Discussion Paper No. 1198

An Evolutionary Approach to Congestion

by
William H. Sandholm*
MEDS-KGSM
Northwestern University
Evanston, IL 60208, USA
e-mail: whs@nwu.edu

First Version: April 30, 1997
This Version: December 7, 1997

* I thank Eddie Dekel, Nolan Miller, Jeroen Swinkels, and a seminar audience at Washington University for their comments. Financial support from a State Farm Companies Foundation Dissertation Fellowship is gratefully acknowledged.

Abstract

Using techniques from evolutionary game theory, we analyze potential games with continuous player sets, a class of games which includes a general model of network congestion as a special case. We concisely characterize both the complete set of Nash equilibria and the set of equilibria which are robust against small disturbances of aggregate behavior. We provide a strong evolutionary justification of why equilibria must arise. We characterize situations in which stable equilibria are socially efficient, and show that in such cases, evolution always increases aggregate efficiency. Applying these results, we construct a parameterized class of congestion tolls under which evolution yields socially optimal play. Finally, we characterize potential games with continuous player sets by establishing that a generalization of these games is precisely the limiting version of finite player potential games (Monderer and Shapley (1996)) which satisfy an anonymity condition.

JEL Classification Numbers: C61, C72, C73, D62, R41, R48.

1. Introduction

Congestion is the bane of motorized society. Each year, automotive traffic squanders six billion hours of drivers' time; a conservative estimate places the dollar value of this loss at $48 billion.[1] While some traffic is unavoidable, delays are often far worse than necessary. An omnipotent social planner could often vastly improve aggregate welfare by enforcing an efficient pattern of route selection among drivers. Unfortunately, because of externalities, such a pattern typically fails to arise on its own: driving on congested highways is a large non-cooperative game; in choosing which route to take, drivers neither consider nor bear the full social costs of the congestion they create.

Congestion is also a major problem in communication networks. For example, when Internet users visit web sites, their utilization of network resources imposes externalities on other users. Because individuals do not weigh the effects of the congestion their actions create, inefficiencies arise.

In this paper, we introduce a new game-theoretic approach to studying congestion which is based on an externality symmetry property which all congestion problems share. We study *potential games with continuous player sets*, the complete class of games satisfying the externality symmetry condition. We characterize the sets of all Nash equilibria and of Nash equilibria robust against small disturbances of behavior, and provide a strong evolutionary justification of why equilibria must arise. We characterize situations in which evolution increases social welfare, and establish that in these cases, the equilibria which are stable under evolutionary dynamics are precisely those which are socially efficient. Lastly, we use these efficiency results to construct a class of congestion tolls under which evolution yields socially optimal play.

We analyze network congestion using *continuous congestion games*. Congestion games were introduced by Rosenthal (1973), who proved the existence of pure strategy equilibria in a finite player setting. Unfortunately, in most models of congestion, the number of players is very large, so the discrete model can prove quite cumbersome. This complexity is mitigated by the fact that in many large population settings, players need not concern themselves with each individual opponent's behavior, as their payoffs only depend on opponents' aggregate behavior. We therefore model congestion using continuous player sets and payoffs

[1] See Arnott and Small (1994, p. 446).

which only depend on the distribution of opponents' actions. In addition to yielding a more tractable model, the continuum assumption allows us to use the tools of marginal analysis. This enables us to investigate the efficiency of aggregate behavior, which seems quite difficult to analyze in a discrete setting.

After specifying our game theoretic model of congestion, we must specify what constitutes a reasonable prediction of play. The condition that is usually imposed is Nash equilibrium. While this condition is standard, using it as a basis for prediction requires a theoretical justification.

One approach to this issue establishes conditions under which fully rational players will coordinate on an equilibrium.[2] In general, the demands that this eductive approach makes on players are enormous: at a minimum, players must have thorough knowledge of opponents' intentions and must be able to respond to them optimally. These requirements, which at best are rarely satisfied, are especially questionable when applied to games played by large populations. Hence, in our current context, an eductive justification of equilibrium is inappropriate.

Evolution provides another basis for justifying equilibrium. Rather than imposing strong assumptions about players' knowledge and abilities, in following the evolutionary paradigm one assumes that players respond myopically to opportunities to improve their payoffs. In contrast to eduction, myopic adjustment makes only modest demands on players' abilities. Awareness of the available options and the ability to choose one which currently seems promising are sufficient; these demands seem almost minimal conditions defining an economic agent.

Unfortunately, this approach often fails to justify equilibrium play. Even in simple normal form games, evolutionary dynamics can converge to closed orbits far from any Nash equilibrium. In certain games, even chaotic behavior is possible. Rarely is global convergence to equilibrium assured.[3] These problems present major difficulties for analyzing large population games: in general, we cannot guarantee that a population's behavior will ever even roughly approximate equilibrium play.

---

2 For models following this approach, see Aumann and Brandenburger (1995) and the references therein.

3 Numerous examples of limit cycles under the replicator dynamics and under the best response dynamics can be found in Weibull (1995) and Gaunersdorfer and Hofbauer (1995), respectively. Cowan (1992) analyzes an example in which fictitious play of a 4 x 4 game leads to chaotic behavior.

What prevents convergence? In evolutionary models, players switch to strategies which increase their payoffs. At first glance, constant adjustment toward better performing strategies might appear to drive the population towards an equilibrium state, where all payoff improvement opportunities are exhausted. Unfortunately, this reasoning is flawed. At each population state, payoffs induce a set of "uphill" directions in which evolution might lead. However, as the state changes, so to do the payoffs accruing to each strategy, and consequently the uphill directions. A wide variety of dynamic behaviors are possible depending on how the landscape adjusts. Indeed, cycling is a common result: in this case, the payoff landscapes behave like a staircase in an Escher lithograph, forming a cycle of inclines which the population perpetually ascends.[4]

Shifting landscapes are a consequence of externalities: when players switch strategies, they not only cause discrete jumps in their own payoffs, but also marginally shift the overall strategy distribution, altering the payoffs to all strategies. It is therefore noteworthy that the externalities which arise in congestion games have a very special structure. As an illustration, consider a traffic flow over a network of streets, and single out two complete routes (i.e., sets of successive streets), A and B. If a newcomer to the city chooses to commute via route A, the effect he has on a driver following route B only depends on the streets routes A and B have in common. Hence, the marginal impact of this new commuter on the route B drivers is the sum of the marginal increases in delay on every common street. But a new route B driver has precisely the same set of streets in common with the route A drivers. Therefore, the marginal effect of a new route A drivers on current route B drivers is precisely the same as the marginal effect of a new route B driver on current route A drivers. This property of congestion games we call *externality symmetry*.

Externality symmetry places a critical restriction on the payoff landscapes a game can induce. While it certainly remains true that the uphill direction varies from state to state, under externality symmetry this variation must occur in a topographically consistent fashion. In particular, one can always find a *potential function*: a real-valued function defined on the state space whose uphill directions at each state represent the reasonable courses of evolution. Since strategy revision always drives the population uphill, closed orbits and chaotic behavior are

[4] The shifting landscapes metaphor was introduced by Binmore and Samuelson (1997).

impossible. Thus, the existence of a potential function provides an intuitive basis for justifying equilibrium prediction.

Before discussing this, we consider how the potential function can be used to characterize equilibrium. In a game with continuous player sets, Nash equilibrium requires that within each role in the game, the payoffs to all utilized strategies are equal and are no less than those to unutilized strategies. For example, at interior equilibria, the payoffs to all strategies belonging to the same population are equal. In this case, there is no uphill direction in which evolution can lead. Hence, such a state must be a critical point of the potential function. More generally, we show that the Nash equilibria of a potential game are precisely the points which satisfy the Kuhn-Tucker first order conditions for a maximizer of the potential function. All local maximizers are therefore Nash equilibria. When congestion is bad, these are the only equilibria: when congestion costs are increasing in utilization, the potential function is concave, and equilibrium is unique. However, congestion may be beneficial; congestion externalities may be positive. In such cases, the potential function will not be concave, and equilibria which do not maximize potential will typically exist: for example, every interior local minimizer of potential satisfies the Kuhn-Tucker conditions for a maximum and is therefore Nash. However, we shall see that evolution selects against states which do not maximize potential.

Our main application of the potential function uses it to analyze continuous time evolutionary dynamics. Rather than specify a particular model of evolution, we instead consider a class of dynamics which satisfy a weak payoff monotonicity condition which we call *net monotonicity*. This condition requires that in an average sense, the growth rates of better performing strategies are higher than the growth rates of those which perform poorly. Any reasonable evolutionary process, in particular the replicator dynamics (Taylor and Jonker (1978)) and the best response dynamics (Gilboa and Matsui (1991), Matsui (1992), Hofbauer (1995)) satisfy this condition. In fact, this condition is weaker than nearly any other condition proposed in the evolutionary literature.[5]

Net monotonicity implies that when it is not at rest, the evolutionary process ascends the potential function. Consequently, once the population is near a peak (i.e., a local maximizer) of potential, it must always remain nearby. We are

---

[5] Nachbar (1990), Friedman (1991), Samuelson and Zhang (1992), Swinkels (1993), and Ritzberger and Weibull (1995) study classes of dynamics whose members satisfy some basic evolutionary desiderata. Of these, the conditions considered by Friedman (1991) and by Swinkels (1993) are both the weakest and the closest to the net monotonicity condition we consider here.

therefore able to identify the states which are dynamically stable under net monotone dynamics with the local maximizers of potential. Since all such states are Nash equilibria, only Nash equilibria can be dynamically stable. We also establish the converse result: all Nash equilibria which do not locally maximize potential must be susceptible to some perturbation of aggregate behavior. Local maximizers of potential are therefore the only sensible equilibrium predictions.

While local stability results provide strong reasons to believe that an equilibrium that is reached will persist, they do not indicate whether one should expect equilibrium to be reached at all. To establish this requires a global convergence result which characterizes the ultimate behavior of the population for all possible initial states. Global convergence results can rarely be proved. Nevertheless, without historical data about a population's behavior, only through such a result can an evolutionary analysis fully justify the prediction of equilibrium play.

Global convergence to equilibrium almost never occurs under the replicator dynamics. The problem is extinction: once a strategy is absent from the population, it never regenerates. All strategy profiles consisting of a single pure strategy are therefore rest points; hence, many rest points are not Nash equilibria. However, in most economic contexts, it is reasonable to expect the presence of a profitable deviation to induce economic agents to change strategies. We introduce a condition on evolutionary dynamics which we call *non-extinction* to reflect this assertion. It requires that all rest points of the dynamics are Nash equilibria. This formalizes the evolutionary desideratum noted above: if a state is not an equilibrium, a profitable deviation exists, and we should expect some enterprising players to avail themselves of it.

With the help of this condition, we can state our global convergence result. We prove that in any potential game, under any evolutionary dynamics satisfying net monotonicity and non-extinction, all evolutionary solution paths converge to Nash equilibria. Regardless of the initial behavior of the population, the end result is equilibrium; closed orbits and chaotic behavior are both impossible. Therefore, in any congestion game, evolution completely justifies equilibrium prediction.

While we can guarantee that equilibrium behavior will eventually emerge, the equilibrium need not exhibit desirable welfare properties. Of course, this is generally true of Nash equilibrium: the dominant strategy equilibrium of the prisoners' dilemma is the archetypal example. One might hope that the properties which lead to convergence in congestion games might also promote efficiency. Unfortunately, externalities often obstruct efficiency. When players choose strategies, they consider only their own payoffs, ignoring the effects of their actions on society at large. For example, roads that are relatively more sensitive to congestion tend to be overused in equilibrium because each driver ignores the marginal increase in congestion that he creates.

This discussion suggests that road networks in which streets are equally sensitive to utilization might have better prospects for efficiency. With this motivation, we consider *isoelastic* congestion games: congestion games in which all streets exhibit isoelastic utilization costs, each with the same elasticity coefficient. More broadly, we define the class of isoelastic potential games, which include isoelastic congestion games as a special case. We show that isoelasticity implies that player's payoffs to each strategy are directly proportional to the marginal social costs of the use of that strategy. Because it ensures that the individual and the society agree on the relative values of his available actions, isoelasticity has strong implications for aggregate efficiency.

The key lemma behind our efficiency results establish that in isoelastic potential games, the potential function is proportional to average payoffs. In conjunction with previous developments, this lemma provides the foundation for our main welfare results. We show that under any net monotone dynamics, average payoffs increase along every evolutionary path. In other words, myopic adjustment increases social welfare. Furthermore, the set of locally efficient states is identical to the set of dynamically stable states. Any state which is robust against small disturbances in aggregate behavior must yield locally optimal aggregate payoffs.

A general global welfare maximization result is clearly impossible. To see why, consider games in which congestion externalities are positive, representing benefits to social coordination. Such games typically possess multiple stable equilibria, each of which may induce different average payoffs. However, when congestion is detrimental to payoffs, it might be hoped that a global efficiency result is possible. We establish that in isoelastic congestion games with non-negative costs, there is a unique connected set of equilibria, and that all equilibria are both globally efficient and globally stable.

While our efficiency results may give reason for optimism (or at least resignation) in the cases in which they apply, the isoelasticity condition which they require is demanding. In most examples, this condition is not met; in such cases, evolution usually fails to generate welfare maximizing outcomes. Inefficiencies arise because drivers overburden the roads which are the most sensitive to use. One might hope that by introducing a system of congestion tolls, a mechanism designer could force drivers to internalize the social costs of their actions, and thereby induce them to select routes in a socially optimal manner.

We construct a parameterized class of congestion tolls which we call *proportional to marginal cost (PMC) tolls*. The tolls are implemented facility by facility, and so may be extracted as the facilities are used.[6] Under every PMC toll, evolution leads to efficient play. Marginal cost pricing, which sets individual costs equal to marginal social costs, occurs as a special case. However, many other noteworthy tolling systems are available, including one which minimizes the size of the tolls while keeping them non-negative, and another which achieves budget balance. Indeed, fixing an efficient state, there is a PMC toll which yields any desired level of revenue. Therefore, aggregate efficiency can be achieved regardless of the budgeting requirements.

Realistically, setting very high tolls will reduce the number of drivers, preventing a social planner from generating arbitrarily large tolling surpluses. More generally, it is often natural to expect the sizes of the populations playing a game to depend on the payoffs the game provides. We therefore extend our model by introducing elastic demand. We show that versions of all of our previous results on equilibrium, evolution, and efficiency continue to hold in this setting, and that different PMC tolls all generate efficient outcomes but lead to different levels of aggregate demand. Thus, by choosing the appropriate PMC tolling scheme, a social planner can strike a balance between the ills of congestion and the dampening effect of tolls on economic activity.

There are many settings in which network allocation issues arise not as a game, but as a decision problem. Consider, for example, a problem faced by a monopoly which runs a utility: to transport goods to all customers at the lowest possible cost. Finding the optimal routing choice often requires solving an enormous mathematical program. Our efficiency results have strong implications for solving such problems. We show how by imposing an internal PMC tolling scheme, a monopolist can implement a decentralized method of establishing and maintaining efficient route allocations.

Our work is closely related to that of Monderer and Shapley (1996), who study potential games with a finite number of players. In these games, as in those studied here, all players' incentives can be captured by a single potential function defined on

---

6 Many cities now use electronic toll collecting systems which could implement such tolling schemes. For discussion of these collection systems, see Gomez-Ibanez and Small (1994).

the set of strategy profiles. Nash equilibria correspond to local maximizers of potential; since myopic better responses increase potential, evolution leads to equilibrium play. Lastly, Monderer and Shapley (1996) establish an equivalence between finite player potential games and the finite player congestion games introduced by Rosenthal (1973).

At least heuristically, the connections between the games studied here and those considered by Monderer and Shapley (1996) are clear. In both cases, all players' incentives are captured by a potential function, guaranteeing convergence to equilibrium. However, examining only the definitions of each class of games, the formal connections between the two are not obvious.

We establish these connections in the penultimate section of the paper. To do so, we must consider infinite player potential games which are augmented by incentive irrelevant payoff shifts. Evolution and equilibrium in the resulting games, termed *quasipotential games*, can still be characterized using potential functions. We also restrict attention to finite-player potential games which admit an anonymous representation, which we call *anonymous finite-player potential games* (hereafter, *AFP games*). To state our limit result, we first define a notion of convergence: we say that a sequence of AFP games converges if the payoff functions and (rescaled versions of) the potential functions converge as the population size grows to infinity. We then prove that the limit of any convergent sequence of AFP games must be a quasipotential game. Conversely, we establish that any quasipotential game can be approximated arbitrarily well by AFP games, formally justifying the use of the continuum of players model to approximate large, finite player games. Using our convergence theorem, we characterize AFP games in which, to a first approximation, evolution yields efficiency.

An early use of a continuous potential function can be found in the work of Beckmann, McGuire, and Winsten (1956). These authors consider a model of traffic which is equivalent to a special case of our elastic demand congestion model. They use a potential function to characterize equilibrium and to establish conditions under which equilibrium is unique. We extend their work by both broadening the class of games to which potential function arguments can be applied and by establishing results on evolutionary dynamics, social efficiency and optimal pricing schemes.[7]

---

[7] Friedman (1996, 1997) considers learning in a class of games with continuous player sets. Building on the earlier paper, Friedman (1997) studies non-atomic externality games. In these games, the aggregate play of each population generates a vector of externalities, with one component

Our conclusions concerning efficiency build on a result from biological evolutionary game theory known as the Fundamental Theorem of Natural Selection. This result, first suggested by Fisher (1930), states that in doubly symmetric normal form games played by a single population of players, average payoffs increase monotonically along every solution path of the replicator dynamics. The symmetry assumption requires that all pairs of entries in the payoff bimatrix are equal. This condition arises naturally in a basic model from population genetics. In this model, individual genes struggle to survive and reproduce, but because of sexual reproduction always compete as members of pairs. The partnership assumption captures the fact that the survival and reproduction of each gene in the depends on the fitness of the pair's host.[8] Our model generalizes the framework of the Fundamental Theorem of Natural Selection in three ways: we allow multiple populations, non-linear payoffs, and a broad range of evolutionary dynamics.

Section 2 introduces congestion games and doubly symmetric games, providing two classes of examples to which our results apply. Section 3 defines potential games and presents the equilibrium characterization result. Sections 4 and 5 investigate evolutionary dynamics and social efficiency, while Section 6 describes congestion tolls which induce efficient play. Section 7 considers potential games with elastic demand. Section 8 explains the formal connections between our work and that of Monderer and Shapley (1996). Section 9 concludes.

2. Examples

2.1 Congestion Games

Continuous congestion games are a simple and tractable framework for studying the utilization of resources arranged in a network. Flows of automobile traffic and electronic data are natural applications. Rosenthal (1973) introduces congestion games with a finite number of players and proves that they possess pure strategy Nash equilibria. Unfortunately, many applications of congestion games involve

---

corresponding to each population. Each (non-identical) player's preferences depend only upon his strategy choice and the externality vector. Using submodularity arguments, Friedman (1997) establishes that convergence to equilibrium play is guaranteed under a number of discrete learning procedures, including adaptive and sophisticated learning (Milgrom and Roberts (1991)) and correlated learning (Foster and Vohra (1996)).

[8] Extending the basic result, Akin and Hofbauer (1982) show that under the assumptions of the fundamental theorem, each solution path converges to a unique limit point. For further discussion of this biological model, see Hofbauer and Sigmund (1988, Chapters 3 and 23).

very large numbers of players, making the discrete formulation difficult to manage. In this section, we introduce congestion games with continuous player sets. We provide examples of these games which serve to illustrate the results which we develop in the remainder of the paper.

Consider a collection of towns connected by a network of streets. We associate each of p pairs of home and work locations with a group of commuters who must travel between them. Each player chooses a route (i.e., a subset of the streets) connecting home to work; his driving time depends upon the traffic on the streets he has chosen.

With this example in mind, we define continuous congestion games. A *congestion model* is a collection $\{P, \{m^p\}_{p\in P}, \{S^p\}_{p\in P}, \Phi, \{\Phi_i^p\}_{i\in S^p, p\in P}, \{c_\phi\}_{\phi\in\Phi}\}$. P is a set of one or more populations, which in the commuting model correspond to the home and work location pairs. The mass of population $p \in P$ is denoted m_p, and $m = \sum_p m^p$ equals the total population mass. The finite set Φ contains the *facilities* which are the building blocks of the players' strategies. Each strategy $i \in S^p$, $p \in P$ entails the use of some subset of the facilities $\Phi_i^p \subset \Phi$. We assume that there are no irrelevant facilities: $\Phi = \bigcup_{p\in P} \bigcup_{i\in S} \Phi_i^p$. Let $\rho^p(\phi) = \{i \in S^p\colon \phi \in \Phi_i^p\}$ denote the set of population p strategies which require facility ϕ. A typical *strategy distribution* is $x = (x_1^1, \dots, x_{\#S^1}^1, x_1^2, \dots, x_{\#S^r}^r)$, where the upper subscript refers to the population and the lower subscript to the strategy. The *utilization* of facility $\phi \in \Phi$ is the total mass of the players whose strategies use that facility:

$$u_\phi(x) = \sum_{p\in S^p} \sum_{i\in\rho^p(\phi)} x_i^p .$$

When a player selects a strategy, he bears the costs of each facility which the strategy requires. The *facility costs*, c_ϕ: $[0, m] \to \mathbf{R}$, are continuously differentiable functions (with one sided derivatives at the endpoints) which report the penalties accruing to users of a facility as a function of its utilization.[9] The *congestion game* derived from a congestion model is identified with its payoff functions:

$$F_i^p(x) = -\sum_{\phi\in\Phi_i^p} c_\phi(u_\phi(x)).$$

---

[9] Later, the differentiability assumption will be dropped.

Since payoffs to the strategies in a congestion game are sums of facility payoffs, the payoffs to any pair of strategies are bound together by the facilities which both use. In the commuting model, the costs of any pair of routes both incorporate the delays on all streets which the routes share. Hence, if we increase the proportion of players from population p using route i, this affects the players taking route $j \in S^q$ through increased traffic on streets in $\Phi_i^p \cap \Phi_j^q$. The marginal effect of this increase is $\frac{\partial F_j^q}{\partial x_i^p} \equiv -\sum_{\phi \in \Phi_i^p \cap \Phi_j^q} c_\phi'$. An increase in the use of route j has an the same marginal effect on route i drivers; hence, $\frac{\partial F_i^p}{\partial x_j^q} \equiv \frac{\partial F_j^q}{\partial x_i^p}$. We call this property *externality symmetry*.

Our conclusions about congestion games fall in three overlapping categories: characterization of equilibrium, convergence of evolutionary dynamics, and efficiency. The backbone of our analysis is an implication of externality symmetry: the existence of a potential function defined on the strategy space. The Nash equilibria of a congestion game are precisely the states which satisfy the Kuhn-Tucker first-order conditions for a maximizer of potential (Proposition 3.1). Under any reasonable evolutionary dynamic, the locally stable states are exactly those which maximize potential (Theorem 4.3). Moreover, the population must eventually settle upon a connected set of Nash equilibria regardless of the initial state (Theorem 4.4).

In general, because players do not bear the full social costs of their actions, Nash equilibria are not efficient. However, if in a congestion game all roads are equally sensitive to traffic, individual and social costs are closely linked: the potential function, which captures individuals' incentives, is proportional to average social costs. Consequently, we are able to show that in these games, which we call isoelastic, aggregate efficiency increases over all evolutionary trajectories (Theorem 5.3), and the locally efficient equilibria are precisely those which are dynamically stable (Theorem 5.4). Finally, if facility costs are also non-negative, equilibrium is essentially unique, globally efficient, and globally stable (Corollary 5.5).

These efficiency results can be illustrated through a simple single population example. A group of commuters must drive from their homes located at O to their offices at D. A network of streets connects the locations. Two of the streets, 2 and 3, go over bridges and are especially susceptible to congestion. Initially, only two routes are connecting O and D are available: route, X, over streets 1 and 2, and route Y, over streets 3 and 4. Two plans for improving the network are proposed. One plan recommends adding a fifth street, while the other recommends expanding streets 2 and 3.

Initially, the time to take either route is 18 minutes plus congestion delays. Fixing the total population mass at one, the time (in hours) required to traverse each street as a function of the proportion of drivers using the street are $c_1(u) = c_4(u) = \frac{1}{5} + u$ and $c_2(u) = c_3(u) = \frac{1}{10} + u^2$. Hence, the total driving times on each route as a function of the drivers' choices, (x, y), are[10] $F_X(x, y) = c_1(x) + c_2(x) = \frac{3}{10} + x + x^2$, and $F_Y(x, y) = c_3(y) + c_4(y) = \frac{3}{10} + y + y^2$. Symmetry considerations imply that the equilibrium distribution is $(x, y) = (\frac{1}{2}, \frac{1}{2})$, yielding an equilibrium driving time of exactly 63 minutes. The equilibrium state is also the most efficient allocation of drivers, minimizing the average driving time.

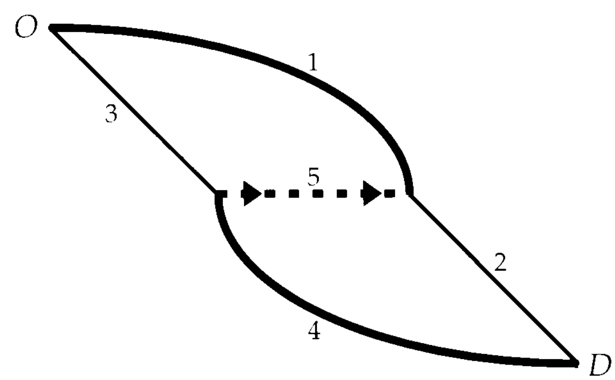

Figure 1

Now suppose that street 5 is constructed, and that to prevent delays at the intersection only west to east traversals are permitted. Driving times on the new street are given by $c_5(u) = \frac{1}{10} + \frac{1}{2}u$. Three routes from O to D are now available: routes X and Y above, and route Z, which uses streets 3, 5, and 2. The costs associated with each route are

$$F_X(x, y, z) = c_1(x) + c_2(x + z) = \tfrac{3}{10} + x + (x+z)^2,$$
$$F_Y(x, y, z) = c_3(y + z) + c_2(y) = \tfrac{3}{10} + (y+z)^2 + y,$$
$$F_Z(x, y, z) = c_3(y + z) + c_5(z) + c_2(x + z) = \tfrac{3}{10} + (y+z)^2 + \tfrac{1}{2}x + (x+z)^2.$$

Of course, adding a route can only improve the efficient average driving time: here, $(x^e, y^e, z^e) \approx (.465, .465, .070)$ yields the efficient average driving time of 62.47 minutes. However, at this state, the driving time on route Z is less than that on the

[10] Notice that these functions are the negations of the payoff functions.

other routes: $60 \cdot F_Z(x^e,\ y^e,\ z^e) = 54.48 < 63.06 = 60 \cdot F_Y(x^e,\ y^e,\ z^e) = 60 \cdot F_Z(x^e,\ y^e,\ z^e)$. Therefore, drivers from routes X and Y would switch to route Z. In fact, it can be shown that under any reasonable model of evolution, regardless of the initial distribution of drivers, the distribution of drivers' choices must ascend the function

$$f(x, y, z) = -\left[\tfrac{1}{3}(x^3 + y^3 + 2z^3) + (x^2 + y^2)(z + \tfrac{1}{2}) + (x + y + \tfrac{1}{4})z^2\right].$$

The population must settle at the unique local maximizer of this function, the state $(x^*, y^*, z^*) = (.419, .419, .162)$. This is the unique equilibrium of the game and yields a common driving time of 63.39 minutes. Comparing this to the original time of 63.00 minutes, we see that, adding street 5 *increases* the equilibrium driving time. This phenomenon is known as *Braess' paradox*.[11] In Section 6, we show how efficiency can be restored to this network by imposing congestion tolls.

Suppose instead that streets 2 and 3 were widened, so that $c_2(u) = c_3(u) = \frac{1}{10} + \frac{1}{4}u$. Symmetry again yields an equilibrium of $(\frac{1}{2}, \frac{1}{2})$, resulting in a driving time of exactly 58.5 minutes. This is the efficient designation of drivers for this configuration of streets. Thus, widening existing roads would prove much more effective than constructing a new road.

Finally, consider the consequences of implementing both projects. In this case, the driving time functions are $F_X(x, y, z) = \frac{3}{10} + \frac{5}{4}x + \frac{1}{4}z$, $F_Y(x, y, z) = \frac{3}{10} + \frac{5}{4}y + \frac{1}{4}z$, and $F_Z(x, y, z) = \frac{3}{10} + \frac{1}{4}x + \frac{1}{4}y + \frac{3}{2}z$. It can be shown that any plausible evolutionary trajectory must ascend the function

$$f(x, y, z) = -[\tfrac{5}{8}(x^2 + y^2) + \tfrac{3}{4}z^2 + \tfrac{1}{4}(x + y)z]$$

and therefore must converge to its unique local maximum of $(x^*, y^*, z^*) \approx (.385, .385, .230)$. This is the unique equilibrium of the game, yielding driving times of 50.30 minutes. Moreover, the function f is an affine transformation of the average cost function

$$\bar{F}(x,y,z) = \tfrac{5}{4}(x^2 + y^2) + \tfrac{3}{2}z^2 + \tfrac{1}{2}(x + y)z + \tfrac{3}{10}.$$

11 The first example of this phenomenon was given by Braess (1968). For further discussion of Braess' paradox, see Pas and Principio (1997) and the references therein.

Hence, the globally stable equilibrium (x^*, y^*, z^*) is also the state which minimizes average driving time. In Section 5, we explain why the cost structure in this example induces equilibrium efficiency.[12]

2.2 Reflexive Externality Games

In many settings, by performing a particular action a player only imposes externalities on opponents who choose the same action. We call games with this property *reflexive externality games.* They are played by a single population and are defined by the property that the payoff to each strategy only depends on the use of that strategy: $F_i(x) = F_i(x_i)$ for all $x \in \mathcal{X}$ and $i \in S$. Such games trivially satisfy externality symmetry. It is worth noting that reflexive externality games are single population congestion games whose facility requirement sets Φ_i form a partition of Φ.

While our discussions have focused on settings in which congestion is bad, congestion is often a boon; externalities can be negative or positive. Importantly, the models we have described can be applied in either case. In congestion games, decreasing facility costs reflect positive externalities, while increasing facility costs indicate negative externalities; in reflexive externality games, positive and negative externalities are represented by increasing and decreasing payoff functions, respectively. We shall see that in the presence of positive externalities multiple equilibria are the norm, but that under negative externalities, equilibria are unique.

2.3 Doubly Symmetric Games

Most work in evolutionary game theory has focused on *single population linear games*, in which $F(x) = Ax$ for some $n \times n$ matrix A (i.e., $F_i(x) = (Ax)_i$ for all $i \in S$). The standard interpretation of these games is based on a random matching environment. Each entry A_{ij} of the payoff matrix expresses the payoffs that a player selecting strategy i would receive in a random match with a player choosing strategy

[12] Monderer and Shapley (1996) prove an equivalence between congestion games and potential games with a finite number of non-anonymous players. However, it is easy to generate examples which show that this result does not extend to the current context. To prove that every potential game is a potential game, Monderer and Shapley (1996) use a construction in which the number of facilities grows (indeed, grows exponentially) in the number of players. When there are a continuum of players such a construction cannot be used. We note one exceptional case: it is easy to show that linear potential games can always be derived from congestion models with linear costs.

j. Thus, when the state is x, the expected return from a random match for a player choosing strategy i is $(Ax)_i$. We can interpret these payoffs as the expected payoffs from a single random match, or the average realized payoffs accruing to players after a very large number of matches. In either case, it is the linearity of expectation which defines the payoffs' functional form. The set of Nash equilibria of the population game coincides with the set of symmetric mixed strategy Nash equilibria of the two player normal form game with payoff bimatrix $\{(A_{ij}, A_{ij}^T)\}_{i,j \in S}$.

If the payoff matrix A is symmetric ($A_{ij} = A_{ji}$ for all $i, j \in S$), it defines a *doubly symmetric game*. Under the random matching interpretation, the two players involved in any match receive the same payoff. Hence, one player's gain is by definition a gain for the opponent he meets. In the Introduction, we explained the importance of these games in biology. We shall soon see that all doubly symmetric games are isoelastic potential games; consequently, all of our results concerning evolution and efficiency apply in this setting.[13]

3. Potential Games

A game with p continuous populations of players is defined by a strategy set for each population and payoff functions for each strategy. The set of populations is denoted $P = \{1, \ldots, r\}$, where $r \geq 1$. The set of strategies for population p is denoted $S^p = \{1, \ldots, n^p\}$, and $n = \sum_p n^p$ gives the total number of pure strategies. Each population has a *mass* m^p, and $m = \sum_p m^p$ denotes the aggregate population mass. By allowing different population masses, we ensure that each infinitesimal player is of the same "size".

The set of strategy distributions of population $p \in P$ is denoted $\mathcal{S}^p = \{x \in \mathbf{R}^{n^p}_+ : \sum_i x_i^p = m^p\}$, while $\mathcal{S} = \{x = (x^1, \ldots, x^r) \in \mathbf{R}^n_+ : x^p \in \mathcal{S}^p\}$ is the set of overall strategy distributions. The continuous payoff functions are given by $F_i^p : \mathcal{S} \to \mathbf{R}$, $i \in S^p$, $p \in P$. Note that the payoffs to a strategy in population p can depend on the strategy distribution within population p itself. We let $F^p : \mathcal{S} \to \mathbf{R}^{n^p}$ refer to the vector of payoff functions for strategies belonging to population p and let $F : \mathcal{S} \to \mathbf{R}^n$ denote

---

[13] It is easy to show that all doubly symmetric games can be represented as single population congestion games with linear facilities costs; the converse representation also holds. We present doubly symmetric games separately because the interpretations of the games are quite different, and the homeomorphism between the two is not so transparent so as to yield a clear connection between the two interpretations.

the vector of all payoff functions. Similar notational conventions are used throughout. However, when considering single population games, we omit superscripts and assume that the population mass equals one.

The best response correspondence, BR: $\mathcal{S} \to \mathcal{S}$, maps each state $x \in \mathcal{S}$ to the set of states whose supports consist entirely of best responses to x. Letting $C^p(x^p) = \{i \in S_p: x_i > 0\}$ denote the support of x^p, we define BR^p and BR by

$$BR^p(x) = \{z^p \in \mathcal{S}^p: \; C^p(z^p) \subset \operatorname*{arg\,max}_{j \in S^p} F_j^p(x)\}, \text{ and}$$
$$BR(x) = \{z \in \mathcal{S}: \; z^p \in BR^p(x) \; \forall \, p \in P\}.$$

A *Nash equilibrium* of game F is a state whose support consists solely of best responses to itself: $C^p(x^p) \subset \operatorname{arg\,max}_j F_j^p(x)$ for all $p \in P$. The Nash equilibria are therefore the fixed points of BR; existence follows from the standard argument.

As we saw in Section 2, in congestion games the marginal effect of adding a player choosing strategy i on the payoffs of players choosing strategy j is the same as the marginal effect of adding a player choosing strategy j on the payoffs of players choosing strategy i. In general, we say that a game F satisfies *externality symmetry* if the following condition holds:

$$\text{(ES)} \qquad \frac{\partial F_i^p}{\partial x_j^q} \equiv \frac{\partial F_j^q}{\partial x_i^p} \text{ for all } i \in S^p, j \in S^q, \text{ and } p, q \in P.$$

Unfortunately, in proposing externality symmetry we create a technical difficulty. To see this most easily, consider a game with a single population. To make sense of the partial derivatives in condition (ES), payoffs need to be defined for all population masses in a small interval around 1. However, since we are only concerned with the proportions of players choosing each strategy, the space of strategy distributions only contains states for which the overall population mass is exactly one. Thus, the partial derivatives in condition (ES) are not well-defined.

In congestion games, this problem is insubstantial. While we have defined congestion games with fixed population masses, there is no difficulty in allowing slight fluctuations: since payoffs are defined by way of facility costs, if we expand the set of states to allow a range of aggregate population masses, payoffs continue to be well-defined. Payoffs outside $\mathcal{S}$ are those which would arise if the total population mass was slightly lower or higher than the benchmark value. In this case, each partial derivative in condition (ES) has a natural interpretation as a measure of the marginal impact of a newcomer.

To make formal sense of condition (ES), we must assume that payoffs are C^1 (continuously differentiable) functions on a wedge of $\mathbf{R}^n$ containing the simplex. More precisely, the *wedge* is defined by $\mathcal{W} = \{x \in \mathbf{R}^n\colon\ 0 \le x_i^p \le m^p\ \forall\ i \in S^p, p \in P,\ m^p - \alpha \le \sum_i x_i^p \le m^p + \alpha\ \ \forall\ p \in P\}$, where α is a positive constant. Externality symmetry then implies that every potential game admits a *potential function*: a scalar valued function $f\colon \mathcal{W} \to \mathbf{R}$ whose gradient, ∇f, is F: That is, condition (ES) implies condition (P):

(P) There exists a C^1 function $f\colon \mathcal{W} \to \mathbf{R}$ such that $\dfrac{\partial f}{\partial x_i^p}(x) = F_i^p(x)$

for all $x \in \mathcal{S}$, $i \in S^p$, and $p \in P$.

We call any game which satisfies condition (P) a *potential game*.

If the potential function f is C^2 (twice continuously differentiable), then the game associated with f satisfies externality symmetry. Thus, condition (ES) is stronger than condition (P). The main advantage of the added generality of condition (P) is that it allows us to consider games with non-differentiable payoff functions. Using this condition also makes extending the payoff functions to the wedge unnecessary.

While one can verify that a set of payoff functions define a potential game by checking externality symmetry directly, it is more fruitful to construct the potential function itself. A congestion game with payoff functions $F_i^p(x) = -\sum_{\phi \in \Phi_i^p} c_\phi(u_\phi(x))$ admits a potential function of the form

$$f(x) = -\sum_{\phi \in \Phi} \int_0^{u_\phi(x)} c_\phi(z)\,dz.$$

A reflexive externality game with payoffs $F_i(x) = F_i(x_i)$ is represented by the potential function

$$f(x) = \sum_{i \in S} \int_0^{x_i} F_i(z)\,dz.$$

A doubly symmetric game with payoffs $F_i(x) = (Ax)_i$ admits the potential function

$$f(x) = \tfrac{1}{2} x \cdot Ax.$$

Finally, any single population, two strategy game is a potential game: given any continuous payoff functions F_1, F_2: $\mathcal{S} \to \mathbf{R}$, a potential function satisfying condition (P) is given by

$$f(x, y) = \int_0^x F_1(z, 1-z)\,dz + \int_0^y F_2(1-z, z)\,dz.$$

The potential function encodes all of the information players need when deciding whether to switch strategies. Consider a state $x \in \mathcal{S}$ at which $F_i^p(x) > F_j^p(x)$ and $x_j^p > 0$. At such a state, a player choosing strategy j would prefer to switch to strategy i. But observe that $\frac{\partial f}{\partial(x_i^p - x_j^p)}(x) = \frac{\partial f}{\partial x_i^p}(x) - \frac{\partial f}{\partial x_j^p}(x) = F_i^p(x) - F_j^p(x) > 0$. Consequently, a profitable strategy change leads to an ascension of the potential function. More generally, we show in Section 4 that the uphill directions of the potential function include all those in which a population of myopically rational players might evolve. This property lies at the heart of our analysis.

Before considering evolution in potential games, it is revealing to characterize their equilibria. The discussion above suggests that the Nash equilibria of the game are related to the local maximizers of the potential function on strategy space. The Kuhn-Tucker first-order necessary conditions for this maximization problem are

(KT1)	$\frac{\partial f}{\partial x_i^p}(x) = \mu^p - \lambda_i^p$	for all $i \in S^p$ and $p \in P$,
(KT2)	$\lambda_i^p x_i^p = 0$	for all $i \in S^p$ and $p \in P$, and
(KT3)	$\lambda \geq 0$,	

where $\mu \in \mathbf{R}^r$ and $\lambda \in \mathbf{R}^n$. We now show that these conditions completely characterize the set of Nash equilibria.

Proposition 3.1: *The state $x \in \mathcal{S}$ is a Nash equilibrium of the potential game F if and only if (x, μ, λ) satisfies* (KT1), (KT2), *and* (KT3) *for some $\lambda \in \mathbf{R}^n$ and $\mu \in \mathbf{R}^r$.*

Proof: If x is a Nash equilibrium of F, then since $F(x) = \nabla f(x)$, the Kuhn-Tucker conditions are satisfied by x, $\mu^p = \max_j F_j^p(x)$, and $\lambda_i^p = \mu^p - F_i^p(x)$.

Conversely, if (x, μ, λ) satisfies the Kuhn-Tucker conditions, then for every $p \in P$, (KT1) and (KT2) imply that $F_i^p(x) = \frac{\partial f}{\partial x_i^p}(x^*) = \mu^p$ for all $i \in C^p(x^p)$. Furthermore, (KT1) and (KT3) imply that $F_j^p(x) = \mu^p - \lambda_j^p \leq \mu^p$ for all $j \in S^p$. Hence, $C^p(x^p) \subset \arg\max_j F_j^p(x)$, and so x is a Nash equilibrium of F. ■

Observe that in equilibrium, the vector of Lagrange multipliers μ gives the equilibrium payoffs in each of the r populations.

Since $\mathcal{S}$ satisfies constraint qualification, satisfaction of the Kuhn-Tucker conditions is necessary but not sufficient for the local maximization of potential. Thus, Proposition 3.1 suggests a division of the Nash equilibria into two classes: the *stable equilibria*, which are local maximizers of the potential function, and the *unstable equilibria*, which are not. This terminology is justified by Theorem 4.3, which establishes that the equilibria which are locally stable under evolutionary dynamics are precisely those which locally maximize potential.

In most potential games, many equilibria exist. However, it follows immediately from Proposition 3.1 that if a game's potential function is concave, all of its equilibria lie in a single convex set. In congestion games the potential function is given by $f(x) = -\sum_{\phi \in \Phi} \int_0^{u_\phi(x)} c_\phi(z)\,dz$; this function is concave whenever the facility costs c_ϕ are increasing. Thus, we are able to show that when externalities are negative, equilibria are essentially unique.

Corollary 3.2: *If the potential function f is concave on $\mathcal{S}$, then all equilibria of the corresponding potential game F are in a single convex set. In particular, these statements are true of all congestion games with increasing facility costs.*

Proof: In the Appendix.

If in addition each strategy uses at least one facility with strictly increasing costs, the potential function is strictly concave and equilibrium is unique.

4. Evolutionary Dynamics

In this section, we characterize evolution in potential games under net monotone dynamics. To accomplish this, we show that the potential function serves as a global Lyapunov function under any such dynamics. The existence of a Lyapunov function rules out both closed orbits and chaotic behavior. Under the additional assumption of non-extinction, it also guarantees the global convergence of solution paths to Nash equilibria.

Throughout this section we use terminology which is standard in dynamical systems and in evolutionary game theory. Formal definitions omitted from the text can be found in the Appendix.

4.1 Net Monotonicity and Non-Extinction

Any evolutionary dynamics associated with a game are described by a C^1 vector field $V\colon \mathcal{S} \to \mathbf{R}^n$ on strategy space. The differential equation $\dot{x} = V(x)$ defines a dynamical system on $\mathcal{S}$ which for each initial condition x_0 admits a unique solution $\{x_t\}_{t \geq 0}$ satisfying $\frac{d}{dt} x_t = V(x_t)$ for all $t \geq 0$. We assume throughout that $\mathcal{S}$ is *forward invariant* under V: solutions do not leave strategy space.

The precise functional forms of the evolutionary dynamics used in the literature reflect specific assumptions about the forces guiding evolution.[14] In particular applications, these precise assumptions may not apply. Consequently, rather than investigating particular equations of motion, it is desirable to study general classes of evolutionary dynamics.

Our main condition on the dynamics is called *net monotonicity.*

Net monotonicity: $\sum_{p \in P} \sum_{i \in S^p} V_i^p(x) F_i^p(x) > 0$ whenever $V(x) \neq \vec{0}$.

Geometrically, net monotonicity requires that except at rest points, the vector field defining the direction of evolution forms acute angles with the payoff vector field. For a more intuitive interpretation of this condition, observe that since the solutions do not leave the strategy space, $\sum_i V_i^p(x) = 0$ for each population p, which implies that $\sum_p \sum_i V_i^p(x) = 0$. Net monotonicity therefore asks that the growth rates of the better performing strategies tend to be higher than those of the worse performing strategies. This need only be true in an average sense: even best responses may become less prevalent if this decline is outweighed by the growth of other strategies receiving high payoffs. Moreover, there is only a single inequality restricting evolution in all r populations. Hence, the growth of poorly performing strategies in some populations can be compensated for by the growth of well performing strategies in others. The most closely related conditions in the evolutionary game theory literature are those of Friedman (1991) and Swinkels (1993), who impose restrictions similar to net monotonicity on *each* population.[15]

---

14 The prime examples are the replicator dynamics (Taylor and Jonker (1978)) and the best response dynamics (Gilboa and Matsui (1991), Matsui (1992), and Hofbauer (1995)).

15 Friedman (1991) considers *weak compatibility*, which combines net monotonicity within each population with *extinction*: $x_i^p = 0$ implies that $V_i^p(x) = 0$. Swinkels (1993) studies *myopic adjustment dynamics*, which satisfy net monotonicity within each population, but with a weak inequality

Since $F(x) = \nabla f(x)$, $F(x)$ is the direction of steepest ascent of the potential function. Hence, stated in terms of the potential function, net monotonicity requires that whenever the population is moving, it is moving uphill. This observation formalizes our claim in Section 3.3 that any reasonable evolutionary process must climb the potential function. It also underlies our main technical lemma. We call a C^1 function $f: \mathcal{S} \to \mathbf{R}$ a *global Lyapunov function* for the dynamical system $\dot{x} = V(x)$ if for every solution trajectory $\{x_t\}_{t \geq 0}$, (*i*) $\frac{d}{dt}f(x_t) \geq 0$ for all t, and (*ii*) $\frac{d}{dt}f(x_t) = 0$ implies that $V(x_t) = 0$. Condition (*i*) requires that the function f is weakly increasing along all solution trajectories, while condition (*ii*) demands that f is strictly increasing except at fixed points of V. Lemma 4.1 establishes that the potential function is a Lyapunov function under any net monotone dynamics, providing a powerful tool for characterizing evolution.

Lemma 4.1: *If F is a potential game and V is net monotone, then the potential function of F is a global Lyapunov function for* $\dot{x} = V(x)$.

Proof: Net monotonicity implies that $\frac{d}{dt}f(x_t) = \nabla f(x_t) \cdot \dot{x}_t = F(x_t) \cdot V(x_t) \geq 0$ and that $V(x_t) = 0$ whenever $\frac{d}{dt}f(x_t) = 0$. ■

Under the replicator dynamics, all faces and vertices of strategy space are forward invariant: once a strategy is extinct, it never resurfaces. We feel that this property of the replicator dynamics is not sensible in most applications. In economic settings, knowledge of a payoff improving alternative should lead players to select this alternative regardless of whether it is currently being played. In biological contexts, rare mutations generate slight forces away from the boundaries of strategy space which prevent this form of extinction; see Binmore and Samuelson (1997). In either case, the abandonment of a strategy only seems reasonable as a consequence of inferior performance, not as an implication of other modeling assumptions.

With this in mind, we introduce a second condition on the dynamics which we call *non-extinction*. It requires that any rest point of the dynamics be a Nash equilibrium.

Non-extinction: $V(x) = \bar{0}$ implies that x is a Nash equilibrium of F.

---

replacing the strict one and with the additional requirement that all Nash equilibria are rest points (although this latter requirement is omitted in some of his results). For other conditions on evolutionary game dynamics which are stronger than net monotonicity, see Nachbar (1990), Samuelson and Zhang (1992), and Ritzberger and Weibull (1995). See Weibull (1995) for a survey.

Non-extinction is essentially a requirement that players avail themselves of opportunities to improve their payoffs. While not satisfied by the replicator dynamics, non-extinction is satisfied by the best response dynamics.

4.2 Evolutionary Stability

We now present our evolutionary stability results, which characterize the dynamics at three levels of proximity to rest points. We first characterize the rest points themselves in terms of the Nash equilibria of the underlying game. Next, we consider local dynamics, studying evolution in small neighborhoods of all rest points. Finally, we characterize the global behavior of the dynamics, proving that all solution paths must converge to equilibrium.

A Nash equilibrium state is defined by the property that no player can improve his payoffs by switching strategies. Thus, strategy revision by an infinitesimal set of agents cannot improve their payoffs. This implies that any change from a Nash equilibrium state must violate net monotonicity, and therefore that all Nash equilibria are dynamic rest points. Under the additional assumption of net monotonicity, *only* Nash equilibria can be rest points of the dynamics. This is the content of Proposition 4.2.[16]

Proposition 4.2: *If V is net monotone, then all Nash equilibria of F are rest points of $\dot{x} = V(x)$. If in addition V satisfies non-extinction, then the set of Nash equilibria of F and the set of rest points of $\dot{x} = V(x)$ coincide.*

Proof: Let x be a Nash equilibrium of F, and let V be some net monotone dynamics. Suppose $D^p(x)$ is the set of strategies in S^p that are in decline at x: $D^p(x) = \{i \in S^p\colon\ V_i^p(x) < 0\}$. Then $D^p(x) \subset C^p(x^p) \subset \operatorname{argmax}_i F_i^p(x)$. But since $\mathcal{Y}$ is forward invariant under V, $V^p(x) \cdot \vec{1} = 0$; therefore, the inclusion implies that $V^p(x) \cdot F^p(x) \leq 0$. Summing over p, we see that $V(x) \cdot F(x) \leq 0$. Net monotonicity then implies that $V(x) = 0$. The proof of the second claim follows immediately from the first claim and the definition of non-extinction. ■

---

[16] It is worth noting that the proof of this proposition does not depend on the existence of a potential function; the result is valid for any game with continuous player sets, not just potential games.

While we are able to identify equilibria with rest points of the dynamics, not all equilibria are stable. The prototypical example of a failure of stability is a mixed equilibrium of a pure coordination game. Consider a population of players involved in the doubly symmetric game with the payoff bimatrix in Figure 2.

1, 1	0, 0
0, 0	2, 2

Figure 2

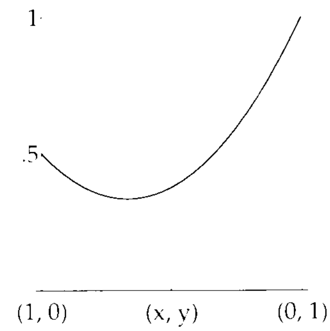

Figure 3

The payoff functions of this game are $F_1(x, y) = x$ and $F_2(x, y) = 2y$. In addition to its two pure equilibria, the game possesses a mixed equilibrium in which $\frac{2}{3}$ of the players choose strategy one. By Proposition 4.2, this equilibrium must be a rest point under any net monotone dynamics. However, it is repelling: for example, a slight increase in the number of players choosing strategy one renders strategy one a unique best response, driving all members of the population to coordinate on this strategy. This is illustrated in Figure 3, which graphs the potential function of this game: $f(x, y) = \frac{1}{2}x^2 + y^2$. Since evolution must proceed uphill, we see that the mixed equilibrium $(\frac{2}{3}, \frac{1}{3})$ is dynamically unstable. Only the equilibria which locally maximize potential constitute credible predictions of play.

The predictions of the model of the behavior of a large population of agents should be robust to small fluctuations in the agents' aggregate behavior. Since we ought to expect aggregate play to exhibit small deviations from the equilibrium distribution, equilibria that are not robust against small disturbances do not constitute reasonable forecasts. It is therefore important to determine which equilibria are resistant to such fluctuations.

There are two main conditions used in evolutionary game theory to characterize the local stability of equilibria. Roughly speaking, an equilibrium is *Lyapunov stable* if no small perturbation can cause the population to leave the vicinity of the equilibrium, even temporarily. The stronger criterion of *asymptotic stability* requires that in addition, the population returns to equilibrium after any small perturbation. Since net monotone dynamics drive the population up the contours of the potential function, it seems plausible that connections exist between its local maximizers and the game's dynamically stable equilibria. These connections are established in Theorem 4.3.

Before stating this result we require some additional definitions. We call a closed set *isolated* if there exists a neighborhood of the set containing no Nash equilibria outside the set. A set $A \subset \mathcal{S}$ is *smoothly connected* if for any points x and y in A there exists a continuous, piecewise differentiable curve γ contained in A whose endpoints are x and y. Closed, connected sets of Nash equilibria satisfy the latter two properties in all but pathological cases. A set $A \subset \mathcal{S}$ is a *local maximizer set* for f on $\mathcal{S}$ if (*i*) A is smoothly connected; (*ii*) f is constant on A; and (*iii*) there exists a neighborhood B of A such that $f(y) < f(x)$ for all $y \in B - A$ and $x \in A$. Since f is continuous, all local maximizer sets are closed. Moreover, by Proposition 3.1, all local maximizer sets consist entirely of Nash equilibria. Finally, the formal definitions of Lyapunov stability and asymptotic stability can be found in the Appendix.

We now state our local stability result.

Theorem 4.3: *If V is net monotone and A is a smoothly connected closed set, then*

(*i.a*) *If A is a local maximizer set, then A is Lyapunov stable;*

(*i.b*) *If A is asymptotically stable, then A is a local maximizer set.*

If in addition V satisfies non-extinction and A is isolated, the following are equivalent:

(*ii.a*) *A is a local maximizer set;*

(*ii.b*) *A is Lyapunov stable;*

(*ii.c*) *A is asymptotically stable.*

Proof: In the Appendix.

Part (*i.a*) of Theorem 4.3 tells us that under net monotone dynamics, all local maximizers of potential are Lyapunov stable: small perturbations in behavior are not enough to move the population away from these sets. Part (*i.b*) provides a partial converse: any set which satisfies the stronger criterion of asymptotic stability must be a local maximizer set. All other equilibria can be upset, though perhaps only slightly, by an appropriate tremble.

If the evolutionary dynamics also satisfy non-extinction, we can say considerably more. Without this condition, it is clear that local maximizers need not be asymptotically stable; even if the set is isolated, there can be non-Nash rest points of the dynamics arbitrarily close by. Once we assume that such rest points cannot exist, asymptotic stability follows. Furthermore, under dynamics satisfying both conditions, local maximization of potential becomes a necessary condition for any sort of stability: non-maximizers cannot even satisfy the weak requirement of Lyapunov stability. In sum, under dynamics satisfying net monotonicity and non-extinction, local maximizers of potential are attractors, and no other equilibria are even moderately stable.

Local stability results are most important once a system reaches equilibrium, as they establish whether we should expect the equilibrium to persist. However, they do not guarantee that equilibrium will ever be reached. In general, evolutionary game dynamics can exhibit closed orbits and chaotic behavior, with solution trajectories perpetually avoiding neighborhoods of rest points. When this occurs, equilibrium prediction is obviously inappropriate. Unless one can guarantee that such behavior is impossible, a characterization of local stability is of decidedly less interest.

Fortunately, we are able to establish that in potential games, under dynamics satisfying net monotonicity and non-extinction, convergence to equilibrium is assured. Let the *limit set* of x, $\omega(x)$, be the set of accumulation points of the solution with initial condition x: $\omega(x) = \{z \in \mathcal{S}\colon \exists \left\{x_{t_k}\right\}_{k=1}^{\infty} \subset \left\{x_t\right\}_{t\geq 0}$ such that $\lim_{k\to\infty} x_{t_k} = z\}$. Using this definition, we state our global convergence result.

Theorem 4.4: *If V is net monotone and satisfies non-extinction, each limit set $\omega(x)$ is a closed, connected set of Nash equilibria. Consequently, the union of all limit sets, $\bigcup_{x\in\mathcal{S}} \omega(x)$, is equal to the set of Nash equilibria, and therefore to the set of fixed points of $\dot{x} = V(x)$.*

Proof: Follows immediately from Lemma 4.1, Proposition 4.2, and Lemma A.1 in the Appendix. ■

The first claim of Theorem 4.4 establishes that solution trajectories starting from each initial condition must converge to a connected set of Nash equilibria. In generic games, all connected sets of equilibria are singletons, so each trajectory converges to a unique limit point. More generally, because of our weak restrictions on the dynamics, it is easy to construct examples in which convergence occurs in a setwise fashion. The second claim shows that the sets of limit points, Nash equilibria, and rest points are identical. In the absence of fluctuations in behavior, each of the three sets describes the possible long run behaviors of the population. Together, these results fully justify Nash equilibrium prediction.

5. Efficiency

While in potential games evolution guarantees equilibrium play, our results thus far make no claims about equilibrium efficiency. We now investigate payoff efficiency in potential games. Our main results show that in isoelastic potential games, average payoffs must increase along all solution paths, and all stable equilibria are efficient. We begin this section by using a simple subclass of congestion games to describe the impediments to efficiency. We then prove our results in the more general context of potential games and discuss the implications of these results for models of congestion.

5.1 Isoelasticity

To illustrate the barriers to efficiency, we consider a single population model of traffic flow. Payoffs, reflecting driving times, are negative: $F_i(x) < 0$ whenever $x_i > 0$. Since adding cars increases travel times, $\frac{\partial F_i}{\partial x_i}(x) \leq 0$ for all $x \in \mathcal{X}$ and all $i \in S$. In this example, we assume that no two routes share a common street. This prohibits cross-effects between the utilization of one route and the payoffs to another: $\frac{\partial F_i}{\partial x_j} \equiv 0$ for all $i \neq j$. By assuming away cross-effects we can make simple, precise statements about efficiency.[17]

Considerable insight into the welfare properties of congestion games can be gained by considering payoff elasticities. We define the *elasticity* of strategy i (more precisely, strategy i's payoff elasticity of utilization) by

$$\varepsilon_i(x) = \frac{x_i \frac{\partial F_i}{\partial x_i}(x)}{F_i(x)}.$$

The elasticity measures how sensitive route i's costs are to changes in the level of driving volume. Since both payoffs and their "on-diagonal" partial derivatives are negative, elasticities are positive.

---

[17] Barro and Romer (1987) consider models of congestion without cross-effects and state conditions under which congestion and equilibrium efficiency are compatible.

Welfare can be measured using the average payoffs in the population at that state. We call a state *locally efficient* if average payoffs $\bar{F}(x) = \sum_i x_i F_i(x)$ reach a local maximum at that state. Consider the impact on social welfare of introducing a new player playing strategy i to the population. Since there are no cross-effects between strategies, the marginal effect of this addition is equal to

$$\begin{aligned} \frac{\partial \bar{F}}{\partial x_i}(x) &= \sum_k \left(x_k \frac{\partial F_k}{\partial x_i}(x) \right) + F_i(x) \\ &= x_i \frac{\partial F_i}{\partial x_i}(x) + F_i(x). \end{aligned}$$

The marginal social cost of increasing the representation of strategy i can thus be split into two terms. The first, $x_i \frac{\partial F_i}{\partial x_i}(x)$, is the purely social component of marginal social cost: it is the aggregate effect of the entrant on current strategy i players. The second, $F_i(x)$ is the payoff received by the new strategy i player, which has a marginal impact on average payoffs because he is infinitesimally small. Substituting, we see that

$$\frac{\partial \bar{F}}{\partial x_i}(x) = (\varepsilon_i(x) + 1)\, F_i(x).$$

Thus, in the absence of cross effects, the marginal social cost created by a new strategy i player is equal to the payoffs received by this player multiplied by one plus strategy i's elasticity.

We now consider two implications of the previous equation. First, even in our simple example without cross-effects, equilibria are not generally efficient. In particular, strategies whose payoffs are more sensitive to their level of utilization are overused in equilibrium. Suppose that z is an equilibrium of the congestion game in which strategies i and j are both used. We can compute the marginal effect on welfare of a switch by a player from strategy j to strategy i as

$$\frac{\partial \bar{F}}{\partial x_i}(z) - \frac{\partial \bar{F}}{\partial x_j}(z) = (\varepsilon_i(z) + 1)\, F_i(z) - (\varepsilon_j(z) + 1)\, F_j(z).$$

Since in equilibrium $F_i(z) = F_j(z) = \mu < 0$, this last expression equals

$$= (\varepsilon_i(z) - \varepsilon_j(z))\, \mu.$$

Hence, if route j is more sensitive to traffic than route i, a shift in behavior from strategy j to strategy i is welfare improving.

Second, consider a state at which the payoff elasticities of two strategies are equal: $\varepsilon_i(x) = \varepsilon_j(x) = \varepsilon$. In this case,

$$\frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_j}(x) = (\varepsilon + 1)\,(F_i(x) - F_j(x)).$$

To interpret this equation, consider a driver who is debating a switch from route j to route i. The previous equation implies that his benefit from making this switch, $F_i(x) - F_j(x)$, is directly proportional to the marginal social benefit of this switch, $\frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_j}(x)$. That is, the driver's incentives are simply a rescaling of social incentives. Hence, if the payoff elasticities of different routes are equal, individual rationality leads to increases in aggregate efficiency.

The computations above take advantage of the absence of cross-effects. Nevertheless, they suggest that if payoffs are equally sensitive to strategy utilization, the prospects for efficiency are improved. In fact, with appropriate assumptions on payoff elasticities we are able to obtain strong efficiency results.

These assumptions are easiest to understand in the context of congestion games. Recall that the payoffs of a congestion game are defined by the facility payoffs c_ϕ. We define the *facility payoff elasticities* by

$$\varepsilon_\phi(u) = \frac{u c_\phi'(u)}{c_\phi(u)}$$

for all levels of utilization for which the facility costs are non-zero. A congestion game is *isoelastic* if there exists an ε such that $\varepsilon_\phi \equiv \varepsilon$ for all $\phi \in \Phi$. When this condition is satisfied, the costs of all facilities are equally sensitive at all levels of use.

More generally, we say a potential game is *isoelastic* of degree $\varepsilon \neq -1$ if its payoff functions are differentiable almost everywhere and it admits a potential function which is homogenous of degree $\varepsilon + 1$ on $\mathcal{X}$. The following proposition justifies our choice of nomenclature.

Proposition 5.1: *Any isoelastic congestion game is an isoelastic potential game.*

Proof: Recall that a potential function for congestion games is given by

$$f(x) = -\sum_{\phi \in \Phi} \int_0^{u_\phi(x)} c_\phi(z)\,dz.$$

Since here the c_ϕ are isoelastic functions with elasticity ε, they must take the form $c_\phi(u) = \alpha_\phi u^\varepsilon$, where the α_ϕ are constants. Moreover, since the c_ϕ must be well-defined at zero, ε must be non-negative. Therefore,

$$\begin{aligned} f(x) &= -\sum_{\phi\in\Phi}\int_0^{u_\phi(x)} \alpha_\phi z^\varepsilon \, dz \\ &= -\sum_{\phi\in\Phi} \tfrac{\alpha_\phi}{\varepsilon+1}(u_\phi(x))^{\varepsilon+1}. \end{aligned}$$

But since each u_ϕ is linear in x, f is a sum of functions which are homogenous of degree $\varepsilon + 1$ in x. Therefore, f is itself homogenous of degree $\varepsilon + 1$. ■

Define the *aggregate payoff function*, $\bar{F}\colon \mathcal{X} \to \mathbf{R}$, by

$$\bar{F}(x) = \sum_p \sum_i x_i^p F_i^p(x)$$

A state is *locally efficient* if it is a local maximizer of $\bar{F}$. The key to establishing our efficiency results is the observation that isoelasticity implies the proportionality of potential to aggregate payoffs.

Lemma 5.2: *If F is isoelastic of degree $\varepsilon \neq -1$, the function $f(x) = \frac{1}{\varepsilon+1}\bar{F}(x)$ is a potential function for F.*

Proof: We need to show that $\nabla f \equiv F$. By the definition of isoelasticity, each payoff function F_i^p is homogenous of degree ε. Hence, if $x \in \mathcal{S}$ is a point of differentiability of F, then externality symmetry and Euler's law imply that

$$\begin{aligned} \frac{\partial f}{\partial x_i^p}(x) &= \frac{\partial}{\partial x_i^p}\left(\tfrac{1}{\varepsilon+1}\bar{F}(x)\right) \\ &= \tfrac{1}{\varepsilon+1}\left(\sum_{q\in P}\sum_{j\in S^q} x_j^q \frac{\partial F_j^q}{\partial x_i^p}(x) + F_i^p(x)\right) \\ &= \tfrac{1}{\varepsilon+1}\left(\sum_{q\in P}\sum_{j\in S^q} x_j^q \frac{\partial F_i^p}{\partial x_j^q}(x) + F_i^p(x)\right) \\ &= \tfrac{1}{\varepsilon+1}\left(\varepsilon F_i^p(x) + F_i^p(x)\right) = F_i^p(x). \end{aligned}$$

Since F is continuous, and since by definition any potential function of F is C^1, $\frac{\partial f}{\partial x_i^p}(x) = F_i^p(x)$ for all $x \in \mathcal{S}$. ■

Notice that the first summation above represents the marginal effect on the payoffs of the original population of introducing a new player choosing strategy $i \in S^p$. Therefore, the computations within the parentheses show that once again, equality of elasticities causes individuals' payoffs to be proportional to marginal social payoffs.

In conjunction with Lemma 4.1, Lemma 5.2 provides us with a powerful tool for investigating efficiency in isoelastic games. According to Lemma 4.1, directions of ascent of the potential function are precisely those in which rational adjustments by individual players can lead. Lemma 5.2 shows that if a potential game is isoelastic, uphill directions of the potential function are those which increase aggregate efficiency. Therefore, in isoelastic potential games, individuals' incentives and social preferences are perfectly aligned; any adjustment in behavior which benefits the adjuster also benefits the population as a whole.

The first implication we draw from this concordance of incentives is that average payoffs must increase over all evolutionary paths.

Theorem 5.3: *If F is an isoelastic potential game with $\varepsilon > -1$ and V is net monotone, then all solutions of the dynamical system $\dot{x} = V(x)$ satisfy $\frac{d}{dt}\overline{F}(x_t) \geq 0$, with equality only at rest points of V. That is, evolution increases aggregate payoffs. If $\varepsilon < -1$, then solutions satisfy $\frac{d}{dt}\overline{F}(x_t) \leq 0$.*

Proof: Follows immediately from Lemma 4.1 and Lemma 5.2. ■

Recall that doubly symmetric games are single population games in which payoffs are determined by a single symmetric matrix: $F_i(x) = (Ax)_i$ for all $i \in S$. It is easily verified that doubly symmetric games are single population isoelastic potential games with $\varepsilon = 1$. Therefore, since the replicator dynamics are net monotone, Theorem 5.3 implies the Fundamental Theorem of Natural Selection.

The second claim of Theorem 5.3 states that in isoelastic games with $\varepsilon < -1$, evolution *decreases* social efficiency. In these games, individual and social incentives are perfectly misaligned. However, it should be noted that this situation cannot arise in isoelastic congestion games, which must have $\varepsilon \geq 0$; for ε to be negative, facility costs would need to be infinite at utilization level zero.

Because evolutionary paths ascend the potential function, dynamically stable equilibria coincide with the local maximizers of potential. But since in isoelastic games the potential function represents average payoffs, the local maximizers are also the locally efficient states. Therefore, the locally efficient states and the dynamically stable equilibria coincide: the equilibria which are resistant to small behavior trembles are precisely those which are socially efficient.

Theorem 5.4: *If F is an isoelastic potential game with $\varepsilon > -1$, then under any net monotone dynamics, all stable equilibria are locally efficient and all locally efficient states are stable equilibria.*

Proof: Follows immediately from Theorem 4.3 and Lemma 5.2. (The precise meaning of stability in the above statement depends on whether non-extinction is also assumed: see Theorem 4.3.) ■

In general, we can only guarantee that stable states will be locally efficient. For example, in congestion games in which congestion is beneficial, multiple strict equilibria will exist; while each may induce different payoffs, each is locally stable. When congestion is detrimental, one might hope for a stronger result. As we saw in Section 3, when facility costs are increasing, the potential function is concave, so all equilibria maximize potential. But in isoelastic congestion games the potential function is proportional to the aggregate payoff function. Hence, since each equilibrium is a global maximizer of potential, each is globally efficient. Moreover, since all plausible evolutionary processes lead to equilibrium play, all must yield global efficiency. These conclusions are summarized in Corollary 5.5.[18]

Corollary 5.5: *In isoelastic congestion games with non-negative costs, all Nash equilibria are in a single convex set, and all are globally efficient. Moreover, under dynamics satisfying net monotonicity and non-extinction, the equilibrium set is both locally and globally stable.*

5.2 Efficiency in Moderately Congested Traffic Flows

Uncongested traffic flows are efficient. If there are no purely social costs to congestion, and all drivers choose the fastest available route, equilibrium clearly maximizes welfare. To see this formally, observe that if each cost function c_ϕ is independent of the level of utilization, then their common isoelasticity is zero, implying efficiency.

---

[18] Using techniques different from ours, Dafermos and Sparrow (1969) show that in congestion games with increasing, isoelastic, convex costs (i.e., $\varepsilon > 1$), Nash equilibria are globally efficient. Corollary 5.5 establishes that convexity is unnecessary for this result.

More interesting is the case of moderately congested traffic flows. Walters (1961) and Vickrey (1969) both suggest that when roads are utilized at levels below capacity, congestion costs are affine functions of utilization. With this in mind, we say that congestion costs are *moderate* if c_ϕ is affine and non-decreasing for all $\phi \in \Phi$. To obtain an efficiency result, we must also assume that the uncongested travel times of taking various routes are the same.[19] This is plausible if all routes between location pairs are of approximately the same length. Formally, we say that a congestion game has *equal fixed costs* if for all $p \in P$, there is a constant k^p such that $\sum_{\phi \in \Phi_i^p} c_\phi(0) = k^p$ for all $i \in S^p$.

If the cost functions include fixed costs, they clearly cannot be isoelastic. However, if within each population the fixed costs of each strategy are equal, drivers' decisions are unaffected by these fixed costs, which must be borne regardless of the route ultimately chosen. Hence, no player's incentives are affected if we simply eliminate all fixed costs from the model.[20] Moreover, because the fixed costs have a uniform impact on aggregate welfare, an efficient state in the new game is also efficient in the original game. Since striking the fixed costs generates a game with non-decreasing, linear costs, Corollary 5.5 implies the following efficiency result.

Corollary 5.6: *In congestion games with moderate congestion costs and equal fixed costs, evolution increases aggregate efficiency and all equilibria are globally stable and globally efficient.*

6. Efficiency through Congestion Tolls

6.1 Proportional to Marginal Cost Pricing

So far, we have characterized evolution in potential games and established that evolution leads to efficiency when costs are isoelastic. Unfortunately, as Braess' paradox illustrates, when costs are not isoelastic inefficient outcomes are possible. Inefficiency arises because players do not consider social costs when selecting strategies. By introducing a system of congestion tolls, we can force players to weigh

---

[19] Indeed, Braess' (1968) original example uses affine cost functions which do not have this property.
[20] This shifting of payoffs is the basic idea behind quasipotential games, which are discussed in Section 7.1.

the social costs of their actions in a manner which leads to the play of welfare maximizing strategy distributions.

We begin by considering this mechanism design problem in the context of potential games. While in any potential game, evolution leads to the ascension of the potential function, f, our mechanism designer prefers that evolution increase overall social welfare as measured by average payoffs, $\overline{F}$. Suppose the mechanism designer can introduce an extra component (e.g., transfer payments) to the payoffs of the game which alters each player's incentives but which is irrelevant to social welfare. If he could introduce these payoffs in such a way that the newly constructed game had potential function $\overline{F}$, then evolution in this new game would lead to increases in efficiency in the original game. In fact, as long as the potential function of the new game is proportional to $\overline{F}$, both functions exhibit the same uphill directions, and evolution enhances aggregate welfare.

Call the potential function of the new game $\hat{f}$. We would like to construct payoffs in such a way that $\hat{f} \equiv k\overline{F}$, where $k > 0$. To accomplish this, we write $\hat{f} \equiv f + t$: that is, we express the new potential function as the sum of the original potential function and a new function, t, which we call the *transfer potential function*. Differentiating the latter identity, we see that we must solve for a set of *transfers* T_i^p such that the augmented payoff functions $\hat{F}_i^p \equiv F_i^p + T_i^p$ define a potential game whose potential function is $\hat{f}$. Observing that $t \equiv k\overline{F} - f$, we see that if the payoff functions are smooth,

$$\begin{aligned} T_i^p(x) &= \tfrac{\partial t}{\partial x_i^p}(x) \\ &= k\tfrac{\partial \overline{F}}{\partial x_i^p}(x) - \tfrac{\partial f}{\partial x_i^p}(x) \\ &= k\left(\sum_{q\in P}\sum_{j\in S^q} x_j^q \tfrac{\partial F_j^q}{\partial x_i^p}(x)\right) + (k-1)F_i^p(x). \end{aligned}$$

The summation in the last expression represents the purely social marginal effect of a player choosing action $i \in S^p$. When $k = 1$, the transfer simply adds the social impact of each action to the individual player's payoffs, implying that the newly constructed payoffs are equal to marginal social payoffs. Thus, forcing the new potential function to equal average costs yields marginal cost pricing. More generally, individual's payoffs in the new game are *proportional* to marginal social payoffs: $\hat{F}_i^p \equiv F_i^p + T_i^p \equiv k\tfrac{\partial \overline{F}}{\partial x_i^p}$. For this reason, we call these transfer schemes *proportional to marginal cost (PMC) pricing*. Under every PMC pricing plan, evolution yields efficiency.

To determine PMC transfer payments in congestion games, recall that payoff functions in these games are given by $F_i^p(x) = -\sum_{\phi \in \Phi_i^p} c_\phi(u_\phi(x))$. Therefore, noting that

$$\frac{\partial u_\phi}{\partial x_j^q}(x) = \begin{cases} 1 & \text{if } \phi \in \Phi_j^q, \\ 0 & \text{otherwise,} \end{cases}$$

we see that

$$\begin{aligned} T_i^p(x) &= k\left(\sum_{q \in P}\sum_{j \in S^q} x_j^q \left(\sum_{\phi \in \Phi_i^p \cap \Phi_j^q}\left(-c_\phi'(u_\phi(x))\right)\right)\right) + (k-1)\left(\sum_{\phi \in \Phi_i^p}\left(-c_\phi(u_\phi(x))\right)\right) \\ &= -\left(k \sum_{\phi \in \Phi_i^p} u_\phi(x) c_\phi'(u_\phi(x)) + (k-1)\sum_{\phi \in \Phi_i^p} c_\phi(u_\phi(x))\right). \end{aligned}$$

This equation implies that PMC transfers can be achieved by placing congestion tolls directly on the individual facilities. Letting $\kappa = \frac{1}{k} - 1$, we see that for each $\kappa > -1$, the following equation defines a set of PMC congestion tolls:

$$\text{(T)} \qquad \tau_\phi^\kappa(u) = \tfrac{1}{\kappa+1}\left(u c_\phi'(u) - \kappa c_\phi(u)\right).$$

The parameter κ, which we call the *elasticity threshold*, has a simple interpretation: when the cost elasticity of a facility is exactly κ, the congestion toll is set to zero; at cost elasticities higher than κ, positive tolls are imposed; at cost elasticities lower than κ, subsidies are paid.

In general, without additional intervention by the mechanism designer, PMC tolling can only guarantee local efficiency. However, if congestion costs are increasing and never too concave (i.e., if $u c_\phi''(u) \geq -2c_\phi'(u)$ for all ϕ and u), average payoffs are concave. In such cases, evolution under PMC tolling leads to global efficiency.

Theorem 6.1: *In congestion games whose cost functions satisfy $u c_\phi''(u) \geq -2c_\phi'(u)$ for all ϕ and u, aggregate payoffs are concave. Consequently, under any PMC toll, evolution from any initial condition yields global efficiency.*

Proof: In the Appendix.

If congestion costs are increasing and convex, the conditions of the theorem are met, and evolution under PMC tolls yields global efficiency.

We illustrate PMC tolling using our example from Section 2. In particular, we reconsider the case in which a fifth street was added, increasing equilibrium travel times. Ignoring fixed costs, which are equal over each complete route, the travel times on the streets are given by $c_1(u) = c_2(u) = u$, $c_2(u) = c_3(u) = u^2$, and $c_5(u) = \frac{1}{2}u$. Therefore, applying the previous formula, we see that the efficient driver allocation of $(x^e, y^e, z^e) \approx (.465, .465, .070)$ can be achieved by introducing congestion tolls of

$$\tau_1(u) = \tau_4(u) = \frac{1-\kappa}{\kappa+1}\,u,\ \ \tau_2(u) = \tau_3(u) = \frac{2-\kappa}{\kappa+1}\,u^2,\ \text{and}\ \tau_5(u) = \frac{1-\kappa}{2(\kappa+1)}\,u$$

for any $\kappa > -1$.

Three values of the elasticity threshold deserve special attention. First, choosing a elasticity threshold of zero, which corresponds to setting $k = \frac{1}{\kappa+1} = 1$, induces marginal cost pricing. The resulting tolls are $\tau_1(u) = \tau_4(u) = u$, $\tau_2(u) = \tau_3(u) = 2u^2$, and $\tau_5(u) = \frac{1}{2}u$: marginal cost pricing imposes tolls on every street.

Intuitively, the inefficiencies that arise in this example are entirely due to congestion on streets 2 and 3. Adding street 5 reduced efficiency because it allowed drivers to use both of the congested arteries: since the drivers do not account for the social impact of their choices, route Z, which includes both street 2 and street 3, is overused in equilibrium. It therefore seems reasonable to hope that efficiency could be restored by imposing tolls solely on these two streets.

This is precisely what happens if we choose a elasticity threshold of one. In this case, $\tau_2(u) = \tau_3(u) = \frac{1}{2}u^2$, and $\tau_1(u) = \tau_4(u) = \tau_5(u) = 0$. Thus, while marginal cost pricing yields efficiency, by using PMC tolling we can achieve efficiency with fewer and smaller tolls.

Finally, consider choosing $\kappa \approx 1.414$, leading to tolls of $\tau_1(u) = \tau_4(u) = -.172\,u$, $\tau_2(u) = \tau_3(u) = .243\,u^2$, and $\tau_5(u) = -.086\,u$. In this case, negative tolls are charged on the less congested streets. Aggregate toll revenue, given by

$$x^e[\tau_1(x^e) + \tau_2(x^e + z^e)] + y^e[\tau_3(y^e + z^e) + \tau_4(y^e)]$$
$$+ z^e[\tau_2(x^e + z^e) + \tau_3(y^e + z^e) + \tau_5(z^e)],$$

equals zero. Thus, by using tolls on the more congested streets to subsidize travel on the less congested ones, we can simultaneously achieve efficiency and budget balance.[21]

This example illustrates a number of general properties of PMC tolling. Under marginal cost pricing ($\kappa = 0$), as long as costs are monotone only positive tolls are charged. However, as the example demonstrated, marginal cost pricing often imposes more drastic tolls than are needed to generate efficiency. To determine the smallest non-negative PMC tolls, we let

$$\varepsilon_{\min} = \min_{\phi} \inf_{u} \varepsilon_{\phi}(u)$$

equal the minimum cost sensitivity of any facility at any level of utilization. By choosing $\kappa = \varepsilon_{\min}$, we achieve efficient allocations with the least intrusive tolls that are guaranteed to be non-negative. Equation (T) implies that tolls are non-negative at $\kappa = \varepsilon_{\min}$, and that choosing a higher value of κ would lead to the imposition of a negative toll. On the other hand, if costs are non-negative and monotone, lower elasticity thresholds generate higher tolls:

$$\tfrac{\partial}{\partial \kappa} \tau_{\phi}^{\kappa}(u) = \frac{-(u c_{\phi}'(u) + c_{\phi}(u))}{(\kappa+1)^2} \leq 0.$$

Hence, setting κ equal to $\varepsilon_{\min}$ yields tolls with the desired properties.[22]

Finally, elasticity thresholds parameterize toll revenues. To see this, note that for any u such that either $c_{\phi}(u)$ or $c_{\phi}'(u)$ is strictly positive, $\lim_{\kappa \downarrow -1} \tau_{\phi}^{\kappa}(u) = \infty$, $\lim_{\kappa \uparrow \infty} \tau_{\phi}^{\kappa}(u) = -\infty$, and $\tfrac{\partial}{\partial \kappa} \tau_{\phi}^{\kappa}(u) < 0$. Hence, if we fix any state $x \in \mathscr{S}$ which induces positive costs for some strategy, there is a one-to-one, decreasing mapping between elasticity thresholds (on the interval $(-1, \infty)$) and aggregate toll revenues (on the real line). Thus, given any efficient state and any desired toll revenue r, there is a unique set of

---

[21] Of course, one could also achieve budget balance by charging a positive toll and then dividing the accumulated surplus among all drivers. The advantage of the method proposed above is that all subsidies are given as the drivers use the routes, obviating the need to determine who the drivers are.

[22] In general, if one cannot ignore fixed costs (as we did in the example), then any facility with a positive fixed cost will have an elasticity which approaches zero as its utilization approaches zero. On the other hand, the population will ultimately allocate itself efficiently. Thus, if negative tolls are undesirable, what is most important is to ensure that they need not be paid at states near the efficient states that will ultimately arise. Consequently, if in computing $\varepsilon_{\min}$ we restrict attention to utilization levels near those of efficient states, then $\varepsilon_{\min}$ will typically be strictly positive. Thus, at the states that are actually likely to arise, tolls will be non-negative and smaller than those resulting from marginal cost pricing.

PMC tolls which yield revenue r at this state. In particular, efficiency and budget balance can always be achieved.

6.2 Decentralized Solution of Network Decision Problems

So far, we have concerned ourselves exclusively with games played by continuous sets of players. However, congestion is also a common problem for monopolists who must transport their products over networks. Consider, for example, a utility company which must provide customers with electricity over a power grid. Given consumer demands, it would like to do so in a way which minimizes total transport costs.

Our results imply that a monopolist can use a simple method to find and maintain an efficient routing for its goods.[23] The monopolist imposes a purely internal PMC congestion toll. Whenever a particular bundle of goods could be transported at a lower toll-augmented cost by switching its route, the routing is gradually adjusted until the improvement opportunities vanish. Theorem 5.3 implies that this myopic improvement process always lowers aggregate transportation costs. If the original cost functions satisfy the conditions of Theorem 6.1, this procedure must lead to a globally optimal routing choice. If not, refining the adjustment procedure can guarantee locally and even globally efficient outcomes.[24] Either way, our efficiency results provide a decentralized method of both finding efficient route allocations and maintaining efficiency in the face of variations in available routes and supplies, transport costs, and customer demands.

7. Potential Games with Elastic Demand

To this point we have implicitly assumed that all players from each population p must choose one of the strategies from S^p. However, if all strategies are quite costly, some players might prefer to opt out of the game entirely. For example, if the game models route selection, lengthy driving times and high tolls might cause drivers to

---

[23] Note that if more than one supply source exists, we can specify populations by destination rather than by origin/destination pairs so as to internalize the choice of sources.

[24] If the firm occasionally makes small, random adjustments in the routing choice, by applying Theorem 5.4 we see that these perturbations will serve to disrupt any inefficient equilibrium, thereby guaranteeing local efficiency. In addition, if the monopolist initially solves for the globally optimal routing, decentralized adjustment can maintain the global optimum so long as changes in the underlying model do not cause dramatic shifts in the cost structure (e.g., the globally optimal routing choice).

abandon the commute. To capture this possibility, we could simply add a strategy whose payoff equals some minimal acceptable value independently of the behavior of the population. But more generally, we would like to allow different players to have different reservation values for playing the game. In this section, we show how elastic demand can be incorporated into potential games, and explain how our previous results can be extended to this case.

To define the demand functions, we let I^p be a closed interval of real numbers representing possible payoffs in population p. Let $Q = (Q^1, \ldots, Q^r)$: $\prod_p I^p \to \prod_p [0, m^p]$ denote the vector of demands in each population as functions of the payoffs in all populations. We assume that Q is C^1 and invertible, and that its derivative matrix is always of full rank. Furthermore, we assume that "cross-price" effects are symmetric: $\frac{\partial Q^i}{\partial \pi^j} \equiv \frac{\partial Q^j}{\partial \pi^i}$ for all $i, j \in P$. Finally, we let $\tilde{Q} = (\tilde{Q}^1, \ldots, \tilde{Q}^r)$: $\prod_p [0, m^p] \to \prod_p I^p$ denote the inverse demand function.

To introduce elastic demand to potential games, we add strategies representing outside options. For each population p, we let the augmented strategy space equal $\tilde{S}^p = \{0\} \cup S^p$, where strategy 0 represents the outside option. We define the payoffs to the new strategy in terms of the inverse demand function:

$$F_0^p(x) = \tilde{Q}^p(m^1 - x_0^1, \ldots, m^r - x_0^r).$$

$\tilde{Q}^p(m^1 - x_0^1, \ldots, m^r - x_0^r)$ is the payoff received by the marginal player who leaves the game. But while $\tilde{Q}^p$ only represents the marginal player's payoffs, using it as the "payoff function" for the outside option is nevertheless the appropriate way to introduce elastic demand. If there exists a strategy $i \in S^p$ such that $F_i^p > \tilde{Q}^p$, there are players currently using the outside option who would prefer to use strategy i. If all strategies $j \in S^p$ satisfy $F_j^p < \tilde{Q}^p$, then there are players in the game who would prefer to switch to the outside option. Only when $F_j^p \leq \tilde{Q}^p$ for all $j \in S^p$ with equality for at least one j is demand calibrated with the available payoffs.

The elastic demand game we have defined admits a potential function. We establish this by checking that externality symmetry holds. First, observe that $\frac{\partial F_i^p}{\partial x_0^q} \equiv 0 \equiv \frac{\partial F_0^q}{\partial x_i^p}$ for all $i \in S^p$ and all $p, q \in P$. Second, since the original game is a potential game, we see that $\frac{\partial F_i^p}{\partial x_j^q} \equiv \frac{\partial F_j^q}{\partial x_i^p}$. Finally, we claim that $\frac{\partial F_0^p}{\partial x_0^q} \equiv \frac{\partial F_0^q}{\partial x_0^p}$. To establish this, let DV_x denote the derivative of the vector field F at the point x. The claim follows from the observations that $D\tilde{Q}_q = (DQ_{\tilde{Q}(q)})^{-1}$ (i.e., that the derivative of the inverse demand function equals the inverse of the derivative of the demand function) and that the inverse of a symmetric matrix is symmetric.

While our results from the previous sections can be extended to this general setting, to simplify the discussion we restrict attention to congestion games and to demand functions which only depend on own-population payoffs (so that $\tilde{Q}^p(m^1 - x_0^1, \ldots, m^r - x_0^r) = \tilde{Q}^p(m^p - x_0^p)$ for all $p \in P$). In congestion games, facility costs are well defined for all population sizes in the interval $[0, m^p]$, and hence payoff functions are well defined on the new strategy space $\tilde{\mathcal{X}}$, which allows for the outside option. The potential function for this elastic demand game is

$$f(x) = -\sum_{o \in \Phi}\left(\int_0^{u_o(x)} c_o(z)\,dz\right) - \sum_{p \in P}\left(\int_0^{m^p - x_0^p} \tilde{Q}^p(z)\,dz\right).$$

It is easily checked that the equilibria of this game are those points which satisfy the Kuhn-Tucker first order conditions for a maximum, and that equilibrium is unique if congestion costs are increasing and the law of demand holds.[25]

All of our results on evolutionary dynamics continue to hold in the elastic demand setting. The obvious extension of net monotonicity to this context has a natural interpretation: it requires that better performing strategies tend to prosper, and that players tend to switch to and from their outside option as their preferences dictate. Our results generalize immediately: the potential function continues to be a Lyapunov function under net monotone dynamics, and the local and global stability results still obtain.

The isoelasticity conditions from Section 5 continue to yield behavior which is efficient *conditional on realized demand*. To see this, consider an equilibrium of an elastic demand game whose payoff functions (other than the inverse demand function) are isoelastic. If we fix demand levels at their equilibrium values, the game that remains is isoelastic; hence, equilibrium behavior must be efficient conditional on these levels of demand. An analogous argument establishes that placing PMC tolls on the facilities yields efficiency conditional on realized demand.

While a social planner designing tolls for a highway network clearly desires conditional efficiency, he also is likely to have preferences concerning realized

[25] Since the demand functions are functions of payoffs rather than prices, we must require them (and hence the inverse demand functions) to be *increasing*.

demand itself. If demand is very high, even the conditionally efficient allocation may result in severe congestion. On the other hand, excessive tolls inordinately depress demand, preventing productive economic activity.

Fortunately, in choosing the elasticity threshold a social planner can influence aggregate demand. Recall that PMC tolls are parameterized by their elasticity threshold. Increasing this threshold reduces tolls, which makes continuing to drive a more desirable option. While under inelastic demand this directly affects aggregate toll revenues, under elastic demand the immediate effect is on the attractiveness of the outside option to the marginal driver. By choosing the elasticity threshold appropriately, a social planner can strike a balance between congestion delays and the tolls' dampening effects.[26]

8. Potential Games as Limits of Finite Player Games

In this section, we establish the connections between potential games with a continuum of players and the finite player potential games of Monderer and Shapley (1996). We begin by defining quasipotential games, which are extensions of continuous potential games that admit a potential function characterization of equilibrium and evolution. We then prove that quasipotential games are precisely the limits of finite player potential games which satisfy an anonymity condition.

For notational convenience, we restrict attention to the case of single population games; all of the results which follow can be extended to the multipopulation case.

8.1 Quasipotential Games

In Section 5.2, we noted that subtracting a constant from all players payoff functions did not affect their incentives: for each fixed profile of opponents' actions, the relative payoffs to a player's strategies are unchanged. More generally, suppose we fix a particular action profile for a player's opponents and add a constant to the player's payoffs at all strategy profiles consistent with the specified profile of opponents' behavior. Once again, the player's incentives are unaffected. Consequently, an analysis of evolutionary dynamics would not be affected by such a change.

---

[26] Since the toll elasticity is scalar valued while the levels of demand in each population form a vector, a social planner can only choose the levels of demand from a one parameter family.

We call any game that can be created by modifying a potential game with such payoff shifts *quasipotential game*. Let a *shift function*, $\sigma: \mathscr{S} \to \mathbf{R}$, be a continuous function from the simplex to the real line. A *shift vector*, $\Sigma: \mathscr{S} \to \mathbf{R}^n$, is a vector valued function whose components are all the same shift function: $\Sigma_i \equiv \sigma$ for all $i \in S$. A population game $\hat{F}$ is a quasipotential game if $\hat{F} \equiv F + \Sigma$ for some potential game F and some shift vector Σ.

Since adding a shift vector to payoffs does not affect the set of Nash equilibria, Proposition 3.1 extends to quasipotential games: the Nash equilibria of the quasipotential game are precisely the points which satisfy the Kuhn-Tucker conditions for a maximum of the potential function. Furthermore, it is easy to see that the potential function f associated with any potential game F will serve as a Lyapunov function under net monotone dynamics for any quasipotential game derived from F. Since $V \cdot \bar{1} \equiv 0$,

$$\hat{F} \cdot V \equiv (F + \Sigma) \cdot V \equiv (F \cdot V) + \sigma(\bar{1} \cdot V) \equiv F \cdot V.$$

Thus, if V is net monotone with respect to $\hat{F}$ and $\dot{x} = V(x)$, then

$$\tfrac{d}{dt} f(x_t) = \nabla f(x_t) \cdot \dot{x}_t = F(x_t) \cdot V(x_t) = \hat{F}(x_t) \cdot V(x_t) \geq 0,$$

with equality only when $V(x_t) = 0$. Hence, by our results in Section 4, the potential function of the potential game F also describes evolution in any quasipotential game derived from F. It is worth noting, however, that the efficiency results from Section 5 need only carry over if the shift function is constant; other shifts perturb average payoffs in a way which is not captured by potential.

8.2 Anonymous Finite-Player Potential Games

We describe a finite player game by a triple $\{I, \{S_\alpha\}_{\alpha \in I}, \{u_\alpha\}_{\alpha \in I}\}$, where $I = \{1, 2, ..., N\}$ denotes the set of players, and S_α and u_α represent player α's pure strategy set and payoff function, respectively. The payoff functions are defined on the set of pure strategy profiles $S = \prod_{\alpha \in I} S_\alpha$. For each $\alpha \in I$, let $S_{-\alpha} = \prod_{\beta \neq \alpha} S_\beta$. Following Monderer and Shapley (1996), we call a game a *finite player potential game* if there exists a function $P: S \to \mathbf{R}$ such that

$$u_\alpha(s'_\alpha, s_{-\alpha}) - u_\alpha(s_\alpha, s_{-\alpha}) = P(s'_\alpha, s_{-\alpha}) - P(s_\alpha, s_{-\alpha})$$

for all s_α, $s'_\alpha \in S_\alpha$, $s_{-\alpha} \in S_{-\alpha}$, and $\alpha \in I$. That is, any unilateral deviation has the same effect on both the deviator's payoffs and the potential function. The potential function serves as proxy for each player's payoff function when the strategies of his opponents are held fixed. It is easily verified that potential games can also be characterized as the class of games which admit the representation

$$u_\alpha(s_\alpha, s_{-\alpha}) = P(s) + a_\alpha(s_{-\alpha}),$$

where for each $\alpha \in I$, a_α is a function from opponents' strategy profiles $S_{-\alpha}$ to the real line.

Since in continuous potential games the players are anonymous, we restrict attention to finite potential games in which players' payoff functions can be expressed as a function of their action and their opponents' aggregate behavior. Furthermore, since we are focusing on continuous potential games with a single population, we only consider games in which all players' payoff functions are identical.[27] More precisely, we first assume that all players strategy sets are identical: $S_\alpha = S_\beta$ for all α, $\beta \in I$. Define the function x to map strategy profiles $s \in S$ (and, with a slight abuse of notation, $s_{-\alpha} \in S_{-\alpha}$) into their distributions in the simplex $\mathcal{S}$. Then the anonymity conditions imply that each player's payoff function can be expressed in the form

$$u_\alpha(s_\alpha, s_{-\alpha}) = P(x(s)) + a(x(s_{-\alpha})),$$

where the potential function is now a map from a discrete version of the simplex to the real line, as is the *adjustment function*, a.

This expression for player α's payoffs only depends on the strategy distribution, $x(s)$, and player α's strategy choice, s_α (which is used to determine $x(s_{-\alpha})$); it does not depend directly on the identity of player p. Consequently, we can represent this game with anonymous payoff functions u_i: $\mathcal{S} \rightarrow \mathbf{R}$ which for each strategy $i \in S$ express the payoffs to strategy $i \in S$ given overall strategy distribution $x \in \mathcal{S}$. Letting e_i denote the i^{th} unit vector, we write

$$u_i(x) = P(x) + a\left(\tfrac{Nx - e_i}{N-1}\right).$$

---

[27] More generally, in the multipopulation case, all players within each population must have identical payoff functions.

We call any finite player game which admits this representation an *anonymous finite-player potential (AFP) game.*

We now show that despite their rather different definitions, quasipotential games are precisely the limits of sequences of AFP games in which the number of players approaches infinity. To prove this result, we need to assume that the sequences of potential functions and payoff functions converge to limit functions. We must therefore extend the domains of definition of each from a discrete version of strategy space to the entirety of strategy space. From here onward, each potential function is assumed to be a C^1 function on $\mathcal{S}$, and each payoff function is continuous on $\mathcal{S}$.

Fixing a strategy set S, let $\left\{\left(\left\{u_i^N\right\}_{i\in S}, P^N\right)\right\}_{N=n}^{\infty}$ be a sequence of AFP games whose population sizes approach infinity. We say that the sequence *converges* to the limit game $\left(\left\{u_i\right\}_{i\in S}, P\right)$ if (i) $\frac{1}{N}P^N \xrightarrow{C^1} P$, (ii) $u_i^N \xrightarrow{C^0} u_i$ for all $i \in S$, and (iii) there exists a K such that all u_i^N and all $\frac{1}{N}\frac{\partial P^N}{\partial x_i}$ are Lipschitz continuous with Lipschitz constant K.[28] Since the potential function in finite player games aggregates payoffs, in condition (i) we rescale potential by the number of players so that we can compare the potential functions of games with different numbers of players. Condition (ii) requires the uniform convergence of the payoff functions. Condition (iii) is a smoothness condition which is satisfied in most examples.[29] Using this natural notion of convergence, we can state our limit result.

Theorem 8.1: *If* $\left\{\left(\left\{u_i^N\right\}_{i\in S}, P^N\right)\right\}_{N=n}^{\infty}$ *is a convergent sequence of AFP games, then its limit,* $\left(\left\{u_i\right\}_{i\in S}, P\right)$*, is a quasipotential game. That is, there exists a shift function* σ *such that for all* $i \in S$, $u_i \equiv \frac{\partial P}{\partial x_i} + \sigma$.

Proof: In the Appendix.

In AFP games, potential functions are only defined on the strategy space, $\mathcal{S}$. But in order for the partial derivative mentioned in the theorem to be well-defined, the limit potential function P must be defined on the wedge $\mathcal{W}$. Therefore, to make

---

[28] The C^0 norm of a continuous function defined on a compact domain is the supremum norm: $\|h\|_{C^0} = \sup_x h(x)$. The C^1 norm of a continuously differentiable function is the sum of its C^0 norm and the C^0 norms of its partial derivatives.

[29] Alternatively, one could use a stronger notion of convergence requiring that the rescaled potential functions converge in C^2 and that the utility functions converge in C^1.

sense of the statement of the theorem, we must extend each P^N to $\mathcal{X}$ in such a way that the extended functions continue to converge in the C^1 norm. The limit function will then be C^1 function on $\mathcal{X}$. Importantly, *any* convergent sequence of extensions yields a quasipotential game as its limit. Our proof establishes that given any such convergent sequence, the function

$$\sigma(x) = \lim_{N \to \infty}\left(a^N(x) + P^N\left(\tfrac{N-1}{N}x\right)\right)$$

exists and is the shift function for the limit quasipotential game. Consequently, the limit games arising from different extensions will differ in their shift functions σ. This observation also follows from noting that the unshifted payoffs $\frac{\partial P}{\partial x_i}$ depend on the extensions of P^N but must ultimately yield the same shifted payoff functions $u_i = \frac{\partial P}{\partial x_i} + \sigma$ regardless of the extension.

The converse of Theorem 8.1 is also true: given a any quasipotential game with potential function P (defined on $\mathcal{X}$) and shift function σ, the proof of Theorem 8.1 implies that the sequence of AFP games defined by setting $P^N(x)$ equal to $NP(x)$ and $a^N(x)$ equal to $\sigma(x) - P(\frac{N-1}{N}x)$ converges to the original quasipotential game. In other words, every quasipotential game is the limit of an appropriate sequence of AFP games, and may therefore be viewed as an approximation to a large, finite player potential game.

Finally, one can combine Theorem 8.1 with our results from Section 5 to establish conditions under which evolution in large AFP games will yield approximate efficiency. Suppose that a sequence of AFP games converges to an isoelastic potential game: that is, we can choose extensions of the potential functions P^N such that the limit potential function is homogenous of positive degree and the limit shift function is identically zero (or, more generally, a constant function). Theorem 8.1 then implies that for any AFP game far enough along the sequence, the potential function P^N will be approximately proportional to average payoffs. Therefore, in these large, finite games, evolution leads to approximate efficiency.

9. Conclusion

We investigate equilibrium, evolution, and efficiency in games satisfying externality symmetry. All such games admit a potential function, whose gradient equals the payoff functions of the game. Since any reasonable evolutionary process must increase potential, we are able to establish results on both the local and global stability of equilibrium. We provide conditions under which evolution yields efficiency, and exhibit pricing schemes which induce efficiency in cases in which it does not naturally arise. Finally, we show that our class of continuous games is the limiting case of the discrete potential games of Monderer and Shapley (1996).

The examples developed throughout this paper illustrate our model's fitness as a framework for studying externalities in large populations. We hope that future research on externalities will benefit from the tools developed here.

Appendix

A.1 Evolutionary Dynamics: Definitions and Auxiliary Results

Most of the definitions we require are stated in Section 4; only those which were omitted are provided here. A *neighborhood* of a closed set $A \subset \mathcal{X}$ is a set which is open relative to $\mathcal{X}$ and contains A. A closed set of fixed points $A \subset \mathcal{X}$ is *Lyapunov stable* if every neighborhood B of A contains a neighborhood B' of A such that $\{x_t\}_{t \geq 0} \subset B$ for all $x \in B'$. That is, solutions starting at all points sufficiently close to A always remain close to A. A closed set of fixed points $A \subset \mathcal{X}$ is *asymptotically stable* if it is Lyapunov stable and there exists a neighborhood B of A such that $\omega(x) \subset A$ for all $x \in B$: in other words, solutions starting from all points close enough to A remain nearby and eventually converge to A. The existence of a global Lyapunov function allows a strong characterization of the limit behavior of a dynamical system, as the next proposition shows. Its proof can be found in Losert and Akin (1983, Proposition 1) and Robinson (1995, Theorem 5.4.1).

Lemma A.1: *If $f: \mathcal{X} \to \mathbf{R}$ is a global Lyapunov function for $\dot{x} = V(x)$, then each limit set $\omega(x)$ is a non-empty, compact, and connected set consisting entirely of fixed points of V and upon which f is constant.*

Global Lyapunov functions also form the basis for a sufficient condition for Lyapunov stability. The following lemma follows immediately from a result of Bhatia and Szegö (1970, Corollary 3.3.5); also see Weibull (1995, Theorem 6.4).

Lemma A.2: *If $f: \mathscr{S} \to \mathbf{R}$ is a global Lyapunov function for $\dot{x} = V(x)$ and A is a local maximizer set of f, then A is Lyapunov stable.*

To prove asymptotic stability results, we need slightly stronger conditions. We call a C^1 function $f: \mathscr{S} \to \mathbf{R}$ a *strict local Lyapunov function* for the set A under $\dot{x} = V(x)$ if (*i*) A is a local maximizer set of f, and (*ii*) there exists a neighborhood B of A such that $\frac{d}{dt} f(x_t) > 0$ whenever $x \in B - A$. The existence of a strict local Lyapunov function for the set A implies its asymptotic stability. The proof of this result can be found in Hale (1969, Theorem 10.1.1); also see Weibull (1995, Theorem 6.4).

Lemma A.3: *If $f: \mathscr{S} \to \mathbf{R}$ is a strict local Lyapunov function for A, then A is asymptotically stable.*

A.2 Proofs Omitted from the Text

Proof of Corollary 3.2: The first claim follows immediately from Proposition 3.1. To prove the second claim, consider a congestion game whose payoff functions are constructed from weakly increasing cost functions c_ϕ. Choose $x, y \in \mathscr{S}$; by the first claim, it is enough to show that the potential function $f(x) = -\sum_{\phi \in \Phi} \int_0^{u_\phi(x)} c_\phi(z)\,dz$ is concave on the line segment joining x and y.

For $\lambda \in [0, 1]$, let $v(\lambda) = \lambda x + (1 - \lambda)y$, and let $v_\phi(\lambda) = u_\phi(\lambda x + (1 - \lambda)y)$. If the cost functions are differentiable, then

$$\frac{d^2}{(d\lambda)^2} f(v(\lambda)) = -\sum_{\phi \in \Phi} c_\phi'(v_\phi(\lambda))\left(v_\phi'(\lambda)\right)^2,$$

which is non-positive since $c_\phi' \geq 0$. If the cost functions are not differentiable, the result is proved in a similar fashion by considering the left and right hand derivatives of c_ϕ. ■

Proof of Theorem 4.3: We first require the following lemma.

Lemma A.4: *The potential function f is constant on any smoothly connected set of Nash equilibria of F.*

Proof: Let A be a smoothly connected set of Nash equilibria of F, and let y and z be elements of A. Then there exists a piecewise smooth function γ: $[0, 1] \to A \subset \mathcal{Y}$ with $\gamma(0) = y$ and $\gamma(1) = z$. Since

$$f(z) - f(y) = \int_\gamma \nabla f \cdot d\bar{x} = \int_0^1 [F(\gamma(t)) \cdot \gamma'(t)]dt,$$

it is sufficient to show that the integrand $F(\gamma(t)) \cdot \gamma'(t)$ equals zero at all points at which γ is differentiable.

Let t be a point of differentiability of γ. For each $p \in P$, define $\mu^p(x) = \max_i F_i^p(x)$ and $\lambda_i^p(x) = \mu^p(x) - F_i^p(x)$ for all $x \in A$. Observe that these functions are continuous, and that $F_i^p(x) = \mu^p(x) - \lambda_i^p(x)$. Hence,

$$\begin{aligned} F(\gamma(t)) \cdot \gamma'(t) &= \sum_{p \in P} \left(F^p(\gamma(t)) \cdot (\gamma^p)'(t) \right) \\ &= \sum_{p \in P} \left(\mu^p(\gamma(t)) [\bar{1} \cdot (\gamma^p)'(t)] + [\lambda^p(\gamma(t)) \cdot (\gamma^p)'(t)] \right). \end{aligned}$$

Since $\gamma \subset \mathcal{Y}$, $\bar{1} \cdot (\gamma^p)'(t) = 0$ for each $p \in P$, so the first term in each summand is identically zero. To prove this is also true of each second term, note that since all points on γ are Nash equilibria, Proposition 3.1 implies that $x_i^p\ \lambda_i^p(x) = 0$ for all $x \in \gamma$. Furthermore, that $\gamma \subset \mathcal{Y}$ implies that $(\gamma_i^p)'(t) = 0$ if $\gamma_i^p(t) = 0$. Consequently, $\lambda_i^p(\gamma(t))$ $(\gamma_i^p)'(t) = 0$ for all $i \in S^p$ and all $p \in P$, completing the proof. ❑

We continue with the proof of Theorem 4.3. Part (*i.a*) follows immediately from Lemma A.2 and Lemma 4.1. To prove part (*i.b*), let A be a smoothly connected asymptotically stable set. Then there is a neighborhood B of A such that $V(x) \neq \bar{0}$ and $\omega(x) \subset A$ for all $x \in B - A$. Net monotonicity implies that $\frac{d}{dt}f(x_t) = \nabla f(x_t) \cdot \dot{x}_t = F(x_t) \cdot V(x_t) > 0$ for all $x_t \in B - A$. Applying Lemma A.4, let c be the unique value taken by f on A. Then there cannot exist a $z \in B - A$ such that $f(z) \geq c$, as this would contradict that $\frac{d}{dt}f(x_t)\big|_{x_t = z} > 0$ and $\omega(z) \subset A$. Therefore, since A is smoothly connected it is a local maximizer set.

To prove the remaining results, we show that $(ii.c) \Rightarrow (ii.b) \Rightarrow (ii.a) \Rightarrow (ii.c)$. The first implication, that asymptotic stability implies Lyapunov stability, is true by definition. To prove the second implication is true, let A be a smoothly connected, isolated, Lyapunov stable set, and suppose to the contrary that A is not a local maximizer set. Since A is isolated, we can find a neighborhood C of A which contains no Nash equilibria outside A. Because C is an open neighborhood of A, we can choose a second neighborhood B of A such that the (Hausdorff) distance between B and $\mathcal{S} - C$ is strictly positive. Finally, since A is Lyapunov stable, there exists a third neighborhood $B' \subset B$ of A such that all solutions starting in B' never leave B.

Applying Lemma A.4 once again, let c be the unique value which f takes on A. Since A is not a local maximizer set, there exists a point $x \in B'$ such that $f(x) > c$. By Lemma A.1, the limit set of x, $\omega(x)$, consists entirely of rest points of the dynamics. No point in A can be in $\omega(x)$, as this would violate net monotonicity. Moreover, C contains no Nash equilibria, and therefore, since V satisfies non-extinction, no rest points. Consequently, $\omega(x) \subset \mathcal{S} - C$. But since B is bounded away from $\mathcal{S} - C$, this contradicts that the orbit of x never leaves B. Hence, A must be a local maximizer set. This proves that $(ii.b) \Rightarrow (ii.a)$.

Finally, we prove that $(ii.a) \Rightarrow (ii.c)$. Let A be an isolated local maximizer set. Since A is isolated, there is a neighborhood B of A such that all Nash equilibria in B are in A. Hence, non-extinction implies that $V(x) \neq 0$ for all $x \in B - A$, and so net monotonicity implies that $\frac{d}{dt} f(x_t) > 0$ for all $x_t \in B - A$. Thus, since A is a local maximizer set of f, f is a strict local Lyapunov function for A. Lemma A.3 then implies that A is asymptotically stable. ■

Proof of Theorem 6.1: We need only prove the first claim, as in its presence the second claim follows from previous results. Assume that each facility cost function c_ϕ satisfies $u c_\phi''(u) \geq -2c_\phi'(u)$ for all utilization levels u. Rewrite the aggregate payoff function as

$$\begin{aligned}
\bar{F}(x) &= \sum_p \sum_i x_i^p F_i^p(x) \\
&= -\sum_p \sum_i x_i^p \sum_{\phi \in \Phi_i^p} c_\phi(u_\phi(x)) \\
&= -\sum_\phi u_\phi(x) c_\phi(u_\phi(x)).
\end{aligned}$$

Choose $x, y \in \mathscr{S}$. It is enough to show that the aggregate payoff is concave on the line segment joining x and y.

For $\lambda \in [0, 1]$, let $v(\lambda) = \lambda x + (1 - \lambda)y$, and let $v_\phi(\lambda) = u_\phi(\lambda x + (1 - \lambda)y)$. Differentiating twice, we see that

$$\frac{d^2}{(d\lambda)^2}\bar{F}(v(\lambda)) = -\sum_{\phi \in \Phi}\left(v_\phi'(\lambda)\right)^2\left(v_\phi(\lambda)c_\phi''(v_\phi(\lambda)) + 2c_\phi'(v_\phi(\lambda))\right).$$

Substituting u for $v_\phi(\lambda)$, we see that this expression is non-positive whenever $uc_\phi''(u) \geq -2c_\phi'(u)$. Hence, $\bar{F}$ is concave. ■

Proof of Theorem 8.1: We are given a convergent sequence of AFP games $\left\{\left(\{u_i^N\}_{i \in S}, P^N\right)\right\}_{N=n}^{\infty}$. To prove our result, we must first extend the potential functions so that they converge in the C^1 sense on $\mathscr{H}$. Any extension will do; for concreteness, one can think of the extension which is homogenous of degree zero.

Let a^N be the adjustment function for the N player AFP game. Since the payoff functions u_i^N and the potential functions P^N are continuous functions on $\mathscr{S}$, we can take the adjustment functions a^N to be continuous on $\mathscr{S}$. For all large enough N, define

$$\begin{aligned} \sigma^N(x) &= a^N(x) + P^N\left(\tfrac{N-1}{N}x\right) \text{ and} \\ \sigma(x) &= \lim_{N \to \infty} \sigma^N(x). \end{aligned}$$

We prove that the limit game $\left(\{u_i\}_{i \in S}, P\right)$ is a quasipotential game by establishing that σ is its shift function: $u_i \equiv \frac{\partial P}{\partial x_i} + \sigma$ for all $i \in S$. We begin by showing that the limit which defines σ exists.

Observe that

$$\begin{aligned} u_i^N(x) &= P^N(x) + a^N\left(\tfrac{Nx - e_i}{N-1}\right) \\ &= P^N(x) - P^N\left(x - \tfrac{e_i}{N}\right) + P^N\left(x - \tfrac{e_i}{N}\right) + a^N\left(\tfrac{Nx - e_i}{N-1}\right) \\ &= P^N(x) - P^N\left(x - \tfrac{e_i}{N}\right) + \sigma^N\left(\tfrac{Nx - e_i}{N-1}\right). \end{aligned} \tag{1}$$

If we let $\sigma_i^N(x) = \sigma^N\left(\frac{Nx - e_i}{N-1}\right)$, then

$$u_i^N(x) = P^N(x) - P^N\left(x - \tfrac{e_i}{N}\right) + \sigma_i^N(x). \tag{2}$$

To prove that the σ_i^N converge, we require the following lemma.

Lemma A.5: *The functions* $\left(P^N(x) - P^N\left(x - \frac{e_i}{N}\right)\right)$ *converge uniformly to* $\frac{\partial P}{\partial x_i}$ *as* $N \to \infty$.

Proof: Since each P^N is C^1, the Mean Value Theorem implies that for some $z_i^N(x)$ on the linear segment connecting x and $x - \frac{e_i}{N}$,

$$P^N(x) - P^N\left(x - \tfrac{e_i}{N}\right) = \nabla P^N(z_i^N(x)) \cdot \tfrac{1}{N} e_i,$$

or, alternatively,

$$P^N(x) - P^N\left(x - \tfrac{e_i}{N}\right) = \tfrac{1}{N} \tfrac{\partial P^N}{\partial x_i}(z_i^N(x)). \tag{3}$$

Since the functions $\frac{1}{N}\frac{\partial P^N}{\partial x_i}$ are uniformly Lipschitz with constant K, and since $\left|z_i^N(x) - x\right| \le \frac{1}{N}$, we see that

$$\left|\tfrac{1}{N}\tfrac{\partial P^N}{\partial x_i}(z_i^N(x)) - \tfrac{1}{N}\tfrac{\partial P}{\partial x}(x)\right| \le \tfrac{K}{N} \text{ for all } x \in \mathscr{X} \tag{4}$$

Equation (3) implies that

$$\begin{aligned}
&\max_{x \in \mathscr{X}}\left|\left(P^N(x) - P^N\left(x - \tfrac{e_i}{N}\right)\right) - \tfrac{\partial P}{\partial x_i}(x)\right| \\
&\quad \le \max_{x \in \mathscr{X}}\left(\left|\tfrac{1}{N}\tfrac{\partial P^N}{\partial x_i}(z_i^N(x)) - \tfrac{1}{N}\tfrac{\partial P^N}{\partial x_i}(x)\right| + \left|\tfrac{1}{N}\tfrac{\partial P^N}{\partial x_i}(x) - \tfrac{\partial P}{\partial x_i}(x)\right|\right),
\end{aligned}$$

But equation (4) and the fact that $\frac{1}{N}\frac{\partial P^N}{\partial x_i} \xrightarrow{C^0} \frac{\partial P}{\partial x_i}$ imply that the right hand side of the above expression approaches zero as N approaches infinity. This completes the proof of the lemma. ❑

We continue by noting the following Lipschitz-like inequality for the functions $\left(P^N(x) - P^N\left(x - \frac{e_i}{N}\right)\right)$. Applying the uniform Lipschitz condition satisfied by the functions $\frac{1}{N}\frac{\partial P^N}{\partial x_i}$ and using the definition of z_i^N above, we see that

$$\begin{aligned}
&\left|\left(P^N(x) - P^N\left(x - \tfrac{e_i}{N}\right)\right) - \left(P^N(y) - P^N\left(y - \tfrac{e_i}{N}\right)\right)\right| \\
&\quad = \left|\tfrac{1}{N}\tfrac{\partial P^N}{\partial x_i}(z_i^N(x)) - \tfrac{1}{N}\tfrac{\partial P^N}{\partial x_i}(z_i^N(y))\right| \\
&\quad \le \left|\tfrac{1}{N}\tfrac{\partial P^N}{\partial x_i}(z_i^N(x)) - \tfrac{1}{N}\tfrac{\partial P^N}{\partial x_i}(x)\right| + \left|\tfrac{1}{N}\tfrac{\partial P^N}{\partial x_i}(x) - \tfrac{1}{N}\tfrac{\partial P^N}{\partial x_i}(y)\right| + \left|\tfrac{1}{N}\tfrac{\partial P^N}{\partial x_i}(y) - \tfrac{1}{N}\tfrac{\partial P^N}{\partial x_i}(z_i^N(y))\right|
\end{aligned}$$

$$\leq K\left(\left|x-y\right|+\tfrac{2}{N}\right). \tag{5}$$

By the definition of convergence for a sequence of AFP games, $u_i^N \xrightarrow{C^0} u_i$. This fact, Lemma A.5, and equation (2) imply that the σ_i^N converge uniformly on $\mathscr{S}$. Call the limit function σ_i.

We claim that $\sigma_i \equiv \sigma_j$ for all $i, j \in S$. To see this, suppose to the contrary that $\sigma_i(z) > \sigma_j(z) + \varepsilon$ for some z and some $\varepsilon > 0$. Then for all large enough N, $\sigma_i^N(z) > \sigma_j^N(z) + \frac{\varepsilon}{2}$. The definition of σ_i^N then implies that for all large enough N,

$$\left|\sigma^N\left(\tfrac{Nz-c_i}{N-1}\right)-\sigma^N\left(\tfrac{Nz-c_j}{N-1}\right)\right| > \tfrac{\varepsilon}{2}. \tag{6}$$

For any x, let $w_k^N(x) = \frac{(N-1)x+c_k}{N}$. Applying equation (1) and inequality (5), and recalling that the u_i^N satisfy a uniform Lipschitz condition, we see that

$$\begin{aligned}
&\left|\sigma^N(x)-\sigma^N(y)\right| \\
&= \Big|\Big(u_k^N(w_k^N(x))-\big(P^N(w_k^N(x))-P^N\big(w_k^N(x)-\tfrac{c_k}{N}\big)\big)\Big) \\
&\qquad -\Big(u_k^N(w_k^N(y))-\big(P^N(w_k^N(y))-P^N\big(w_k^N(y)-\tfrac{c_k}{N}\big)\big)\Big)\Big| \\
&\leq \left|u_k^N(w_k^N(x))-u_k^N(w_k^N(y))\right| + \Big|\big(P^N(w_k^N(x))-P^N\big(w_k^N(x)-\tfrac{c_k}{N}\big)\big) \\
&\qquad -\big(P^N(w_k^N(y))-P^N\big(w_k^N(y)-\tfrac{c_k}{N}\big)\big)\Big| \\
&\leq 2K\left(\left|w_k^N(x)-w_k^N(y)\right|+\tfrac{1}{N}\right) \\
&\leq 2K\left(\left|x-y\right|+\tfrac{2\sqrt{2}+1}{N}\right).
\end{aligned}$$

This contradicts inequality (6). Therefore, the σ_i must all identical: $\sigma_i \equiv \sigma$ for all $i \in S$.

It is easily verified that σ is the shift function defined at the beginning of the proof. To complete the proof, we need to show that for all $i \in S$, $u_i \equiv \frac{\partial P}{\partial x_i} + \sigma$. But in light of our computations above, this identity follows from taking the limit of equation (2) as N approaches infinity. ■

References

Akin, E., and J. Hofbauer (1982). "Recurrence of the Unfit," *Math. Biosci.* **61**, 51-62.

Arnott, R., and K. A. Small (1994). "The Economics of Traffic Congestion," *American Scientist* **82**, 446-455.

Aumann, R., and A. Brandenburger (1995). "Epistemic Conditions for Nash Equilibrium," *Econometrica* **63**, 1161-1180.

Barro, R. J., and P. M. Romer (1987). "Ski-Lift Pricing, with Applications to Labor and Other Markets," *Amer. Econ. Rev.* **77**, 875-890.

Beckmann, M., C. B. McGuire, and C. B. Winsten (1956). *Studies in the Economics of Transportation*. New Haven: Yale University Press.

Bhatia, N. P., and G. P. Szegö (1970). *Stability Theory for Dynamical Systems*. Berlin: Springer-Verlag.

Binmore, K., and L. Samuelson. (1997). "Evolutionary Drift and Equilibrium Selection," mimeo, University College London and University of Wisconsin.

Braess, D. (1968). "Über ein Paradoxen der Verkehrsplanung," *Unternehmens-forschung* **12**, 258-268.

Cowan, S. (1992). "Dynamical Systems Arising from Game Theory," Ph.D. dissertation, University of California, Berkeley.

Dafermos, S., and F. T. Sparrow (1969). "The Traffic Assignment Problem for a General Network," *J. Res. Nat. Bur. Standards B* **73**, 91-118.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.

Foster, D., and R. Vohra (1996). "Calibrated Learning and Correlated Equilibrium," mimeo, University of Pennsylvania.

Friedman, D. (1991). "Evolutionary Games in Economics," *Econometrica* **59**, 637-666.

Friedman, E. J. (1996). "Dynamics and Rationality in Ordered Externality Games," *Games Econ. Behav.* **16**, 65-76.

Friedman, E. J. (1997). "Learnability in a Class of Non-Atomic Games Arising on the Internet," mimeo, Rutgers University.

Gaunersdorfer, A., and J. Hofbauer (1995). "Fictitious Play, Shapley Polygons, and the Replicator Equation," *Games Econ. Behav.* **11**, 279-303.

Gilboa, I., and A. Matsui (1991). "Social Stability and Equilibrium," *Econometrica* **59**, 859-867.

Gomez-Ibanez, J. A., and K. A. Small (1994). *Road Pricing for Congestion Management: A Survey of International Practice.* Washington: National Academy Press.

Hale, J. (1969). *Ordinary Differential Equations.* New York: Wiley.

Hofbauer, J. (1995). "Stability for the Best Response Dynamics," mimeo, Universität Wien.

Hofbauer, J., and K. Sigmund (1988). *Theory of Evolution and Dynamical Systems.* Cambridge: Cambridge University Press.

Losert, V., and E. Akin (1983). "Dynamics of Games and Genes: Discrete versus Continuous Time," *J. Math. Biology* **17**, 241-251.

Matsui, A. (1992). "Best Response Dynamics and Socially Stable Strategies," *J. Econ. Theory* **57**, 343-362.

Maynard Smith, J. (1974). "The Theory of Games and the Evolution of Animal Conflicts," *J. Theor. Biol.* **47**, 209-221.

Maynard Smith, J., and G. Price (1973). "The Logic of Animal Conflict," *Nature* **246**, 15-18.

Monderer, D., and L. S. Shapley (1996). "Potential Games," *Games Econ. Behav.* **14**, 124-143.

Milgrom, P. and J. Roberts (1991). "Adaptive and Sophisticated Learning in Normal Form Games," *Games Econ. Behav.* **58**, 82-100.

Nachbar, J. H. (1990). "'Evolutionary' Selection Dynamics in Games: Convergence and Limit Properties," *Int. J. Game Theory* **19**, 59-89.

Pas, E. I., and S. L. Principio (1997). "Braess' Paradox: Some New Insights," *Transpn. Res. B* **31**, 265-276.

Ritzberger, K., and J. W. Weibull (1995). "Evolutionary Selection in Normal Form Games," *Econometrica* **63**, 1371-1399.

Robinson, C. (1995). *Dynamical Systems.* Boca Raton, Fl.: CRC Press.

Rosenthal, R. W. (1973). "A Class of Games Possessing Pure Strategy Nash Equilibria," *Int. J. Game Theory* **2**, 65-67.

Samuelson, L., and J. Zhang (1992). "Evolutionary Stability in Asymmetric Games," *J. Econ. Theory* **57**, 363-391.

Swinkels, J. M. (1993). "Adjustment Dynamics and Rational Play in Games," *Games Econ. Behav.* **5**, 455-484.

Taylor, P., and L. Jonker (1978). "Evolutionary Stable Strategies and Game Dynamics," *Math. Biosci.* **40**, 145-156.

Vickrey, W. S. (1969). "Congestion Theory and Transport Investment," *Amer. Econ. Rev., Papers and Proceedings* **59**, 251-260.

Walters, A. A. (1961). "The Theory and Measurement of Private and Social Cost of Highway Congestion," *Econometrica* **29**, 676-699.

Weibull, J. (1995). *Evolutionary Game Theory*. Cambridge: MIT Press.