# Network Topology and the Efficiency of Equilibrium*

IGAL MILCHTAICH

Department of Economics, Bar-Ilan University,
Ramat-Gan 5290002, Israel

Email: igal.milchtaich@biu.ac.il
Web: https://sites.google.com/view/milchtaich/

*Games and Economic Behavior* 57 (2006), 321–346

**Abstract.** Different kinds of networks, such as transportation, communication, computer, and supply networks, are susceptible to similar kinds of inefficiencies. These arise when congestion externalities make the cost for each user depend on the other users' choice of routes. If each user chooses the least expensive (e.g., the fastest) route from the users' common point of origin to the common destination, the result may be Pareto inefficient in that an alternative choice of routes would reduce the costs for *all* users. Braess's paradox represents an extreme kind of inefficiency, in which the equilibrium costs may be reduced by *raising* the cost curves. As this paper shows, this paradox occurs in an (undirected) two-terminal network if and only if it is *not* series-parallel. More generally, Pareto inefficient equilibria occur in a network if and only if one of three simple networks is embedded in it.
*JEL Classification*: C72, R41.

**Key Words:** Congestion, externalities, topological efficiency, nonatomic games, network topology, Braess's paradox, transportation networks, Wardrop equilibrium

* The author thanks Grant Keady, Bruce Calvert, Claude Penchina, and Hideo Konishi for helpful discussions. This research was supported by the Israel Science Foundation (grant no. 749/02).

# 1 Introduction

In transportation and other kinds of physical networks with large numbers of users, congestion externalities are a potential source of inefficiency. A remarkable example of this, known as Braess's paradox (Braess, 1968; Murchland, 1970; Arnott and Small, 1994), is shown in Figure 1. Cars arrive at a constant rate at vertex $o$ of the depicted network and leave it at vertex $d$. The network consists of three fast roads ($e_1$, $e_4$ and $e_5$) and two slow ones ($e_2$ and $e_3$). The travel time on each road is an increasing function of the flow on it, or the average number of vehicles passing a fixed point in the road per unit of time. (This is a reasonable assumption if the density of vehicles on the road is relatively low, so that the flow is well below the road's capacity. See Sheffi, 1985, Chapter 13 and Figure 1.8.) However, regardless of the flow, the travel time on the route consisting of the three fast roads is shorter than on any of the alternative routes. Therefore, at (the Wardrop) equilibrium, when the entire flow passes on the fastest routes, all vehicles use this one. The travel time on the network (as computed in the caption to Figure 1) is then 21 minutes. Suppose, however, that the transverse road, $e_5$, is closed, or its physical condition deteriorates to the point at which the travel time on it becomes similar to that on each of the two slow roads. The road's new cost curve is higher than the old one: for any flow on $e_5$, the travel time is longer than before. As a result of the change in costs, the old equilibrium is replaced by a new one, in which the transverse road is not used at all. Paradoxically, the new travel time is shorter than before: 20 minutes. The longer previous travel time is due to the motorists' concern for their own good only, which results in overuse of the fast roads and consequently an inefficient equilibrium. As pointed out by Newell (1980) and Sheffi (1985), traffic engineers have long known that more restricted travel choices and reduced capacity may improve the flow in the network as a whole. For instance, this is the underlying principle behind many traffic control schemes, such as ramp metering on freeway entrances (Sheffi, 1985, p. 77).

Braess's paradox is not limited to transportation networks. The potential occurrence of this or similar paradoxes has been demonstrated for such diverse networks as computer and telecommunication networks, electrical circuits, and mechanical systems. Remarkably, much of this literature (e.g., Frank, 1981; Cohen and Horowitz, 1991; Cohen and Jeffries, 1997) is concerned with essentially the same network as in Braess's (1968) original example, the Wheatstone network, shown in Figure 1. As it turns out, there is a good reason for this. As this paper shows, this is, in a sense, the *only* two-terminal network in which Braess's paradox can occur. More precisely, a necessary and sufficient condition for the existence of *some* cost function exhibiting the paradox is that the network has an embedded Wheatstone network. In networks without this property, called series-parallel networks, Braess's paradox never occurs. Several alternative characterizations of series-parallel networks are given below.

Braess's paradox is not the only kind of inefficiency caused by congestion externalities. Consider, for example, the series-parallel network in Figure 2(a), which represents the alternatives faced by weekend visitors to a certain seaside town where the only attractions are the two nearby beaches. The two edges joining $o$ and $v$ represent the alternatives of going to the North Beach ($e_1$) or the South Beach ($e_2$) on Saturday. The two edges joining $v$ and $d$ represent the same two alternatives on Sunday. The South Beach is more remote, and the cost of getting there is two units more than for the North Beach. On the other hand, it is

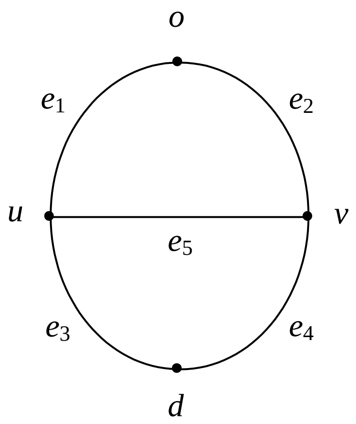


**Figure 1 Braess's paradox. A continuum of users travels from $o$ to $d$ on the two-terminal *Wheatstone network* shown. The travel time on each edge $e$ is an increasing function of the fraction $x$ of the total flow from $o$ to $d$ that passes on $e$. The travel times, in minutes, are given by $1 + 6x$ for $e_1$ and $e_4$, and $15 + 2x$ for $e_2$ and $e_3$. If the travel time on $e_5$ is also given by $1 + 6x$, then, at equilibrium, the entire flow passes on $e_1$, $e_5$ and $e_4$, which constitute the fastest route from $o$ to $d$. The total travel time is then $3 \times (1 + 6 \times 1) = 21$ minutes. If, however, the travel time on $e_5$ is longer, and given by $15 + 2x$, then, at equilibrium, there is no flow on that edge: half the users go though $e_1$ and $e_3$ and half though $e_2$ and $e_4$. The equilibrium travel time is then *shorter*: $(1 + 6 \times ½) + (15 + 2 \times ½) = 20$ minutes.**

a longer beach, and therefore does not get crowded as fast. However, the additional pleasure of spending the day on an uncrowded beach never exceeds the difference in travel costs. Therefore, at equilibrium, all the visitors go to the North Beach, both on Saturday and on Sunday. Crowding then costs each person four units of pleasure. However, if people had taken turns, half of them going to the South Beach on Saturday and the other half on Sunday, then the cost for all individuals would be lower, and equal to 3½. Thus, this arrangement, which is not an equilibrium, represents a strict Pareto improvement over the equilibrium. The difference between this example and the one in Figure 1 is that, in the case of Braess's paradox, Pareto improvement results from raising the costs of certain facilities (e.g., increasing the travel time on the transverse road in Figure 1), thereby creating a new equilibrium that is better for everyone. By contrast, in the present example it is not possible to make everyone better off by raising the costs of facilities (e.g., charging congestion-dependent entry fees to beaches). Since the networks in Figure 2 are series-parallel, Braess's paradox cannot occur. Hence, any Pareto improvement necessarily involves non-equilibrium behavior, i.e., use of certain routes for which less costly alternatives exist.

One of the main results of this paper is that the three networks in Figure 1 and Figure 2 essentially represent the only kinds of two-terminal network topologies in which congestion externalities may lead to Pareto inefficient equilibria. For example, such inefficiencies never arise in networks such as in Figure 3. The crucial difference between this network and those previously mentioned is that the routes in it are *linearly independent*, in the sense that each one includes at least one edge that is not part of any other route. This is equivalent to the following condition: none of the routes in the network has a pair of edges, each of which also belongs to some other route that does not include the other edge. In a similar, but not identical, context, Holzman and Law-Yone (1997, 2003) show this to be a necessary and sufficient condition for weak Pareto efficiency of the equilibria for all systems of nonnegative,

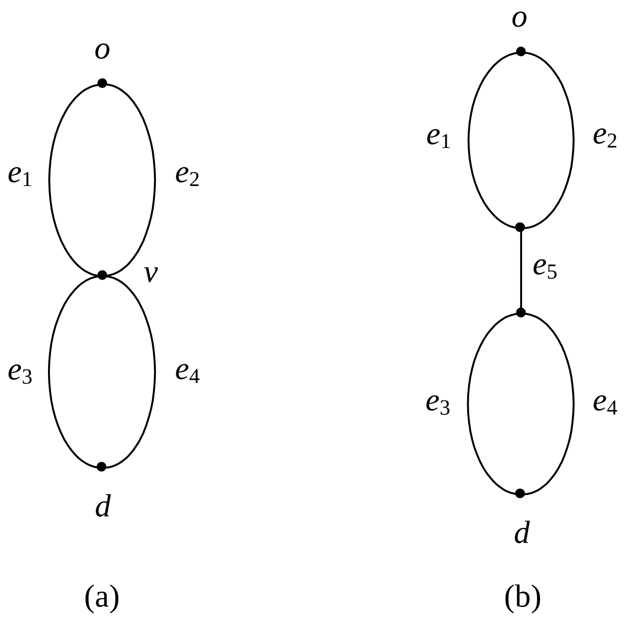


**Figure 2 Another kind of inefficiency caused by congestion externalities. The cost of each edge $e$ in network (a) is an increasing function of the fraction $x$ of the total flow from $o$ to $d$ that passes on $e$. For $e_1$ and $e_3$, the cost is given by $2x$. For $e_2$ and $e_4$, it is $2 + x$. At equilibrium, only $e_1$ and $e_3$ are used, and the equilibrium cost is $(2 \times 1) + (2 \times 1) = 4$. However, this outcome is Pareto inefficient. Splitting the flow, so that half of it goes through $e_1$ and $e_4$ and half through $e_2$ and $e_3$, would reduce the cost to $(2 \times ½) + (2 + ½) = 3½$. (Further reduction is not possible, since it can be shown that the *mean* cost in this example cannot be less than $3½$.) A similar phenomenon occurs in network (b). Indeed, since all routes from $o$ to $d$ pass through the middle edge $e_5$, the cost of this edge is irrelevant.**

nondecreasing edge costs. Their work differs from this paper mainly in referring to networks with a finite number of users, each of whom has a non-negligible effect on the others. Here, by contrast, there is a continuum of users, which may be viewed as a mathematical idealization of a very large population of individuals, each with an almost negligible effect on the others. As this paper shows, the cases of finite and infinite populations differ in a number of ways. Most importantly, in the latter but not the former case, the connection between linear independence of the routes and efficiency of the equilibria also holds for *heterogeneous* populations, in which users have different cost functions. With a finite number of non-identical users, cost functions with Pareto inefficient equilibria exist for *any* nontrivial network, i.e., one with two or more routes. With a continuum of users, an assignment of cost functions with a Pareto inefficient equilibrium exists if and only if the routes in the network are not linearly independent.

The emphasis in this paper is on *topological efficiency*: network topologies for which cost functions giving rise to inefficiencies do not exist. Most related papers put the emphasis on the cost functions themselves. For example, formulas yielding, under certain conditions, the change in cost induced by the creation of additional routes were obtained by Steinberg and Zangwill (1983) and Dafermos and Nagurney (1984). In principle, these formulas can be used to determine whether a form of Braess's paradox occurs in the network. They are, however, rather complicated. An interesting (and, as shown by Calvert and Keady, 1993, rather unique) case, in which the topology of the network is irrelevant, is that of edge costs that are

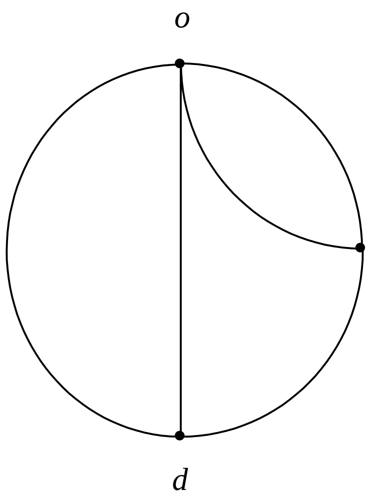


**Figure 3 A network with linearly independent routes. In such a network, equilibria are always Pareto efficient.**

given by homogeneous functions of the same degree, i.e., each of them has the form $\alpha x^{\beta}$, with $\alpha, \beta > 0$, and $\beta$ is the same for all edges. In this case, not only are the equilibria Pareto efficient but they are even *socially optimal* in that the *mean* (equivalently, aggregate) *cost* incurred by the users is minimized. Intuitively, this is because, in this case, users switching from one route to another make proportional changes to their own and the social costs (see Milchtaich, 2004).

The topology of a network does not indicate whether equilibria are socially optimal. Even for very simple networks, it is also necessary to know the functional form of the cost functions (Milchtaich, 2004). The same is true for the so-called *price of anarchy*, which is the ratio between the equilibrium cost and the mean cost incurred by the users at a social optimum. For general cost functions, this ratio is unbounded, but for various natural classes of such functions, bounds do exist (Roughgarden, 2003; Roughgarden and Tardos, 2004). For example, the maximum price of anarchy with linear edge costs is $4/3$ (thus, greater than the $4\colon 3½$ ratio between the equilibrium cost and the mean cost at the social optimum in the example in Figure 2). Under weak assumptions on the class of allowable cost functions, a network with only two parallel edges suffices to achieve the worst possible ratio. Thus, as for social optimality, the network topology does not play a role in determining the price of anarchy (Roughgarden, 2003). This contrasts sharply with the situation for Pareto efficiency, in which, as this paper shows, the network topology matters. Several other properties of the equilibria in route selection games occupy an intermediate position between these two extremes in that they depend on certain "global" properties of the network. For example, the maximum reduction in equilibrium cost achievable by removing one (as in the original Braess's paradox) or more edges in a two-terminal network is positively related with the number of vertices (Roughgarden, 2004). It is also positively related with the number of edges removed (Lin, Roughgarden and Tardos, 2004).

The rest of the paper is organized as follows. The next section presents in some detail the needed graph-theoretic definitions and results. With few exceptions (a network with linearly independent routes is a notable one), standard terminology is used (which, however, does not always have the exact same meaning in all sources). Flows, cost functions, and related terms are defined in Section 3. The definition of equilibrium with identical users, and results about its efficiency, are given in Section 4. The first result links the non-occurrence of Braess's paradox with series-parallel networks, and the second one links (the stronger property of) Pareto efficiency of the equilibria with (the smaller class of) networks with linearly independent routes. Section 5 deals with non-identical users, and shows that, in this case, both properties of the equilibria are linked with networks with linearly independent routes. In Section 6, some of the definitions and assumptions underlying these results are discussed. In particular, the advantages of dealing with undirected rather than directed networks are explained, and the similarities and differences between the topological conditions for efficiency and *uniqueness* of the equilibrium (Milchtaich, 2005) are described. The proofs of all the propositions and theorems in this paper are given in Section 7.

## 2 Graph-Theoretic Preliminaries

An *undirected multigraph* consists of a finite *vertex set* $\mathcal{V}$ and a finite *edge set* $\mathcal{E}$. Each edge $e$ joins two distinct vertices, $u$ and $v$, which are referred to as the *end vertices* of $e$. Thus, loops are not allowed, but more than one edge can join two vertices. An edge $e$ and a vertex $v$ are said to be *incident* with each other if $v$ is an end vertex of $e$. A *path* of length $n$ ($n \geq 0$) is an alternating sequence $p$ of vertices and edges $v_0 e_1 v_1 \cdots v_{n-1} e_n v_n$, beginning and ending with vertices, in which each edge is incident with the two vertices immediately preceding and following it and all the vertices (and necessarily all the edges) are distinct. Because of the latter assumption, each vertex and each edge in $p$ either *precedes* or *follows* each of the other vertices and edges. The first and last vertices, $v_0$ and $v_n$, are called the *initial* and *terminal* vertices in $p$, respectively. The path $v_n e_n v_{n-1} \cdots v_1 e_1 v_0$, which includes the same vertices and edges as $p$ but passes through them in reverse order, is denoted by $-p$. If $q$ is a path of the form $v_n e_{n+1} v_{n+1} \cdots v_{m-1} e_m v_m$, the initial vertex of which is the same as the terminal vertex of $p$ but all the other vertices and edges are not in $p$, then $v_0 e_1 v_1 \cdots v_{n-1} e_n$ $v_n e_{n+1} v_{n+1} \cdots v_{m-1} e_m v_m$ is also a path, denoted by $p + q$. A *section* of $p$ is any path $s$ of the form $v_{n_1} e_{n_1+1} v_{n_1+1} \cdots v_{n_2-1} e_{n_2} v_{n_2}$, with $0 \leq n_1 \leq n_2 \leq n$. Each section is uniquely identified by its initial vertex $u$ and terminal vertex $v$, and may therefore be denoted by $p_{uv}$. If the length of $s$ is zero, i.e., it does not include any edges, then $u$ and $v$ coincide. If the length is one, i.e., the section has a single edge $e$, then $u$ and $v$ are the two end vertices of $e$. In this case, $s$ is called an *arc*, and may be viewed as a specification of the *direction* in which $p$ passes through $e$.

A *two-terminal network* (*network*, for short) is an undirected multigraph together with a distinguished ordered pair of distinct vertices, an *origin* $o$ and a *destination* $d$, such that each vertex and each edge belong to at least one path with the initial vertex $o$ and terminal vertex $d$. Any path $r$ with these initial and terminal vertices will be called a *route*. The set of all routes in a network, denoted by $\mathcal{R}$, is its *route set*.

Two networks $G'$ and $G''$ will be said to be *isomorphic* if there is a one-to-one correspondence between their vertex sets and between the edge sets such that (i) the incidence relation is preserved and (ii) the origin and destination in $G'$ are paired with the origin and destination in $G''$, respectively. A network $G'$ is a *sub-network* of a network $G''$ if the former is isomorphic to a network derived from the latter by deleting a subset of its edges

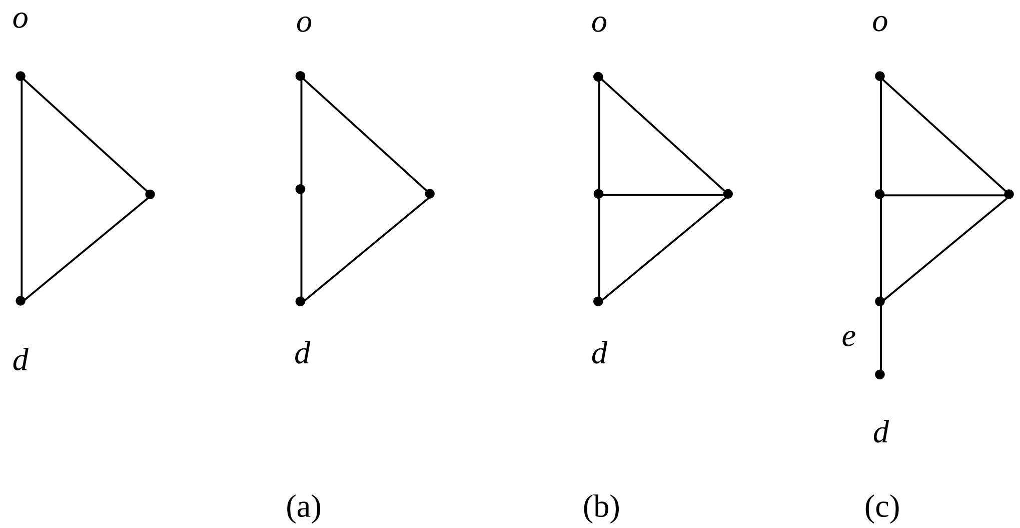


**Figure 4 Embedding. The left network is embedded in each of the other three, which are obtained from it by: (a) subdividing an existing edge, (b) adding a new edge, and, finally, (c) extending the destination.**

and vertices, which does not include $o$ or $d$. A network $G'$ is *embedded* in a network $G''$ if $G''$ is isomorphic to $G'$ or to a network derived from $G'$ by applying the following operations any number of times in any order (see Figure 4):

(a) The *subdivision of an edge*: its replacement by two edges with a single common end vertex.

(b) The *addition of a new edge* joining two existing vertices.

(c) The *extension of a terminal vertex*: addition of a new edge $e$ joining $o$ or $d$ with another, new vertex, which becomes the new origin or destination, respectively.

It can be shown that, for any network $G'$, addition and subdivision of edges always give a new network, with the same origin–destination pair. By using both operations, complete new paths, which only have their initial and terminal vertices in $G'$, can be added to it. This is done by first joining the two vertices by a new edge $e$, and then subdividing $e$ one or more times. Because each vertex and each edge in a network are in some route, this shows that $G'$ is embedded in every network $G''$ of which it is a sub-network.

Two networks $G'$ and $G''$ with the same origin–destination pair, but without any other common vertices or edges, may be connected *in parallel*. The vertex and edge sets in the resulting network $G$ are the unions of those in $G'$ and $G''$, and the origin and destination are the same as in these networks. Two networks $G'$ and $G''$ with a single common vertex (and, hence, without common edges), which is the destination in $G'$ and the origin in $G''$, may be connected *in series*. The vertex and edge sets in the resulting network $G$ are the unions of those in $G'$ and $G''$, the origin coincides with the origin in $G'$, and the destination with that in $G''$. It is not difficult to see that, when two networks $G'$ and $G''$ are connected in parallel or in series, each of them is embedded in the resulting network $G$.

A network is said to be *series-parallel* if two routes never pass through any edge in opposite directions. The two networks in Figure 2 are series-parallel. The Wheatstone network in Figure 1 is not series-parallel, since there are two routes passing through $e_5$ in opposite directions. In fact, as the following proposition shows, the Wheatstone network is embedded in *any* network that is not series-parallel. (Theorem 1 of Duffin, 1965, makes the same assertion. However, since that paper uses somewhat different definitions, an explicit proof is needed here, which is given in Section 7.)

**Proposition 1.** For a two-terminal network $G$, the following conditions are equivalent:

(i) $G$ is series-parallel.

(ii) For every pair of distinct vertices $u$ and $v$, if $u$ precedes $v$ in some route $r$ containing both vertices, then $u$ precedes $v$ in all such routes.

(iii) The network in Figure 1 is *not* embedded in $G$.

As noted by Riordan and Shannon (1942), series-parallel networks can also be defined recursively. A network is series-parallel if and only if it can be constructed from single edges by carrying out the operations of connecting networks in series or in parallel any number of times. (Hence the term "series-parallel.") This can be stated as follows.

**Proposition 2.** A two-terminal network $G$ is series-parallel if and only if

(i) it has a single edge only,

(ii) it is the result of connecting two series-parallel networks in parallel, or

(iii) it is the result of connecting two series-parallel networks in series.

One corollary of Proposition 2 is that every series-parallel network is *planar* and, moreover, remains so when a new edge, joining $o$ and $d$, is added to it. Equivalently (see Harary, 1969), every series-parallel network can be embedded in the plane in such a way that $o$ and $d$ lie on the exterior face, or boundary. Using Proposition 2, this corollary can easily be proved by induction on the number of edges. Another corollary of Proposition 2 is the following result, which may help in verifying that a given network is series-parallel.

**Proposition 3.** A two-terminal network G is series-parallel if and only if the vertices can be indexed in such a way that, along each route, they have increasing indices.

A network will be said to have *linearly independent routes* if each route has at least one edge that does not belong to any other route. (The reason for this name is given by Proposition 6 below.) The simplest such network is *parallel network*, which consists of one or more edges connected in parallel. Another example is shown in Figure 3. Every network with linearly independent routes is series-parallel but the converse is false. The two networks in Figure 2 are series-parallel but they do not have linearly independent routes. Indeed, the next proposition implies that at least one of these two networks is embedded in *any* series-parallel network that does not have linearly independent routes. The proposition also gives two other characterizations of networks with linearly independent routes. The first characteristic property is that pairs of routes never merge only in the middle: any common section extends to either the origin or the destination. The second property is that the route set does not contain a *bad configuration* (Holzman and Law-Yone, 1997, 2003), which is defined as a triplet of routes, the first of which includes some edge $e_1$ that does not belong to the second route, the second route includes some edge $e_2$ that does not belong to the first one, and the third route includes both $e_1$ and $e_2$.

**Proposition 4.** For a two-terminal network $G$, the following conditions are equivalent:

(i) $G$ has linearly independent routes.

(ii) For every pair of routes $r$ and $r'$ and every vertex $v$ common to both routes, either the section $r_{ov}$ (which consists of $v$ and all the vertices and edges preceding it in $r$) is equal to $r'_{ov}$, or $r_{vd}$ is equal to $r'_{vd}$.

(iii) A triplet of routes constituting a bad configuration does not exist.

(iv) None of the networks in Figure 1 and Figure 2 is embedded in $G$.

The following recursive characterization of networks with linearly independent routes, which differs from that for series-parallel ones (Proposition 2) only in having a more restrictive part (iii), reveals another facet of the difference between these two kinds of networks. This characterization is essentially a corollary of Theorem 1 of Holzman and Law-Yone (2003) (which, however, relates to *directed* networks).

**Proposition 5.** A two-terminal network $G$ has linearly independent routes if and only if

(i) it has a single edge only,

(ii) it is the result of connecting two networks with linearly independent routes in parallel, or

(iii) it is the result of connecting in series a network with linearly independent routes and one with a single edge (or, equivalently, extending the origin or the destination in the first network).

Since, in every network $G$, each route $r$ has a unique set of edges, it is represented by a unique binary vector, in which 1 is assigned to each edge $e$ that belongs to $r$ and 0 to any other edge. This vector can be viewed as an element of the vector space $\mathbb{F}_2^{|\mathcal{E}|}$, so-called the *edge space* of $G$ (Diestel, 2000), where $|\mathcal{E}|$ is the number of edges in $G$ and $\mathbb{F}_2$ is the field of the integers modulo 2. Thus, each collection of routes in $G$ corresponds to a set of vectors in the edge space. These vectors are linearly independent if and only if it is not possible to write any one of them as the (component-wise) sum modulo 2 of some of the others. As the following proposition shows, networks with linearly independent routes are characterized by the property that the collection of *all* routes corresponds to a linearly independent set in the edge space.[1]

[1] Note that linear independence is defined with respect to $\mathbb{F}_2$, not (the real field) $\mathbb{R}$. For example, the Wheatstone network does not have linearly independent routes. Although the binary vectors representing its four routes are linearly independent in $\mathbb{R}^{|\mathcal{E}|}$, they are not so in $\mathbb{F}_2^{|\mathcal{E}|}$, since each of them is equal to the sum modulo 2 of the other three.

**Proposition 6.** A two-terminal network $G$ is a network with linearly independent routes if and only if its route set $\mathcal{R}$ corresponds to a linearly independent set of vectors in the edge space.

## 3 The Model

A *flow vector* for a network $G$ is a nonnegative vector $\mathbf{f} = (f_r)_{r\in\mathcal{R}}$ specifying the flow $f_r$ on each route $r$. The flow $f_p$ on each *path* $p$ is defined as the total flow on all the routes containing $p$:

$$f_p = \sum_{\substack{r\in\mathcal{R} \\ p \text{ is a section of } r}} f_r. \tag{1}$$

If $p$ is a path of length zero, consisting of a single vertex $v$, then $f_p$ is a *junction flow*: it gives the total flow on all the routes passing through $v$. In particular,

$$f_o = \sum_{r\in\mathcal{R}} f_r$$

(which is clearly equal to $f_d$) represents the *total origin–destination flow*. If $p$ is a path of length one, consisting of a single edge $e$ and its two end vertices, then $f_p$ is an *arc flow*: it gives the total flow on $e$ in the direction specified by $p$. Each edge is associated with a pair of arc flows, one for each direction. In a series-parallel network, in which all routes pass through an edge in the same direction, only one of these flows can be positive. In a network with linearly independent routes, in which each route includes at least one edge that is not in any other route, the arc flows uniquely determine the flow vector.

A *cost function* for a network $G$ is a vector-valued function $\mathbf{c}$ specifying the cost $c_p(\mathbf{f})$ of each path $p$ as a function of the (entire) flow vector $\mathbf{f}$.[2] The following monotonicity condition is assumed to hold: For every path $p$ and every pair of flow vectors $\hat{\mathbf{f}}$ and $\tilde{\mathbf{f}}$, if $\hat{f}_s \geq \tilde{f}_s$ and $\hat{f}_{-s} \geq \tilde{f}_{-s}$ for all sections $s$ of $p$, then $c_p(\hat{\mathbf{f}}) \geq c_p(\tilde{\mathbf{f}})$. This implies, in particular, that the cost of a path only depends on the flow on each of its sections and the flows in the opposite directions.[3] A cost function will be said to be *increasing* if it satisfies the following additional

[2] Thus, the vectors in the domain and range of **c** have different dimensions: the former equals the number of routes in *G* and the latter the number of paths. Note that the costs are not assumed to be nonnegative. However, they may well be *thought of* as such. Indeed, the assumption that the costs are nonnegative is implicit in the definition of equilibrium (in the next section). This definition only considers routes, which by definition do not pass through any vertex more than once.

[3] Because this monotonicity condition involves a potentially long list of premises, the restriction it puts on the allowable cost functions is relatively weak. Stronger conditions, e.g., a requirement that the cost of a path can increase only if one of the relevant *arc* flows increases, could be used instead. However, weaker conditions here and in the next definition are preferable since they make for stronger results. See also the discussion in Section 6. Because of the dimensionality issue mentioned in the previous footnote, the present monotonicity conditions are technically incomparable with *monotonicity* and *strict monotonicity* as defined, e.g., by Nagurney (1999). Note, however, that the latter are wider in that they allow for "crosstalk," i.e., the cost of a route may be influenced by the flow on a parallel route.

condition: For every route $r$ and every pair of flow vectors $\hat{\mathbf{f}}$ and $\tilde{\mathbf{f}}$, if $\hat{f}_s \geq \tilde{f}_s$ and $\hat{f}_{-s} \geq \tilde{f}_{-s}$ for all sections $s$ of $r$, and there is at least one section $s$ of length one for which $\hat{f}_s > \tilde{f}_s$, then $c_r(\hat{\mathbf{f}}) > c_r(\tilde{\mathbf{f}})$. A cost function $\mathbf{c}$ is (additively) *separable* if the equality $c_{r_{ov}}(\mathbf{f}) = c_{r_{ou}}(\mathbf{f}) + c_{r_{uv}}(\mathbf{f})$ holds for every route $r$, every pair of distinct vertices $u$ and $v$ such that $u$ precedes $v$ in $r$, and every flow vector $\mathbf{f}$. In other words, separability means that the cost of each route is the sum of the costs of its arcs.

## 4 Efficiency of Equilibrium

A flow vector $\mathbf{f}^*$ is said to be an *equilibrium* if the entire flow in the network is on minimal-cost routes, that is,

$$\text{for all routes } r \text{ with } f_r^* > 0,\ c_r(\mathbf{f}^*) = \min_{q \in \mathcal{R}} c_q(\mathbf{f}^*). \tag{2}$$

For an equilibrium $\mathbf{f}^*$, the minimum in (2), denoted by $c(\mathbf{f}^*)$, is the *equilibrium cost*. In the transportation literature, a flow vector satisfying (2) is known as Wardrop, or user equilibrium. This condition expresses the principle that, at equilibrium, the travel time on all used routes is the same, and less than or equal to that of a single vehicle on any unused route (Wardrop, 1952). The equilibrium condition (2) can also be given a variational inequality formulation (Nagurney, 1999, Theorem 4.5): For every flow vector $\mathbf{f}$ with the same total origin–destination flow as $\mathbf{f}^*$,

$$\sum_{r \in \mathcal{R}} c_r(\mathbf{f}^*)(f_r^* - f_r) < 0.$$

If the cost function is continuous, then, by standard results (e.g., Nagurney, 1999, Theorem 1.4), it follows from this formulation that for any $f_o \geq 0$ there is at least one equilibrium with a total origin–destination flow of $f_o$.

Definition 1. *Braess's paradox* occurs in a network $G$ if there are two separable cost functions $\hat{\mathbf{c}}$ and $\tilde{\mathbf{c}}$ such that $\hat{c}_r(\mathbf{f}) \geq \tilde{c}_r(\mathbf{f})$ for all routes $r$ and flow vectors $\mathbf{f}$, but for every equilibrium[4] $\hat{\mathbf{f}}$ with respect to $\hat{\mathbf{c}}$ with a total origin–destination flow of unity and every equilibrium[4] $\tilde{\mathbf{f}}$ with respect to $\tilde{\mathbf{c}}$ with a similar total origin–destination flow, the equilibrium costs satisfy $\hat{\mathbf{c}}(\hat{\mathbf{f}}) < \tilde{\mathbf{c}}(\tilde{\mathbf{f}})$.

Thus, Braess's paradox occurs if raising the edge costs can *lower* the equilibrium cost. This is the case in the example in Figure 1, in which a higher cost for the transverse edge $e_5$ results in a shorter equilibrium travel time on the network. By Proposition 1, the network in Figure 1 is embedded in every network that is not series-parallel. This implies that Braess's paradox occurs in all such networks. The following theorem shows that it occurs *only* in these networks.

**Theorem 1.** Braess's paradox does *not* occur in a two-terminal network $G$ if and only if $G$ is series-parallel.

[4] This definition may be changed a little by replacing "every equilibrium" with "some equilibrium." This change does not affect any of the results below.

Theorem 1 confirms unproven assertions made by Murchland (1970) and Calvert and Keady (1993). Murchland asserts that deletion of one or more edges from a series-parallel network cannot be beneficial. Calvert and Keady present a theorem (Theorem 11) stating that Braess's paradox cannot occur in a series-parallel physical network in which the potential difference between the two end vertices of each edge is determined as an increasing function by the quotient of the flow on the edge and an edge-specific conductivity factor. This refers to a version of Braess's paradox occurring when the total power loss in the network can be decreased by reducing the conductivity of some edge, with the total origin–destination flow held constant.

Even though series-parallel networks never exhibit Braess's paradox, they do not always have efficient equilibria. This is demonstrated by the example in Figure 2, in which the equilibrium flow can be rearranged in such a way that the costs of all used routes are below the equilibrium cost. As the next theorem shows, the reason inefficient equilibria occur in the networks in Figure 2 is that their routes are not linearly independent.

DEFINITION 2. For given network $G$ and cost function $\mathbf{c}$, an equilibrium $\mathbf{f}^*$, with equilibrium cost $c(\mathbf{f}^*)$, is *weakly Pareto efficient* if, for every flow vector $\mathbf{f}$ with the same total origin–destination flow as $\mathbf{f}^*$, there is some route $r$ with $f_r > 0$ for which $c_r(\mathbf{f}) \geq c(\mathbf{f}^*)$. The equilibrium is *Pareto efficient* if, for every flow vector $\mathbf{f}$ with the same total origin–destination flow as $\mathbf{f}^*$, either $c_r(\mathbf{f}) = c(\mathbf{f}^*)$ for all $r$ with $f_r > 0$ or there is some route $r$ with $f_r > 0$ for which $c_r(\mathbf{f}) > c(\mathbf{f}^*)$.

**Theorem 2.** For a two-terminal network $G$, the following conditions are equivalent:

(i) For any cost function, all equilibria are weakly Pareto efficient.

(ii) For any increasing cost function, all equilibria are Pareto efficient.

(iii) $G$ has linearly independent routes.

The weak Pareto efficiency of the equilibria in networks with linearly independent routes implies that, in such networks, the equilibrium cost is uniquely determined by the cost function and the total origin–destination flow, and can only increase or remain unchanged if the former or the latter increase. With separable cost functions, this is also true for general series-parallel networks (see Lemma 4 below). However, for a *non*-separable cost function in a series-parallel network that does not have linearly independent routes, the equilibrium cost may not be unique, or may decrease rather than increase with rising costs. For example, suppose that, in Figure 2(a), a toll is charged for using the equilibrium route $oe_1ve_3d$. Increasing the toll from zero to $1½$ *decreases* the (unique) equilibrium cost linearly from $4$ to $3½$. (The lower equilibrium cost is also achievable by the following turning restriction, which is equivalent to infinite toll: traffic emerging from $e_1$ is not allowed to turn into $e_3$.) This example shows that Theorem 1 would not hold if, in the definition of Braess's paradox, the assumption of separable cost functions were dropped. This contrasts with the situation in Theorem 2, which does not assume separability.

Theorem 2 parallels an earlier result of Holzman and Law-Yone (1997, 2003) for route selection games in *directed* networks with a *finite* number of players. If the set of all (directed) routes in a directed network does not contain a bad configuration (as defined in Section 2), then, for any (finite) number of players and any nonnegative separable cost function, all the equilibria are weakly Pareto efficient and, moreover, *strong* in the sense that no group of players can make all its members better off by changing their route choices. Conversely, if a triplet of routes constituting a bad configuration exists, then, for any number of players, there is a nonnegative separable cost function for which none of the equilibria is weakly Pareto efficient. For directed networks, the absence of a bad configuration is a stronger condition than linear independence of the routes (cf. Proposition 4). For this reason, Theorem 2, in the form given above, does not hold for such networks. For example, the three routes in the directed Wheatstone network are linearly independent in the sense that each of them has a directed edge that is not in any other route. Nevertheless, the example in Figure 1 shows that equilibria in this network are not always weakly Pareto efficient.

## 5 Non-Identical Users

The most significant difference between the finite route selection games considered by Holzman and Law-Yone (1997, 2003) and the present model is that, in the former but not in the latter, *heterogeneity* is a potential source of inefficiency. The population of users is heterogeneous if there are differences in the intrinsic quality they assign to routes or the degree to which they are affected by congestion. For example, some motorists may be concerned primarily with the travel time, and others with the distance traveled. In finite populations (Milchtaich, 1996), such differences may lead to inefficient equilibria. This is demonstrated by the simple two-person game in which there are two parallel routes, each favored by a different person. If sharing a route with the other user is very costly, then there are two pure-strategy Nash equilibria, and the one in which both persons use their favorite routes strictly Pareto dominates that in which each of them uses the other route. This example can easily be extended to any nontrivial network (i.e., one with more than one route), which shows that Holzman and Law-Yone's result cannot be extended to heterogeneous finite populations. By contrast, it is shown below that Theorem 2 can be extended. In particular, with a continuum of non-identical users, a Nash equilibrium may be strictly Pareto dominated by another equilibrium only in a network with an embedded network as in Figure 1 or Figure 2. In networks without this property, i.e., with linearly independent routes, the equilibria are always weakly Pareto efficient, for both heterogeneous and homogeneous populations.

Even with a continuum of users, however, there are differences between the cases of identical and non-identical users. In particular, the result that, in a series-parallel network, Braess's paradox cannot occur (Theorem 1) does not extend to the case of non-identical users. For example, in Figure 2(a), if half the users were charged a hefty toll for using edge $e_1$ and the other half for using $e_3$, they would not use these edges (the former would take $e_2$ and $e_3$ and the latter $e_1$ and $e_4$), and consequently the equilibrium costs for all users would decrease from 4 to 3½. More precisely, this is an example of a natural generalization of Braess's paradox, which is defined below. This paradox does not occur in networks with linearly independent routes, in which the equilibria are weakly Pareto efficient. As Theorem

3 below shows, with non-identical users, these are, in fact, the *only* networks in which Braess's paradox does not occur.

Dropping the assumption that all users of the same route incur the same costs leads to the following modified version of the model described in Section 3. The *population of users* is an infinite set $I$ (e.g., the unit interval $[0,1]$), endowed with a nonatomic probability measure (e.g., Lebesgue measure), which assigns values between zero and one to a $\sigma$-algebra of subsets of $I$, the *measurable sets*. These values are interpreted as the set sizes relative to the total population. For a network $G$, a *strategy profile* is a mapping $\sigma: I \to \mathcal{R}$ (from users to routes) such that, for each route $r$, the set of all users $i$ with $\sigma(i) = r$ is measurable. The measure of this set, denoted by $f_r(\sigma)$, is the flow on route $r$. Thus, for every strategy profile $\sigma$, there is a corresponding flow vector $\mathbf{f}(\sigma)$ with a total origin–destination flow of unity. For each user $i$, a *cost function* $\mathbf{c}^i$ specifies the cost $c_p^i(\mathbf{f})$ of each path $p$ in $G$ as a function of the flow vector $\boldsymbol{f}$. As in the case of identical users, $\mathbf{c}^i$ is assumed to satisfy the condition that, for every path $p$ and every pair of flow vectors $\hat{\mathbf{f}}$ and $\tilde{\mathbf{f}}$, if $\hat{f}_s \geq \tilde{f}_s$ and $\hat{f}_{-s} \geq \tilde{f}_{-s}$ for all sections $s$ of $p$, then $c_p^i(\hat{\mathbf{f}}) \geq c_p^i(\tilde{\mathbf{f}})$. The definitions of increasing and separable cost functions are also similar to those for the case of a homogeneous population of users. A strategy profile $\sigma$ is a *Nash equilibrium* if each of the routes is a minimal-cost route for its users, that is,

$$\text{for each user } i,\ c^i_{\sigma(i)}(\mathbf{f}(\sigma)) = \min_{q \in \mathcal{R}} c^i_q(\mathbf{f}(\sigma)). \tag{3}$$

In this case, the minimum in (3), denoted by $c^i(\mathbf{f}(\sigma))$, is the *equilibrium cost* for user $i$. In the special case in which all users have the same cost function, this definition essentially reduces to (2).

Definition 3. For given network $G$ and assignment of cost functions $\mathbf{c}^i$, a strategy profile $\sigma$ is *weakly Pareto efficient* if, for every strategy profile $\tau$, there is some user $i$ for which $c^i_{\tau(i)}(\mathbf{f}(\tau)) \geq c^i_{\sigma(i)}(\mathbf{f}(\sigma))$. A strategy profile $\sigma$ is *Pareto efficient* if, for every strategy profile $\tau$, either $c^i_{\tau(i)}(\mathbf{f}(\tau)) = c^i_{\sigma(i)}(\mathbf{f}(\sigma))$ for all users $i$ or there is some $i$ for which $c^i_{\tau(i)}(\mathbf{f}(\tau)) > c^i_{\sigma(i)}(\mathbf{f}(\sigma))$. A strategy profile $\sigma$ is *hyper-efficient* (Milchtaich, 2004) if, for every strategy profile $\tau$,

$$\begin{array}{l}\text{either } c^i_{\tau(i)}(\mathbf{f}(\tau)) = c^i_{\sigma(i)}(\mathbf{f}(\sigma)) \text{ for all users } i \text{ or} \\ \text{there is some } i \text{ with } \tau(i) \neq \sigma(i) \text{ for which } c^i_{\tau(i)}(\mathbf{f}(\tau)) > c^i_{\sigma(i)}(\mathbf{f}(\sigma)).\end{array} \tag{H}$$

*Braess's paradox with non-identical users* occurs in a network $G$ if it is possible to assign two separable cost functions $\hat{\mathbf{c}}^i$ and $\tilde{\mathbf{c}}^i$ for each user $i$ such that $\hat{c}^i_r(\mathbf{f}) \geq \tilde{c}^i_r(\mathbf{f})$ for all users $i$, routes $r$ and flow vectors $\mathbf{f}$, but for every Nash equilibrium $\sigma$ with respect to the first assignment and every Nash equilibrium $\tau$ with respect to the second, the equilibrium costs for each user $i$ satisfy $\hat{c}^i(\mathbf{f}(\sigma)) < \tilde{c}^i(\mathbf{f}(\tau))$.

Except for the notion of hyper-efficiency, Definition 3 is a straightforward generalization of Definitions 1 and 2. Hyper-efficiency, which is meaningful also for identical uses, means that any effective change of route choices is harmful to some of those changing routes. Clearly, any hyper-efficient strategy profile $\sigma$ is both Pareto efficient and a Nash equilibrium. Indeed, it is a strong, and even strictly strong,[5] equilibrium. This means that deviations are never profitable, not only for individuals but also for groups of users: Any deviation that makes someone in the group better off must leave someone else in it worse off. The following theorem shows that, in a network with linearly independent routes, if all the cost functions are increasing, the converse is also true. That is, under these conditions, *every* Nash equilibrium is hyper-efficient and, hence, Pareto efficient and a strictly strong equilibrium. Clearly, a similar result also holds in the special case of identical users.

**Theorem 3.** For a two-terminal network $G$, the following conditions are equivalent:

(i) For any assignment of cost functions, all Nash equilibria are weakly Pareto efficient.

(ii) For any assignment of increasing cost functions, all Nash equilibria are hyper-efficient.

(iii) Braess's paradox with non-identical users does not occur in $G$.

(iv) $G$ has linearly independent routes.

## 6 Discussion

Some properties of the equilibria in route selection games with a continuum of users are virtually independent of the network topology, and others strongly depend on it. Social optimality of the equilibria and the price of anarchy are among the former (see the introduction). Non-occurrence of Braess's paradox and Pareto efficiency of the equilibria (this paper) and (with non-identical users) uniqueness of the equilibrium (Milchtaich, 2005; and see below) are among the latter. Dependence of a property on the network topology means that it holds for *all* allowable cost functions if and only if the network belongs to some specified non-trivial class. As this paper shows, with identical users, the two-terminal networks in which Braess's paradox never occurs are the series-parallel ones (Theorem 1). Those in which only Pareto efficient equilibria occur are the networks with linearly independent routes (Theorem 2). With non-identical users, each of these two properties is guaranteed to hold if and only if the network has linearly independent routes (Theorem 3). These conditions for *topological efficiency* are different from those for *topological uniqueness*. A two-terminal network is said to have the latter property if, for any assignment of separable cost functions, the flow on each arc is the same in all Nash equilibria. (This refers to a heterogeneous population of users. With a homogeneous population, the equilibrium is always essentially unique for any network.) Two equivalent characterizations of the class of all two-terminal networks with the topological uniqueness property are given in Milchtaich (2005). This class is incomparable with the two classes considered in this paper: Linear independence of the routes is neither a sufficient nor a necessary condition for topological uniqueness, and the same is true for series-parallel network. For example, multiple equilibria, which differ even in the mean cost incurred by the (non-identical) users, may exist in the network with linearly independent routes shown in Figure 3, but not in the

[5] A strategy profile $\sigma$ is a *strictly strong equilibrium* (Voorneveld at al., 1999) if, for every strategy profile $\tau$, $c^i_{\tau(i)}(\mathbf{f}(\tau)) \geq c^i_{\sigma(i)}(\mathbf{f}(\sigma))$ for all users $i$ or there is some $i$ with $\tau(i) \neq \sigma(i)$ for which $c^i_{\tau(i)}(\mathbf{f}(\tau)) > c^i_{\sigma(i)}(\mathbf{f}(\sigma))$.

Wheatstone network (Figure 1). A parallel network (which consists of several edges connected in parallel) belongs to all three classes, and the network obtained by connecting a single edge in parallel with the Wheatstone network does not belong to any of them.

An essential feature of the models in this paper and Milchtaich (2005) is that networks are *undirected*. This diverges from the common practice of modeling transportation and other kinds of networks by means of directed networks (e.g., Sheffi, 1985; Bell and Iida, 1997; Nagurney, 1999), so that each edge can be traversed in only one direction. (Thus, for example, a two-way highway is described by a *pair* of edges.) In most of the related literature, the description of a network involves these two kinds of data: A directed graph, which describes both the physical network and the direction of travel on each edge; and a corresponding system of edge costs, which gives the cost of each directed edge as a function of the flow on it. A different kind of transportation model, described by Beckmann et al. (1956), assumes all roads to be two-way, with the same travel costs in both directions. These authors moreover assume that the costs depend only on the sum of the flows in all routes passing through the road in either direction. (The two kinds of models coincide in the special case of series-parallel networks. In this case, all routes pass through an edge in the same direction, and so the cost of passing through it in the opposite direction and the effect of the opposite flow on the cost are irrelevant.) The model presented in this paper subsumes both these kinds of models. Here, a separable cost function associates with each edge a pair of edge costs—one for each direction. In each direction, the cost is a function of the flows on the edge in that and the opposite direction, and possibly also the junction flows. The first kind of model described above corresponds to the special case in which the cost of passing through each edge in a particular direction is prohibitively high. The second kind corresponds to a case in which the two costs are equal, and only depend on the sum of the arc flows in both directions. This shows, in particular, that using an undirected network does not preclude directionality. It only makes it part of the cost function rather than the network topology, and thus allows it to vary. In other words, an undirected network only precludes *predetermined* directionality. Thus, topological efficiency essentially means that, regardless of how the edges in the network are directed and the edge costs, inefficient equilibria do not exist. Similarly, topological uniqueness refers to absence of multiple equilibria, regardless of directionality and costs. Similar properties may also be defined for directed networks. However, there seem to be no known results linking the topology of directed networks with the efficiency or uniqueness of the equilibria in nonatomic congestion games, other than those that can be derived from the results in this paper and Milchtaich (2005) as special cases.

The definition of cost function in this paper is rather wide. It does not assume separability, which means that turning restrictions, for example, can be incorporated simply by assigning very high (effectively, infinite) costs to certain routes (see the example in Section 4). It also allows route costs to be affected by the flows on their vertices (which may represent, for example, a crude measure of congestion at four-way stop junctions). In Milchtaich (2005), a more standard definition is used, which requires cost functions to be: (i) increasing, (ii) nonnegative, and (iii) separable, with (iv) the cost for each user of each edge $e$ in each direction depending only on the flow on $e$ in that direction (and not on the opposite flow or the flows on the end vertices). Adopting the same restrictive definition here, in either the homogeneous or the heterogeneous case, would not affect any of the theorems. Indeed, inspection of the proofs of Theorems 1, 2 and 3 shows them to be also valid if the definition of cost function is augmented with any subset of (i)–(iv), to which the requirement of *continuity* of the payoff function may be added. This is mainly because all five properties hold for the cost functions in the two examples given in the introduction. Thus, in particular, both the possibility of non-separable cost functions and the possible effects of junction flows are not crucial elements of the present model. The notion of *embedding in the wide sense* used in Milchtaich (2005) is also different from the present notion of embedding. The former is wider than the latter, but more complicated, and in the present context, it does not offer any advantages. However, it could be used, as both Propositions 1 and 4 can be shown to also hold with "embedding in the wide sense" replacing "embedding."

In this paper, flow is always assumed to originate in a single vertex $o$ and terminate in a single vertex $d$. Multiple origin–destination pairs are not allowed. This restriction can be partially circumvented by connecting all sources to a single, fictitious, vertex, from which all flow is assumed to originate, and similarly for the sinks. However, such a construction substantially alters the network topology. This leaves open the question, how does the results in this paper change when there are more than one origin or destination.

## 7 Proofs

This section gives the proofs for the results presented in this paper.

*Proof of Proposition 1.* (i) ⇒ (iii). The network in Figure 1 has an edge, $e_5$, through which two routes pass in opposite directions. It is easy to see that this property is preserved under the three operations that define embedding. Therefore, a network in which the one in Figure 1 is embedded is not series-parallel.

(ii) ⇒ (i). This follows from the special case in which $u$ and $v$ are the two end vertices of an edge $e$.

(iii) ⇒ (ii). Suppose that condition (ii) does not hold for $G$: There are two routes $r$ and $r'$ and two vertices $u$ and $v$ common to both routes, such that $u$ precedes $v$ in $r$ but follows it in $r'$. Suppose that these vertices are chosen in such a way that the length of the section $r_{uv}$ is maximal. Then, any vertex $u'$ common to $r$ and $r'$ that precedes $u$ in $r$ must precede $v$ in $r'$, and any vertex $v'$ common to both routes that follows $v$ in $r$ must follow $u$ in $r'$ (see Figure 5). Let $u'$ be the *last* vertex before $u$ in $r$ that is also in $r'$ (possibly, $u' = o$), and $v'$ the *first* vertex after $v$ in $r$ that is also in $r'$ (possibly, $v' = d$). All the edges in $r_{u'u}$, and all the vertices in this section of $r$ with the exception of the initial and terminal ones, do *not* belong to $r'$, and the same is true for $r_{vv'}$. This implies that the network in Figure 1 is embedded in the sub-network of $G$ consisting of all the vertices and edges in $r_{u'u}$, $r_{vv'}$ and $r'$. Hence, it is also embedded in $G$. ∎

The following lemma, which is essentially part of Theorem 3.3 of Harary (1969), is used in the proof of Proposition 2.

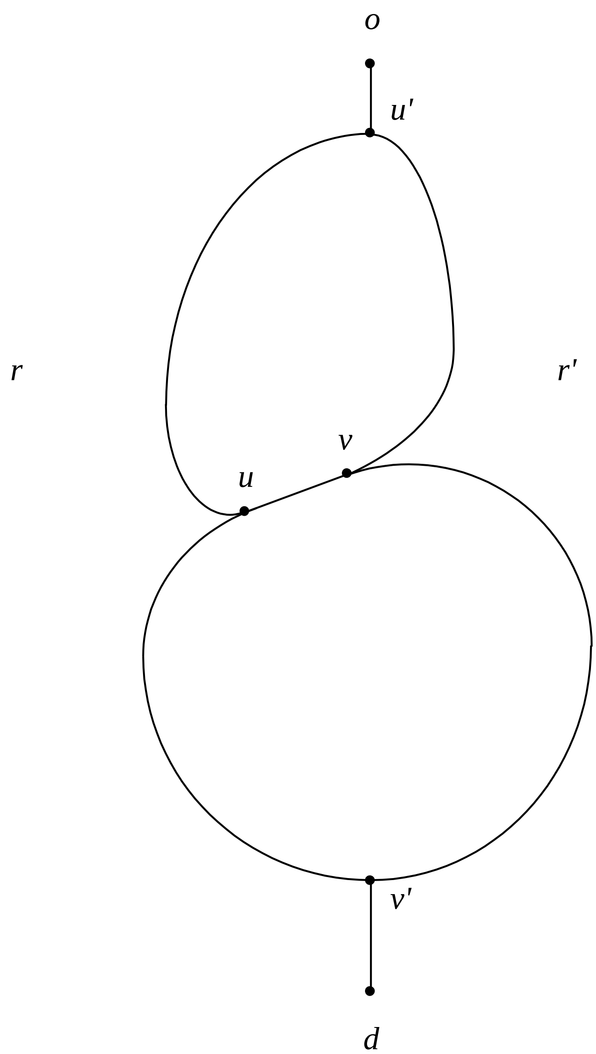


Figure 5

**Lemma 1.** A network $G$ can be obtained by connecting two other networks in series if and only if every two routes in $G$, distinct or identical, have at least one vertex in common, other than $o$ and $d$.

*Proof*. The necessity of this condition is clear. To prove sufficiency, suppose that the condition holds, and consider the set of all triplets $(p, q, v)$ consisting of two distinct routes $p$ and $q$ and a vertex $v$ common to both routes such that: (i) $p_{ov}$ and $q_{ov}$ do not have common vertices other than $o$ and $v$ (which implies that $v \neq d$, since, by assumption, routes $p$ and $q$ do have at least three common vertices), and (ii) $p_{vd} = q_{vd}$ (which implies that $v \neq o$, since $p \neq q$). If this set is empty, then there is only one edge incident with $o$, which implies that $G$ is the result of connecting two networks in series, one of which only has that single edge. Suppose, then, that the above set is nonempty, and choose an element $(p, q, v)$ such that the length of $p_{vd}$ $(= q_{vd})$ is minimal.

CLAIM 1. The vertex $v$ belongs to all routes in $G$.

Suppose the contrary, that $v$ does not belong to some route $r$. In that route, let $v'$ be the first vertex that is also in $p_{vd}$, and $u$ the last vertex before $v'$ that is also in $p$ or in $q$ (see Figure 6).

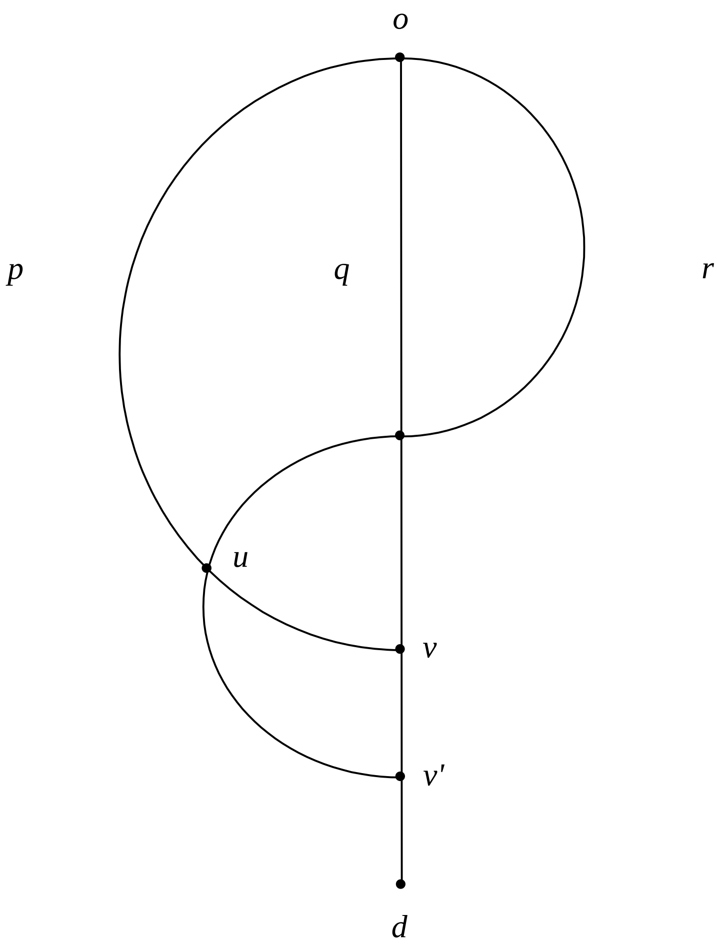


Figure 6

Without loss of generality, it may be assumed that $u$ is in $p$. Consider the route $p' = p_{ou} + r_{uv'} + p_{v'd}$. Clearly, (i) $p'_{ov'}$ and $q_{ov'}$ do not have common vertices other than $o$ and $v'$, and (ii) $p'_{v'd} = q_{v'd}$. However, the section $p'_{v'd}$ is shorter than $p_{vd}$, which contradicts the way the triplet $(p, q, v)$ was chosen. This contradiction proves Claim 1.

CLAIM 2. Any vertex $u$ that precedes $v$ in *some* route in $G$ also precedes it in every other route to which $u$ belongs.

Suppose the contrary, that there are two routes $r$ and $r'$ such that some vertex $u \neq v$ common to both routes precedes $v$ in $r$ but follows it in $r'$. Choosing $u$ to be the *first* such vertex in $r$ guarantees that $r_{ou}$ and $r'_{ud}$ do not have common vertices other than $u$ (see Figure 5). Clearly, the route $r_{ou} + r'_{ud}$ does not include $v$. This contradicts Claim 1, and thus proves Claim 2.

It follows from Claims 1 and 2 that $G$ is the result of connecting two networks in series: the network $G'$ consisting of $v$ (as destination) and all the vertices and edges that precede it in some route in $G$, and the network $G''$ consisting of $v$ (as the origin) and all the vertices and edges that follow it in some route in $G$. ■

*Proof of Proposition 2.* Observe, first, that if a network $G$ is the result of connecting two networks $G'$ and $G''$ in series or in parallel, then $G$ is series-parallel if and only if both $G'$ and $G''$ are series-parallel. From this observation, it immediately follows that if $G$ is series-parallel and has some route with only one edge, then $G$ satisfies conditions (i) or (ii). In the rest of the proof, it will be assumed that $G$ is series-parallel and all the routes in it have at least two edges. It is to be shown that $G$ is the result of connecting two networks in series or in parallel.

CLAIM. In the route set of $G$, "routes $p$ and $q$ have a vertex in common, other than $o$ and $d$" is an equivalence relation.

Reflexivity of the relation follows from the assumption that each route in $G$ has at least two edges. Symmetry is obvious. It remains to show that the relation is transitive. That is, if $p$, $q$ and $r$ are three routes such that there is some vertex $v \neq o, d$ that is common to $p$ and $q$ and some vertex $u \neq o, d$ that is common to $p$ and $r$, then there is also some vertex, other than $o$ and $d$, common to $q$ and $r$. Suppose this is not so. Without loss of generality, it may be assumed that $u$ precedes $v$ in $p$ and that none of the other vertices in $p_{uv}$ belongs to either $q$ or $r$ (see Figure 6). This assumption implies that the network in Figure 1 is embedded in the sub-network of $G$ consisting of all the vertices and edges in $p_{uv}$, $q$ and $r$. However, by Proposition 1, this contradicts the assumption that $G$ is series-parallel. This contradiction proves the claim.

Two cases are possible. Either the above equivalence relation holds for all pairs of routes in $G$, or there are two or more equivalence classes. In the former case, it follows from Lemma 1 that $G$ is the result of connecting two networks in series. In the latter case, choose one of the equivalence classes and consider the sub-network $G'$ consisting of all the vertices and edges that belong to at least one route in this equivalence class, as well as the sub-network $G''$ consisting of all the vertices and edges that belong to at least one route *not* in the class. Each vertex $v$, other than $o$ and $d$, belongs to one, and only one, of these two sub-networks. (Otherwise, $v$ would belong to two routes in two different equivalence classes, which is impossible by definition of the equivalence relation.) Therefore, each edge also belongs to one, and only one, of them. This implies that $G$ is the result of connecting $G'$ and $G''$ in parallel. ■

*Proof of Proposition 3.* The condition in the proposition is equivalent to the following one:

> There is a one-to-one function $Z$ from the vertex set to the integers such that, for every pair of distinct vertices $u$ and $v$, if $u$ precedes $v$ in some route $r$, then $Z(u) < Z(v)$.

This condition is sufficient for $G$ to be series-parallel, since it implies that the same end vertex of each edge $e$ precedes the other in every route passing through $e$. The necessity of the condition can be proved by induction on the number of edges in $G$. If there is only one edge, the condition holds trivially. Suppose that $G$ has more than one edge, and that the condition holds for every series-parallel network with a smaller number of edges than $G$. By Proposition 2, $G$ is the result of connecting two series-parallel networks, $G'$ and $G''$, in series or in parallel. Since they have less edges than $G$, both networks satisfy the above condition. Thus, a function $Z'$ as above exists for $G'$ and another one $Z''$ for $G''$. It is not difficult to see that $Z'$ and $Z''$ can be chosen in such a way that, for every vertex $u$ in $G'$ and every vertex $v$ in $G''$, $Z'(u) = Z''(v)$ if and only if $u = v$. This implies that the functions $Z'$ and $Z''$ have a unique common extension $Z$ on the union of the two vertex sets, which is the vertex set in $G$. This proves that $G$ also satisfies the above condition. ■

*Proof of Proposition 4.* (i) ⇒ (iv). Suppose that one of the networks in Figure 1 and Figure 2 is embedded in $G$. In each of these networks, there is a route, every edge of which is also in some other route. It is easy to see that this property is preserved under the three operations that define embedding, and so it also holds for $G$. Hence, $G$ is not a network with linearly independent routes.

(iii) ⇒ (i). Suppose that $G$ is not a network with linearly independent routes: There is some route $r$, every edge of which is also in some other route. It has to be shown that a bad configuration exists in $G$. Let $q \neq r$ be a route with the greatest number of $r$'s edges. Since only $r$ itself can have *all* the edges in $r$, there is some edge $e_1$ in $r$ that is not in $q$. By assumption, $e_1$ belongs to some route $p \neq r$. By construction, $p$ does not have more edges in common with $r$ than $q$, and so there is at least one edge $e_2$ common to $r$ and $q$ that is not in $p$. This implies that $p$, $q$ and $r$ constitute a bad configuration.

(ii) ⇒ (iii). Suppose that a bad configuration exists: There are three routes $p$, $q$ and $r$ such that some edge $e_1$ belongs to $p$ but not to $q$, another one $e_2$ belongs to $q$ but not to $p$, and both edges are in $r$, with $e_1$ preceding $e_2$, say. It has to be shown that condition (ii) does not hold. If the network is not series-parallel, this follows immediately from Proposition 1, since condition (ii) does not hold for the network in Figure 1 (in which it is violated by the two routes that include both $u$ and $v$). Suppose, then, that the network is series-parallel. This implies that $p$ and $q$ pass through $e_1$ and $e_2$, respectively, in the same directions as $r$. Consider the set $\mathcal{V}'$ of all vertices in $r$ that follow $e_1$ but precede $e_2$. Let $u$ be the first vertex in $q$ that is also in $\mathcal{V}'$, and $v$ the last vertex in $p$ that is also in $\mathcal{V}'$ (see Figure 7). If $q_{ou}$ and $p_{vd}$ have a common vertex $w$, then $p_{ow} \neq q_{ow}$ and $p_{wd} \neq q_{wd}$. The former inequality holds since $e_1$ is in the section $p_{ow}$ but not in $q$, and the latter inequality holds since $e_2$ is in the section $q_{wd}$ but not in $p$. These two inequalities imply that condition (ii) does not hold. Suppose, then, that $q_{ou}$ and $p_{vd}$ do not have a common vertex. This implies that $u$ precedes $v$ in $r$, because if $u$ followed $v$, then $q_{ou} + (-r_{vu}) + p_{vd}$ would be a route passing through the edges in $r_{vu}$ in opposite direction to $r$, which contradicts the assumption that the network is series-parallel. Consider the two routes $r$ and $r' = q_{ou} + r_{uv} + p_{vd}$. They satisfy $r_{ov} \neq r'_{ov}$ and $r_{vd} \neq r'_{vd}$, since $e_1$ and $e_2$ are in $r_{ov}$ and $r_{vd}$, respectively, but not in $r'$. These inequalities imply that condition (ii) does not hold.

(iv) ⇒ (ii). Suppose that condition (ii) does not hold for $G$: There are two routes $r$ and $r'$ and a vertex $v$ common to both routes such that $r_{ov} \neq r'_{ov}$ and $r_{vd} \neq r'_{vd}$. It is to be shown that condition (iv) also does not hold. If $G$ is not series-parallel, this follows from Proposition 1. Suppose, then, that $G$ is series-parallel. Let $e_1$ and $e_2$ be edges in $r_{ov}$ and $r_{vd}$, respectively, that are not in $r'$. Without loss of generality, it may be assumed that $v$ is the *last* vertex between $e_1$ and $e_2$ in $r$ that is also in $r'$. Let $u$ (possibly, $u = v$) be the *first* such vertex (see Figure 7). Let $u'$ be the last vertex before $u$ in $r$ that is also in $r'$, and $v'$ the first vertex after $v$ in $r$ that is also in $r'$. In the sections $r_{u'u}$ and $r_{vv'}$, only the initial and terminal vertices, and

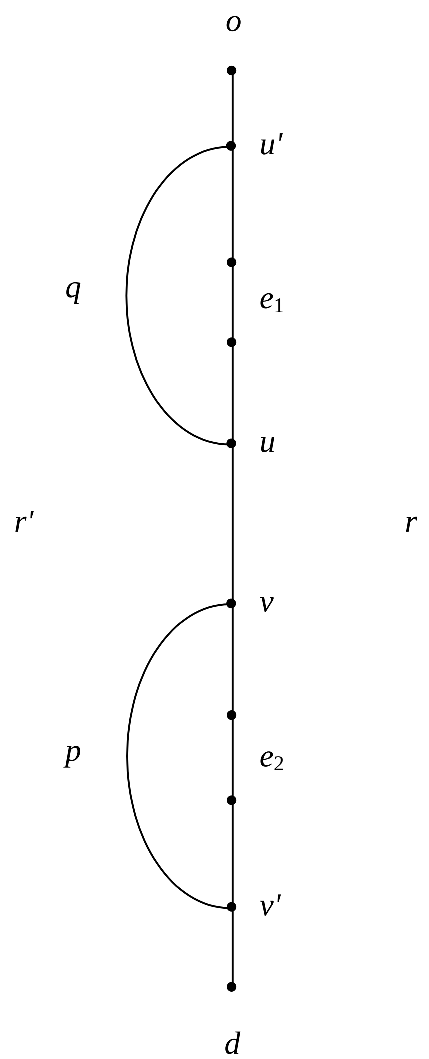


Figure 7

none of the edges, are in $r'$. By Proposition 1, $v'$ and $u$ follow $v$ and $u'$, respectively, also in $r'$. In addition, either $v$ coincides with $u$, or it follows it also in $r'$. In the former case, the network in Figure 2(a) is embedded in the sub-network of $G$ consisting of all the vertices and edges in $r_{u'u}$, $r_{vv'}$ and $r'$. In the latter case, the network in Figure 2(b) is embedded in this sub-network. This implies that one of the networks in Figure 2 is embedded in $G$. ■

*Proof of Proposition 5.* One direction is obvious: If $G$ satisfies (i), (ii) or (iii), then it is a network with linearly independent routes. To prove the converse, suppose that $G$ is a network with linearly independent routes.

CLAIM 1. If two routes in $G$ have a vertex in common, other than $o$ and $d$, then they have the same first or last edge.

This follows from condition (ii) in Proposition 4.

CLAIM 2. In the route set of $G$, “routes $p$ and $q$ have an edge in common” is an equivalence relation.

This follows from condition (iii) in Proposition 4: If route $r$ shares an edge $e_1$ with route $p$ and an edge $e_2$ with route $q$, then at least one of the edges $e_1$ and $e_2$ must be common to $p$ and $q$, otherwise the three routes would constitute a bad configuration.

Two cases are possible. Either the equivalence relation in Claim 2 holds for all pairs of routes in $G$, or there are two or more equivalence classes. In the latter case, choose one of the equivalence classes, and consider the sub-network $G'$ consisting of all the vertices and edges that belong to at least one route in this equivalence class, as well as the sub-network $G''$ consisting of all the vertices and edges that belong to at least one route *not* in the class. It follows from Claim 1 and the definition of the equivalence relation that every vertex in $G$, other than $o$ and $d$, belongs to one, and only one, of these two networks, and the same is true for every edge. This implies that $G$ is the result of connecting $G'$ and $G''$ in parallel. Since G is a network with linearly independent routes, the same is clearly true for $G'$ and $G''$. Hence, $G$ satisfies (ii).

In the rest of this proof, it will be assumed that the equivalence relation in Claim 2 holds for all pairs of routes in $G$, i.e., every route has an edge in common with every other route.

CLAIM 3. All the routes in $G$ have the same first edge, or they all have the same last edge.

To prove this, suppose that there are two routes $p$ and $q$ that do *not* have the same last edge. Since, by assumption, they have *some* edge in common, it follows from Claim 1 that their first edge is the same. Call this edge $e$. By a similar argument, any route $r$ that does not have $e$ as its first edge must have the same last edge as $p$. Similarly, it must have the same last edge as $q$. However, $p$ and $q$ do *not* have the same last edge. This contradiction proves that $e$ must be the first edge in every route, which completes the proof of Claim 3.

It follows from Claim 3 that there is some edge $e$, with $o$ or $d$ as one of its end vertices, that belongs to all routes. This implies that $e$ is the only edge incident with $o$ or $d$, respectively. Therefore, either $G$ has a single edge, or it is the result of connecting in series the network $G'$ consisting of $e$ and its two end vertices and the network $G''$ consisting of the end vertex of $e$ which is not $o$ or $d$ and the remaining vertices and edges in $G$. Clearly, the latter is a network with linearly independent routes. Hence, $G$ satisfies (i) or (iii). ■

*Proof of Proposition 6.* If $G$ is a network with linearly independent routes, then by definition each route $r$ in $G$ has an edge $e$ that is not in any other route. In the binary vector representing $r$, the coordinate corresponding to $e$ is $1$, and in the vectors representing the other routes, it is $0$. This implies that the former is not a linear combination of the latter, and so the vectors representing $G$'s routes are linearly independent. If $G$ is not a network with linearly independent routes, then by Proposition 4, one of the networks in Figure 1 and Figure 2 is embedded in it. In each of these three networks, the sum modulo 2 of the binary vectors representing the four routes in the network is zero, since an even number of routes pass through each edge. It is easy to see that the existence of four such routes is preserved under the three operations that define embedding. Therefore, also in the edge space of $G$, the vectors representing the routes are not linearly independent. ■

The following three lemmas are used in the proof of Theorem 1.

**Lemma 2.** Let $G$ be a series-parallel network, and $\hat{\mathbf{f}}$ and $\tilde{\mathbf{f}}$ two flow vectors. If the total origin–destination flows satisfy $\hat{f}_o \geq \tilde{f}_o$ and $\hat{f}_o > 0$, then there is some route $r$ such that, for all sections $s$ of $r$ of length zero or one (i.e., those consisting of only one vertex, or one edge and its two end vertices), $\hat{f}_s \geq \tilde{f}_s$ and $\hat{f}_s > 0$. If $\hat{f}_o > \tilde{f}_o$, then a similar result holds with the last pair of inequalities replaced by $\hat{f}_s > \tilde{f}_s$.

*Proof.* The proof of the lemma proceeds by induction on the number of edges. For a network with a single edge, the result is trivial. Consider, then, a series-parallel network $G$ with two or more edges. The induction hypothesis is that the assertion of the lemma holds for any two flow vectors in every series-parallel network with a smaller number of edges than $G$. By Proposition 2, $G$ is the result of connecting two series-parallel networks, $G'$ and $G''$, in series or in parallel. Consider, first, the case in which $G'$ and $G''$ are connected in series, so that the destination in $G'$, $v$, coincides with the origin in $G''$. The route sets of $G$ and $G'$, $\mathcal{R}$ and $\mathcal{R}'$, are connected by $\mathcal{R}' = \{r_{ov} \mid r \in \mathcal{R}\}$. (Thus, the elements of the latter are paths in $G$.) Every flow vector $\mathbf{f}$ for $G$ induces a flow vector $\mathbf{f}'$ for $G'$, which is given by $\mathbf{f}' = (f_{r'})_{r' \in \mathcal{R}'}$. By definition, the flow $f'_p$ on each path $p$ in $G'$ is the total flow on all the routes in $G'$ containing $p$, that is,

$$f'_p = \sum_{\substack{r' \in \mathcal{R}' \\ p \text{ is a section of } r'}} f_{r'}. \tag{4}$$

It is not difficult to see that this is equal to $f_p$, the flow on $p$ in $\mathbf{f}$. Therefore, if $\hat{\mathbf{f}}$ and $\tilde{\mathbf{f}}$ are two flow vectors for $G$ satisfying the pair of inequalities $\hat{f}_o \geq \tilde{f}_o$ and $\hat{f}_o > 0$, or the single inequality $\hat{f}_o > \tilde{f}_o$, then similar inequalities or inequality hold for the induced flow vectors $\hat{\mathbf{f}}'$ and $\tilde{\mathbf{f}}'$. It then follows from the induction hypothesis that there is a route $r'$ in $G'$ such that, for all sections $s$ of $r'$ of length zero or one, $\hat{f}_s \geq \tilde{f}_s$ and $\hat{f}_s > 0$, or $\hat{f}_s > \tilde{f}_s$, respectively. By similar considerations, there is a route $r''$ in $G''$ such that similar inequalities or inequality hold for all sections $s$ of $r''$ of length zero or one. Therefore, the same is true for the route $r = r' + r''$ in $G$. This completes the proof for the case in which $G$ is the result of connecting two series-parallel networks in series.

Next, suppose that $G$ is the result of connecting $G'$ and $G''$ in parallel, so that $o$ and $d$ are also the origin and destination, respectively, in $G'$ and $G''$. The route sets of $G'$ and $G''$, $\mathcal{R}'$ and $\mathcal{R}''$, are disjoint, and the route set $\mathcal{R}$ of $G$ is their union. Each flow vector $\mathbf{f}$ for $G$ induces flow vectors $\mathbf{f}'$ and $\mathbf{f}''$ for $G'$ and $G''$, which are given by $\mathbf{f}' = (f_{r'})_{r' \in \mathcal{R}'}$ and $\mathbf{f}'' = (f_{r''})_{r'' \in \mathcal{R}''}$. The induced flow vectors give (by (4) or a similar equation) the flow $f'_p$ or $f''_p$ on each path $p$ in $G'$ or $G''$, respectively. If $p$ is *not* one of the zero-length paths $o$ and $d$, then $f'_p = f_p$ or $f''_p = f_p$. However (unlike the case previously considered), the total origin–destination flows in $G'$ and $G''$ do not satisfy similar equalities. In fact, $f'_o + f''_o = \sum_{r' \in \mathcal{R}'} f_{r'} + \sum_{r'' \in \mathcal{R}''} f_{r''} = \sum_{r \in \mathcal{R}} f_r = f_o$. It follows from this equation that if $\hat{\mathbf{f}}$ and $\tilde{\mathbf{f}}$ are flow vectors satisfying $\hat{f}_o \geq \tilde{f}_o$ and $\hat{f}_o > 0$, then the inequalities $\hat{f}'_o \geq \tilde{f}'_o$ and $\hat{f}'_o > 0$, *or* $\hat{f}''_o \geq \tilde{f}''_o$ and $\hat{f}''_o > 0$, hold. By the induction hypothesis, there is then some route $r$, either in $G'$ or $G''$ (hence, in $G$), such that, for all sections $s$ of $r$ of length zero or one, $\hat{f}_s \geq \tilde{f}_s$ and $\hat{f}_s > 0$. By a similar argument, if $\hat{f}_o > \tilde{f}_o$, there is a route $r$ in $G'$ or $G''$ such that, for all sections $s$ of $r$ of length zero or one, $\hat{f}_s > \tilde{f}_s$. ∎

**Lemma 3.** Let $G$ be a series-parallel network, $\mathbf{c}$ a separable cost function, and $\mathbf{f}^*$ a corresponding equilibrium. For every route $r$, the following conditions are equivalent:

(i) The route $r$ is a minimal-cost route (i.e., its cost equals the equilibrium cost $c(\mathbf{f}^*)$).

(ii) Every edge in $r$ is in some minimal-cost route.

*Proof.* Let $r$ be a route satisfying (ii). It is to be shown that $r$ satisfies (i), or equivalently, that for some minimal cost route $q$, $c_r(\mathbf{f}^*) = c_q(\mathbf{f}^*)$. Clearly, it suffices to prove the following (stronger) claim.

CLAIM. For every minimal-cost route $q$ and every vertex $v$ common to $r$ and $q$, $c_{r_{ov}}(\mathbf{f}^*) = c_{q_{ov}}(\mathbf{f}^*)$.

For $v = o$, the claim is trivial. Proceeding by induction on the length of $r_{ov}$, suppose that $v \neq o$ and that, for every minimal-cost route $p$ that includes the vertex $u$ immediately preceding $v$ in $r$, $c_{r_{ou}}(\mathbf{f}^*) = c_{p_{ou}}(\mathbf{f}^*)$. Since $r$ satisfies (ii), there is a minimal-cost route $p$ that includes the edge $e$ in $r$ joining $u$ and $v$. Thus, $r_{uv} = p_{uv}$, and hence $c_{r_{uv}}(\mathbf{f}^*) = c_{p_{uv}}(\mathbf{f}^*)$. By separability of $\mathbf{c}$ and the induction hypothesis, this implies that $c_{r_{ov}}(\mathbf{f}^*) = c_{p_{ov}}(\mathbf{f}^*)$. To complete the proof of the claim, it remains to show that, for every pair of minimal-cost routes $p$ and $q$ with a common vertex $v$, $c_{p_{ov}}(\mathbf{f}^*) = c_{q_{ov}}(\mathbf{f}^*)$. By symmetry, it suffices to show that $c_{p_{ov}}(\mathbf{f}^*) \geq c_{q_{ov}}(\mathbf{f}^*)$, or equivalently $c_{p_{ov}}(\mathbf{f}^*) + c_{q_{vd}}(\mathbf{f}^*) \geq c_q(\mathbf{f}^*)$. The sum on the left-hand side of the latter inequality is the cost of the route $p_{ov} + q_{vd}$. (By (ii) in Proposition 1, $p_{ov}$ and $q_{vd}$ do not have common vertices other than $v$.) This cost cannot be less than that of the minimal-cost route $q$, which proves the above inequality. ■

**Lemma 4.** Let $G$ be a series-parallel network, $\hat{\mathbf{c}}$ and $\tilde{\mathbf{c}}$ separable cost functions such that $\hat{c}_r(\mathbf{f}) \geq \tilde{c}_r(\mathbf{f})$ for all routes r and flow vectors $\mathbf{f}$, and $\hat{\mathbf{f}}$ and $\tilde{\mathbf{f}}$ corresponding equilibria such that the total origin–destination flows satisfy $\hat{f}_o \geq \tilde{f}_o$. Then, the equilibrium costs satisfy $\hat{\mathbf{c}}(\hat{\mathbf{f}}) \geq \tilde{\mathbf{c}}(\tilde{\mathbf{f}})$.

*Proof.* If $\hat{\mathbf{f}} = \tilde{\mathbf{f}}$, then the inequality $\hat{\mathbf{c}}(\hat{\mathbf{f}}) \geq \tilde{\mathbf{c}}(\tilde{\mathbf{f}})$ follows trivially from the assumption that $\hat{c}_r(\mathbf{f}) \geq \tilde{c}_r(\mathbf{f})$ for all routes $r$ and flow vectors $\mathbf{f}$. Suppose, then, that $\hat{\mathbf{f}} \neq \tilde{\mathbf{f}}$, which clearly implies that $\hat{f}_o > 0$. By Lemma 2, there is some route $r$ such that, for all sections $s$ of $r$ of length zero or one, $\hat{f}_s \geq \tilde{f}_s$ and $\hat{f}_s > 0$. Since $G$ is series-parallel and $\hat{\mathbf{c}}$ is separable, the weak inequalities ($\hat{f}_s \geq \tilde{f}_s$) imply that $\hat{c}_r(\hat{\mathbf{f}}) \geq \hat{c}_r(\tilde{\mathbf{f}})$. The strict inequalities ($\hat{f}_s > 0$) imply that every edge $e$ in $r$ is in some route $q$ with $\hat{f}_q > 0$. By the equilibrium condition (2) and Lemma 3, this implies that $\hat{c}_r(\hat{\mathbf{f}}) = \hat{c}(\hat{\mathbf{f}})$. Hence, $\hat{c}(\hat{\mathbf{f}}) \geq \hat{c}_r(\tilde{\mathbf{f}})$. By the assumption concerning the cost functions, $\hat{c}_r(\tilde{\mathbf{f}}) \geq \tilde{c}_r(\tilde{\mathbf{f}})$, and by definition of equilibrium cost, $\tilde{c}_r(\tilde{\mathbf{f}}) \geq \tilde{c}(\tilde{\mathbf{f}})$. This proves that $\hat{\mathbf{c}}(\hat{\mathbf{f}}) \geq \tilde{\mathbf{c}}(\tilde{\mathbf{f}})$. ■

*Proof of Theorem 1.* Lemma 4 shows that, in a series-parallel network, Braess's paradox does not occur. It remains to prove the converse. By Proposition 1, if a network is not series-parallel, then the Wheatstone network (Figure 1) is embedded in it. As shown, in that particular network, Braess's paradox does occur. The same is true for any network in which the Wheatstone network is embedded. This can easily be seen by considering the following cost-assignment rules for the edges created by the three operations that define embedding. For each of the two edges created by subdividing an existing edge, the cost is one-half that of the original edge. For a new edge joining two existing vertices, the cost is very high (effectively, infinite). For the edge $e$ created by extending the origin or the destination, the cost is an arbitrary increasing function of the flow on $e$. ■

The following lemma is used in the proof of Theorem 2.

**Lemma 5.** A series-parallel network has linearly independent routes if and only if, for every pair of distinct flow vectors $\hat{\mathbf{f}}$ and $\tilde{\mathbf{f}}$ with $\hat{f}_o \geq \tilde{f}_o$, there is some route r such that $\hat{f}_r > \tilde{f}_r$, and $\hat{f}_s \geq \tilde{f}_s$ for all sections $s$ of $r$.

*Proof.* For either of the series-parallel networks in Figure 2, consider the following two flow vectors $\hat{\mathbf{f}}$ and $\tilde{\mathbf{f}}$. For the route $r_1$ passing through the edges $e_1$ and $e_3$, $\hat{f}_{r_1} = 0$ and $\tilde{f}_{r_1} = 2$. For the route $r_2$ passing through $e_2$ and $e_4$, $\hat{f}_{r_2} = 0$ and $\tilde{f}_{r_2} = 0$. For the route $r_3$ passing through $e_2$ and $e_3$, $\hat{f}_{r_3} = 1$ and $\tilde{f}_{r_3} = 0$. For the route $r_4$ passing through $e_1$ and $e_4$, $\hat{f}_{r_4} = 1$ and $\tilde{f}_{r_4} = 0$. The only routes $r$ with $\hat{f}_r > \tilde{f}_r$ are $r_3$ and $r_4$. However, $r_3$ includes $e_3$ and $r_4$ includes $e_1$, and on these two edges, the flow in $\hat{\mathbf{f}}$ is *less* than in $\tilde{\mathbf{f}}$. In view of Propositions 1 and 4, this example shows that for *every* series-parallel network in which the routes are not linearly independent, there is a pair of distinct flow vectors $\hat{\mathbf{f}}$ and $\tilde{\mathbf{f}}$ with equal total origin–destination flows such that a route *r* as above does not exist.

The converse, that the condition in the lemma holds for every network $G$ with linearly independent routes, will be proved by induction on the number of edges. If $G$ only has one edge, the condition clearly holds. Suppose that $G$ has two or more edges. The induction hypothesis is that the condition holds for any network with linearly independent routes and a smaller number of edges than $G$. Let $\hat{\mathbf{f}}$ and $\tilde{\mathbf{f}}$ be two distinct flow vectors for $G$ with $\hat{f}_o \geq \tilde{f}_o$. By Proposition 5, $G$ is the result of connecting two series-parallel networks, $G'$ and $G''$, in series or in parallel. Moreover, in the former case, one of these, say $G''$, has only one edge, $e$. In this case, by similar arguments to those in the proof of Lemma 2, the induction hypothesis implies that there is some route $r'$ in $G'$ such that $\hat{f}_{r'} > \tilde{f}_{r'}$, and $\hat{f}_s \geq \tilde{f}_s$ for all sections $s$ of $r'$. The route $r$ in $G$ obtained by appending $e$ and $d$ to $r'$ has the same properties. This is because every section $s$ of $r$ is: (i) a section of $r'$, (ii) the zero-length path consisting of $d$ alone, or (iii) the result of appending $e$ and $d$ to some section $s'$ of $r'$—in which case the flows on $s$ and $s'$ are always equal. This completes the proof for the case in which $G$ is the result of connecting $G'$ and $G''$ in series. If $G$ is the result of connecting the two networks in parallel, then, again by similar arguments to those in the proof of Lemma 2, it follows from the induction hypothesis that there is some route $r$ in $G'$ or $G''$ such that $\hat{f}_r > \tilde{f}_r$, and $\hat{f}_s \geq \tilde{f}_s$ for all sections $s$ of $r$. Since $r$ is also a route in $G$, this completes the proof. ■

*Proof of Theorem 2.* Suppose that $G$ is a network with linearly independent routes. Let $\mathbf{c}$ be a cost function and $\mathbf{f}^*$ a corresponding equilibrium, with equilibrium cost $c(\mathbf{f}^*)$. If $\mathbf{f}$ is another flow vector with the same total origin–destination flow as $\mathbf{f}^*$, then by Lemma 5, there is some route $r$ such that $f_r > f_r^*$, and $f_s \geq f_s^*$ for all sections $s$ of $r$. Since $G$ has linearly independent routes, there is some section $s$ of $r$ of length one such that $f_s = f_r > f_r^* = f_s^*$. It follows that $c_r(\mathbf{f}) \geq c_r(\mathbf{f}^*)$, and if $\mathbf{c}$ is increasing, then $c_r(\mathbf{f}) > c_r(\mathbf{f}^*)$. Since $f_r > 0$ and, by definition of equilibrium cost, $c_r(\mathbf{f}^*) \geq c(\mathbf{f}^*)$, this proves that $\mathbf{f}^*$ is weakly Pareto efficient. Moreover, if $\mathbf{c}$ is increasing, then $\mathbf{f}^*$ is Pareto efficient.

Suppose now that the routes in $G$ are not linearly independent. By Proposition 4, one of the networks in Figure 1 and Figure 2 is embedded in $G$. As shown in the introduction, for each of these three networks, there is an increasing, nonnegative separable cost function and a corresponding equilibrium that is not even weakly Pareto efficient. The same is true for every network in which one of these networks is embedded. The proof of this is based on the same arguments as in the proof of Theorem 1. ■

*Proof of Theorem 3.* It is not difficult to see that condition (i) implies (iii). Therefore, it suffices to show that if $G$ has linearly independent routes, then (i) and (ii) hold, and if the routes are not linearly independent, then (ii) and (iii) do not hold.

Suppose that $G$ has linearly independent routes. For a given assignment of cost functions $\mathbf{c}^i$, let $\sigma$ be a Nash equilibrium and $\tau$ another strategy profile. If $\mathbf{f}(\sigma) = \mathbf{f}(\tau)$, then it follows from the equilibrium condition (3) that $c^i_{\tau(i)}(\mathbf{f}(\tau)) \geq c^i_{\sigma(i)}(\mathbf{f}(\sigma))$ for all users $i$, and equality holds if $\tau(i) = \sigma(i)$. Hence, condition (H) holds. Suppose, then, that $\mathbf{f}(\tau) \neq \mathbf{f}(\sigma)$. By Lemma 5, there is some route $r$ such that $f_s(\tau) \geq f_s(\sigma)$ for all sections $s$ of $r$, with strict inequality for $r$ itself and, since the routes are linearly independent, also for some section $s$ of $r$ of length one. It follows that, for every user $i$, $c^i_r(\mathbf{f}(\tau)) \geq c^i_r(\mathbf{f}(\sigma)) \geq \min_{q \in \mathcal{R}} c^i_q(\mathbf{f}(\sigma)) = c^i_{\sigma(i)}(\mathbf{f}(\sigma))$, and if $i$'s cost function is increasing, then the first inequality is strict. Since $f_r(\tau) > f_r(\sigma) \geq 0$, there is a nonempty (indeed, positive-measure) set of users $i$ with $\tau(i) = r$. This proves that the Nash equilibrium $\sigma$ is weakly Pareto efficient. Moreover, there is a nonempty set of users $i$ with $\tau(i) = r \neq \sigma(i)$. If any of these users $i$ has an increasing cost function, then, as shown above, $c^i_{\tau(i)}(\mathbf{f}(\tau)) > c^i_{\sigma(i)}(\mathbf{f}(\sigma))$, and hence (H) holds. This proves that if *all* the users have increasing cost functions, then the Nash equilibrium $\sigma$ is hyper-efficient.

Suppose now that the routes in $G$ are not linearly independent. It follows from Theorem 2 that condition (ii) does not hold. It remains to show that Braess's paradox with non-identical users occurs in $G$. By Proposition 4, one of the networks in Figure 1 and Figure 2 is embedded in $G$. If this is the one in Figure 1, then it follows from Theorem 1 that Braess's paradox (with identical users) occurs in $G$. Suppose, then, that the embedded network is one of those in Figure 2. As shown above, Braess's paradox with non-identical users occurs in the network in Figure 2(a). Essentially the same example shows that it also occurs in the network in Figure 2(b). This implies that the paradox also occurs in *G*. The argument is similar to that in the proof of Theorem 1. ■

## References


Arnott, R., Small, K., 1994. The economics of traffic congestion. *American Scientist* **82**, 446–455.

Beckmann, M., McGuire, C. B., Winsten, C. B., 1956. Studies in the Economics of Transportation. Yale University Press, New Haven, CT.

Bell, M. G. H., Iida, Y., 1997. Transportation Network Analysis. Wiley, Chichester, UK.

Braess, D., 1968. Über ein paradoxon aus der verkehrsplanung. *Unternehmensforschung* **12**, 258–268.

Calvert, B., Keady, G., 1993. Braess's paradox and power-law nonlinearities in networks. *Journal of the Australian Mathematical Society, Series B* **35**, 1–22.

Cohen, J. E., Horowitz, P., 1991. Paradoxical behaviour of mechanical and electrical networks. *Nature* **352**, 699–701.

Cohen, J. E., Jeffries, C., 1997. Congestion resulting from increased capacity in single-server queuing networks. *IEEE/ACM Transactions on Networking* **5**, 305–310.

Dafermos, S., Nagurney, A., 1984. On some traffic equilibrium theory paradoxes. *Transportation Research* **18B**, 101–110.

Diestel, R., 2000. Graph Theory, Second Edition. Graduate Texts in Mathematics, Vol. 173. Springer-Verlag, New York.

Duffin, R. J., 1965. Topology of series-parallel networks. *Journal of Mathematical Analysis and Applications* **10**, 303–318.

Frank, M., 1981. The Braess paradox. *Mathematical Programming* **20**, 283–302.

Harary, F., 1969. Graph Theory. Addison-Wesley, Reading, MA.

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

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

Lin, H., Roughgarden, T., Tardos, É., 2004. A stronger bound on Braess's Paradox. *Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms*, 340–341.

Milchtaich, I., 1996. Congestion games with player-specific payoff functions. *Games and Economic Behavior* **13**, 111–124.

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

Milchtaich, I., 2005. Topological conditions for uniqueness of equilibrium in networks. *Mathematics of Operations Research* **30**, 225–244.

Murchland, J. D., 1970. Braess's paradox of traffic flow. *Transportation Research* **4**, 391–394.

Nagurney, A., 1999. Network Economics: A Variational Inequality Approach, Second Edition. Kluwer, Boston.

Newell, G. F., 1980. Traffic Flow on Transportation Networks. MIT Press, Cambridge, MA.

Riordan, J., Shannon, C. E., 1942. The number of two-terminal series-parallel networks. *Journal of Mathematics and Physics* **21**, 83–93.

Roughgarden, T., 2003. The price of anarchy is independent of the network topology. *Journal of Computer and System Sciences* **67**, 341–364.

Roughgarden, T., 2004. On the severity of Braess's paradox: Designing networks for selfish users is hard. Forthcoming in the *Journal of Computer and System Sciences*. Preliminary version in *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, 472–481.

Roughgarden, T., Tardos, É., 2004. Bounding the inefficiency of equilibria in nonatomic congestion games. *Games and Economic Behavior* **47**, 389–403.

Sheffi, Y., 1985. Urban Transportation Networks. Prentice-Hall, Englewood Cliffs, NJ.

Steinberg, R., Zangwill, W. I., 1983. The prevalence of the Braess' paradox. *Transportation Science* **17**, 301–318.

Voorneveld, M., Borm, P., van Megen, F., Tijs, S., Facchini, G., 1999. Congestion games and potentials reconsidered. *International Game Theory Review* **1**, 283–299.

Wardrop, J. G., 1952. Some theoretical aspects of road traffic research. *Proceedings of the Institute of Civil Engineers,* Part II **1**, 325–378.