $5 + (6+5)\alpha \geq 6 + (4+6)\alpha$ holds iff $\alpha \geq 1$, we can conclude that the (appropriately adapted) selfishness level for this game is 1.

Acknowledgement

We acknowledge initial discussions with Po-An Chen.

References

[1] R. Axelrod. *The Evolution of Cooperation*. Basic Books, 1984.

[2] K. Basu. The traveler's dilemma: paradoxes of rationality in game theory. *American Economic Review*, 84(2):391–395, 1994.

[3] K. Binmore. *Playing for Real: A Text on Game Theory*. Oxford University Press, Oxford, 2007.

[4] S. Bowles. *Microeconomics: Behavior, Institutions, and Evolution*. Princeton University Press, Princeton, 2004.

[5] I. Caragiannis, C. Kaklamanis, P. Kanellopoulos, M. Kyropoulou, and E. Papaioannou. The impact of altruism on the efficiency of atomic congestion games. In *Proc. 5th Symposium on Trustworthy Global Computing*, 2010.

[6] P.-A. Chen, B. de Keijzer, D. Kempe, and G. Schäfer. Robust price of anarchy for atomic games with altruistic players. Technical Report, CWI Amsterdam, 2011. Available from `http://oai.cwi.nl/oai/asset/18084/18084D.pdf`.

[7] P.-A. Chen and D. Kempe. Altruism, selfishness, and spite in traffic routing. In *Proc. 10th ACM Conference on Electronic Commerce*, pages 140–149, 2008.

[8] J. Elias, F. Martignon, K. Avrachenkov, and G. Neglia. Socially-aware network design games. In *Proc. INFOCOM 2010*, pages 41–45, 2010.

[9] M. Hoefer and A. Skopalik. Altruism in atomic congestion games. In *Proc. 17th European Symposium on Algorithms*, 2009.

[10] E. Koutsoupias and C. H. Papadimitriou. Worst case equilibria. In *Annual IEEE Symposium on Theoretical Aspects of Computer Science*, pages 404–413, 1999.

[11] J. Ledyard. Public goods: A survey of experimental resesarch. In J. Kagel and A. Roth, editors, *Handbook of Experimental Economics*, pages 111–194. Princeton University Press, 1997.

[12] G. D. Marco and J. Morgan. Slightly altruistic equilibria in normal form games. Working paper No 185, Center for Studies in Economics and Finance, University of Salerno, Italy, 2007. Available from `http://www.csef.it/WP/wp185.pdf`.

[13] D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behaviour*, 14:124–143, 1996.

[14] M. J. Osborne. *An Introduction to Game Theory*. Oxford University Press, Oxford, 2005.

[15] R. W. Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, (2):65–67, 1973.

[16] T. Roughgarden. Intrinsic robustness of the price of anarchy. In *Proc. 41st Annual ACM Symposium on Theory of Computing*, pages 513–522, 2009.

[17] A. S. Schulz and N. E. S. Moses. On the performance of user equilibria in traffic networks. In *SODA*, pages 86–87, 2003.

[18] É. Tardos and V. J. Vazirani. Basic solution concepts and computational issues. In N. Nisan, T. Roughgarden, É. Tardos, and V. J. Vazirani, editors, *Algorithmic Game Theory*, chapter 1, pages 3–28. Cambridge University Press, 2007.