Internalization of Social Cost in Congestion Games

Igal Milchtaich[*]

Economic Theory 71 (2021), 717–760

Congestion models may be studied from either the users' point of view or the social one. The first perspective examines the incentives of individual users, who are only interested in their own, personal payoff or cost and ignore the negative externalities that their choice of resources creates for the other users. The second perspective concerns social goals such as the minimization of the mean travel time in a transportation network. This paper studies a more general setting, in which individual users attach to the social cost some weight r that is not necessarily 0 or 1. It examines the comparative-statics question of whether higher r necessarily means higher social welfare at equilibrium.

Keywords: Social cost, congestion games, altruism, price of anarchy, stability of equilibrium, potential game.

JEL Classification code: C72.

1 Introduction

Negative externalities are a potential source of inefficiency. In congested systems, such as road networks, inefficiency arises when users do not take into account the externalities that their own use of the public resources create for others. For example, slow traffic flow may be the result of drivers only considering the effect of their choice of routes on their own driving time and disregarding its effect on the other drivers. Advising means that would compel users to internalize the social effects of their decisions is a classic theme in economics. Tolls, in particular, may be used for this purpose. However, as a matter of fact, it is not uncommon for people to take social welfare into consideration even in the absence of a material impetus to do so. A simple way to describe such attitudes is to model decision makers as maximizers of a weighted sum of their personal, material payoff and the social payoff. This raises the question of what effect does the weight r attached to the latter has on the actual social welfare at equilibrium. As shown elsewhere (Milchtaich 2012), the effect may actually be negative. In strategic contexts, or games (even symmetric ones), if everyone attaches the same small but positive weight r to each of the other players' payoff, the result may paradoxically be lower personal payoffs than in the case of complete selfishness, $r = 0$. The main question this paper asks is whether and under what circumstances such an outcome is possible also in congestion games, where players affect others only through their use of common resources.

[*] Department of Economics, Bar-Ilan University, Ramat Gan 5290002, Israel
igal.milchtaich@biu.ac.il https://sites.google.com/view/milchtaich/

Two lines of research lead to the question outlined above. One line, already mentioned, is the general question of the *comparative statics* of altruism and spite (Milchtaich 2006a, 2012), that is, the material consequences of attaching, respectively, an increasingly positive or negative weight r, called the altruism coefficient, to the social payoff. The other line of research concerns the effect of altruism or spite on the *price of anarchy* in congestion games (Chen and Kempe 2008, Hoefer and Skopalik 2013, and others).[1] The price of anarchy in a game is the ratio between the largest social cost at any equilibrium and the smallest social cost at any strategy profile (that is, the social optimum). The social cost most commonly considered is the aggregate cost, the sum of the players' personal costs. While the aggregate cost associated with any given strategy profile is unaffected by the altruism coefficient r, the price of anarchy is affected because the set of equilibria generally changes as r changes. A surprising finding of this research is that the price of anarchy may actually increase with increasing r (Caragiannis et al. 2010, Chen et al. 2014). That is, altruism may lead to system performance deterioration. Superficially, at least, this finding resembles the possibility of "negative" comparative statics mentioned above. However, the two approaches to the study of altruism and spite are fundamentally different. Comparative statics do not concentrate on the worst case, but either compare any two equilibria, each corresponding to a different level of r, or trace the effect of a gradual change of the altruism coefficient on the equilibrium behavior starting at a single, particular equilibrium. These two varieties are referred to as global and local comparative statics, respectively. Another difference between the two approaches is that the price of anarchy typically concerns not a single game but a family of games, such as all nonatomic congestion games with linear cost functions, for which the price of anarchy is defined as the maximum over all games in the family. For different values of the altruism coefficient r, the maximum may be attained at different games. Comparative statics, by contrast, always concern a single underlying game, with only the weight r that players attach to the social cost changing. Therefore, they necessarily reflect actual behavioral consequences of that change. The differences between the two approaches raise the question of whether their predictions concerning the material consequences of a change in r are similar or conflicting. As this paper shows, these predictions depend very much on the class of congestion games looked at.

Three broad classes of games are considered below. The most basic one, studied in Section 3, is unweighted congestion games. Each of the N players chooses a subset of the finite set E of resources, and derives from each resource e a (positive or negative) payoff that depends negatively on the total number of players that include e in their choice. The game is not necessarily symmetric, as different players may have different collections of allowable subsets of E. However, even in the symmetric case, and with linear connections between numbers of users and payoffs, global comparative statics are not necessarily positive. In fact, it is possible for non-socially optimal equilibria to exist *only* for large r. This finding parallels the result alluded to above, that with linear cost functions, the price of anarchy is actually an increasing function of r. The main result in Section 3 is the

[1] A related literature examines *malicious* users in congestion games (e.g., Karakostas and Viglas 2007, Gairing 2009). Malice is similar to spite except that it is restricted to a small number of exceptional users rather than reflecting a "bad" social norm.

identification of a condition that "amends" this and guarantees that, for unweighted congestion games that satisfy it, global comparative statics are positive and the price of anarchy is nonincreasing. The condition, dubbed the flow monotonicity property, does not concern the functional form of the cost functions but is a combinatorial condition on the players' allowed choices of resources. In symmetric network congestion games, where the choice is between the different routes connecting the network's origin and destination vertices, the flow monotonicity property spells a particular restriction on the network topology. Section 3.3 examines this restriction and identifies concrete networks that satisfy it, and consequently exhibit positive global comparative statics in all games defined on them.

The second class of congestion games is games with splittable flow, where a player i is not restricted to choosing a single subset of resources but can split his specified weight w_i among several allowable subsets (e.g., several alternative routes). Section 4 studies such games with linear cost functions. The main result is that for every $-1 \leq r \leq 1$ the aggregate payoff at equilibrium has a unique value, which in different subintervals of $[-1,1]$ is either a constant or a strictly increasing function of r. Thus, unlike their "unsplittable" kin, these games always exhibit positive comparative statics and nonincreasing price of anarchy.

The same applies to the (local and global) comparative statics and the price of anarchy in the class of nonatomic congestion games (Section 5), where the finite set of players is replaced with a continuum, to model the limit case of a very large population of users who individually have only a very small effect on the others. Moreover, the results here do not require linear cost functions or splittable flow.

A central insight emerging from the study of comparative statics of altruism and spite is that a negative relation between the altruism coefficient and the equilibrium level of the social payoff is associated with unstable equilibria (Milchtaich 2012). Conversely, stable equilibria guarantee positive, "normal" comparative statics. This finding moreover holds for any choice of social payoff function, not only the aggregate payoff or (the negative of the) cost. It refers to a general notion of static stability that is applicable to any game and generalizes a number of more special stability concepts (Milchtaich 2020). Section 6 presents this notion and several general results that connect it with comparative statics. The stability condition is not generally implied by the equilibrium condition (and vice versa). However, in some classes of games, the implication does hold, and in such games, which have the property that every equilibrium is automatically stable, altruism can only have a positive effect and spite only a negative effect on social welfare. As it turns out, the congestion games studied in the preceding sections have this property. This finding provides a general context to the above results (although it does not go as far as actually implying them), and also suggests a possible extension.

The extension, explored in Sections 6.1 and 6.2, uses general results concerning stability and comparative statics in games where strategies are (real) vectors. In particular, these results are applicable to congestion games with splittable flow, including those with non-linear cost functions. A two-player example is presented which shows that even a very simple such game may exhibit both positive and negative comparative statics. That is, depending on the equilibria looked at, increasing altruism either increases or decreases the material payoffs of both players. The former equilibria as stable, and the latter are unstable. The example was originally presented as a Cournot duopoly game. However, as shown in Section 6.3, the Cournot oligopoly model (that is, quantity competition) is a special case of congestion games with splittable flow.

2 Preliminaries

A strategic game h specifies a (finite or infinite) set of players and, for each player i, a strategy set X_i and a payoff function $h_i: X \longrightarrow \mathbb{R}$, where $X = \prod_i X_i$ is the set of all strategy profiles. A *social payoff function* is any function $f: X \longrightarrow \mathbb{R}$. An important example in a game with a finite number of players is the *aggregate payoff*, $f = \sum_i h_i$. The *altruism coefficient* is an exogenously given parameter $r \leq 1$ that specifies a common degree of internalization of the social payoff by all players, reflecting, for example, a shared social norm. The coefficient defines the *modified game* h^r, where the payoff of each player i is not the original, *personal* (or material) payoff h_i but the *modified* (or perceived) payoff

$$h_i^r = (1 - r)h_i + rf. \tag{1}$$

Comparative statics concern the connection between r and the value of f at the (Nash) equilibria of the modified game h^r. In the special case where f is the aggregate payoff, so

$$h_i^r = h_i + r \sum_{j \neq i} h_j,$$

this connection expresses the effect on the actual equilibrium aggregate personal payoff when each of the players attaches the same weight r to the personal payoff of each of the other players. Thus, r in this case is the ratio between the marginal contributions of any other individual's material utility and a player's own material utility to the latter's perceived utility. Positive, negative or zero r may be interpreted as expressing altruism, spite or complete selfishness, respectively.[2]

A (symmetric) *population game* is specified by a single strategy space[3] X that is a convex set in a (Hausdorff real) linear topological space (for example, the unit simplex in a Euclidean space $\mathbb{R}^n$) and a payoff function $g: X^2 \longrightarrow \mathbb{R}$ that is continuous in the second argument (Milchtaich 2012). However, a population game is interpreted not as an interaction between two specific players but as one involving an (effectively) infinite population of individuals who are "playing the field". This means that an individual's payoff $g(x, y)$ depends only on his own strategy x and on a suitability defined *population strategy* y, which in an element of X that encapsulates the choices of strategies in the population. For example, y may be the population's *mean* strategy with respect to some nonatomic (population) measure, or it may describe the *distribution* of (pure) strategies played.

---

[2] In a more general setting, different individuals i may have different altruism coefficients r_i (see Chen and Kempe 2008, Hoefer and Skopalik 2013) or attach different weights r_{ij} to the personal payoffs of different other persons j (see Rahn and Schäfer 2013, Anagnostopoulos et al. 2015). The present work does not cover these extensions.

[3] A strategy *space* is a strategy set that is endowed with a particular topology, with respect to which continuity and related terms are defined.

A social payoff function for a population game is any continuous function $\varphi: \hat{X} \longrightarrow \mathbb{R}$ whose domain $\hat{X}$ is the *cone* of the strategy space X. The cone consists of all elements of the form tx, with $x \in X$ and $t > 0$. Note that, unlike for "normal" games, where the argument of the social payoff function is a strategy profile, here the argument is a single strategy. The difference reflects an assumption that, in a population game, the social payoff $\varphi(y)$ depends only on the population strategy y; the effect of any single player's action on it is negligible. This assumption necessitates a reinterpretation of the notion of internalization of social payoff, which is understood here as consideration for the *marginal* effect of one's actions on φ (Chen and Kempe 2008). Thus, an individual's concern is not with the effect of a unilateral adoption of a strategy x (which is null) but with the effect that adoption by a small but significant (and representative) proportion p of the population would have. To formalize this idea, suppose that φ has a directional derivative in every direction $\hat{x} \in \hat{X}$, which depends continuously on the point $\hat{y} \in \hat{X}$ at which it is evaluated. In other words, the assumption is that the *differential* $d\varphi: \hat{X}^2 \longrightarrow \mathbb{R}$, defined by (the one-sided derivative)

$$d\varphi(\hat{x}, \hat{y}) = \frac{d}{dt}\bigg|_{t=0^+} \varphi(t\hat{x} + \hat{y}),$$

exists and is continuous in the second argument (Milchtaich 2012). For altruism coefficient $r \leq 1$, the modified game g^r is defined by

$$g^r(x, y) = (1 - r)g(x, y) + r\, d\varphi(x, y). \tag{2}$$

Comparative statics concern the connection between the altruism coefficient r and the value of the social payoff function φ at the equilibrium strategies in the modified game g^r, that is, those (population) strategies $y \in X$ satisfying

$$g^r(y, y) \geq g^r(x, y), \qquad x \in X. \tag{3}$$

2.1 Potential games

A game h is a *potential game* (Monderer and Shapley 1996) if it admits an (*exact*) *potential*, that is, a function $P: X \longrightarrow \mathbb{R}$ such that, whenever a single player i changes his strategy, the resulting change in i's payoff equals the change in P:

$$h_i(y \mid x_i) - h_i(y) = P(y \mid x_i) - P(y), \qquad x_i \in X_i, y \in X,$$

where $y \mid x_i$ denotes the strategy profile that differs from y only in that player i uses strategy x_i.

The concept of potential game may be adapted to population games, essentially by replacing the difference between the two values of the potential with a derivative (Milchtaich 2012). Specifically, a continuous function $\Phi: X \longrightarrow \mathbb{R}$ is a *potential* for a population game g with a strategy space X if, for all $x, y \in X$ and $0 < p < 1$, the derivative on the left-hand side of the following equality exists and the equality holds:

$$\frac{d}{dp}\Phi(px + (1 - p)y) = g(x, px + (1 - p)y) - g(y, px + (1 - p)y). \tag{4}$$

A useful fact (Milchtaich 2020) is that any function $\Phi: X \longrightarrow \mathbb{R}$ that can be extended to a continuous function on the cone of X, such that for the extended function (which is also denoted by Φ) the differential $d\Phi$ exists and is continuous in the second argument, satisfies the identity

$$\frac{d}{dp}\Phi(px + (1-p)y) = d\Phi(x, px + (1-p)y) - d\Phi(y, px + (1-p)y), \qquad x, y \in X, 0 < p < 1. \tag{5}$$

Therefore, a sufficient condition for such Φ to be a potential for g is that

$$d\Phi(x, y) = g(x, y), \qquad x, y \in X. \tag{6}$$

For a game h with a potential P, and any social payoff function f and altruism coefficient r, the modified game h^r is also a potential game, as it is easy to see that the function

$$P^r = (1-r)P + rf \tag{7}$$

is a potential for h^r. Similarly, for a population game g with a potential Φ, and any social payoff function φ and altruism coefficient r,

$$\Phi^r = (1-r)\Phi + r\varphi$$

is a potential for g^r. This follows from the fact Φ^0 $(= \Phi)$ is a potential for g^0 $(= g)$ and, by the identity obtained by replacing Φ with φ in (5), Φ^1 $(= \varphi)$ is a potential for g^1 $(= d\varphi)$.

The next two results are useful for studying the comparative statics of potential games. As their proofs are very similar, only that of the first result is presented.

Proposition 1. For a game h with a potential P, a social payoff function f, and altruism coefficients r and s with $r < s \le 1$, if two distinct strategy profiles y^r and y^s are (global) maximum points of P^r and P^s, respectively, then

$$f(y^r) \le f(y^s).$$

If moreover y^s is a strict (equivalently, unique) maximum point, then the inequality is strict.

Proof. The proof is based on the following identity, which follows immediately from (7):

$$(1-r)(P^s(x) - P^s(y)) + (1-s)(P^r(y) - P^r(x)) = (s-r)(f(x) - f(y)).$$

The identity implies that the difference $f(x) - f(y)$ has the same sign as the expression on the left-hand side. The latter is nonpositive for $x = y^r$ and $y = y^s$ if these strategy profiles maximize P^r and P^s, respectively, and it is moreover negative if y^s is a strict maximum point. ■

Proposition 2. For a population game g with a potential Φ, a social payoff function φ, and altruism coefficients r and s with $r < s \le 1$, if two distinct strategies y^r and y^s are (global) maximum points of Φ^r and Φ^s, respectively, then

$$\varphi(y^r) \le \varphi(y^s).$$

If moreover y^s is a strict maximum point, then the inequality is strict.

In a game h with a potential P, a strategy profile x is an equilibrium if and only if no single player i can increase P by unilaterally deviating from his strategy x_i. In particular, every global maximum point of P is an equilibrium. In a population game g with a potential Φ, even a local maximum point of Φ is an equilibrium strategy, and if the function Φ is concave or strictly concave, then the converse also holds, in fact, an equilibrium strategy is necessarily a *global* maximum or strict global maximum point of Φ, respectively. These assertions easily follow from the following identity, which is a corollary of (4):

$$\left.\frac{d}{dp}\right|_{p=0^+} \Phi(px + (1-p)y) = g(x,y) - g(y,y).$$

2.2 The price of anarchy

The price of anarchy is a measure of the distance of the equilibria from the social optimum, which is defined as the maximum of a specified social payoff function f. Assuming that f is always nonpositive, equivalently, that the *social cost* $-f$ is nonnegative, the PoA in a game is the ratio between (i) the value of the social payoff function (or the social cost) at the equilibrium with the lowest such payoff (or the highest cost) and (ii) the optimal value. In this paper, the focus is on the effect that internalization of the social payoff or cost has on this ratio, in other words, its dependence on the altruism coefficient r. This dependence may be studied in a single game or, more generally, for a specified family of games $\mathfrak{H}$, providing that f is meaningful for all games in $\mathfrak{H}$. Thus, the price of anarchy PoA^r corresponding to altruism coefficient $r \leq 1$ is defined by

$$\mathrm{PoA}^r = \sup\left\{ \frac{f(y^r)}{f(x)} \;\middle|\; \begin{array}{l} h \in \mathfrak{H},\ x \text{ is a strategy profile in } h, \\ y^r \text{ is an equilibrium in } h^r \end{array} \right\},$$

where the quotient is interpreted as 1 if the numerator and denominator are both 0 and as ∞ if only the denominator is 0. A similar definition applies to a family of population games $\mathfrak{G}$, with f replaced by a specified nonpositive social payoff function φ that is meaningful for all games in $\mathfrak{G}$.

3 Congestion Games

In a (weighted) *congestion game*, there is a finite number N of players, numbered from 1 to N, and a finite set E of resources (represented, for example, by the edges in a graph). Each player i has a *weight* $w_i > 0$ and a finite strategy set $S_i \subseteq \mathbb{R}^E$, each element of which is a non-zero vector $x_i = (x_{ie})_{e\in E}$ where each component x_{ie} is either w_i (indicating that using resource e is part of the strategy) or 0 (indicating that the strategy does not include resource e). For a strategy profile $x = (x_1, x_2, \dots, x_N)$, the total use of resource e is expressed by the *flow* (or load) on it, $x_e = \sum_{i=1}^N x_{ie}$. Two strategy profiles x' and x'' are *equivalent* if $x'_e = x''_e$ for all e. The flow on e determines the cost per unit weight of using the resource, which is given by $c_e(x_e)$, where $c_e\colon [0,\infty) \longrightarrow \mathbb{R}$ is a nondecreasing (resource-specific) *cost function*. The total cost for each player i is the sum of the (per unit) costs of the resources included in i's strategy multiplied by the player's weight, and the payoff is the negative of the cost. Thus, player i's payoff is given by

$$h_i(x) = -\sum_{e\in E} x_{ie} c_e(x_e).$$

(Note that, as the cost may be positive or negative, so does the payoff.) The aggregate cost may be viewed as the social cost, and its negative, the aggregate payoff, as the social payoff function f. Thus,

$$f(x) = -\sum_{e\in E} x_e c_e(x_e). \tag{8}$$

With altruism coefficient $r \le 1$, the modified payoff of player i is given by

$$h_i^r(x) = -\sum_{e\in E}\big((1-r)x_{ie} + r x_e\big)\, c_e(x_e) = -\sum_{e\in E}\left(x_{ie} + r\sum_{j\ne i} x_{je}\right) c_e(x_e). \tag{9}$$

3.1 Unweighted congestion games

An *unweighted* congestion game is one where players have unity weights, $w_i = 1$ for all i. Each strategy is thus a binary vector and the flow on each resource e is simply the number of its users, which entails that, effectively, a cost function c_e is completely specified by its restriction to the positive integers. An unweighted congestion game is *symmetric* if all players have the same strategy set.

An unweighted congestion game h is always a potential game (Rosenthal 1973), with the potential

$$P(x) = -\sum_{e\in E}\sum_{t=1}^{x_e} c_e(t). \tag{10}$$

(If $x_e = 0$, the inner sum is interpreted as zero.) As in any potential game, the set of equilibria in h includes the (nonempty) set of global maximum points of P. The following definition and Proposition 3 below identify a condition under which the reverse inclusion also holds, so the two sets are equal. The definition is inspired by the concept of "flow substitution" of Reshef Meir and David C. Parkes (personal communication).

Definition 1. An unweighted congestion game has the *flow monotonicity property* if for every two non-equivalent strategy profiles x' and x'' there is a strategy profile x that differs from x' only in the strategy of a single player and satisfies $\min(x'_e, x''_e) \le x_e \le \max(x'_e, x''_e)$ for all e.

The flow monotonicity property is so called because it is equivalent to the possibility of changing the strategies prescribed by x' sequentially, one player at a time, in such a way that the flow on each resource changes monotonically to that in x''.

Lemma 1. The flow monotonicity property holds if and only if for every two strategy profiles x' and x'' there are strategy profiles $x^0, x^1, \dots, x^N$ such that $x^0 = x'$, each of the following strategy profiles x^i is either equal to its predecessor x^{i-1} or differs from it only in the strategy of player i, the strategy profiles x^N and x'' are equivalent, and the sequence of flows $x_e^0, x_e^1, \dots, x_e^N$ is monotone for every resource e.

Proof. The sufficiency of the condition ("if") is proved by setting $x = x^i$, where i is the smallest index for which $x^i \neq x'$ (which must exist if x' and x'' are non-equivalent). To prove necessity ("only if"), construct for given strategy profiles x' and x'' a finite sequence $\bar{x}^0, \bar{x}^1, \ldots$ as follows: $\bar{x}^0 = x'$, and for $k \geq 1$, $\bar{x}^k$ is a strategy profile that satisfies the conditions specified for x in Definition 1 except that x' there is replaced with $\bar{x}^{k-1}$. Note that $\bar{x}_e^k \neq \bar{x}_e^{k-1}$ for at least one resource e and that, for each e, $\bar{x}_e^0 \leq \bar{x}_e^1 \leq \cdots \leq x_e''$ or $\bar{x}_e^0 \geq \bar{x}_e^1 \geq \cdots \geq x_e''$. These facts imply that the sequence cannot be extended indefinitely; for some $K \geq 0$, the strategy profiles $\bar{x}^K$ and x'' must be equivalent. Now, for $0 \leq i \leq N$, define x^i as the strategy profile that prescribes the same strategies as $\bar{x}^K$ to all players with index i or lower, and otherwise agrees with x'. It is not difficult to see that, because $(x_e' =)\bar{x}_e^0, \bar{x}_e^1, \ldots, \bar{x}_e^K (= x_e'')$ is monotone for each e, the same is true for $(x_e' =)x_e^0, x_e^1, \ldots, x_e^N (= x_e'')$. ■

The significance of the flow monotonicity property lies in the following fact.

Lemma 2. In an unweighted congestion game with the flow monotonicity property, for every two strategy profiles x and y there is a strategy profile z that is equivalent to x and satisfies

$$P(x) - P(y) \leq \sum_{i=1}^{N} \big(P(\, y \mid z_i\,) - P(y)\big). \tag{11}$$

Proof. For $x' = y$ and $x'' = x$, let $x^0, x^1, \ldots, x^N$ be as in Lemma 1, and set $z = x^N$. Thus, x and x^N are equivalent and $y = x^0$, so

$$P(x) - P(y) = P(x^N) - P(x^0) = \sum_{i=1}^{N} \Big(P(x^i\,) - P(x^{i-1})\Big).$$

To prove (11), it suffices to show that

$$P(x^i) - P(x^{i-1}) \leq P(\,x^0 \mid x_i^N\,) - P(x^0), \qquad 1 \leq i \leq N.$$

As P is a potential, both sides of the inequality express the effect on player i's payoff of a unilateral deviation from x_i^{i-1} $(= x_i^0)$ to x_i^i $(= x_i^N)$. The difference is whether the other players are playing according to x^{i-1} or x^0. In the first case (the left-hand side of the inequality), the change in payoff is given by

$$-\sum_{e \in E} \Big(x_{ie}^i c_e(x_e^i) - x_{ie}^{i-1} c_e(x_e^{i-1})\Big),$$

and in the second case (the right-hand side), the summand is replaced with $x_{ie}^i c_e(x_e^0 + x_{ie}^i - x_{ie}^{i-1}) - x_{ie}^{i-1} c_e(x_e^0)$. It is therefore sufficient to show that for all e

$$x_{ie}^i c_e(x_e^i) - x_{ie}^{i-1} c_e(x_e^{i-1}) \geq x_{ie}^i c_e(x_e^0 + x_{ie}^i - x_{ie}^{i-1}) - x_{ie}^{i-1} c_e(x_e^0).$$

If $x_{ie}^i = x_{ie}^{i-1}$, both sides of the last inequality are zero. If $x_{ie}^i = 1$ and $x_{ie}^{i-1} = 0$, then it follows from the monotonicity of $x_e^0, x_e^1, \ldots, x_e^i$ that $x_e^0 \leq x_e^{i-1} = x_e^i - 1$, and if $x_{ie}^i = 0$ and $x_{ie}^{i-1} = 1$, then by a similar argument $x_e^0 \geq x_e^{i-1} = x_e^i + 1$. Since c_e is nondecreasing, in both cases the inequality holds. ■

If y is an equilibrium, then the right-hand side of (11) is nonpositive for any z. This fact immediately gives the following.

Proposition 3. In an unweighted congestion game with the flow monotonicity property, a strategy profile is an equilibrium if and only if it maximizes the potential P.

The implications of this result are studied in the next subsection.

3.2 Comparative statics of unweighted congestion games

For an unweighted congestion game h, consider the social payoff function f defined as the aggregate payoff (and the social cost defined as the aggregate cost $-f$). It follows from (7), (8) and (10) that, for altruism coefficient $r \leq 1$, the modified game h^r has the potential

$$P^r(x) = -\sum_{e\in E}\left((1-r)\sum_{t=1}^{x_e} c_e(t) + r\,x_e\,c_e(x_e)\right) = -\sum_{e\in E}\sum_{t=1}^{x_e}\hat{c}_e^r(t), \tag{12}$$

where the function $\hat{c}_e^r$ is defined by

$$\hat{c}_e^r = (1-r)\,c_e + r\,MC_e,$$

with MC_e denoting (the discrete version of) the *marginal social cost*,

$$MC_e(t) = t\,c_e(t) - (t-1)\,c_e(t-1), \qquad t = 1,2,\dots.$$

The modified game h^r is generally not a congestion game. However, comparison of (10) and (12) shows that P^r is also the potential for *some* congestion game, namely, the unweighted congestion game $\hat{h}^r$ that differs from (the original, unmodified one) h only in that the cost function c_e of each resource e is replaced with $\hat{c}_e^r$. A sufficient condition for this function to be nondecreasing (hence, a legitimate cost function) is that this is so for MC_e and $0 \leq r \leq 1$. This observation leads to the following comparative-statics result.

Theorem 1. For an unweighted congestion game h with nondecreasing marginal social costs, the aggregate payoff as the social payoff function f, and altruism coefficients r and s with $0 \leq r < s \leq 1$, let y^r and y^s be equilibria in the modified games h^r and h^s, respectively. If h has the flow monotonicity property, then

$$f(y^r) \leq f(y^s). \tag{13}$$

If, in addition, $s = 1$, then moreover $f(x) \leq f(y^s)$ for every strategy profile x.

Proof. As indicated, the modified games h^r and h^s share their potentials, P^r and P^s, and therefore also their sets of equilibria, with the unweighted congestion games $\hat{h}^r$ and $\hat{h}^s$, respectively. Each of the latter differs from h only in the cost functions, and therefore has the flow monotonicity property if and only if h has it. In this case, it follows from Proposition 3 that the equilibria y^r and y^s maximize P^r and P^s, respectively. Inequality (13) now follows from Proposition 1. If $s = 1$, then $P^s = f$, so y^s is a maximum point of the social payoff function. ■

For a game h with nonnegative costs, Theorem 1 can be reformulated as follows. If the flow monotonicity property holds, then the price of anarchy (see Section 2.2) weakly decreases as the players increasingly internalize the social cost of their choice of strategies, and reaches the minimum value of 1 with complete internalization. Combining this observation with a simple bound on the price of anarchy gives the following corollary.

Corollary 1. For the family $\mathfrak{H}$ of all unweighted congestion games with the flow monotonicity property for which the cost functions are nonnegative and the marginal social costs are nondecreasing, as well as every subfamily or single game in $\mathfrak{H}$, the price of anarchy PoA^r is nonincreasing for $0 \leq r \leq 1$ and satisfies $1 \leq \mathrm{PoA}^r \leq 1/r$ (hence, $\mathrm{PoA}^1 = 1$).

Proof. Consider any r and s with $0 \leq r \leq s \leq 1$, a game $h \in \mathfrak{H}$, a strategy profile x in h, and equilibria y^r and y^s in h^r and h^s, respectively. By (8), (10) and the nonnegativity and monotonicity of the cost functions, $f \leq P \leq 0$, and therefore $f \leq P^r \leq rf$ by (7). Since, as shown in the proof of Theorem 1, y^r maximizes P^r, these inequalities give

$$f(x) \leq P^r(x) \leq P^r(y^r) \leq rf(y^r).$$

Therefore, by (13), $f(y^s)/f(x) \leq f(y^r)/f(x) \leq 1/r$. From theses inequalities the result easily follows. (By definition of PoA^r, the first inequality gives $f(y^s)/f(x) \leq \mathrm{PoA}^r$.) ■

The conclusion in Corollary 1 strongly contrasts with a previous result concerning a different family of unweighted congestion games, which also have nonnegative cost functions and nondecreasing marginal social costs, namely, the family of all unweighted congestion games with cost functions that are *linear* (more precisely, affine) with nonnegative coefficients. For this family (as well as for the subfamily where $c_e(t) = t$ for all resources e), Caragiannis et al. (2010) and Chen et al. (2014) showed that for $0 \leq r \leq 1$ the price of anarchy (with the aggregate cost as the social cost function) is given by

$$\mathrm{PoA}^r = \frac{5 + 4r}{2 + r},$$

and is thus a strictly *increasing* function of r. The discrepancy shows the significance of the flow monotonicity property. In a game without this property, the price of anarchy is not necessarily nonincreasing,[4] and it may be greater than 1 at $r = 1$. As the next example shows, this may be so even in the special case of a symmetric unweighted congestion game with cost functions that are linear with positive coefficients.

Example 1. A symmetric unweighted congestion game h has two players, four resources and four strategies, which are: resource e_1 alone, and any two of resources e_2, e_3 and e_4. The cost functions are $c_{e_1}(t) = t + 9$, $c_{e_2}(t) = t + 2.5$ and $c_{e_3}(t) = c_{e_4}(t) = 2t + 1$. A strategy profile minimizes the aggregate cost if and only if one player uses e_2 and e_3 and the other player uses e_2 and e_4. Such a socially optimal strategy profile is also an equilibrium in the modified game for every $0 \leq r \leq 1$, and it is moreover the unique kind of equilibrium if

[4] Obviously, the price of anarchy cannot be *strictly* increasing for any single game or a finite family of games. With finitely many games, there are only finitely many possible values for the social cost, which means that PoA^r can be at most nondecreasing.

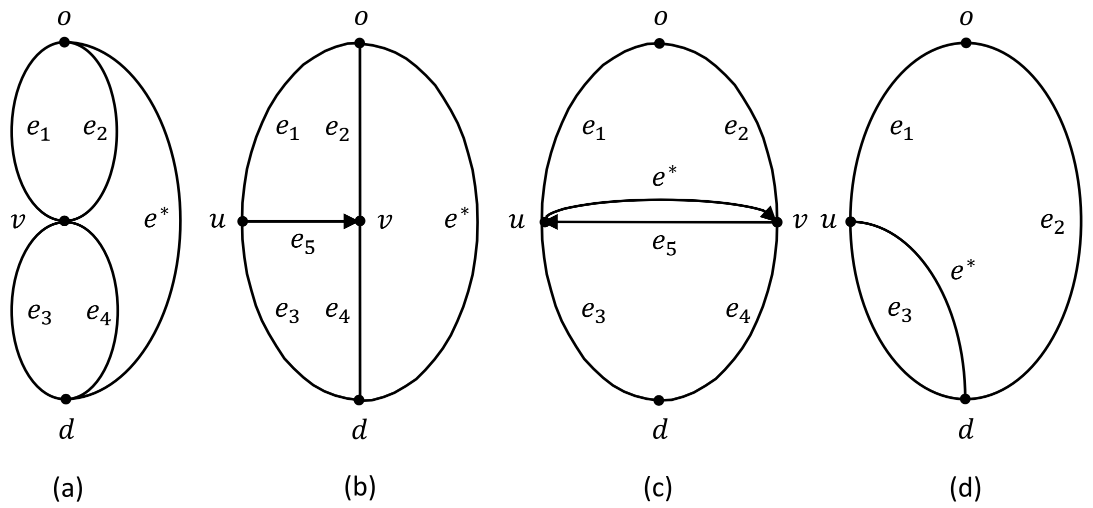

Figure 1. Symmetric unweighted network congestion games. N players (with weight 1) choose routes from o to d. The cost functions are: (a) $c_{e_1}(t) = c_{e_4}(t) = t^2$, $c_{e_2}(t) = c_{e_3}(t) = 3t$ and $c_{e^*}(t) = 5t$, (b) the same, and $c_{e_5}(t) = 0.5t$, (c) $c_e(t) = t$ for all edges e, and (d) any cost functions for which the marginal social costs are nondecreasing. The edges' directions are indicated only where they are ambiguous.

$r < 0.75$. However, for $0.75 \leq r \leq 1$ there is a second kind of equilibrium, where one player uses e_1 and the other uses e_3 and e_4. Therefore, the price of anarchy satisfies $\mathrm{PoA}^r = 1$ for $0 \leq r < 0.75$ but is greater than 1 (specifically, $= 16/15$) for $0.75 \leq r \leq 1$.

The game in Example 1 does not have the flow monotonicity property. For a strategy profile of the first kind (that is, a socially optimal one) the vector of edge flows is $(0,2,1,1)$, and for the second kind, the vector is $(1,0,1,1)$. It is not possible to move a single player from his first strategy to some other strategy in such a way that the flow on each resource lies between its first and second values (because any move changes the flow on e_3 or e_4).

3.3 Network congestion games

Whether a symmetric unweighted congestion game has the flow monotonicity property only depends on the number of players and the combinatorial structure of the strategy set. In a symmetric unweighted *network* congestion game, the combinatorial structure reflects the topology of a specified directed two-terminal network, whose edges represent the resources (see Figure 1). The network has designated origin and destination vertices, o and d, and each of the edges and vertices belongs to at least one *route*, which is a (simple) path that starts at o and ends at d. Each route corresponds to a strategy, which is represented by the binary vector assigning 1 to every edge traversed by the route and 0 to every other edge.
A symmetric unweighted network congestion game can equivalently be described in terms of an *undirected* two-terminal network G, where each vertex and each edge belongs to at least one route,[5] and an assignment of an *allowable direction* to each edge e in G, which must be that in which some route in G traverses e (Milchtaich 2013). The set of strategies

[5] Since a route in an undirected network may traverse an edge in either direction, the meaning of this "non-redundancy" condition is slightly different than in the case of a directed network. However, it is shown in the Appendix that an undirected network satisfies the condition if and only if it is the undirected version of some directed network as above.

consists of all *allowable routes* in G, that is, routes that traverse all their edges in the allowable directions. It is assumed that at least one such route exists. This alternative formulation enables classification according to the underlying undirected network; games on the same undirected network G that differ in the number of players, costs functions or edge directions are grouped together. Such a classification may be useful for identifying certain "topological" properties of network congestion games, that is, properties of individual games that are necessary consequences of the topology of the undirected network, regardless of the other data. This agenda motivates the following definition.

Definition 2. An undirected two-terminal network G has the *flow monotonicity property* if all symmetric unweighted network congestion games on G have the flow monotonicity property.

The monotonicity property is preserved under connection of networks in series.

Lemma 3. If an undirected two-terminal network G is obtained by connecting two or more such networks in series, then G has the flow monotonicity property if and only if each of the constituent networks has the property.

Proof. Assigning allowable directions to the edges in G is equivalent to assigning directions in each of the constituent networks $\bar{G}$, and a route in G is allowable if and only if the section that passes through each $\bar{G}$ is allowable there. Therefore, G has the flow monotonicity property if and only if, for every assignment of allowable directions, every number of players and every pair of non-equivalent strategy profiles x' and x'', there is a strategy profile x that differs from x' only in the strategy of a single player, who only changes his route in one constituent network $\bar{G}$, such that $\min(x'_e, x''_e) \leq x_e \leq \max(x'_e, x''_e)$ for all e in $\bar{G}$. It is easy to see that this condition holds if and only if every $\bar{G}$ have the flow monotonicity property. ■

A similar result does not hold for the connection of networks in parallel. Indeed, the flow monotonicity property does not generally hold even for *series-parallel* networks, which are those that can be constructed from single edges using only the operations of connecting networks in series or in parallel. Moreover, as the next example shows, for symmetric unweighted network congestion games with nonnegative cost functions and nondecreasing marginal social costs on a series-parallel network, it is not even true that $\mathrm{PoA}^1 = 1$ (cf. Corollary 1).

By definition of the modified game with $r = 1$, a strategy profile is an equilibrium in this game if and only if it is "locally" social optimal, in the sense that any single-player deviation either increases the aggregate cost or leaves it unchanged. $\mathrm{PoA}^1 = 1$ would mean that this condition automatically implies "global" social optimality, that is, the social cost cannot decrease even if two or more players deviate simultaneously. In an unpublished technical report, Singh (2008) found that this implication holds for games with cost functions as above on any series-parallel network. However, the example refutes this finding.

Example 2. The series-parallel network in Figure 1a does not have the flow monotonicity property. For $N = 2$, a strategy profile x as in Definition 1 does not exist for a strategy profile x' where one player uses the route $e_1 e_3$ and the other uses $e_2 e_4$ and a strategy profile x'' where the routes are e_1e_4 and e^*. Moreover, such x' and x'' are equilibria in the game specified in Figure 1a, and since in each of them no edge is used by more than one player, they are also "local" social optima. However the aggregate cost in x' is higher than in x'', 8 instead of 7. An almost identical pair of equilibria exists in the game in Figure 1b (where the network is not series-parallel). The only difference is that, in x'', the first route is $e_1e_5e_4$. Similarly, in Figure 1c (again, a non-series-parallel network), only the equilibria where the two players' routes are e_1e_3 and e_2e_4 are socially optimal. With the routes $e_1e^*e_4$ and $e_2e_5e_3$, the aggregate cost is higher, yet no player can unilaterally decrease it (or his own cost) by choosing a different route.

A similar example does not exist for the network in Figure 1d, for any number of players N and any cost functions with nondecreasing marginal social costs. The reason is that, unlike the other networks in Figure 1, this one is an undirected *extension-parallel network* (Holzman and Law-Yone 2003).[6] This subclass of the series-parallel networks is defined in a similar recursive manner except for the proviso that the connection in series of two networks is allowed only if one of them has only one edge (which makes their connection equivalent to the extension of a terminal vertex, o or d, in the other network). Extension-parallel networks satisfy a condition that is even stronger than the flow monotonicity property. A network G has the *augmented flow monotonicity property* (AFMP) if the *augmented network* G^* obtained by connecting G in parallel with a single edge e^* has the flow monotonicity property. The edge e^* may be interpreted as representing a common outside option, an alternative to using the network G.

Lemma 4. An undirected two-terminal network has the augmented flow monotonicity property if and only if it is extension-parallel.

Proof. One direction of the proof is by induction on the structure of extension-parallel networks. For a single-edge network, the AFMP clearly holds. If a network G is obtained by connecting in series an extension-parallel network $\bar{G}$ that has the AFMP and a single edge e, then G also has the AFMP. This is because the set of routes (equivalently, the allowable routes; see Corollary A2 in the Appendix, condition (ii)) in G is obtained from that in $\bar{G}$ simply by appending e (and its non-terminal end vertex) to each route. Therefore, for every strategy profile in the augmented network G^*, the flow on e and the flow on e^* sum up to the number of players N, which implies that the pair of inequalities in Definition 1 holds automatically for e if it holds for e^*. It remains to consider a network G that is obtained by connecting in parallel two extension-parallel networks $\bar{G}$ and $\tilde{G}$ that have the AFMP, and show that the augmented network G^* (consisting of $\bar{G}$, $\tilde{G}$ and a single edge e^* all connected in parallel) has the flow monotonicity property.

[6] The original meaning of extension-parallel network concerned *directed* networks. An alternative term for the undirected version of these networks, which is the version considered here, is networks with *linearly independent routes*. These undirected two-terminal networks are characterized by the property that each route includes at least one edge that does not belong to any other route (Milchtaich 2006b).

For any number of players N, consider two non-equivalent strategy profiles x' and x'', where the players use routes in G^*. Let $\bar{x}'$ and $\bar{x}''$ be the strategy profiles obtained from x' and x'' by replacing each route that lies outside $\bar{G}^*$ (that is, in $\tilde{G}$) with the single-edge route e^*. In particular, for x', the flow $\bar{x}'_{e^*}$ gives the number of players who are not using a route in $\bar{G}$, and similarly for x'' and $\bar{x}''_{e^*}$. Since $\bar{G}$ has the AFMP, if $\bar{x}'$ and $\bar{x}''$ are not equivalent, then applying Definition 1 to them (instead of x' and x'') gives a strategy profile $\bar{x}$, which also involves only routes in $\bar{G}^*$. For $\tilde{G}$, the strategy profiles $\tilde{x}'$ and $\tilde{x}''$ and (if the last two are not equivalent) a strategy profile $\tilde{x}$ are constructed analogously. Using these ingredients, a strategy profile x satisfying the conditions in Definition 1 is constructed below.

Note that $\bar{x}'$ and $\bar{x}''$ are not equivalent or $\tilde{x}'$ and $\tilde{x}''$ are not equivalent. Otherwise, both the flow on each edge in $\bar{G}$ and the total number of players using routes in that network (which is N minus the number of players not doing so) would be the same for x' and x'', and similarly for $\tilde{G}$. However, the conclusion implies that the number of players using e^* is also the same for x' and x'', which contradicts their assumed non-equivalence. By symmetry, it suffices to consider the case where $\bar{x}'$ and $\bar{x}''$ are not equivalent, so a strategy profile $\bar{x}$ as above exists. Let i be the unique player with $\bar{x}_i \neq \bar{x}'_i$.

Suppose, first, that $\bar{x}'_i$ is a route in $\bar{G}$ (hence, $\bar{x}'_i = x'_i$). If the same is true for $\bar{x}_i$, then define x by $x_i = \bar{x}_i$. (The other players' strategies in x are as in x'.) Otherwise, that is, if $\bar{x}_i$ is e^*, then by definition of the flow monotonicity property $\bar{x}''_{e^*} > \bar{x}'_{e^*}$. If also $\tilde{x}''_{e^*} \geq \tilde{x}'_{e^*}$, then necessarily $x''_{e^*} > x'_{e^*}$, and in this case, set x_i to be e^*. If $\tilde{x}''_{e^*} < \tilde{x}'_{e^*}$, then $\tilde{x}'$ and $\tilde{x}''$ are not equivalent and, for the unique player j with $\tilde{x}_j \neq \tilde{x}'_j$, strategy $\tilde{x}_j$ is not e^* (but rather a route in $\tilde{G}$). In this case, define x by $x_j = \tilde{x}_j$ if $\tilde{x}'_j$ is also a route in $\tilde{G}$ and by $x_i = \tilde{x}_j$ if $\tilde{x}'_j$ is e^*. (The assignment $x_i = \tilde{x}_j$ means that player i moves to the route that player j, who may or may not be the same person, was supposed to move to.)

Suppose now that $\bar{x}'_i$ is e^*, which means that $\bar{x}_i$ is a route in $\bar{G}$ and, by definition of the flow monotonicity property, $\bar{x}''_{e^*} < \bar{x}'_{e^*}$. If also $\tilde{x}''_{e^*} \leq \tilde{x}'_{e^*}$, then necessarily $x''_{e^*} < x'_{e^*}$, and in this case, choose any player j for whom x'_j is e^* and define x by $x_j = \bar{x}_i$. If $\tilde{x}''_{e^*} > \tilde{x}'_{e^*}$, then $\tilde{x}'$ and $\tilde{x}''$ are not equivalent and, for the unique player j with $\tilde{x}_j \neq \tilde{x}'_j$, strategy $\tilde{x}'_j$ is not e^* (but rather a route in $\tilde{G}$). In this case, define x by $x_j = \tilde{x}_j$ if $\tilde{x}_j$ is also a route in $\tilde{G}$ and by $x_j = \bar{x}_i$ if $\tilde{x}_j$ is e^*.

It is easy to verify that, in all cases, the strategy profile x satisfies the inequalities in Definition 1 for all edges e in G^*. This completes the first part of the proof: all undirected extension-parallel networks have the augmented flow monotonicity property.

The reason why these networks are the *only* ones having the AFMP is that, by Proposition 4 in Milchtaich (2006b), one (or more) of three specific undirected networks is embedded in every undirected network G that is not extension-parallel (in other words, a network with linearly *dependent* routes; see footnote 6). These three networks are: the figure-eight network, which is the network in Figure 1a minus edge e^*, the "expanded" figure-eight network obtained from the previous network by replacing the vertex v with two vertices connected by an edge, and the *Wheatstone network*, which is the (undirected) network in Figure 1b minus e^*. Therefore, connecting G in parallel with a single edge e^* results in an undirected network G^* for which there exists a two-player game as in Example 2, that is, one that does not have the flow monotonicity property. ∎

The (weaker) flow monotonicity property itself is not limited to undirected extension-parallel networks and those, like the figure-eight network, that are obtained by connecting several such networks in series. It also holds for certain networks that are not even series-parallel, in particular, the Wheatstone network. This can be seen by comparing (i) the directed Wheatstone network obtained by deleting edge e_5 in Figure 1c with (ii) the network in Figure 1d. The latter is extension-parallel, and therefore has the flow monotonicity property. The operation of appending a new vertex v and a new edge e_4 to routes in network (ii) that include either e_2 or e^* defines a one-to-one correspondence between the set of routes in that network and the set of allowable routes in network (i). As the flow on e_4 is always equal to the number of players N minus the flow on e_3, it is easy to conclude that the undirected version of (i), the Wheatstone network, also has the flow monotonicity property.

The last conclusion, Lemmas 3 and 4, Proposition 3 and Theorem 1 together give the following proposition, which is relevant, in particular, to the extension-parallel network in Figure 1d and contrasts with the "bad" Example 2. The part of the proposition concerning the potential is a slightly stronger version of a result originally proved by Fotakis (2010, Lemma 4; the author remarks that this result is also implicit in the work of Holzman and Law-Yone 1997).

Proposition 4. Let G be an undirected extension-parallel network, the Wheatstone network, or any network that is obtained by connecting two or more of these networks in series. Every symmetric unweighted network congestion game h on G has the flow monotonicity property. Therefore, every equilibrium in h maximizes the potential, and if the marginal social costs are nondecreasing, then every equilibrium in the modified game h^1, with the aggregate payoff as the social payoff function, maximizes the latter.

Undirected extension-parallel networks are special also with respect to two alternative notions of social optimality, or efficiency, and with respect to particular dynamics in symmetric unweighted network congestion games.

An undirected two-terminal network satisfies *strong-Nash equivalence* if in every symmetric unweighted network congestion game on it every equilibrium is a *strong equilibrium*. Holzman and Law-Yone (1997, 2003) show that a network[7] has this property if and only if it is extension-parallel. An equivalent way of stating this result is that, for a undirected two-terminal network, being extension-parallel is a necessary and sufficient condition for *weak Pareto efficiency* of all equilibria in all games on a network, meaning that it is never possible to make everyone better off by altering the players' equilibrium route choices. The equivalence holds because an equilibrium is strong if and only if the strategy choices of

[7] These authors actually establish their results for *directed* networks (and nonnegative costs). However, it is not very difficult to conclude from these results that they hold also for undirected networks, for example, by using Proposition A1 and Corollary A2 in the Appendix.

every group of players constitute a weak Pareto efficient equilibrium in the subgame defined by fixing the strategies of the remaining players. That subgame is itself a symmetric unweighted network congestion game on the same network.

Epstein et al. (2009) call an undirected two-terminal network *efficient* if in every symmetric unweighted network congestion game on it every equilibrium minimizes the *makespan*, that is, it is optimal with respect to the maximum (rather than the sum) of the players' costs. They show that an undirected network is efficient if and only if it is extension-parallel.

Kuniavsky and Smorodinsky (2013) study the relation between the set of equilibria in a symmetric unweighted congestion game and its set of *greedy strategy profiles*. The latter set consists of all strategy profiles that can be obtained in a dynamic setting where players join the game sequentially according to some pre-determined order and each player, upon arrival, myopically chooses a strategy that best responds to his predecessors' choices. Their main result (Theorem 4.1) implies that the above two sets coincide in all symmetric unweighted network congestion games on an undirected two-terminal network if and only if it is extension-parallel.

4 Congestion Games with Splittable flow

A *congestion game with splittable flow* is obtained from a (weighted) congestion game as in Section 3 (hereafter, a game with *unsplittable* flow) by replacing the strategy set S_i of each player i with its convex hull $\operatorname{conv} S_i$. Thus, a strategy $x_i = (x_{ie})_{e\in E}$ is now any convex combination of elements of S_i (hereafter, *pure* strategies), with coefficients that express the fraction of the player's weight w_i "shipped" on each pure strategy. The total amount of use by player i of each resource e is expressed by the corresponding component x_{ie} ($\in [0, w_i]$).

An important special case is that of linear cost functions, where for each resource e

$$c_e(t) = a_e t + b_e,$$

with $a_e \geq 0$ (but not necessarily so for b_e). If $a_e > 0$, the cost function is strictly increasing, and if $a_e = 0$, it is constant. Two strategy profiles x and y are *essentially equivalent* if $x_e = y_e$ for every e with $a_e > 0$ and *essentially equal* if $x_{ie} = y_{ie}$ for every i and e with $a_e > 0$.

Linearity of the cost functions implies that the game is a potential game, as it is not difficult to check that the function P defined as follows (Fotakis et al. 2005) is an (exact) potential:

$$P(x) = -\sum_{e\in E}\left(a_e \frac{x_e^2 + \sum_{i=1}^N x_{ie}^2}{2} + b_e x_e\right).$$

With the social payoff function f defined by (8) (i.e., the aggregate payoff) and altruism coefficient r, the modified game has the following potential (see (7)):

$$P^r(x) = -\sum_{e\in E}\left(a_e \frac{(1+r)x_e^2 + (1-r)\sum_{i=1}^N x_{ie}^2}{2} + b_e x_e\right) \tag{14}$$

$$= f(x) + (1-r)\sum_{e\in E} a_e \sum_{\substack{i,j\\ i<j}} x_{ie}x_{je}.$$

Lemma 5. For $-1 \le r \le 1$, the potential P^r is a concave function on the set of strategy profiles, and it is moreover strictly concave if the cost functions are strictly increasing and $r \neq 1$. For every $-1 \le r \le 1$ and strategy profiles x and y,

$$P^r(x) - P^r(y) \le \sum_{i=1}^{N} \frac{d}{d\lambda}\Big|_{\lambda=0^+} P^r(\, y \mid \lambda x_i + (1-\lambda) y_i\,), \tag{15}$$

and equality holds if and only if (i) x and y are essentially equal or (ii) they are essentially equivalent and $r = 1$.

Proof. The left-hand side of (15) is equal to

$$-\sum_{e\in E}\left(a_e \frac{(1+r)(x_e^2 - y_e^2) + (1-r)\sum_{i=1}^{N}(x_{ie}^2 - y_{ie}^2)}{2} + b_e(x_e - y_e)\right). \tag{16}$$

The one-sided derivative on the right-hand side of (15) can be computed by replacing x in (16) with the strategy profile $y \mid \lambda x_i + (1-\lambda)y_i$ and using the definition

$$\frac{d}{d\lambda}\Big|_{\lambda=0^+} P^r(\, y \mid \lambda x_i + (1-\lambda) y_i\,) = \lim_{\lambda\to 0^+} \frac{P^r(y \mid \lambda x_i + (1-\lambda)y_i) - P^r(y)}{\lambda}. \tag{17}$$

This computation gives that the right-hand side of (15) is equal to

$$-\sum_{i=1}^{N}\sum_{e\in E}(a_e(1+r)y_e + a_e(1-r)y_{ie} + b_e)(x_{ie} - y_{ie}). \tag{18}$$

The difference between (18) and (16) simplifies to

$$\sum_{e\in E} a_e\left(\frac{1+r}{2}(x_e - y_e)^2 + \frac{1-r}{2}\sum_{i=1}^{N}(x_{ie} - y_{ie})^2\right).$$

This sum is nonnegative, and is zero only in the two cases indicated. The concavity or strict concavity of P^r now follows quite easily from the fact that for every y expression (18) (hence, the right-hand side of (15)) is affine as a function of x and vanishes at $x = y$. ■

Proposition 5. For a congestion game with splittable flow h where the cost functions are linear, the aggregate payoff as the social payoff function, and altruism coefficient $-1 \le r \le 1$, the set of equilibria in the modified game h^r is nonempty and coincides with the set of maximizers of the potential P^r. If $r \neq 1$, all the equilibria in h^r are essentially equal, and if in addition the cost functions are strictly increasing, then there is in fact only one equilibrium y^r and the mapping $r \mapsto y^r$ is continuous on $[-1,1)$.

Proof. The function P^r is continuous and its domain, the set of strategy profiles, is compact. Therefore, the set of all maximizers of P^r is nonempty. By Lemma 5, this set $\mathcal{X}$ includes the set $\mathcal{Y}$ of all strategy profiles y for which the right-hand side of (15) (equivalently, expression (18)) is nonpositive for all x. On the other hand, $\mathcal{Y}$ includes a superset of $\mathcal{X}$, namely, the set $\mathcal{Z}$ of all equilibria in h^r. This is because, if y is an equilibrium in h^r, then the nominator on the right-hand side of (17) is nonpositive for all i, as it gives the change in player i's payoff resulting from a unilateral deviation to $\lambda x_i + (1-\lambda)y_i$. The inclusions prove that $\mathcal{X} = \mathcal{Y} = \mathcal{Z}$. It follows that, for equilibria x and y in h^r, the left-hand side of (15) is zero and its right-hand side is nonpositive, which implies that this inequality holds as equality and therefore conditions (i) or (ii) in Lemma 5 hold. If $r \neq 1$ and all the cost functions are strictly increasing, the conclusion means that $x = y$, which proves that h^r has a unique equilibrium, denoted y^r. To prove that $r \mapsto y^r$ is a continuous function on $[-1,1)$, it suffices to show that the set

$$\mathcal{A} = \{\,(r, y) \mid -1 \leq r \leq 1, y \text{ is an equilibrium in } h^r\,\}$$

is compact. Expression (18) is continuous (indeed, a polynomial) as a function of x, y and r. Therefore, for every strategy profile x, the set

$$\mathcal{A}_x = \{\,(r, y) \mid -1 \leq r \leq 1, \text{expression (18) is nonpositive}\,\}$$

is compact. Since, as shown at the beginning of the proof,

$$\mathcal{A} = \bigcap_x \mathcal{A}_x, \tag{19}$$

$\mathcal{A}$ is compact too.

For later reference, note that, because for any r and y expression (18) is affine as a function of x, that expression is nonpositive for all strategy profiles x if and only if this is so for all profiles of *pure* strategies. Therefore, the intersection in (19) would not be affected if the index set were replaced by the set of all pure strategy profiles, which is a finite set. ■

With complete internalization of social welfare, $r = 1$, the potential P^1 coincides with the social payoff function. Proposition 5 therefore shows that, in this extreme case, all the equilibria in the modified game maximize social welfare. The next theorem extends this observation by also examining the connection between lower levels of altruism and the corresponding equilibrium levels of the aggregate payoff. As it shows, this connection (which is expressed in the theorem by the function π) is positive.

Theorem 2. For a congestion game with splittable flow h where the cost functions are linear, and the aggregate payoff as the social payoff function f, there is a function $\pi: [-1,1] \longrightarrow \mathbb{R}$ such that, for every $-1 \leq r \leq 1$, all the equilibria y^r in the modified game h^r satisfy

$$\pi(r) = f(y^r).$$

The function π is absolutely continuous, and there is a partition of $[-1,1]$ into finitely many intervals such that in each interval π is either constant or strictly increasing.

Proof. By Proposition 5, any two equilibria x and y in h^r satisfy $P^r(x) = P^r(y)$, and if $r \neq 1$, they are also essentially equal. It follows, by (14), that $f(x) = f(y)$. The conclusion proves that the projection on the first two coordinates of the set

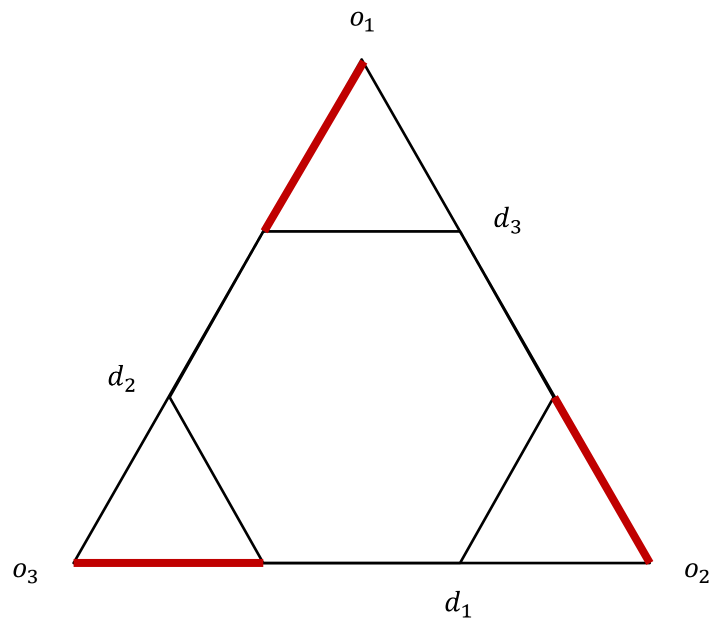

Figure 2. A three-player congestion game with splittable flow. Each player i has to ship a unit flow from o_i to d_i through one or both connecting paths. The cost per unit weight for each edge is equal to the total flow on it, except for the three edges marked by thick lines, where there is an additional (per unit) fixed cost of 2.

$$\mathcal{A}^* = \{\, (r, \alpha, y) \mid -1 \le r \le 1,\ \alpha = f(y),\ y \text{ is an equilibrium in } h^r \,\}$$

is the graph of a function, π: $[-1,1] \longrightarrow \mathbb{R}$. By the Tarski–Seidenberg theorem, a sufficient condition for π to be continuous and semialgebraic is that the set $\mathcal{A}^*$ is closed (equivalently, compact) and semialgebraic. This condition holds because $\alpha = f(y)$ is a polynomial equality and, as shown in the proof of Proposition 5, y is an equilibrium in h^r if and only if the polynomial (18) is nonpositive for every pure strategy profile x. Thus, π is a continuous semialgebraic function.

The last result implies that for every $-1 \le r \le 1$ there is some $\epsilon > 0$, a pair of positive integers p_R and p_L and a pair of analytic functions ξ_R, ξ_L: $(-2\epsilon, 2\epsilon) \longrightarrow \mathbb{R}$ such that $\pi(s) = \xi_R((s-r)^{1/p_R})$ for all $s \in [r, \min(r + \epsilon^{p_R}, 1)]$ and $\pi(s) = \xi_L((r-s)^{1/p_L})$ for all $s \in [\max(r - \epsilon^{p_L}, -1)\,, r]$ (see Bochnak et al. 1998, Proposition 8.1.12). By decreasing ϵ, if necessarily, it can be guaranteed that each of the two analytic functions ξ_R and ξ_L is either constant or strictly monotone in $[0, \epsilon]$. The collection of all intervals of the form $(r - \epsilon^{p_L}, r + \epsilon^{p_R})$, where r varies over all points in $[-1,1]$ and ϵ, p_R and p_L vary correspondingly, constitutes an open cover of $[-1,1]$. Any finite subcover yields a finite set of points, $-1 = r_0 < r_1 < \cdots < r_L = 1$, such that for all $1 \le l \le L$ the restriction of π to the interval $[r_{l-1}, r_l]$ is constant or strictly monotone and has a continuous derivative in (r_{l-1}, r_l). These properties imply that the (continuous) function π is absolutely continuous in the interval. It remains to observe that, by Proposition 1, the function cannot be strictly decreasing there. ■

The next example illustrates the theorem.

Example 3. In a three-player congestion game with splittable flow, E is the set of all edges in the undirected graph shown in Figure 2. The weight of each player i is 1, and the player's two pure strategies are the short, three-edge path connecting i's origin vertex o_i with the destination d_i and the longer, four-edge path that reaches d_i from the opposite direction.[8] The cost of each edge e is equal to the flow on it x_e, except for the three "thick" edges, for which the cost is $x_e + 2$. The minimal total cost of 9 is achieved when all three players use their short paths. If the altruism coefficient satisfies $1/3 \leq r \leq 1$, then this strategy profile is also the unique equilibrium in the modified game, where the payoffs are given by (9). However, for a lower altruism coefficient, $-1 \leq r < 1/3$, it is not an equilibrium. Instead, the unique equilibrium is for every player to ship only $(9-r)/(10-4r)$ of the weight on the short path and the rest on the long path. The corresponding social payoff $\pi(r)$, which is the negative of the aggregate cost, is given by

$$\pi(r) = -\left(169\left(\frac{1-r}{10-4r}\right)^2 + 8\right).$$

As r increases from -1 to $1/3$, $\pi(r)$ increases to the maximum social payoff. Put differently, as the players increasingly internalize the social cost, the latter decreases, until it reaches its minimum and no longer changes. Thus, the price of anarchy is nonincreasing.

Theorem 2 shows that nonincreasing price of anarchy with respect to the aggregate payoff actually holds for any game with linear, nonnegative cost functions (that is, $b_e \geq 0$ for every resource e). The following corollary combines this finding with an extension of the result that, without altruism or spite (that is, for $r = 0$), the price of anarchy for such games does not exceed $3/2$ (Cominetti et al. 2009).

Corollary 2. For the family $\mathfrak{H}$ of all congestion games with splittable flow where the cost functions are linear and nonnegative, as well as every subfamily or single game in $\mathfrak{H}$, the price of anarchy PoA^r is nonincreasing for $-1 \leq r \leq 1$ and satisfies

$$1 \leq \mathrm{PoA}^r \leq \frac{1}{1+r} + \frac{1}{2}. \tag{20}$$

Proof. Consider, more generally, any N-player congestion game with splittable flow h where each cost function c_e is nonnegative and differentiable. It follows as an immediate conclusion from Theorem 3.1 and Proposition 3.2 in Cominetti et al. (2009) that the (nonpositive) aggregate payoff f satisfies

$$f(x) \leq (1 - \max_{e \in E} B^N(c_e)) f(y) \tag{21}$$

for every strategy profile x, every equilibrium y and any assignment of a number $B^N(c_e)$ to each cost function c_e such that

$$t c_e(t) B^N(c_e) \geq s\,(c_e(t) - c_e(s)) + (s^2/4 - (t - s/2)^2/N)\, c_e'(t), \qquad s, t \geq 0. \tag{22}$$

A straightforward generalization of the proofs of these results yields a similar conclusion in the more general case where y is an equilibrium in the modified game h^r, for any $-1 \leq r \leq 1$. The only difference is that the right-hand side of the inequality in (22) is

[8] Note that this (asymmetric) game differs from the network congestion game considered in Section 3.3 in that edges can be traversed in both directions.

replaced with

$$s\,(c_e(t) - c_e(s)) + \big(r(s-t)t + (1-r)(s^2/4 - (t-s/2)^2/N)\big)\, c_e'(t).$$

In the linear case, where the derivative c_e' is constant and therefore $tc_e(t) \geq t^2 c_e'(t)$, the modified inequality holds for

$$B^N(c_e) = \sup_{s,t>0} \frac{s\,(t-s) + r(s-t)t + (1-r)(s^2/4 - (t-s/2)^2/N)}{t^2}$$
$$= \sup_{\gamma>0}\big(\gamma\,(1-\gamma) + r(\gamma-1) + (1-r)(\gamma^2/4 - (1-\gamma/2)^2/N)\big) = \frac{1-r}{3+r+4/(N-1)}.$$

Substitution in (21) and rearrangement give

$$\left(\frac{1}{1+r+(1-r)/N} + \frac{1}{2}\right) f(x) \leq f(y).$$

This inequality proves (20). (It also slightly improves the upper bound for fixed N.) ■

Corollary 2 points to a striking difference between splittable and unsplittable flow. As indicated in Section 3.2, for unweighted congestion games with unsplittable flow that have linear, nonnegative cost functions, PoA^r is strictly *increasing* for $0 \leq r \leq 1$, from $\mathrm{PoA}^0 = 5/2$ to $\mathrm{PoA}^1 = 3$.

5 Nonatomic Congestion Games

A *nonatomic congestion game* has an infinite set, or population, of players I (modeled, e.g., as the unit interval $[0,1]$), endowed with a nonatomic probability measure μ, the population measure (e.g., Lebesgue measure). The players share a finite set E of resources and a finite strategy set $S \subseteq \{0,1\}^E$ whose elements (referred to in the following as *pure* strategies) are non-zero binary vectors. A *strategy profile* is any assignment of a (pure) strategy $\sigma(i) = (\sigma_e(i))_{e\in E} \in S$ to each player i such that, for each resource e, the set of all players i with $\sigma_e(i) = 1$ is measurable. The measure (or "size") y_e of this set, which can be written as

$$y_e = \int \sigma_e(i)\, d\mu(i),$$

is the flow on resource e. The flow vector $y = (y_e)_{e\in E} = \int \sigma(i)\, d\mu(i)$, which represents the population's mean strategy, is called the *population strategy*. It lies in $\operatorname{conv} S$, the convex hull of S. The population strategy determines the cost of using each resource e, which is given by $c_e(y_e)$, where $c_e\colon [0,\infty) \longrightarrow \mathbb{R}$ is a strictly increasing and continuously differentiable cost function. The total cost for each player i is the sum of the costs of the resources e with $\sigma_e(i) = 1$, and the player's payoff is the negative of the cost. A natural, linear extension assigns to each $x = (x_e)_{e\in E} \in \operatorname{conv} S$ the payoff

$$g(x,y) = -\sum_{e\in E} x_e\, c_e(y_e).$$

This equation defines a population game g (see Section 2), where the strategy space is

$X = \operatorname{conv} S$. A corresponding social payoff function φ is given by the *mean payoff* $g(y, y)$:

$$\varphi(y) = -\sum_{e \in E} y_e \, c_e(y_e). \tag{23}$$

(Note that this expression is meaningful also for elements of $\hat{X}$, the cone of X.) With altruism coefficient r, the modified game g^r (which is defined by (2)) is given by

$$g^r(x, y) = -\sum_{e \in E} x_e \left(c_e(y_e) + r \, y_e \, c_e'(y_e) \right) = -\sum_{e \in E} x_e \, \hat{c}_e^r(y_e), \tag{24}$$

where the function $\hat{c}_e^r$ is defined by

$$\hat{c}_e^r = (1 - r) \, c_e + r \, MC_e,$$

with MC_e denoting the *marginal social cost*,

$$MC_e(t) = \frac{d}{dt} t c_e(t) = c_e(t) + t \, c_e'(t), \qquad t \geq 0. \tag{25}$$

If MC_e is strictly increasing and continuously differentiable for every resource e, and $0 \leq r \leq 1$, then g^r may also be viewed as a population game representing a nonatomic congestion game, namely, the one that differs from the original game in that each cost function c_e is replaced with $\hat{c}_e^r$. A sufficient condition for MC_e to be strictly increasing in any interval is that c_e is convex there.

The population game g defined above is a potential game, with the (well known; see Milchtaich 2004, Section 5) potential $\Phi \colon X \longrightarrow \mathbb{R}$ defined by

$$\Phi(x) = -\sum_{e \in E} \int_0^{x_e} c_e(t) \, dt.$$

This is because the extension of Φ to $\hat{X}$, which is defined by the same formula, satisfies condition (6). Since the cost functions c_e are strictly increasing, Φ is strictly concave. If the marginal social costs are strictly increasing in $[0,1]$, then the mean payoff φ is also strictly concave. With $0 \leq r \leq 1$, the same then holds for $\Phi^r = (1 - r)\Phi + r\varphi$, which as indicated (see Section 2.1) is a potential for g^r. This continuous and strictly concave function necessarily has a unique maximum point in X.

Proposition 6. Consider a nonatomic congestion game where for each resource e the marginal social cost MC_e is strictly increasing in $[0,1]$. For the corresponding population game g, and the mean payoff as the social payoff function, for every $0 \leq r \leq 1$ the unique maximizer y^r of the potential Φ^r is the unique equilibrium strategy in the modified game g^r, and the mapping $r \mapsto y^r$ is continuous.

Proof. Since, as indicated, Φ^r is strictly concave, by the remarks that follow Proposition 2 its unique maximum point y^r is also the unique equilibrium strategy in g^r. To prove that the function $r \mapsto y^r$ is continuous, it suffices to show that its graph is closed. A point (r, y), with $0 \leq r \leq 1$, lies on the graph if and only if the inequality in (3) holds for every strategy x.

Since, for every x, both sides of the inequality are continuous as functions of r and y (see (24)), the set of all pairs (r, y) satisfying the inequality is closed, which implies the same for the above graph.

For later reference, note that, since $g^r(x, y)$ is linear in the first argument x, the requirement that the inequality in (3) holds for every strategy x can be replaced by the requirement that it holds for every *pure* strategy, as the latter automatically implies the former. ■

The next theorem shows that, like the congestion games with splittable flow with linear cost functions studied in Section 4, nonatomic congestion games with increasing marginal social costs always exhibit positive comparative statics.

Theorem 3. For the population game g corresponding to a nonatomic congestion game with increasing marginal social costs as in Proposition 6, the mean payoff as the social payoff function φ, and altruism coefficients r and s with $0 \leq r < s \leq 1$, let y^r and y^s be the equilibrium strategies in the modified games g^r and g^s, respectively. If $y^r \neq y^s$, then

$$\varphi(y^r) < \varphi(y^s).$$

If $s = 1$, then moreover $\varphi(x) < \varphi(y^s)$ for every strategy $x \neq y^s$.

Proof. Since, by Proposition 6, each of the two equilibrium strategies is a strict maximum point of the corresponding potential, Φ^r or Φ^s, the first inequality follows from Proposition 2 and the second inequality follows from the fact that $\Phi^1 = \varphi$. ■

The next result provides a more specific description in a special case.

Theorem 4. Consider a nonatomic congestion game where for each resource e the cost function c_e is a polynomial and the marginal social cost MC_e is strictly increasing in $[0,1]$. For the corresponding population game g, and the mean payoff as the social payoff function φ, define the function $\pi: [0,1] \longrightarrow \mathbb{R}$ by

$$\pi(r) = \varphi(y^r),$$

where y^r is the unique equilibrium strategy in the modified game g^r. The function π is absolutely continuous, and there is a partition of $[0,1]$ into finitely many intervals such that in each interval π is either constant or strictly increasing.

Proof. The graph of π is the projection on the first two coordinates of the set

$$\mathcal{A}^* = \{\, (r, \alpha, y) \mid 0 \leq r \leq 1,\ \alpha = \varphi(y),\ y \text{ is an equilibrium strategy in } g^r \,\}.$$

As indicated in the proof of Proposition 6, y is an equilibrium in g^r if and only if the inequality in (3) holds for all pure strategies x. As this requirement involves a finite number of polynomial inequalities (see (24)), and $\alpha = \varphi(y)$ is a polynomial equality (see (23)), $\mathcal{A}^*$ is a semialgebraic compact set. The rest of the proof is similar to the proof of Theorem 2, except that it relies on Theorem 3 rather than Proposition 1. ■

A slightly more special case is considered in the following corollary. The proof is similar to that of Corollary 1 expect that it employs Theorem 3 rather than Theorem 1.

Corollary 3. Consider the family of all nonatomic congestion games where for each resource e the cost function c_e is a (non-constant) polynomial with nonnegative coefficients. For the corresponding family of population games $\mathfrak{G}$, as well as every subfamily or single game in $\mathfrak{G}$, the price of anarchy PoA^r is nonincreasing for $0 \leq r \leq 1$ and satisfies $1 \leq \mathrm{PoA}^r \leq 1/r$ (hence, $\mathrm{PoA}^1 = 1$).

The bound $\mathrm{PoA}^r \leq 1/r$, which was first established by Chen and Kempe (2008), can actually be improved upon. These authors (see also Meir and Parkes 2015) show that, for the subfamily where the cost functions are polynomials of degree at most k (≥ 1), the price of anarchy can be determined precisely. Specifically, for $0 < r \leq 1$,

$$\mathrm{PoA}^r = \frac{1}{1 + rk - k\left(\frac{1 + rk}{1 + k}\right)^{1+\frac{1}{k}}}.$$

(This expression for the price of anarchy is obtained in the limit $a \to 0^+$ in the Pigou example, where there are two resources, a pure strategy is choosing either resource, and the cost functions are $c_{e_1}(t) = t^k$ and $c_{e_2}(t) = 1 + rk + at$.) As k increases, the above expression monotonically increases and tends to the limit

$$\frac{1}{r(1 - \ln r)}$$

(and not to $1/r$, as Chen and Kempe 2008 and Chen et al. 2014 mistakenly assert). The latter is therefore the least upper bound, hence the PoA^r, for all polynomial cost functions with nonnegative coefficients.

6 Comparative Statics and Stability

Comparative statics are closely linked with the stability or instability of the concerned equilibria or equilibrium strategies in the corresponding modified games (Milchtaich 2012). More specifically, stability is associated with “normal”, positive comparative statics, whereby an increase in the altruism coefficient increases social welfare, and (definite) instability is associated with negative comparative statics, in which the opposite relation holds. This association, which is detailed below, provides a general context to the results obtained in the previous sections, as it does not depend on the existence of a potential.

The notion of stability relevant for comparative statics is *static stability* (Milchtaich 2020). This concept differs from dynamic stability in not involving an extraneous law of motion. Instead, stability is completely determined by the game itself, that is, by the players’ payoffs. In symmetric games, static stability generalizes a number of more special stability concepts such as evolutionarily stable strategy, or ESS. The definition of static stability in asymmetric N-player games can be based on that in symmetric games, with the link between the two definitions provided by the concept of symmetrization of an asymmetric game (Milchtaich 2012). However, that definition is equivalent to the following direct one (Milchtaich 2020), where Π denotes the set of all permutation of $(1,2,\ldots,N)$ and, for a set of players $S \subseteq \{1,2,\ldots,N\}$, $y \mid x_S$ is the strategy profile in which the players in S play according to the strategy profile x and everyone else plays according to y.

Definition 3. A strategy profile y in an N-player game h is *stable*, *weakly stable* or *definitely unstable* if it has a neighborhood where, for every strategy profile $x \neq y$, the expression

$$\frac{1}{N!}\sum_{\pi\in\Pi}\sum_{j=1}^{N}\left(h_{\pi(j)}(y \mid x_{\{\pi(1),\pi(2),\ldots,\pi(j)\}}) - h_{\pi(j)}(x \mid y_{\{\pi(j),\pi(j+1),\ldots,\pi(N)\}})\right) \tag{26}$$

is negative, nonpositive or positive, respectively. If a similar condition holds for *all* strategy profiles $x \neq y$, then y is *globally* stable, weakly stable or definitely unstable, respectively.

The stability condition means that the players' overall incentive to move from y is negative. That is, if they change their strategies one-by-one, with each player i moving from strategy y_i to a nearby alternative strategy x_i, the movers *on average* lose (with the average taking into consideration all N players and all $N!$ orders of moves). As the next lemma shows, this requirement can also be interpreted as the condition that, when players only play according to x or according to y, those doing the former fare worse on average. For strategy profiles x and y, define the *payoff of x players when playing against y players* as the quantity

$$\mathcal{H}(x,y) = \sum_{j=1}^{N}\left[\frac{1}{\binom{N}{j}}\sum_{\substack{S\\|S|=j}}\bar{h}_S(y \mid x_S)\right] = \sum_{S}\frac{1}{\binom{N}{|S|}}\bar{h}_S(y \mid x_S) = \sum_{S}\frac{1}{\binom{N}{|S|}}\bar{h}_{S^{\mathrm{C}}}(x \mid y_S),$$

where $|S|$ is the number of players in a set S and $\bar{h}_S = (1/|S|)\sum_{i\in S} h_i$ is their average payoff, which is defined as 0 if $S = \emptyset$. (The third equality is obtained by replacing the summation variable S with the complementary set S^{C} and using the identity $y \mid x_{S^{\mathrm{C}}} = x \mid y_S$.)

Lemma 6 (Milchtaich 2020). Expression (26) is equal to $\mathcal{H}(x,y) - \mathcal{H}(y,x)$.

Static stability and instability are local concepts, defined with respect to a specified topology on each player's strategy set X_i. The product topology on the set $X = \prod_i X_i$ of all strategy profiles gives a meaning to a neighborhood of a strategy profile x: it is any set of strategy profiles whose interior includes x. In principle, any topologies on the strategy sets may be chosen. However, in many games, there are unique natural ones, usually determined by the Euclidean distance between strategies. Global stability and instability correspond to the choice of the *trivial topologies*, where the only neighborhood of any strategy profile is the entire set X. These strong properties imply stability or instability with respect to *any* topologies. Global stability (and global weak stability) of a strategy profile moreover implies that it is an equilibrium, which is not generally true for stability.

The concept of static stability has a somewhat different meaning in population games, where it applies to (population) strategies rather than strategy profiles (Milchtaich 2020). This difference from N-player games reflects the different interpretation of population games as describing a large crowd of interacting individuals.

Definition 4. A strategy y in a population game g is *stable*, *weakly stable* or *definitely unstable* if it has a neighborhood where, for every strategy $x \neq y$, the integral

$$\int_0^1 \big(g(x, px + (1-p)y) - g(y, px + (1-p)y)\big)\, dp$$

is negative, nonpositive or positive, respectively. If a similar condition holds for *all* strategies $x \neq y$, then y is *globally* stable, weakly stable or definitely unstable, respectively.

In the special case of potential games, stability and instability have particularly simple characterizations in terms of the extremum points of the potential.

Theorem 5 (Milchtaich 2020). A strategy profile y in an N-player game with a potential P is stable, weakly stable or definitely unstable if and only if y is, respectively, a strict local maximum, local maximum or strict local minimum point of P. A *global* maximum point of P is both globally weakly stable (and if it is a strict global maximum point, globally stable) and an equilibrium.

Theorem 6 (Milchtaich 2020). A strategy y in a population game with a potential Φ is stable, weakly stable or definitely unstable if and only if y is, respectively, a strict local maximum, local maximum or strict local minimum point of Φ. In the first two cases, y is in addition an equilibrium strategy. If the potential Φ is strictly concave, then an equilibrium strategy is necessarily *globally* stable, and it is therefore the game's unique stable strategy.

Theorem 5 implies that the equilibria considered in Theorems 1 and 2 are globally weakly stable, and Theorem 6 implies that the equilibrium strategies in Theorems 3 and 4 are globally stable. The significance of these findings is elucidated by the next four theorems, which present fundamental links between static stability and the effect of altruism and spite on social welfare in general games. The first two theorems are similar to Theorems 1 and 3 in that they concern *global* comparative statics (Milchtaich 2012, Sections 6 and 7.1), that is, comparison between two strategy profiles in two modified games corresponding to different altruism coefficients r and s, without assuming that the strategy profiles or the coefficients are close or that it is possible to connect them in a continuous manner. The other two theorems are similar to Theorems 2 and 4 in that they concern *local* comparative statics, which involve small, continuous changes to the altruism coefficient r and the corresponding strategies or strategy profiles, and may be thought of as tracing the players' evolving behavior as they respond to the changing r. Note that, like Propositions 1 and 2, the four theorems hold for any choice of social payoff functions.

Theorem 7 For a game h, a social payoff function f, and altruism coefficients r and s with $r < s \leq 1$, if two distinct strategy profiles y^r and y^s are globally weakly stable in the modified games h^r and h^s, respectively, then

$$f(y^r) \leq f(y^s).$$

If moreover y^s is globally stable, then the inequality is strict. A strategy profile that is globally weakly stable or globally stable in h^1 is a maximum or strict maximum point, respectively, of f in the set of all strategy profiles.

Proof. The proof uses the following identity, which holds for all (r, s and) strategy profiles x and y:

$$
\begin{aligned}
(1-r)\sum_S \frac{1}{\binom{N}{|S|}}\Big(\bar{h}_S^s(y \mid x_S) - \bar{h}_{S^c}^s(y \mid x_S)\Big) + (1-s)\sum_S \frac{1}{\binom{N}{|S|}}\Big(\bar{h}_{S^c}^r(y \mid x_S) - \bar{h}_S^r(y \mid x_S)\Big) \\
= (1-r)(1-s)\sum_S \frac{1}{\binom{N}{|S|}}\Big((\bar{h}_S(y \mid x_S) - \bar{h}_{S^c}(y \mid x_S)) + (\bar{h}_{S^c}(y \mid x_S) - \bar{h}_S(y \mid x_S))\Big) \\
+(1-r)s\,(f(x)-f(y)) + (1-s)r\,(f(y)-f(x)) = (s-r)\big(f(x)-f(y)\big).
\end{aligned}
$$

The identity implies that the difference $f(x) - f(y)$ is nonpositive or negative if the first term on the left-hand side is nonpositive or negative, respectively, and the second term is nonpositive. It follows from Lemma 6 that this condition holds with $x = y^r$ and $y = y^s$ if the latter strategy profile is globally weakly stable or globally stable, respectively, in h^s and the former is globally weakly stable in h^r. For $s = 1$, the condition also holds with any other $x \neq y^s$. ∎

Theorem 8. For a population game g, a social payoff function φ, and altruism coefficients r and s with $r < s \leq 1$, if two distinct strategies y^r and y^s are globally weakly stable in the modified games g^r and g^s, respectively, then

$$\varphi(y^r) \leq \varphi(y^s).$$

If moreover y^s is globally stable, then the inequality is strict. A strategy that is globally weakly stable or globally stable in g^1 is a maximum or strict maximum point, respectively, of φ in the set of all strategies.

Proof. The proof uses the following identity, which holds for all (r, s and) strategies x and y:

$$
\begin{aligned}
(1-r)\int_0^1 \big(g^s(x, px+(1-p)y) - g^s(y, px+(1-p)y)\big)\,dp \\
+ (1-s)\int_0^1 \big(g^r(y, py+(1-p)x) - g^r(x, py+(1-p)x)\big)\,dp \\
= (1-r)(1-s)\int_0^1 \big(g(x, px+(1-p)y) - g(y, px+(1-p)y)\big)\,dp \\
+ (1-s)(1-r)\int_0^1 \big(g(y, px+(1-p)y) - g(x, px+(1-p)y)\big)\,dp \\
+ (s-r)\int_0^1 \big(d\varphi(x, px+(1-p)y) - d\varphi(y, px+(1-p)y)\big)\,dp \\
= (s-r)\int_0^1 \frac{d}{dp}\varphi(px+(1-p)y)\,dp = (s-r)\big(\varphi(x)-\varphi(y)\big).
\end{aligned}
$$

(The first equality employs a change of integration variable, from p to $1-p$, in the second integral. The second equality uses (5), with Φ replaced by φ.) The identity implies that the difference $\varphi(x) - \varphi(y)$ is nonpositive or negative if the first term on the left-hand side is nonpositive or negative, respectively, and the second term is nonpositive. This condition holds with $x = y^r$ and $y = y^s$ if the latter strategy is globally weakly stable or globally stable, respectively, in g^s and the former is globally weakly stable in g^r. For $s = 1$, the condition also holds with any other $x \neq y^s$. ∎

Theorem 9 (Milchtaich 2012). For a game h and a social payoff function f such that the payoff functions and the social payoff function are Borel measurable,[9] and altruism coefficients r_0 and r_1 with $r_0 < r_1 \leq 1$, suppose that there is a continuous and finitely-many-to-one[10] function assigning to each $r_0 \leq r \leq r_1$ a strategy profile y^r such that the function $\pi: [r_0, r_1] \longrightarrow \mathbb{R}$ defined by

$$\pi(r) = f(y^r)$$

is absolutely continuous.[11] If the strategy profile y^r is stable, weakly stable or definitely unstable in the modified game h^r for every $r_0 < r < r_1$, then π is strictly increasing, nondecreasing or strictly decreasing, respectively.

Theorem 10 (Milchtaich 2012). For a population game g and a social payoff function φ such that the payoff function and $d\varphi$ are Borel measurable, and altruism coefficients r_0 and r_1 with $r_0 < r_1 \leq 1$, suppose that there is a continuous and finitely-many-to-one function assigning to each $r_0 \leq r \leq r_1$ a strategy y^r such that the function $\pi: [r_0, r_1] \longrightarrow \mathbb{R}$ defined by

$$\pi(r) = \varphi(y^r)$$

is absolutely continuous. If the strategy y^r is stable, weakly stable or definitely unstable in the modified game g^r for every $r_0 < r < r_1$, then π is strictly increasing, nondecreasing or strictly decreasing, respectively.

The last four theorems present a broader perspective than Theorems 1, 2, 3 and 4 do. First, they consider almost completely general games and social payoff functions. Second, they are not restricted to equilibria or equilibrium strategies. Correspondingly, the last two theorems, which concern local comparative statics, consider, not specific relations between altruism coefficients and strategies or strategy profiles, but any mapping $r \mapsto y^r$ that is reasonably "smooth" (the continuity and absolute continuity conditions) and "responsive" (the finitely-many-to-one condition). In this general setting, they show that social welfare necessary (weakly) increases with increasing r if each y^r is (respectively, weakly) stable, and decreases in case of definite instability. As indicated, for the strategy profiles and strategies considered in Theorems 2 and 4, weak stability or stability, respectively, holds automatically. It is shown below that the same is true also in some other games of interest, which include kinds of congestion games not covered by the previous sections.

6.1 Games with differentiable payoffs

Consider an N-player game h where the strategy space of each player i is a set in a Euclidean space $\mathbb{R}^{n_i}$ (where the topology is given by the Euclidean distance). With strategies written as column vectors, a strategy profile $x = (x_1, x_2, \ldots, x_N)$ is an n-dimensional column vector, where $n = \sum_i n_i$. It is an *interior* strategy profile if each x_i is an interior strategy in the sense that it lies in the *relative interior* of player i's strategy space. The gradient with respect to the components of player i's strategy is denoted ∇_i and is written as an n_i-dimensional row

[9] A sufficient condition for Borel measurability of a function is that it is continuous.

[10] A function is finitely-many-to-one if the inverse image of every point is a finite set.

[11] A sufficient condition for absolute continuity is that the function is continuously differentiable.

vector (of first-order differential operators). Correspondingly, for each i and j, $\nabla_i^{\mathrm{T}}\nabla_j$ is an $n_i \times n_j$ matrix (of second-order differential operators). In particular, $\nabla_i^{\mathrm{T}}\nabla_i h_i$ is the Hessian matrix of player i's payoff function with respect to the player's own strategy. These Hessian matrices are the diagonal blocks in the $n \times n$ block matrix

$$H = \begin{pmatrix} \nabla_1^{\mathrm{T}}\nabla_1 h_1 & \cdots & \nabla_1^{\mathrm{T}}\nabla_N h_1 \\ \vdots & \ddots & \vdots \\ \nabla_N^{\mathrm{T}}\nabla_1 h_N & \cdots & \nabla_N^{\mathrm{T}}\nabla_N h_N \end{pmatrix}.$$

The value that the matrix H attains when its entries are evaluated at a strategy profile x is denoted $H(x)$.

Theorem 11 (Milchtaich 2020). If y is an interior equilibrium at which the players' payoff functions are twice continuously differentiable,[12] then a sufficient condition for it to be stable is that the matrix $H(y)$ is negative definite.[13] If the players' strategy spaces are convex, then the same is true also with the qualifier 'interior' dropped.

In view of the general connection between static stability and the effect of altruism and spite on social welfare established above, the last result suggests a link between positive comparative statics and negative definiteness of the matrix H^r obtained from H by replacing the payoff function h_i of each player i with the modified payoff h_i^r. The next result, which is a generalization of Proposition 8 in Milchtaich (2012), makes good on this suggestion by establishing such a *direct* link. As it shows, the effect of changing the altruism coefficient r is determined by the negative of a quadratic form whose matrix is H^r. Therefore, if H^r is negative definite at the corresponding equilibria, and the latter change at all when r increases or decreases, then social welfare necessarily also increases or decreases, respectively.

Theorem 12. For a game h where the strategy space of each player is a set in a Euclidean space, a social payoff function f, and altruism coefficients r_0 and r_1 with $r_0 < r_1 \leq 1$, suppose that there is a continuously differentiable function assigning to each $r_0 < r < r_1$ an interior equilibrium y^r in the modified game h^r at which the payoff functions and social payoff function are twice continuously differentiable. The corresponding rate of change of the social payoff is given by

$$\frac{d}{dr} f(y^r) = -(1-r)\left(\frac{dy^r}{dr}\right)^{\mathrm{T}} H^r(y^r)\,\frac{dy^r}{dr}. \tag{27}$$

Proof. For each of the N players i, whose strategy space is a subset of $\mathbb{R}^{n_i}$, let v_i be any vector in the subspace of $\mathbb{R}^{n_i}$ that is spanned by all vectors of the form $x_i - y_i$, where x_i and y_i are strategies of i. (Note that, for all r, dy_i^r / dr lies in this subspace.) Since y_i^r is an

[12] Technically, this assumption means that there is an extension of each payoff function to an open neighborhood of y in the n-dimensional Euclidean space where it has continuous second-order partial derivatives.

[13] A square matrix A is said to be positive or negative definite if the symmetric matrix $(1/2)(A + A^{\mathrm{T}})$ has the same property, equivalently, if the latter's eigenvalues are all positive or negative, respectively.

interior strategy, for sufficiently small $\epsilon > 0$ the expressions $(1/\epsilon)\big(h_i(y^r \mid y_i^r + \epsilon v_i) - h_i(y^r)\big)$ and $-(1/\epsilon)\big(h_i(y^r \mid y_i^r - \epsilon v_i) - h_i(y^r)\big)$ are well defined. Since y^r is an equilibrium in h^r, the first expression is nonpositive and the second one is nonnegative. However, as $\epsilon \to 0$, both expressions tend to $\nabla_i h_i^r(y^r)\, v_i$, and therefore this expression must be zero:

$$\nabla_i h_i^r(y^r)\, v_i = \nabla_i((1-r)h_i + rf)(y^r)\, v_i = 0. \tag{28}$$

Total derivation with respect to r gives

$$\frac{d}{dr}\nabla_i h_i^r(y^r)\, v_i = \nabla_i(-h_i + f)(y^r)\, v_i + \sum_{j=1}^{N}\left(\frac{dy_j^r}{dr}\right)^{\mathrm{T}} \nabla_j^{\mathrm{T}}\nabla_i h_i^r(y^r)\, v_i = 0.$$

Summation over i of both sides of the right equality, choice of $v_i = (1-r)\, dy_i^r / dr$ and use of the equality $\nabla_j^{\mathrm{T}}\nabla_i h_i^r = (\nabla_i^{\mathrm{T}}\nabla_j h_i^r)^{\mathrm{T}}$ (which holds because the first-order partial derivatives commute) give

$$\sum_{i=1}^{N}\nabla_i(-(1-r)h_i + (1-r)f)(y^r)\frac{dy_i^r}{dr} + (1-r)\left(\frac{dy^r}{dr}\right)^{\mathrm{T}} H^r(y^r)\,\frac{dy^r}{dr} = 0.$$

It follows from the second equality in (28) that the first term on the left-hand side (the sum) is equal to $df(y^r)/\,dr$. ■

6.2 Games with splittable flow revisited

The results presented above can be used for extending the study of congestion games with splittable flow beyond the linear case examined in Section 4. The extension adds all cost functions c_e that are twice continuously differentiable and satisfy

$$-c_e'(t) < \frac{1}{2}t\, c_e''(t) < \frac{N+1}{N-1}c_e'(t), \qquad t \geq 0. \tag{29}$$

This pair of inequalities can be written also as

$$0 < (MC_e)' < \frac{4N}{N-1}c_e', \tag{30}$$

where MC_e is the marginal social cost defined in (25). In the proof of the following result, player i's strategies are viewed as vectors in $\mathbb{R}^{E_i}$ (rather than $\mathbb{R}^E$), where E_i is the set of all resources e that player i may actually use, in the sense that $x_{ie} > 0$ for some strategy x_i. Thus, irrelevant coordinates are ignored.

Proposition 7. For a congestion game with splittable flow h where the cost functions are twice continuously differentiable and satisfy (29), the aggregate payoff as the social payoff, and altruism coefficient $0 \leq r < 1$, the matrix $H^r(x)$ is negative definite for every strategy profile x and the modified game h^r has at most one equilibrium, which is necessarily stable.

Proof. By (9), for a strategy profile x and players $1 \leq i, j \leq N$

$$\nabla_j h_i^r(x) = -\left(\left((1-r)1_{i=j} + r\right) c_e(x_e) + ((1-r)x_{ie} + r x_e)\, c_e'(x_e)\right)_{e\in E_j},$$

$$\nabla_i^{\mathrm{T}} \nabla_j h_i^r(x) = -\left(\left(\left((1-r)1_{i=j} + r + 1\right) c_e'(x_e) + ((1-r)x_{ie} + r x_e)\, c_e''(x_e)\right) 1_{e'=e}\right)_{e'\in E_i, e\in E_j}.$$

Therefore, for any column vector $a = (a_1, a_2, \ldots, a_N)$ with $a_i \in \mathbb{R}^{E_i}$ for all i,

$$a^{\mathrm{T}} H^r(x)\, a = -\sum_{i,j=1}^{N} \sum_{e\in E_i\cap E_j} a_{ie} a_{je} \left(\left((1-r)1_{i=j} + r + 1\right) c_e'(x_e) + ((1-r)x_{ie} + r x_e)\, c_e''(x_e)\right)$$
$$= -(1-r) \sum_e c_e'(x_e) \sum_{i,j\in I_e} a_{ie} a_{je} \left(1_{i=j} + \frac{1+r}{1-r} + 2\left(x_{ie} + \frac{r}{1-r} x_e\right) \gamma_e\right),$$

where I_e is the set of all players i with $e \in E_i$ and $\gamma_e = (1/2)c_e''(x_e)/c_e'(x_e)$. (Condition (29) implies that $c_e'(x_e) > 0$.) It follows that $H^r(x)$ is negative definite if and only if for every resource $e \in \bigcup_i E_i$ the matrix

$$\left(1_{i=j} + 1 + 2r\frac{1 + x_e\gamma_e}{1-r} + 2x_{ie}\gamma_e\right)_{i\in I_e, j\in I_e} \tag{31}$$

is positive definite. The (symmetric) matrix with all entries equal to $2r(1 + x_e\gamma_e)/(1-r)$ is positive semidefinite, because $1 + x_e\gamma_e > 0$ by (29). Therefore, a sufficient condition for positive definiteness of (31) is that all eigenvalues of the symmetric matrix

$$C = \left(1 + (x_{ie} + x_{je})\gamma_e\right)_{i\in I_e, j\in I_e}$$

are greater than -1. Since every vector $v \in \mathbb{R}^{I_e}$ orthogonal to both $(1,1,\ldots,1)$ and $(x_{ie})_{i\in I_e}$ satsifies $Cv = 0$, the matrix C has at most two eigenvectors with non-zero eigenvalues. If $\gamma_e = 0$ or $x_{ie} = x_e/N_e$ for all $i \in I_e$, where $N_e = |I_e|$, then there is in fact no more than one such eigenvector, with eigenvalue $N_e + 2x_e\gamma_e$, which (since $2 + 2x_e\gamma_e > 0$) is greater than -1. Assume, then, that neither of these conditions holds, and consider for a variable λ the (non-zero) vector $v = (N_e x_{ie} + \lambda/\gamma_e - x_e)_{i\in I_e}$. The condition that v is an eigenvector with eigenvalue λ (that is, $Cv = \lambda v$) simplifies to $Q(\lambda) = 0$, where

$$Q(\lambda) = \lambda(\lambda - 2x_e\gamma_e - N_e) - x_e^2\gamma_e^2 Z$$

and $Z = N_e\left(\sum_{i\in I_e} x_{ie}^2\right)/x_e^2 - 1 = N_e \sum_{i\in I_e}(x_{ie}/x_e - 1/N_e)^2$. By the assumption that $x_{ie} \neq x_e/N_e$ for some $i \in I_e$, $Z > 0$, so $Q(0) < 0$. Therefore, the quadratic, convex polynomial Q has one positive root and one negative root, which are the remaining two eigenvalues of C. The negative one is greater than -1 if and only if $Q(-1) > 0$. Solving this inequality for $x_e\gamma_e$ gives

$$-\frac{N_e + 1}{\sqrt{1 + (N_e+1)Z} + 1} < x_e\gamma_e < \frac{N_e + 1}{\sqrt{1 + (N_e+1)Z} - 1}. \tag{32}$$

As, clearly, $Z \leq N_e - 1$, the right-hand side in (32) is greater than or equal to

$$\frac{N_e + 1}{\sqrt{1 + (N_e+1)(N_e-1)} - 1} = \frac{N_e + 1}{N_e - 1} \geq \frac{N + 1}{N - 1}$$

and the left-hand side is smaller than or equal to -1. Therefore, (32) is implied by (29). This proves that $H^r(x)$ is negative definite.

By Theorem 11, the negative definiteness and the fact that the strategy spaces are convex imply that every equilibrium in h^r is stable. The same two facts also imply that the game has at most one equilibrium (Rosen 1965, Milchtaich 2020). ∎

Inspection of the proof of Proposition 7 shows that a very similar result holds for a somewhat larger class of cost functions, namely, those that are twice continuously differentiable outside the origin and satisfy (30) there (equivalently, satisfy the pair of inequalities in (29) for $t > 0$). The only difference is that, in this case, 'strategy profile' and 'equilibrium' need to be replaced with 'interior strategy profile' and 'interior equilibrium'. In a congestion game with splittable flow, a strategy of a player i is an interior strategy if and only if it is "completely mixed" in the sense that it can be expressed as a convex combination with positive coefficients of *all* of i's pure strategies.

In the special case of cost functions of the form

$$c_e(t) = a_e t^{d_e} + b_e,$$

with $a_e, d_e > 0$, the first inequality in (30) holds automatically outside the origin and the second one holds there if and only if

$$d_e < \frac{3N+1}{N-1}.$$

This is the same bound on the exponent obtained by Altman el at. (2002) as a sufficient condition for the *uniqueness* of equilibrium.[14] (Cf. Proposition 7.) It holds not only in the linear case, where $d_e = 1$ for all resources e, but also in the quadratic, cubic and other cases. This is a significant extension, as a game where even only one cost function is non-linear is generally not a potential game.

Corollary 4. If the game h in Theorem 12 is a congestion game with splittable flow with cost functions that are twice continuously differentiable outside the origin and satisfy (30) there, and the social payoff function f is the aggregate payoff, then the latter's rate of change (27) is nonnegative at every nonnegative $r_0 < r < r_1$ and is moreover positive where $dy^r/\, dr \neq 0$.

With cost functions that do not satisfy (30), comparative statics are not necessarily positive.

Example 4 (Milchtaich 2012). A symmetric congestion game with splittable flow has two resources and two players, whose identical weights are some large number and whose two pure strategies are choosing either resource. One resource e has cost function

$$c_e(t) = -\frac{1}{\big((t+0.4)\ln(t+1.4)\big)^{1.5}}$$

[14] Because of a typo, the condition is presented in the paper in a slightly altered, stronger form (Nahum Shimkin, personal communication).

(negative, and thus contributing positively to the payoff) and the other's is identically zero. Computation shows that, for every value of r in some interval straddling the origin (but longer on the positive side), the corresponding modified game has precisely two equilibria, x^r and y^r, which are both symmetric, so the two players ship an identical quantity on resource e. Both flows on e, x_e^r and y_e^r, are higher than the socially optimal flow. However, the first flow is smaller than the second one, and correspondingly the players' (identical) payoff in x^r is higher than in y^r. As r increases along the interval, x_e^r continuously decreases, and the corresponding payoff increases, while for y^r the opposite trends hold. Thus, local comparative statics are positive for the "good" equilibria x^r and negative for the "bad" ones y^r. This shows that, even in such a very simple game, both kinds of local comparative statics may occur, global comparative statics need not be positive, and the payoff at the worst equilibrium may strictly decrease with increasing r.

Theorem 9 can be used to infer that the bad equilibria in Example 4 are unstable. Computation shows that the good equilibria are stable: they satisfy the condition in Theorem 11, with h replaced by the modified game h^r. (For an equilibrium y in h^r, the negative definiteness condition is equivalent to $y_e c_e''(y_e)/c_e'(y_e) > -(3+r)/(1+r)$.)

6.3 Cournot oligopoly

Example 4 is presented in Milchtaich (2012, Example 1) not as a congestion game with splittable flow but as a (symmetric) Cournot competition. The alternative presentation used here is made possible by the fact that quantity competition among N firms producing an identical good is a special kind of congestion game with splittable flow. Specifically, suppose that each firm i has a maximum output of w_i and a convex cost function $C_i\colon [0, w_i] \longrightarrow \mathbb{R}_+$ with $C_i(0) = 0$, and has to choose an output level $0 \leq q_i \leq w_i$. The price of the good is $P(Q)$, where $Q = \sum_{i=1}^N q_i$ is the aggregate output and $P\colon [0, \sum_{i=1}^N w_i] \longrightarrow \mathbb{R}_+$ is a nonincreasing price (or inverse demand) function. The profit of firm i is therefore $q_i P(Q) - C_i(q_i)$, or $q_i(P(Q) - AC_i(q_i))$, where AC_i is the firm's average cost function (with $AC_i(0) = C_i'(0)$). The corresponding congestion game has $N+2$ resources, numbered from 0 to $N+1$. Firm i's weight is its maximum output w_i and it has two pure strategies: choosing resource 0, which is interpreted as being idle, and choosing resources i and $N+1$, which represents full production. The cost function of resource 0 is identically zero, that of $N+1$ is given by $c_{e_{N+1}}(t) = -P(t)$, and for $1 \leq i \leq N$ the cost is $c_{e_i}(t) = AC_i(\min(t, w_i))$. (Example 4 corresponds to the special case of production at zero cost, for which resources 1 through N are redundant.) When each firm i ships quantity q_i on "full production" and $w_i - q_i$ on "idle", the payoff of firm i, which is the negative of its total cost, is $q_i(P(Q) - AC_i(q_i))$, the profit.

In this congestion game, the marginal social costs of resources 1 through N have a simple meaning. Namely, they coincide with the marginal costs of production for the corresponding firms (in the relevant ranges). The marginal social cost $MC_{e_{N+1}}$ is the negative of the marginal aggregate revenue. Thus, it is nondecreasing if and only if the aggregate revenue is a concave function of the aggregate output. This fact and the findings in the previous subsection hint at a possible link between such concavity and positive local comparative statics with respect to the firms' aggregate profit. In fact, in the special case of a Cournot competition between two identical firms, concavity is known to guarantee an even stronger property, namely, positive global comparative statics. (This result, too, has to do with global weak stability.) In view of the alternative presentation described above, the result can also be interpreted as concerning (a very special kind of) congestion games with splittable flow.

Proposition 8 (Milchtaich 2012, Section 6.2)**.** For a Cournot oligopoly game h as above with only two, identical firms, the aggregate profit as the social payoff function f, and altruism coefficients r and s with $-1 < r < s \leq 1$, let y^r and y^s be equilibria in the modified games h^r and h^s, respectively. If the aggregate revenue is a concave function of the aggregate output, then

$$f(y^r) \leq f(y^s).$$

7 Summary and Discussion

Users of common resources often exert negatives externalities on the other users, and are similarly affected by them. Their own experience with the ill effects of congestion may make it easier for people to sympathize with its other sufferers. It may seem only logical that if this sympathy led all users to internalize the others' welfare and to exert a certain, common degree of altruism in their choice of resources, everyone would be better off. However, as this paper shows, this is so only in some circumstances. Internalization of social welfare does not *have* to be socially beneficial. In fact, it may make everyone worse off.

The analysis presented here only concerns the expected material consequences (hence, the social desirability) of internalization of social welfare, that is, its effect on the actual level of the social payoff at equilibrium. The questions of the origin of such attitudes, their prevalence and evolution (e.g., Heifetz el al. 2007) lie outside the scope of the present work. In some contexts, altruism may have a biological basis. If the individuals involved are related, then the effect of the use of a common resource on the *inclusive fitness* of individual i is measured by the effect on i's own fitness plus r times the effect on each of the other individuals j, where r is the coefficient of relatedness between i and j (Milchtaich 2006a). In a social context, a similar parameter r may express individual i's level of empathy towards j. Edgeworth (1881) called this parameter the coefficient of *effective sympathy*. More generally, the altruism coefficient r may specify the weight a person attaches to some completely general social payoff, like the degree of global warming, that is not necessarily expressible in terms of the individuals' personal, material payoffs but is a function of everyone's actions. Some of the results in this paper apply to such general social payoff functions while others are specific to a particular function, the aggregate personal payoff (or the negative of the aggregate cost). In addition, some results also apply to negative r, which may be interpreted as spite.

A strong simplifying assumption made in this paper is that all the individuals involved are equally altruistic: they have the same r. This may be so, for example, if all users of the common resources share similar values or a similar upbringing. A second assumption, which is implicit in the functional form of the modified payoff, is that the latter incorporates the social payoff linearly. A non-linear effect would require adding higher-order terms to the function. However, these terms may be expected to be relatively unimportant if the effect is weak, that is, the personal payoff predominates. In this case, a linear relation with a small coefficient r may be a good approximation. Many of the phenomena described in this paper are evident already for arbitrarily small r.[15]

The simplest scenario where the use of common resources gives rise to negative externalities is that of an unweighted congestion game. As shown, it is generally not true in this setting that altruism necessarily has a positive effect on the aggregate personal payoff. In fact, even with linear cost functions, inefficient equilibria may exist *only* for altruism coefficients close to 1, which reflect complete or almost complete selflessness. In other words, as the altruism coefficient r increases, the price of anarchy in the game may (weakly) increase. The main results in Section 3 concern the characterization of the games for which this is never so. A sufficient condition for nonincreasing price of anarchy is the flow monotonicity property. This is because this property guarantees that equilibria must maximize the game's potential function. The flow monotonicity property only concerns the "congestion game form"; the cost functions are irrelevant. For symmetric unweighted network congestion games, this fact translates to a condition on the topology of the undirected network on which the game is defined. In particular, if that network is extension-parallel, then the flow monotonicity property automatically holds for the game, and therefore altruism can only increase social welfare or leave it unchanged. The same it not true for general series-parallel networks.

Allowing users to split their demand among several strategies has a significant effect on comparative statics. In particular, with linear cost functions, the aggregate personal payoff at equilibrium is uniquely determined by the altruism coefficient r and is nondecreasing in $[-1,1]$. That is, an increasingly positive or negative r can only increase or decrease, respectively, the social welfare or leave it unchanged. The highest aggregate payoff is attained with complete selflessness, $r = 1$.

Internalization of social payoff may occur even if a single person's ability to affect the social payoff is practically nonexistent. For example, people may care about their carbon footprint even while acknowledging its insignificance on a global scale. They may justify their attitude by saying that the effect would not be insignificant if sufficiently many people acted on these concerns. Put differently, the emphasis is on the *marginal* effect of an act on the social payoff, that is, the effect per unit mass of actors. (In the special case where the total effect is proportional to the size of the set of actors, the marginal effect of an action expresses the outcome if everybody switched to that action.) In this context, the altruism coefficient r is the weight attached to the marginal effect, and comparative statics concern the relation between this weight and the actual level of the social payoff at equilibrium. For congestion games, the relevant model is that of a population game arising from a nonatomic congestion game, where the set of players is modeled as a continuum. The social payoff function is the mean payoff, which is analogous to the aggregate personal payoff in games with a finite number of players. As shown, comparative statics are qualitatively quite similar to those in

[15] The modified payoff is linear also in r. However, this linearity does not represent a very strong assumption because multiplying the expression on the right-hand side of (1) by any positive-valued function of r would not affect the preferences expressed by it. In other words, only the ratio $(1 - r){:}\, r$, which gives the marginal rate of substitution of personal payoff for social payoff, matters.

the case of a finite number of users with splittable flow. For r between 0 and 1, the mean payoff at equilibrium is uniquely determined by the altruism coefficient as a nondecreasing function, which reaches its peak at $r = 1$.

The results presented above rely to a large extent on the fact that the games considered are potential games and the equilibria or equilibrium strategies involved are necessarily maximizers of the potential. However, Section 6 points to a deeper link underlying the connection between comparative statics and the potential function. Specifically, it shows that a strict maximum point of the potential is necessarily (statically) stable, and that stability or definite instability, respectively, necessarily brings about positive or negative comparative statics. Moreover, the last connection holds not only for the aggregate or mean payoff but for general social payoff functions. This connection may be used for determining the effect of altruism in congestion games to which the previous results do not apply, in particular, congestion games with splittable flow that are not potential games because the cost functions are not linear (but, say, quadratic or cubic). A nice fact is that such games can model Cournot oligopoly. The meaning of negative comparative statics in this context is that two duopolists that internalize their competitor's profit to some degree may have lower profits at equilibrium than they would have if they only cared about their own profit. This finding is surprising in view of the fact that *complete* internalization would turn the duopoly into a monopoly, and drive the profits up.

References

Altman, E., Basar, T., Jimenez, T. and Shimkin, N. (2002). Competitive routing in networks with polynomial costs. *IEEE Transactions on Automatic Control* 47, 92–96.

Anagnostopoulos, A., Becchetti, L., de Keijzer, B. and Schäfer, G. (2015). Inefficiency of games with social context. *Theory of Computing Systems* 57, 782–804.

Bochnak, J., Coste, M. and Roy, M.-F. (1998). *Real Algebraic Geometry*. Springer-Verlag, Berlin.

Caragiannis, I., Kaklamanis, C., Kanellopoulos, P., Kyropoulou, M. and Papaioannou, E. (2010). The impact of altruism on the efficiency of atomic congestion games. *Lecture Notes in Computer Science* 6084, 172–188.

Chen, P.-A., de Keijzer, B., Kempe, D. and Schäfer, G. (2014). Altruism and its impact on the price of anarchy. *ACM Transactions on Economics and Computation* 2 Issue 4, Article No. 17.

Chen, P.-A. and Kempe, D. (2008). Altruism, selfishness, and spite in traffic routing. *Proceedings of the 9th ACM conference on Electronic commerce*, pp. 140–149.

Cominetti, R., Correa, J. R. and Stier-Moses, N. E. (2009). The impact of oligopolistic competition in networks. *Operations Research* 57, 1421–1437.

Edgeworth, F. Y. (1881). *Mathematical Psychics*. Kegan Paul, London.

Epstein, A., Feldman, M. and Mansour, Y. (2009). Efficient graph topologies in network routing games. *Games and Economic Behavior* 66, 115–125.

Fotakis, D. (2010). Congestion games with linearly independent paths: Convergence time and price of anarchy. *Theory of Computing Systems* 47, 113–136.

Fotakis, D., Kontogiannis, S. and Spirakis, P. (2005). Selfish unsplittable flows. *Theoretical Computer Science* 348, 226–239.

Gairing, M. (2009). Malicious Bayesian congestion games. *Lecture Notes in Computer Science* 5426, 119–132.

Heifetz, A., Shannon C. and Spiegel, Y. (2007). The dynamic evolution of preferences. *Economic Theory* 32, 251–286.

Hoefer, M. and Skopalik, A. (2013). Altruism in atomic congestion games. *ACM Transactions on Economics and Computation* 1 Issue 4, Article No. 21.

Holzman, R. and Law-Yone, N. (1997). Strong equilibrium in congestion games. *Games and Economic Behavior* 21, 85–101.

Holzman, R. and Law-yone (Lev-tov), N. (2003). Network structure and strong equilibrium in route selection games. *Mathematical Social Sciences* 46, 193–205.

Karakostas, G. and Viglas, A. (2007). Equilibria for networks with malicious users. *Mathematical Programming* 110, 591–613.

Kuniavsky, S. and Smorodinsky, R. (2013). Greediness and equilibrium in congestion games. *Economics Letters* 121 499–503.

Meir, R. and Parkes, D. (2015). Playing the wrong game: Smoothness bounds for congestion games with behavioral biases. *ACM SIGMETRICS Performance Evaluation Review* 43, 67–70.

Milchtaich, I. (2004). Social optimality and cooperation in nonatomic congestion games. *Journal of Economic Theory* 114, 56–87.

Milchtaich, I. (2006a). Comparative statics of games between relatives. *Theoretical Population Biology* 69, 203–210.

Milchtaich, I. (2006b). Network topology and the efficiency of equilibrium. *Games and Economic Behavior* 57, 321–346.

Milchtaich, I. (2012). Comparative statics of altruism and spite. *Games and Economic Behavior* 75, 809–831.

Milchtaich, I. (2013). Representation of finite games as network congestion games. *International Journal of Game Theory* 42, 1085–1096.

Milchtaich, I. (2015). Network topology and equilibrium existence in weighted network congestion games. *International Journal of Game Theory* 44, 515–541.

Milchtaich, I. (2020). Static stability in games. Bar-Ilan University, Research Institute for Econometrics and Economic Theory, Working Paper 1-20, April 2020.

Monderer, D. and Shapley, L. S. (1996). Potential games. *Games and Economic Behavior* 14, 124–143.

Rahn, M. and Schäfer, G. (2013). Bounding the inefficiency of altruism through social contribution games. *Lecture Notes in Computer Science* 8289, 391–404.

Rosen, J. B. (1965). Existence and uniqueness of equilibrium points for concave N-person games. *Econometrica* 33, 520–534.

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

Singh, Ch. (2008). Marginal cost pricing for atomic network congestion games. Technical report.

Appendix: Networks

A directed or undirected two-terminal network is, respectively, a directed or undirected multigraph (where a pair of vertices may be joined by more than one edge) that is endowed with two terminal vertices, o and d, and satisfies the following "non-redundancy" condition: each edge e and each vertex v is included in at least one route in the network. In the undirected case, a route is defined as a (simple) path of the form $oe_1v_1 \cdots v_{n-1}e_nd$ ($n \geq 1$), that is, one that begins at the origin vertex o and ends at the destination vertex d. (An alternative, simpler notation is $e_1e_2 \cdots e_n$.) In the directed case, a route also has to traverse each of its edges e in the direction assigned to e, that is, the tail and head vertices of e must be its immediate predecessor and successor, respectively, in the route. This difference means that the set of routes in a directed two-terminal network is a subset of the set of routes in the network's *undirected version*, which is obtained by ignoring the edge directions.

An arbitrary assignment of directions to the edges of an undirected two-terminal network G does not necessarily give a directed network, as the non-redundancy condition may fail to hold. If the non-redundancy condition does hold, the resulting directed two-terminal network is said to be a *directed version* of G. Such a version always exists.

Proposition A1. Every undirected two-terminal network has at least one directed version.

Proof. The proof is by induction on the number of edges in the network. If there is only one edge, the assertion is trivial. To establish the inductive step, consider a network G with more than one edge, and some edge e incident with the origin o. If no other edge is incident with o, then by the induction hypothesis there exists at least one directed version of the network obtained from G by contracting e, that is, eliminating the edge and its non-terminal vertex v and replacing v with o as the terminal vertex of all the edges originally incident with v. Any such directed version gives a directed version of G, in which e is directed from o to v. Suppose, then, that e is not the only edge incident with o, and consider the subnetwork G_e of G whose edges and vertices are all those that belong to some route in G where the first edge is e.

Claim. If a route r in G includes a non-terminal vertex u that is in G_e, then every edge (hence, also every vertex) that follows u in r is also in G_e.

It clearly suffices to consider the edge e' that immediately follows u in r. Suppose that e' is not in G_e. Let v the first vertex in r that follows e' and is in G_e. (Possibly, $v = d$.) All the edges and vertices in r between u and v are not in G_e. Adding them to it creates a new subnetwork of G, since the addition is equivalent to adding to G_e a single edge with end vertices u and v and then subdividing the edge, if needed, one or more times (Milchtaich 2015, Section 2.2). By definition, there exists in the new subnetwork a route r' that includes e'. Necessarily, r' also includes e. The conclusion contradicts the assumption that e' is not in G_e, and thus proves the claim.

Consider now the collections of all edges and vertices that are obtained from those in G by (i) eliminating all edges that are in G_e and (ii) identifying with d all non-terminal vertices that are in G_e. It follows from the Claim that these collections constitute a network G' (with the terminal vertices o and d). To see this, consider any edge or non-terminal vertex in them and a route r in G that includes it. It follows from the Claim that the inclusion still holds if r is replaced with its section r_{ou}, which is the path obtained from r by deleting all edges and vertices that follow u, where u is the first vertex in r, other than o, that is in G_e. As u is one of the vertices identified with d, r_{ou} is a route in G'.

By the induction hypothesis, both G_e and G' have directed versions, which together specify directions for all edges in G. It remains to show that these directions define a directed version of G, that is, the non-redundancy condition holds. Any edge or vertex that is or is not in G_e is included in some route r in the directed version of G_e or some route r' in the directed version of G', respectively. By construction, r is automatically a route also in G and it honors the directions specified for its edges. Only the second assertion is necessarily true for r', which is a path that starts at o, ends at some vertex u in G_e, and does not include any other vertex or edge that is in G_e. However, it is possible to extend r' to a route in G that honors the edges' directions by replacing the vertex u with the section r''_{ud} of some route r'' in the directed version of G_e that includes u, where r''_{ud} is obtained from r'' by deleting all the edges and vertices that precede u. ■

Corollary A1. The collection of the undirected versions of all directed two-terminal networks coincides with the collection of all undirected two-terminal networks.

Clearly, an assignment of directions to the edges of an undirected two-terminal network G defines a directed version of G only if it satisfies the condition mentioned in Section 3.3, which is that the direction of each edge e coincides with the direction in which some route in (the undirected network) G traverses e. However, this necessary condition is not sufficient. For example, it is not difficult to see that an assignment satisfying the condition may render a non-terminal degree-two vertex the head vertex of both edges incident with it, which means that the non-redundancy condition for directed networks is violated. However, it follows from Proposition A1 that if there is only *one* assignment of directions that satisfies the above condition, then that assignment necessarily does give a directed version of G. Such uniqueness is rather special. Indeed, the following result can easily be deduced from Proposition 1 in Milchtaich (2006b).

Corollary A2. For an undirected two-terminal network G, the following conditions are equivalent:

(i) For every edge e, all routes in G that include e traverse it in the same direction.

(ii) For every assignment of allowable directions to the edges in G such that the direction of each edge e is that in which some route in G traverses e, all routes in G are allowable.

(iii) G is series-parallel.