Selfishness Level of Strategic Games (Short Version)

Krzysztof R. Apt* Guido Schäfer†

May 1, 2013

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 offered to each player to achieve that a social optimum is realized in a pure Nash equilibrium. The selfishness level is unrelated to the price of stability and the price of anarchy and in contrast to these notions is invariant under positive linear transformations of the payoff functions. Also, it naturally applies to other solution concepts and other forms of games.

We study the selfishness level of several well-known strategic games. This allows us to quantify the implicit tension within a game between players' individual interests and the impact of their decisions on the society as a whole. Our analysis reveals that the selfishness level often provides more refined insights into the game than other measures of inefficiency, such as the price of stability or the price of anarchy.

In particular, the selfishness level of finite games that have a generalized ordinal potential games is finite, while that of weakly acyclic games can be infinite. We derive explicit bounds on the selfishness level of fair cost sharing games and linear congestion games, which depend on specific parameters of the underlying game but are independent of the number of players. Further, we show that the selfishness level of the n-players Prisoner's Dilemma is $1/(2n-3)$, that of the n-players public goods game is $(1-\frac{c}{n})(c-1)$, where c is the public good multiplier, and that of the Traveler's Dilemma game is $\frac{1}{2}$. Finally, the selfishness level of Cournot competition (an example of an infinite potential game), Tragedy of the Commons, and Bertrand competition is infinite.

*Centre for Mathematics and Computer Science (CWI) and University of Amsterdam, The Netherlands. Address: Centrum Wiskunde & Informatica, Science Park 123, 1098 XG, Amsterdam, the Netherlands. Email: k.r.apt@cwi.nl.

†Centre for Mathematics and Computer Science (CWI) and VU University Amsterdam, The Netherlands. Address: Centrum Wiskunde & Informatica, Science Park 123, 1098 XG, Amsterdam, the Netherlands. Email: g.schaefer@cwi.nl.

1 Selfishness level

1.1 Definition

A ***strategic game*** (in short, a game) $G = (N, \{S_i\}_{i\in N}, \{p_i\}_{i\in N})$ is given by a set $N = \{1, \dots, n\}$ of $n > 1$ 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 $p_i : S_1 \times \cdots \times S_n \to \mathbb{R}$. The players choose their strategies simultaneously and 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 $s = (s_1, \dots, s_n)$.

We call $s \in S_1 \times \cdots \times S_n$ a ***joint strategy***, denote its ith element by s_i, denote $(s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n)$ by s_{-i} and similarly with S_{-i}. Further, we write (s'_i, s_{-i}) for $(s_1, \dots, s_{i-1}, s'_i, s_{i+1}, \dots, s_n)$, 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 strategic game $G = (N, \{S_i\}_{i\in N}, \{p_i\}_{i\in N})$ is ***symmetric*** if all players have the same set of strategies and the payoff for playing a particular strategy only depends on the strategies played by the other players (but not on their identities); more formally, $S_i = S_j$ for every $i, j \in N$, $i \neq j$, and for every joint strategy $s = (s_1, \dots, s_n)$, for every $i \in N$ and every permutation π of $\{1, \dots, n\}$, we have $p_i(s_1, \dots, s_n) = p_{\pi(i)}(s_{\pi(1)}, \dots, s_{\pi(n)})$.

A joint strategy s is a ***Nash equilibrium*** if for all $i \in \{1, \dots, n\}$ 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 $SW(s) := \sum_{i=1}^{n} p_i(s)$ the ***social welfare*** of s. When the social welfare of s is maximal we call s a ***social optimum***.

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 pure Nash equilibrium of $G(\alpha)$ is a social optimum of $G(\alpha)$. Then we say that G is α***-selfish***. We define the selfishness level of G as

$$\inf\{\alpha \in \mathbb{R}_+ \mid G \text{ is } \alpha\text{-selfish}\}. \tag{1}$$

Here we adopt the convention that the infimum of an empty set is ∞. Further, we stipulate that the selfishness level of G is denoted by α^+ iff the selfishness level of G is $\alpha \in \mathbb{R}_+$ but G is *not* α-selfish (equivalently, the infimum does not belong to the set). We show below (Theorem 2) that pathological infinite games exist for which the selfishness level is of this kind; none of the other studied games is of this type.

We give some remarks before we proceed.

1. 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 apply similarly 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 $SC(s) := \sum_{i=1}^{n} c_i(s)$.

2. Other definitions of an altruistic version of a game are conceivable and, depending on the underlying application, might seem more natural than the one we use here.

However, we show in Section 1.3 that our definition is equivalent to several other models of altruism that have been proposed in the literature.

3. The selfishness level refers to the smallest α such that *some* Nash equilibrium in $G(\alpha)$ is also a social optimum. Alternatively, one might be interested in the smallest α such that *every* Nash equilibrium in $G(\alpha)$ corresponds to a social optimum. However, as argued in Section 1.2, this alternative notion is not very meaningful.

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.

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 various 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 ∞.

Example 4. Game with a bad Nash equilibrium The following game results from equipping each player in the Matching Pennies game with a third strategy E (for edge):

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

Its unique Nash equilibrium is (E, E). It is easy to check that the selfishness level of this game is ∞. (This is also an immediate consequence of Theorem 4 (iii) below.)

Example 5. Game with no Nash equilibrium Consider a game G on the left and the resulting game $G(\alpha)$ for $\alpha = 1$ on the right.

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

	C	D
C	$6, 6$	$4, 2$
D	$6, 3$	$3, 3$

The game G has no Nash equilibrium, while in the game $G(1)$ the social optimum, (C, C), is also a Nash equilibrium. As in the Prisoner's Dilemma game 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 the game G is 1.

1.2 Properties

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. The ***price of anarchy*** is defined as the ratio $SW(s)/SW(s')$ where s is a social optimum and s' is a Nash equilibrium with the lowest social welfare in G.

So the price of stability of G is 1 iff its selfishness level 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, which we denote by $PD(\alpha, \beta)$, of the Prisoner's Dilemma game G with $x = \frac{\alpha}{\alpha+1}$:

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

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 defined the selfishness level of a game as the smallest α such that the price of stability of $G(\alpha)$ is 1. Alternatively, one might want to define the selfishness level as the smallest α such that the price of anarchy of $G(\alpha)$ is 1. But as it turns out, this notion is not very informative. Namely, as noted already above, the social welfare of a joint strategy s in $G(\alpha)$ equals $(1+\alpha n)SW(s)$. So if for two joint strategies s and s' we have $SW(s) < SW(s')$ in G, then this inequality remains valid when using the social welfare in $G(\alpha)$. Hence, for no α the joint strategy s can be a social optimum in $G(\alpha)$. In particular, this also holds if s and s' are Nash equilibria. Thus, with this alternative definition, the only possible levels would be 0 (when all Nash equilibria of G have the same social welfare) or ∞ (otherwise).

Further, in contrast to the price of stability and the price of anarchy the notion of the selfishness level is invariant under simple payoff transformations. It is a direct consequence of the following observation, where given a game G and a value a we denote by $G+a$ (respectively, aG) the game obtained from G by adding to each payoff function the value a (respectively, by multiplying each payoff function by a).

Proposition 1. *Consider a game G and $\alpha \geq 0$.*

(i) For every a, G is α-selfish iff $G + a$ is α-selfish.

(ii) For every $a > 0$, G is α-selfish iff aG is α-selfish.

Proof. (i) It suffices to note that $r[a]_i(s) = r_i(s) + \alpha an + a$, where r_i and $r[a]_i$ are the payoff functions of player i in the games $G(\alpha)$ and $(G+a)(\alpha)$. So for every joint strategy s

- s is a Nash equilibrium of $G(\alpha)$ iff it is a Nash equilibrium of $(G + a)(\alpha)$,
- s is social optimum of $G(\alpha)$ iff it is a social optimum of $(G + a)(\alpha)$.

(ii) It suffices to note that for every $a > 0$, $r[a]_i(s) = ar_i(s)$, where this time $r[a]_i$ is the payoff function of player i in the game $(aG)(\alpha)$, and argue as above. □

In particular, for symmetric games Proposition 1 implies that the selfishness level is invariant under affine transformations of the payoff functions.

Note that the selfishness level is not invariant under a multiplication of the payoff functions by a value $a \leq 0$. Indeed, for $a = 0$ each game aG has the selfishness level 0. For $a < 0$ take the game G from Example 4 whose selfishness level is ∞. In the game aG the joint strategy (E, E) is both a Nash equilibrium and a social optimum, so the selfishness level of aG is 0.

The above proposition also allows us to frame the notion of selfishness level in the following way. Suppose the original n-player game G is given to a game designer who has a fixed budget of $SW(s)$ for each joint strategy s and that the selfishness level of G is $\alpha < \infty$. How should the game designer then distribute the budget of $SW(s)$ for each joint strategy s among the players such that the resulting game has a Nash equilibrium that coincides with a social optimum? By scaling $G(\alpha)$ by the factor $a := 1/(1 + \alpha n)$ we ensure that for each joint strategy s its social welfare in the original game G and in $aG(\alpha)$ is the same. Using Proposition 1, we conclude that α is the smallest non-negative real such that $aG(\alpha)$ has a Nash equilibrium that is a social optimum. The game $aG(\alpha)$ can then be viewed as the intended transformation of G. That is, each payoff function p_i of the game G is transformed into the payoff function

$$r_i(s) := \frac{1}{1+\alpha n} p_i(s) + \frac{\alpha}{1+\alpha n} SW(s).$$

Let us return now to the 'borderline case' of the selfishness level that we denoted by α^+. We have the following result.

Theorem 2. *For every $\alpha \geq 0$ there exists a game whose selfishness level is α^+.*

Proof. We first prove the result for $\alpha = 0$. That is, we show that there exists a game that is α-selfish for every $\alpha > 0$, but is not 0-selfish. To this end we use the games $PD(\alpha, \beta)$ defined in the proof of Theorem 1.

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 $\alpha > 0$.

To deal with the general case we prove two claims that are of independent interest.

Claim 1. *For every game G and $\alpha \geq 0$ there is a game G' such that $G'(\alpha) = G$.*

Proof. We define the payoff of player i in the game G' by

$$p'_i(s) := p_i(s) - \frac{\alpha}{1+n\alpha} SW(s),$$

where p_i is his payoff in the game G. Denote by $SW'(s)$ the social welfare of a joint strategy s in the game G' and by r'_i the payoff function of player i in the game $G'(\alpha)$. Then

$$\begin{aligned} r'_i(s) &= p'_i(s) + \alpha SW'(s) \\ &= p_i(s) - \frac{\alpha}{1+n\alpha} SW(s) + \alpha \left(SW(s) - \frac{n\alpha}{1+n\alpha} SW(s) \right) \\ &= p_i(s) + \left(\alpha - \frac{\alpha}{1+n\alpha} - \frac{n\alpha^2}{1+n\alpha} \right) SW(s) \\ &= p_i(s). \end{aligned}$$

□

Claim 2. *For every game G and $\alpha, \beta \geq 0$*

$$G(\alpha + \beta) = G(\alpha)\left(\frac{\beta}{1 + n\alpha}\right).$$

Proof. Denote by $SW'(s)$ the social welfare of a joint strategy s in the game $G(\alpha)$, by p_i, r_i and r' the payoff functions of player i in the games G, $G(\alpha)$, and $G(\alpha)(\frac{\beta}{1+n\alpha})$. Then

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

so

$$\begin{aligned} r_i'(s) &= r_i(s) + \frac{\beta}{1 + n\alpha} SW'(s) \\ &= p_i(s) + \alpha SW(s) + \frac{\beta}{1 + n\alpha}(SW(s) + n\alpha SW(s)) \\ &= p_i(s) + \left(\alpha + \frac{\beta}{1 + n\alpha} + \frac{\beta n \alpha}{1 + n\alpha}\right) SW(s) \\ &= p_i(s) + (\alpha + \beta) SW(s), \end{aligned}$$

which proves the claim. □

To prove the general case fix $\alpha \geq 0$ and $\beta > 0$ and take a game G whose selfishness level is 0^+. By Claim 1 there is a game G' such that $G'(\alpha) = G$. Then G' is not α-selfish, since G is not 0-selfish.

Further, by Claim 2

$$G'(\alpha + \beta) = G'(\alpha)\left(\frac{\beta}{1 + n\alpha}\right) = G\left(\frac{\beta}{1 + n\alpha}\right).$$

But by its choice the game G is $\frac{\beta}{1+n\alpha}$-selfish, so G' is $\alpha + \beta$-selfish, which concludes the proof. □

1.3 Alternative definitions

Our definition of the selfishness level depends on the way the altruistic versions of the original game are defined. Three other models of altruism were proposed in the literature. As before, let $G := (N, \{S_i\}_{i \in N}, \{p_i\}_{i \in N})$ be a strategic game. Consider the following four definitions of altruistic versions of G:

Model A ([6]): For every $\alpha \geq 0$, $G(\alpha) := (N, \{S_i\}_{i \in N}, \{r_i^\alpha\}_{i \in N})$ with

$$r_i^\alpha(s) = p_i(s) + \alpha SW(s) \quad \forall i \in N. \tag{2}$$

Model B ([5]): For every $\beta \in [0, 1]$, $G(\beta) := (N, \{S_i\}_{i \in N}, \{r_i^\beta\}_{i \in N})$ with

$$r_i^\beta(s) = (1 - \beta) p_i(s) + \frac{\beta}{n} SW(s) \quad \forall i \in N. \tag{3}$$

Model C ([4]): For every $\gamma \in [0,1]$, $G(\gamma) := (N, \{S_i\}_{i\in N}, \{r_i^\gamma\}_{i\in N})$ with

$$r_i^\gamma(s) = (1-\gamma)p_i(s) + \gamma SW(s) \quad \forall i \in N. \tag{4}$$

Model D ([3]): For every $\delta \in [0,1]$, $G(\delta) := (N, \{S_i\}_{i\in N}, \{r_i^\delta\}_{i\in N})$ with

$$r_i^\gamma(s) = (1-\delta)p_i(s) + \delta(SW(s) - p_i(s)) \quad \forall i \in N. \tag{5}$$

Our selfishness level notion for Model A extends to Models B, C and D in the obvious way: We say that G is β-selfish for some $\beta \in [0,1]$ iff a pure Nash equilibrium of the altruistic version $G(\beta)$ is also a social optimum. The selfishness level of G with respect to Model B is then defined as the infimum over all $\beta \in [0,1]$ such that G is β-selfish. The respective notions for Models C and D are defined analogously.

The following theorem shows that the selfishness level of a game with respect to Models A, B, C and D relate to each other via simple transformations. (Note that for Model D this transformation only applies for $\delta \in [0, \frac{1}{2}]$.)

Theorem 3. *Consider a strategic game $G := (N, \{S_i\}_{i\in N}, \{p_i\}_{i\in N})$ and its altruistic versions defined according to Models A, B, C and D above.*

(i) G is α-selfish with $\alpha \in \mathbb{R}_+$ iff G is β-selfish with $\beta = \frac{\alpha n}{1+\alpha n} \in [0,1]$.

(ii) G is α-selfish with $\alpha \in \mathbb{R}_+$ iff G is γ-selfish with $\gamma = \frac{\alpha}{1+\alpha} \in [0,1]$.

(iii) G is α-selfish with $\alpha \in \mathbb{R}_+$ iff G is δ-selfish with $\delta = \frac{\alpha}{1+2\alpha} \in [0, \frac{1}{2}]$.

Proof. We prove the following more general claim. Fix $x, y > 0$. For every $\lambda \in [0, \frac{1}{x}]$, define $G(\lambda) := (N, \{S_i\}_{i\in N}, \{r_i^\lambda\}_{i\in N})$ with

$$r_i^\lambda(s) = (1 - x\lambda)p_i(s) + \frac{\lambda}{y} SW(s). \tag{6}$$

We show that G is α-selfish for $\alpha \geq 0$ iff G is λ-selfish for $\lambda = \frac{\alpha y}{1+\alpha xy} \in [0, \frac{1}{x}]$.

By substituting $\lambda = \frac{\alpha y}{1+\alpha xy}$ in (6), we obtain

$$r_i^\lambda(s) = \frac{1}{1+\alpha xy} p_i(s) + \frac{\alpha}{1+\alpha xy} SW(s) = \frac{1}{1+\alpha xy} r_i^\alpha(s).$$

As a consequence, since $\frac{1}{1+\alpha xy} > 0$ for every $\alpha \geq 0$ the pure Nash equilibria and social optima, respectively, of $G(\lambda)$ and $\frac{1}{1+\alpha xy} G(\alpha)$ coincide. Thus, G is λ-selfish iff $\frac{1}{1+\alpha xy} G$ is α-selfish. Also, it follows from Proposition 1 that $\frac{1}{1+\alpha xy} G$ is α-selfish iff G is α-selfish.

Further, note that

$$\lim_{\alpha\to\infty} \frac{\alpha y}{1+\alpha xy} = \frac{1}{x}\left(1 - \lim_{\alpha\to\infty} \frac{1}{1+\alpha xy}\right) = \frac{1}{x}.$$

That is, the selfishness level of G with respect to Model A is ∞ iff the selfishness level of G with respect to $G(\lambda)$ is $\frac{1}{x}$.

Now, (i) follows from the above with $x = 1$ and $y = n$, (ii) follows with $x = y = 1$ and (iii) follows with $x = 2$ and $y = 1$. □

2 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.

It will turn out that to determine the selfishness level of a game we need to consider deviations from its stable social optima. Consider a deviation s'_i of player i from a stable social optimum s. If player i is better off by deviating to s'_i, then by definition the social welfare decreases, i.e., $SW(s_i, s_{-i}) - SW(s'_i, s_{-i}) > 0$. If this decrease is small, while the gain for player i is large, then strategy s'_i is an attractive and socially acceptable option for player i. We define player i's ***appeal factor*** of strategy s'_i given the social optimum s as

$$AF_i(s'_i, s) := \frac{p_i(s'_i, s_{-i}) - p_i(s_i, s_{-i})}{SW(s_i, s_{-i}) - SW(s'_i, s_{-i})}.$$

In what follows we shall characterize the selfishness level in terms of bounds on the appeal factors of profitable deviations from a stable social optimum. First, note the following properties of social optima.

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 $s'_i \in U_i(s)$, $\alpha \geq AF_i(s'_i, s)$, where

$$U_i(s) := \{s'_i \in S_i \mid p_i(s'_i, s_{-i}) > p_i(s_i, s_{-i})\}. \tag{7}$$

The set $U_i(s)$, with the "$>$" sign replaced by "$\geq$", is called an *upper contour set*, see, e.g., [12, page 193]. Note that if s is a stable social optimum, then $s'_i \in U_i(s)$ implies that $SW(s_i, s_{-i}) > SW(s'_i, s_{-i})$.

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 $s'_i \in S_i$

$$p_i(s_i, s_{-i}) + \alpha SW(s_i, s_{-i}) \geq p_i(s'_i, s_{-i}) + \alpha SW(s'_i, s_{-i}). \tag{8}$$

If $p_i(s_i, s_{-i}) \geq p_i(s'_i, s_{-i})$, then (8) 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 (8) holds for all $i \in N$ and $s'_i \in S_i$ iff

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

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

This leads us to the following result.

Theorem 4. *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* $\alpha(s) := \max_{i \in N, s'_i \in U_i(s)} AF_i(s'_i, s)$ *is finite.*

(ii) If the selfishness level of G *is finite, then it equals* $\min_{s \in SSO} \alpha(s)$*, where* SSO *is the set of stable social optima.*

(iii) 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.*

(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 Lemma 1, and (iii) by (i). □

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 5. *For each function* $f : \mathbb{N} \to \mathbb{R}_+$ *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 4 (ii) the selfishness level equals $f(n)$. □

3 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 [10]) for which the selfishness level is finite.

3.1 Potential games

Given a game $G := (N, \{S_i\}_{i \in N}, \{p_i\}_{i \in N})$, a function $P : S_1 \times \cdots \times S_n \rightarrow \mathbb{R}$ is called

- an ***exact potential*** for G if for all $i \in N$, $s_{-i} \in S_{-i}$ and $s_i, s'_i \in S_i$,

$$p_i(s_i, s_{-i}) - p_i(s'_i, s_{-i}) = P(s_i, s_{-i}) - P(s'_i, s_{-i}),$$

- a ***generalized ordinal potential*** for G if for all $i \in N$, $s_{-i} \in S_{-i}$ and $s_i, s'_i \in S_i$,

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

A game that possesses a generalized ordinal potential is called a ***generalized ordinal potential game***.

Theorem 6. *Every finite generalized 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 4 (ii). □

In particular, every finite congestion game has a finite selfishness level as by the result of [13] these games have an exact potential.

We shall derive explicit bounds for two special cases of these games in Sections 3.3 and 3.4.

3.2 Weakly acyclic games

Following [10], given a game $G := (N, \{S_i\}_{i \in N}, \{p_i\}_{i \in N})$, a ***path*** in $S_1 \times \cdots \times S_n$ is a sequence $(s^1, s^2, \ldots)$ of joint strategies such that for every $k > 1$ there is a player i such that $s^k = (s'_i, s^{k-1}_{-i})$ for some $s'_i \neq s^{k-1}_i$. A path is called an ***improvement path*** if it is maximal and for all $k > 1$, $p_i(s^k) > p_i(s^{k-1})$, where i is the player who deviated from s^{k-1}. A game G has the ***finite improvement property*** (***FIP***) if every improvement path is finite. Following [9, 15], a game G is called ***weakly acyclic*** if for every joint strategy there exists a finite improvement path that starts at it.

By the result of [10] finite games that have the FIP coincide with the games that have a generalized ordinal potential. So by Theorem 6 these games have a finite selfishness level. In contrast, the selfishness level of a weakly acyclic game can be infinite. Indeed, the following game is easily seen to be weakly acyclic:

	H	T	E
H	$1, -1$	$-1, \ \ 1$	$-1, -0.5$
T	$-1, \ \ 1$	$1, -1$	$-1, -0.5$
E	$-0.5, -1$	$-0.5, -1$	$-0.5, -0.5$

Yet, on the account of Theorem 4 (iii), its selfishness level is infinite.

3.3 Fair cost sharing games

In a fair cost sharing game, see, e.g., [1], 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 $N = \{1, \dots, n\}$ 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 $c_e \in \mathbb{R}_+$ is the cost of facility $e \in E$. It is called a *singleton* cost sharing game if for every $i \in N$ and for every $s_i \in S_i$: $|s_i| = 1$. For a joint strategy $s \in S_1 \times \cdots \times S_n$ let $x_e(s)$ be the number of players using facility $e \in E$, i.e., $x_e(s) = |\{i \in N \mid e \in s_i\}|$. The cost of a facility $e \in E$ is evenly shared among the players using it. That is, the cost of player i is defined as $c_i(s) = \sum_{e \in s_i} c_e / x_e(s)$. The social cost function is given by $SC(s) = \sum_{i \in N} c_i(s)$.

We first consider singleton cost sharing games. Let $c_{\max} = \max_{e \in E} c_e$ and $c_{\min} = \min_{e \in E} c_e$ refer to the maximum and minimum costs of the facilities, respectively.

Proposition 2. *The selfishness level of a singleton cost sharing game is at most* $\max\{0, \frac{1}{2} \frac{c_{\max}}{c_{\min}} - 1\}$*. Moreover, this bound is tight.*

3.4 Linear congestion games

In a congestion game $G := (N, E, \{S_i\}_{i \in N}, \{d_e\}_{e \in E})$ we are given a set of players $N = \{1, \dots, n\}$, a set of facilities E with a delay function $d_e : \mathbb{N} \to \mathbb{R}_+$ 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., $x_e(s) = |\{i \in N \mid e \in s_i\}|$. 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 $SC(s) = \sum_{i \in N} c_i(s)$. Here we call a congestion game *symmetric* if there is some common strategy set $S \subseteq 2^E$ such that $S_i = S$ for all i. It is *singleton* if every strategy $s_i \in S_i$ is a singleton set, i.e., for every $i \in N$ and for every $s_i \in S_i$, $|s_i| = 1$. 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 $a_e, b_e \in \mathbb{R}_+$ are non-negative real numbers.

We first derive a bound on the selfishness level for symmetric singleton linear congestion games. As it turns out, a bound similar to the one for singleton cost sharing games does not extend to symmetric singleton linear congestion games. Instead, the crucial insight here is that the selfishness level depends on the *discrepancy* between facilities in a stable social optimum. We make this notion more precise.

Let s be a stable social optimum and let x_e refer to $x_e(s)$. Define the *discrepancy* between two facilities e and e' with $a_e + a_{e'} > 0$ under s as

$$\delta(x_e, x_{e'}) = \frac{2a_e x_e + b_e}{a_e + a_{e'}} - \frac{2a_{e'} x_{e'} + b_{e'}}{a_e + a_{e'}}. \tag{9}$$

We show below that $\delta(x_e, x_{e'}) \in [-1, 1]$. Define $\delta_{\max}(s)$ as the maximum discrepancy between any two facilities e and e' under s with $a_e + a_{e'} > 0$ and $\delta(x_e, x_{e'}) < 1$; more formally, let

$$\delta_{\max}(s) = \max_{e, e' \in E} \{\delta(x_e, x_{e'}) \mid a_e + a_{e'} > 0 \text{ and } \delta(x_e, x_{e'}) < 1\}.$$

Let $\delta_{\max}$ be the maximum discrepancy over all stable social optima, i.e., $\delta_{\max} = \max_{s \in SSO} \delta_{\max}(s)$. Further, let $\Delta_{\max} := \max_{e \in E}(a_e + b_e)$ and $\Delta_{\min} := \min_{e \in E}(a_e + b_e)$. Moreover, let $a_{\min}$ be the minimum non-zero coefficient of a latency function, i.e., $a_{\min} = \min_{e \in E: a_e > 0} a_e$.

Proposition 3. *The selfishness level of a symmetric singleton linear congestion game is at most*

$$\max\left\{0, \frac{1}{2}\frac{\Delta_{\max} - \Delta_{\min}}{(1 - \delta_{\max})a_{\min}} - \frac{1}{2}\right\}.$$

Moreover, this bound is tight.

Proposition 4. *The selfishness level of a linear congestion game with non-negative integer coefficients is at most* $\max\{0, \frac{1}{2}(L\Delta_{\max} - \Delta_{\min} - 1)\}$. *Moreover, this bound is tight.*

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

Intuitively, this means that when the number of players in the Prisoner's Dilemma game increases, a smaller share of the social welfare is needed to resolve the underlying conflict. That is, its 'acuteness' diminishes with the number of players. The formal reason is that the appeal factor of each unilateral deviation from the social optimum is inversely proportional to the number of players.

Proof. In this game $s = \mathbf{1}$ is the unique social optimum, with for each $i \in N$, $p_i(s) = 2(n-1)$ and $SW(s) = 2n(n-1)$. Consider now the joint strategy (s'_i, s_{-i}) in which player i deviates to the strategy $s'_i = 0$. We have then $p_i(s'_i, s_{-i}) = 2(n-1) + 1$ and $SW(s'_i, s_{-i}) = 2(n-1) + 1 + 2(n-1)(n-2)$. Hence $AF_i(s'_i, s) = \frac{1}{2n-3}$. The claim now follows by Theorem 4 (ii). □

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.

3.6 Public goods

We consider the public goods game with n players. Every player $i \in N = \{1, \ldots, n\}$ chooses an amount $s_i \in [0, b]$ that he contributes to a public good, where $b \in \mathbb{R}_+$ is the budget. The game designer collects the individual contributions of all players, multiplies their sum by $c > 1$ and distributes the resulting amount evenly among all players. The payoff of player i is thus $p_i(s) := b - s_i + \frac{c}{n}\sum_{j \in N} s_j$.

Proposition 6. *The selfishness level of the n-players public goods game is* $\max\left\{0, \frac{1-\frac{c}{n}}{c-1}\right\}$.

In this game, every player has an incentive to "free ride" by contributing 0 to the public good (which is a dominant strategy). This is exactly as in the n-players Prisoner's Dilemma game. However, the above proposition reveals that for fixed c, in contrast to the Prisoner's Dilemma game, this temptation becomes stronger as the number of players increases. Also, for a fixed number of players this temptation becomes weaker as c increases.

Proof of Proposition 6. Note that $SW(s) = bn + (c-1)\sum_{i \in N} s_i$. The unique social optimum of this game is therefore $s = \mathbf{b}$ with $p_i(s) = cb$ for every $i \in N$ and $SW(s) = cbn$. Suppose player i deviates from s by choosing $s'_i \in [0, b)$. Then $p_i(s'_i, s_{-i}) = cb + (1 - \frac{c}{n})(b - s'_i)$. Thus,

$$p_i(s'_i, s_{-i}) - p_i(s) = (1 - \tfrac{c}{n})(b - s'_i) \quad \text{and} \quad SW(s) - SW(s'_i, s_{-i}) = (c-1)(b - s'_i).$$

If $1 - \frac{c}{n} \leq 0$ then $U_i(s) = \emptyset$ and the selfishness level is zero. Otherwise, $1 - \frac{c}{n} > 0$ and $U_i(s) = [0, b)$. We conclude that in this case $AF_i(s'_i, s) = (1 - \frac{c}{n})/(c-1)$ for every $s'_i \in U_i(s)$. The claim now follows by Theorem 4 (ii). □

3.7 Traveler's dilemma

This is a strategic game discussed in [2] with two players $N = \{1, 2\}$, strategy set $S_i = \{2, \ldots, 100\}$ 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 7. *The selfishness level of the Traveler's Dilemma game is* $\frac{1}{2}$.

Proof. The unique social optimum of this game is $s = (100, 100)$, while $(2, 2)$ is its unique Nash equilibrium. If player i deviates from s 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 $AF_i(s'_i, s) = (s'_i - 98)/(200 - 2s'_i)$. The maximum, $\frac{1}{2}$, is reached when $s'_i = 99$. So the claim follows by Theorem 4 (ii). □

3.8 Tragedy of the Commons

Assume that each player $i \in N = \{1, \ldots, n\}$ 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 [11, Exercise 63.1] and [14, pages 6–7]): $p_i(s) := \max(0,\ s_i(1 - \sum_{j \in N} s_j))$. 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 8. *The selfishness level of the n-players Tragedy of the Commons 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 $t := \sum_{j \in N} s_j$. 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 $\sum_{j \in N} s_j = \frac{1}{2}$. Fix $i \in \{1, \ldots, n\}$. 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

$$p_i(x, s_{-i}) = x(a + \tfrac{1}{2} - x) \quad \text{and} \quad SW(x, s_{-i}) = (\tfrac{1}{2} - a + x)(1 - \tfrac{1}{2} + a - x) = \tfrac{1}{4} - (a-x)^2.$$

Also $x \neq a$. Hence

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

Since $p_i(x, s_{-i}) - p_i(a, s_{-i}) = (a-x)(x - \frac{1}{2})$ 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 $\max_{a < x < \frac{1}{2}} AF_i(x, s) = \infty$. But s was an arbitrary stable social optimum, so the claim follows by Theorem 4 (i). □

3.9 Cournot competition

We consider Cournot competition for n firms with a linear inverse demand function and constant returns to scale, see, e.g., [8, pages 174–175]. So we assume that each player $i \in N = \{1, \ldots, n\}$ has a strategy set $S_i = \mathbb{R}_+$ and payoff function $p_i(s) := s_i(a - b \sum_{j \in N} s_j) - cs_i$ for some given a, b, c, where $a > c \geq 0$ and $b > 0$.

The price of the product is represented by the expression $a - b \sum_{j \in N} s_j$ and the production cost corresponding to the production level s_i by cs_i. In what follows we rewrite the payoff function as $p_i(s) := s_i(d - b \sum_{j \in N} s_j)$, where $d := a - c$. Note that the payoffs can be negative, which was not the case in the tragedy of the commons game. Still the proofs are very similar for both games.

Proposition 9. *The selfishness level of the n-players Cournot competition game is ∞.*

Proof. We first determine the stable social optima of this game. Fix a joint strategy s and let $t := \sum_{j \in N} s_j$. 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 $\sum_{j \in N} s_j = \frac{d}{2b}$. Fix $i \in N$. Let $u := \sum_{j \neq i} s_j$. For every strategy z of player i

$$p_i(z, s_{-i}) = -bz^2 + (d - bu)z \quad \text{and} \quad SW(z, s_{-i}) = -bz^2 + (d - 2bu)z + u(d - bu).$$

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,

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

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

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

Since $p_i(x, s_{-i}) - p_i(y, s_{-i}) = b(y - x)(x - \frac{d}{2b})$ 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 $\max_{y < x < \frac{d}{2b}} AF_i(x, s) = \infty$. But s was an arbitrary stable social optimum, so the claim follows by Theorem 4 (i). □

This proof shows that for every stable social optimum s, for every player there exist deviating strategies with an arbitrary high appeal factor. In fact, $\lim_{x \to y^+} AF_i(x, s) = \infty$, i.e., the appeal factor of the deviating strategy x converges to ∞ when it converges from the right to the original strategy y in s.

3.10 Bertrand competition

Next, we consider Bertrand competition, a game concerned with a simultaneous selection of prices for the same product by two firms, see, e.g., [8, pages 175–177]. The product is then sold by the firm that chose a lower price. In the case of a tie the product is sold by both firms and the profits are split. We assume that each firm has identical marginal costs $c > 0$ and no fixed cost, and that each strategy set S_i equals $[c, \frac{a}{b})$, where $c < \frac{a}{b}$. The payoff function for player $i \in \{1, 2\}$ is given by

$$p_i(s_i, s_{3-i}) := \begin{cases} (s_i - c)(a - bs_i) & \text{if } c < s_i < s_{3-i} \\ \frac{1}{2}(s_i - c)(a - bs_i) & \text{if } c < s_i = s_{3-i} \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 10. *The selfishness level of the Bertrand competition game is ∞.*

Proof. Let $d := \frac{a+bc}{2b}$. If $SW(s) > 0$, then $SW(s) = (s_0 - c)(a - bs_0)$, where $s_0 := \min(s_1, s_2)$. Note that $d \in (c, \frac{a}{b})$, since by the assumption $bc < a$. Hence s is a social optimum iff $\min(s_1, s_2) = d$.

If s is a social optimum with $s_1 \neq s_2$, then player i with the larger s_i can profitably deviate to s_{3-i} (that equals d), while (s_{3-i}, s_{3-i}) remains a social optimum. So the only stable social optimum is (d, d).

Fix $i \in \{1, 2\}$. Note that if s_i is slightly lower than d, then $p_i(s_i, d) > p_i(d, d)$. Further,

$$\lim_{s_i \to d^-} (p_i(s_i, d) - p_i(d, d)) = \frac{1}{2}(d - c)(a - bd), \quad \text{while} \quad \lim_{s_i \to d^-} (SW(d, d) - SW(s_i, d)) = 0$$

and $SW(d, d) - SW(s_i, d) \neq 0$ for $s_i \neq d$. Hence

$$\max_{c < s_i < d} \frac{p_i(s_i, d) - p_i(d, d)}{SW(d, d) - SW(s_i, d)} = \infty.$$

The claim now follows by Theorem 4 (i). □

References

[1] E. Anshelevich, A. Dasgupta, J. Kleinberg, E. Tardos, T. Wexler, and T. Roughgarden. The price of stability for network design with fair cost allocation. In *Proc. 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 295–304, 2004.

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

[3] 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*, pages 172–188, 2010.

[4] P.-A. Chen, B. de Keijzer, D. Kempe, and G. Schäfer. The robust price of anarchy of altruistic games. In *Proc. 7th Workshop on Internet and Network Economics*, pages 383–390, 2011.

[5] 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.

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

[7] A. Fiat, A. Karlin, E. Koutsoupias, and A. Vidali. Approaching utopia: strong truthfulness and externality-resistant mechanisms. In *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, pages 221–230. ACM, 2013.

[8] G. Jehle and P. Reny. *Advanced Microeconomic Theory.* Addison Wesley, New York, NY, third edition, 2011.

[9] I. Milchtaich. Congestion games with player-specific payoff functions. *Games and Economic Behaviour*, 13:111–124, 1996.

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

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

[12] K. Ritzberger. *Foundations of Non-cooperative Game Theory.* Oxford University Press, Oxford, 2002.

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

[14] 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.

[15] H. P. Young. The evolution of conventions. *Econometrica*, 61(1):57–84, 1993.