PRELIMINARY AND INCOMPLETE
Please do not cite without permission.

A NETWORK CENTRALITY CHARACTERIZATION OF ROBUST COOPERATION

MATT ELLIOTT* AND BENJAMIN GOLUB**

Abstract. In global climate talks, negotiations are multilateral, and reducing emissions is individually costly but provides (different) benefits to (some) other countries. Similarly, in teams, multilateral deviations are a natural concern and a team member's efforts are individually costly but provide (different) benefits to (some) other team members. In these environments, are there equilibria of repeated interaction that are robust to potential coalitional deviations? If so, how can they be characterized in terms of the network of benefits that agents can confer on each other? Are these outcomes efficient? To address these problems, we set up an infinitely repeated simultaneous-move game in which agents can take actions that are individually costly but provide benefits to other agents, allowing any group of agents to collude and form a deviating coalition. We find the answers to these questions are deeply tied to questions and results from classical general equilibrium theory. Equilibria do exist, they are all efficient, and a class of such equilibria can be found by reducing the problem to a well-studied general equilibrium problem and characterizing the competitive equilibria in terms of a single spectral condition related to a network centrality measure.

Keywords: general equilibrium, core, competitive equilibrium, repeated games, coalitional deviations, strong Nash equilibrium, spectral graph theory, centrality.

1. Introduction

Seminal work by Aumann (1959) defined *strong Nash equilibria* as ones robust to coalitional deviations and noted that their analysis and even their existence presented substantially harder problems than the analogous ones under unilateral-deviation equilibrium concepts. To our knowledge, despite some notable advances discussed below, the subsequent work on characterizing coalitionally robust outcomes has been less extensive than the analogous folk theorem literature for Nash equilibria of repeated games.

Seeking a richer understanding coalitionally robust solutions, we study strong Nash equilibria of the infinite repetition of a stage game in which each agent chooses an action from a continuum. Without any restrictions, Aumann showed that strong Nash equilibria of such a repeated game need not exist. To supply some structure that ensures

Date Printed. September 28, 2011.
*Microsoft Research and Department of Economics, Stanford University.
Email: matt.elliott@stanford.edu, web: http://www.stanford.edu/~mle1.
**Department of Economics, MIT and Graduate School of Business, Stanford University.
Email: bgolub@mit.edu, web: http://www.stanford.edu/~bgolub.
We thank Kyle Bagwell, Matt Jackson, Muriel Niederle, Andy Skrzypacz, Moshe Tennenholtz and Bob Wilson for their comments and suggestions. Elliott gratefully acknowledges financial support from the Davidow fellowship administered through SEIPR. Golub gratefully acknowledges financial support from the Lieberman fellowship at Stanford and the Prize Fellowships in Economics, History, and Politics at Harvard University.

their existence, we assume that increasing one's action is costly, but provides (heterogeneous) benefits to other agents. There are many important environments in which this holds, and where coalition formation is natural: examples include collaboration on a joint project by members of a team, as well as negotiations among countries about the provision of a public good such as tariff reductions or pollution limits. In each of these cases, the benefits of an agent's contribution may be distributed heterogeneously among the others, depending, e.g., on the underlying production function, trade flows, or geography.

By combining results of Aumann (1959), Shapley-Shubik (1969), and Rubinstein (1980), we can show that, under these assumptions, strong Nash equilibria of the repeated game exist and, moreover, can be sustained by threats that are immune to coalitional renegotiation. We then seek to characterize the equilibria, as well as the efficient outcomes, in ways that are useful theoretically, computationally, and empirically.

Our characterizations are built on applying spectral theory to a certain network of benefit flows. The heterogeneous marginal benefits yielded by different agents' effort investments naturally induce a network structure among the agents: the weight of a link from one agent to another captures how much the former can benefit the latter per unit of cost at the margin. The first theorem asserts that, under some conditions, actions are on the Pareto frontier if and only if the matrix of link weights evaluated at those actions (called the *gift matrix*) has a largest eigenvalue (spectral radius) of exactly 1. This condition corresponds to a natural intuition. Suppose an agent increases his contribution slightly, and all agents who benefit "pass forward" all the resulting benefits by increasing their actions – and so on. If more benefits are eventually returned to the first agent than her up-front investment of effort, then a Pareto improvement has been constructed. This occurs precisely when the system is exploding at the margin – when its largest eigenvalue is greater than 1. Efficient outcomes are obtained by exhausting all such money pumps.

To analyze coalitionally robust outcomes, we use results from Shapley and Shubik (1969) to reduce our problem to a general equilibrium (GE) problem, showing that competitive equilibria of a certain auxiliary economy are strong Nash equilibria of our repeated game. Going beyond Shapley and Shubik (1969), we then construct a class of strong Nash equilibria by chracterizing the competitive equilibria of the GE problem – this time, through a condition on a special eigen*vector* (an eigenvector centrality) of the gift matrix. While competitive equilibria can (by definition) be characterized by a system of equations, the key contribution here is to show that, in our setting, these equations boil down to a single condition that can be interpreted and applied in several important ways. This result also connects – to our knowledge, for the first time – the study of outcomes that are robust to coalitional deviations with the study of network centrality.

A recent literature on network centrality has related Nash equilibria in one-shot games to spectral and centrality conditions under linear-quadratic utilities; key papers in this literature include Balaster, Calvó-Amengol, and Zenou (2006) and Bramoullé, Kranton, and d'Amours (2010). By studying a different problem with group deviations, we have a new structure to exploit, and find different spectral conditions characterizing robust outcomes of repeated, rather than one-shot, interactions. Furthermore, to obtain the spectral conditions in this different setting, we need only concave utility functions instead of the linear-quadratic setup previously utilized. This is because a mapping to a general equilibrium problem permits us to formally reduce the problem to a price-theoretic one focused on benefits evaluated locally, at the margin.

2. Model and Definitions

2.1. **The Game.** We begin by defining a stage game Γ, in which each member of a set $N = \{1, 2, \ldots, n\}$ of players simultaneously chooses an action[1] $a_i \in \mathbb{R}_{\geq 0}$, the set of nonnegative real numbers. Player i's stage game payoff is denoted[2] $u_i(\boldsymbol{a})$. As a normalization, we assume that $\boldsymbol{u}(\mathbf{0}) = \mathbf{0}$. The repeated game $\Gamma^*(\delta)$ is one in which each agent takes an action a_{ti} in each of infinitely many discrete periods t, and where the utility U_i of player i is the discounted sum of his stage game payoffs: $U_i \equiv \sum_{t=0}^{\infty} \delta^t u_i(\boldsymbol{a}_t)$, for some $\delta < 1$. We assume all players have complete and symmetric information.

2.2. **Equilibrium Concept and Sustainable Actions.** Our focus will be on outcomes of the repeated game that are robust to the possibility of coalitional deviations. To make this formal, we define the standard notion of a *strong Nash equilibrium* (Aumann 1959).

Definition. A *strong Nash equilibrium* of a game G played by the set of players N is a strategy profile $\boldsymbol{\sigma}$ of the game G such that there is no coalition $M \subseteq N$ and no other strategy profile $\boldsymbol{\sigma}'$ of this game so that:

(i) $\sigma_i' = \sigma_i$ for all $i \notin M$;
(ii) each $i \in M$ strictly prefers $\boldsymbol{\sigma}'$ to $\boldsymbol{\sigma}$.

Following Rubinstein (1980), we call a strategy profile $\boldsymbol{\sigma}$ a *strong perfect Nash equilibrium* of an extensive-form game G if it is a strong Nash equilibrium of every subgame of G.

The key object of study in this paper is the set of *sustainable* action vectors, which we define as follows.

Definition. An action vector $\boldsymbol{a}$ of the stage game Γ is *sustainable* if, for high enough $\delta < 1$, the infinite repetition of $\boldsymbol{a}$ occurs on the path of play of some strong perfect Nash equilibrium $\boldsymbol{\sigma}$ of the repeated game $\Gamma^*(\delta)$.

2.3. **Assumptions on Utilities.** In general, sustainable action vectors do not always exist (Aumann 1959). To make the study of these objects tractable, some assumptions on the game are useful. We assume that $u_i : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function and make five further assumptions on the stage game payoffs u_i.

Assumption (Costly Actions)**.** For every $i \in N$, agent i's finds it costly, at the margin, to invest effort: $\frac{\partial u_i}{\partial a_i}(\boldsymbol{a}) < 0$.

Assumption (Positive Externalities)**.** Increasing any player's action level weakly benefits all other players: $\frac{\partial u_i}{\partial a_j} \geq 0$ for all $j \neq i$.

Assumption (Strictly Concave Payoffs)**.** Each utility function $u_i : \mathbb{R}^n_{\geq 0} \to \mathbb{R}$ is strictly concave.

Assumption (Existence of Interior Optima)**.** For any $\boldsymbol{\theta} \in \mathbb{R}^n_{\geq 0}$, the function defined by $\mathbf{a} \mapsto \sum_i \theta_i u_i(\mathbf{a})$ achieves a maximum over the domain $\mathbb{R}^n_{>0}$ of entrywise positive action vectors.

[1]The analysis can be extended to multiple actions with only bookkeeping complexities.
[2]We sometimes refer to vectors of actions and utilities, so that in the stage game actions $\boldsymbol{a}$ yield payoffs $\boldsymbol{u}(\boldsymbol{a})$.

Assumption (Connectedness of Benefit Flows)**.** There do not exist an action profile $\boldsymbol{a} \in \mathbb{R}^n_{\geq 0}$ and a nonempty $M \subseteq N$ so that for every $i \in M$ and $j \notin M$, we have $\frac{\partial u_i}{\partial a_j}(\boldsymbol{a}) = 0$.

We view the positive externalties assumption as the most restrictive. Nonetheless, there is a broad class of economic applications for which this restriction is reasonable. When team members increase their effort levels, they provide (different) benefits to the other team members; when countries agree to limit their greenhouse gas emissions, they typically provide benefits to other countries. The World Trade Organization (WTO) provides the institutional setting for multilateral tariff reduction negotiations. In a companion paper (Elliott and Golub 2011) we show how multiple institutional features of the WTO combine to ensure that any country's tariff reductions typically only benefit other countries.

We do not view the assumption on the connectedness of benefit flows as restrictive. The purpose of the assumption is to ensure that a key matrix in our analysis is irreducible. Were this matrix to be reducible at equilibrium actions levels, the statements of our results would change only slightly: we would need to state conditions on each irreducible block of the matrix.

2.4. **Key Notions for the Analysis of Sustainable Actions.**

2.4.1. *The Jacobian and Scaling-Indifference.*

To analyze sustainable action profiles, some additional definitions are useful. We define the Jacobian $\boldsymbol{J}(\boldsymbol{a})$ of the game Γ to be the n-by-n matrix whose (i, j) element is

$$J_{ij}(\boldsymbol{a}) = \frac{\partial u_i}{\partial a_j}(\boldsymbol{a}).$$

Definition. The game Γ satisfies *scaling-indifference* at an action profile $\boldsymbol{a}$ if $\boldsymbol{J}(\boldsymbol{a})\,\boldsymbol{a} = \boldsymbol{0}$.

Why is this condition called scaling-indifference? Note that the vector $\boldsymbol{J}(\boldsymbol{a})\,\boldsymbol{v}$ gives the marginal changes in utilities when actions are changed from $\boldsymbol{a}$ to $\boldsymbol{a} + \epsilon\boldsymbol{v}$ for some vector $\boldsymbol{v} \in \mathbb{R}^n$ and some small real number ϵ. That is, to a first-order approximation, $\boldsymbol{u}(\boldsymbol{a}+\epsilon\boldsymbol{v}) \approx \boldsymbol{u}(\boldsymbol{a})+\epsilon\boldsymbol{J}(\boldsymbol{a})\boldsymbol{v}$. Suppose now that actions are scaled by $1+\epsilon$, for some small real number ϵ (this corresponds to setting $\boldsymbol{v} = \boldsymbol{a}$). If the scaling-indifference condition holds and $\boldsymbol{J}(\boldsymbol{a})\,\boldsymbol{a} = \boldsymbol{0}$, all agents are indifferent, at the margin, to this small proportional perturbation in everyone's actions.

2.4.2. *The Gift Matrix and Virtual Costs.*

Our main results can be stated using only the notions above (see Section 4.2). However, the proofs and certain intuitive interpretations are aided by two additional notions. The first of these is the *gift matrix*, a close relative of the Jacobian. The second is an alternative way of measuring actions.

The *gift matrix* $\boldsymbol{G}$ is defined as follows:

$$G_{ij}(\boldsymbol{a}) = \begin{cases} \frac{J_{ij}(\mathbf{a})}{-J_{ii}(\mathbf{a})} & \text{if } i \neq j \\ 0 & \text{otherwise.} \end{cases}$$

The quantity G_{ij} captures the marginal contribution j makes to i's pleasure per unit of j's pain.[3]

[3]One might be concerned that the definition of this matrix is affected by economically meaningless rescalings of some individuals' utilities; it turns out, reassuringly, that such rescalings do not change any conditions in the results.

Finally, we define the *virtual cost* of agent i at an action profile $\boldsymbol{a}$ as

$$v_i(\boldsymbol{a}) = -a_i \cdot \frac{\partial u_i}{\partial a_i}(\boldsymbol{a}).$$

One interpretation of this is as follows. If i were a price-taker choosing his labor supply and being paid in utiles, then his total labor income at action profile $\boldsymbol{a}$ would be $v_i(\boldsymbol{a})$.

3. Pareto Frontier

Theorem 1. Under the assumptions above, the action profile $\boldsymbol{a}$ is a Pareto-efficient outcome of the stage game if and only if 1 is a largest eigenvalue of $\boldsymbol{G}(\boldsymbol{a})$.

Some intuition for Theorem 1 can be obtained by considering the directed graph representation of $\boldsymbol{G}(\boldsymbol{a})^T$, where the superscript T represents that $\boldsymbol{G}(\boldsymbol{a})$ is transposed. In the directed graph, the players are the vertices and the weights of directed edges are given by the elements of $\boldsymbol{G}(\boldsymbol{a})^T$. The (i, j) element of $\boldsymbol{G}(\boldsymbol{a})^T$ reflects the normalized marginal benefits that i provides to j by increasing his action, and this is represented by a weighted directed edge from i to j.

Consider now the following exercise. Suppose agent i increases her action at the margin. This imposes a marginal cost on i. However, the increase in i's action also generates benefits at the margin for other agents, as captured by the directed graph. Suppose these agents pass along all these benefits by taking higher actions such that they are just as well-off as they were before i increased her action. Repeat this process *ad infinitum*, with the exception that agent i does not pass on any further benefits and instead accumulates the benefits that flow back to her. If this process generates a flow of marginal benefits to i greater than her initial investment, then a Pareto improvement has been constructed: i's utility has increased and no one else's utility has decreased. This positive net flow to i can be viewed as a "money pump" at the margin. Person i puts some "money" into the system and more "money" is returned to him, without anyone else having to make a net contribution. If the marginal benefits were held constant as actions changed, then i could generate infinite utility by utilizing this money pump. The system explodes at the margin in this way precisely when the largest eigenvalue of $\boldsymbol{G}(\boldsymbol{a})$ is greater than one. Unfortunately, by the assumption that interior optima exist, the money pump must eventually be exhausted – precisely when the largest eigenvalue of $\boldsymbol{G}(\boldsymbol{a})$ is exactly 1. A similar intuition shows that a Pareto improvement is possible by some agent reducing her implemented action if the largest eigenvalue of $\boldsymbol{G}(\boldsymbol{a})$ is less than one.

4. Sustainable Outcomes

Having characterized the efficient outcomes, we can now proceed with the study of ones that are sustainable even under the possibility of coalitional deviations. The results can be stated in two essentially equivalent ways. One way uses the gift matrix, and the other way uses the notion of scaling-indifference.

4.1. Gift Matrix Formulation.

Theorem 2. Under the assumptions above:

(i) If $\boldsymbol{G}(\boldsymbol{a})\,\boldsymbol{v}(\boldsymbol{a}) = \boldsymbol{v}(\boldsymbol{a})$ then $\boldsymbol{a}$ is sustainable.
(ii) There exists an $\boldsymbol{a} \in \mathbb{R}^n_{\geq 0}$ so that $\boldsymbol{G}(\boldsymbol{a})\,\boldsymbol{v}(\boldsymbol{a}) = \boldsymbol{v}(\boldsymbol{a})$.
(iii) Any sustainable action profile is Pareto efficient.

Part (i) of Theorem 2 identifies a sufficient condition for a sustainable outcome and part (ii) shows that this sufficient condition is not vacuous. This sufficient condition requires that the virtual costs be a right eigenvector of the gift matrix with an associated eigenvalue of 1. From Theorem 1 we therefore know that this sustainable outcome is on the Pareto frontier. Part (iii) of Theorem 2 asserts that all sustainable outcomes are on the Pareto frontier. This last part is not especially surprising, as at every point not on the Pareto frontier there is profitable (static) deviation for the grand coalition.

To gain some intuition for the sufficient condition identified in Theorem 2, note that, if this condition holds, then for each agent i we have:

$$v_i = \sum_{j \neq i} G_{ij} v_j. \tag{1}$$

Equation 1 shows that at the sustainable outcome we identify, each agent's virtual cost is a weighted sum of the other agents' virtual costs, where these the weight on v_j is proportional to the marginal benefits that j provides to i. This emphasizes that our sufficient condition is a centrality condition – the vector of virtual costs is a (right-hand) eigenvector centrality of the gift matrix. While centrality conditions have been used in applications such as Google's PageRank, and in sociological indices of prestige (Jackson 2008), to the best of our knowledge we are the first to identify outcomes robust to coalitional deviations in a game using a centrality condition. The other applications of centrality conditions in game theory we are aware of have characterized equilibria in games with only individual deviations and under the restriction that payoff functions are linear-quadratic. In contrast, we require no parametric restrictions for our result, though we do need some assumptions on the utility functions (recall Section 2.3).

4.2. Scaling-Indifference Formulation.

Additional intuition can be gained by doing a little algebra to show that $\boldsymbol{G}(\boldsymbol{a})\,\boldsymbol{v}(\boldsymbol{a}) = \boldsymbol{v}(\boldsymbol{a})$ if and only if $\boldsymbol{J}(\boldsymbol{a})\,\boldsymbol{a} = \boldsymbol{0}$. This alternative representation of our sufficient condition is the scaling-indifference condition we define above in Section 2.4.1. At such an action vector, all agents are indifferent to a proportional increase (or decrease) in everyone's effort levels at the margin. In Section 6, it is shown that this can be used to find a gradual path to a sustainable outcome.

Moreover, it is easy to see by following the proof of Theorem 2 that the Pareto-efficiency of an action profile is equivalent to $\boldsymbol{J}(\boldsymbol{a})$ having an eigenvalue of 0, which yields an alternative spectral condition for efficiency.

5. Proof of the Main Results by Reduction to a General Equilibrium Problem

5.1. Proof of Theorem 1.

Consider the Pareto problem. A social planner can implement any set of actions and places the vector of nonnegative Pareto weights θ_i, not all equal to 0, on different agent's payoffs. The Pareto problem is then:

$$\max_{\boldsymbol{a} \in \mathbb{R}^n_{\geq 0}} \sum_{i=1}^{n} \theta_i u_i(\boldsymbol{a}).$$

It is well known that concavity of all the u_i guarantees that the set of Pareto-efficient points coincides with the set of solutions to this problem for different vectors $\boldsymbol{\theta}$.

Suppose that an action vector $\boldsymbol{a}^*$ solves this problem. The assumptions of (i) concave utilities and (ii) the existence of interior optima (as stated in Section 2.3) guarantee that this problem has a unique solution $\boldsymbol{a}^*$, which is interior and is characterized by the first order condition, namely $\boldsymbol{\theta J}(\boldsymbol{a}^*) = \mathbf{0}$. Using the definition of $\boldsymbol{G}$ this is equivalent to:

$$\boldsymbol{\theta}\left(\boldsymbol{G}(\boldsymbol{a}^*) - \mathbf{I}\right) = \mathbf{0}$$
$$\boldsymbol{\theta G}(\boldsymbol{a}^*) = \boldsymbol{\theta}$$

where $\mathbf{I}$ is the identity matrix. Thus $\boldsymbol{G}(\boldsymbol{a}^*)$ has an eigenvalue of 1. Since $\boldsymbol{G}(\boldsymbol{a}^*)$ is irreducible, has only nonnegative elements and the eigenvector $\boldsymbol{\theta}$ is nonnegative, the Perron-Frobenius Theorem guarantees that 1 is a largest eigenvalue of $\boldsymbol{G}(\boldsymbol{a}^*)$.

Conversely, if $\boldsymbol{G}(\boldsymbol{a}^*)$ has a largest eigenvalue of 1, then the Perron-Frobenius Theorem guarantees the existence of a nonnegative eigenvector $\boldsymbol{\theta}$ such that $\boldsymbol{\theta G}(\boldsymbol{a}^*) = \boldsymbol{\theta}$. Tracing backwards through the equivalences in the previous argument (recall that the first-order conditions are sufficient to show an optimum by concavity of the problem), we find that $\boldsymbol{a}^*$ solves the Pareto problem for Pareto weights $\boldsymbol{\theta}$.

5.2. **Proof of Theorem 2.** The proof proceeds in four steps, using some new concepts and an auxiliary game that will be defined in the corresponding sections.

(1) Without loss of generality, the word "perfect" may be dropped from the definition of sustainable action profiles without changing the set of sustainable action profiles (Rubinstein 1980).
(2) Once the word "perfect" has been dropped, the set of sustainable action profiles is the same as the β-core of the game Γ (Aumann 1959).
(3) The β-core of Γ is the same as the core of an associated artificial competitive economy E in which all the external benefits produced by agents are separately tradeable, rather than delivered automatically to the beneficiaries (Shapley-Shubik 1969).
(4) Competitive equilibria of the economy E exist and are in the core (by standard facts about competitive equilibrium – see, e.g. Mas-Colell, Whinston and Green 1995, Chapter 18). These competitive equilibria are characterized by $\boldsymbol{G}(\boldsymbol{a})\,\boldsymbol{v}(\boldsymbol{a}) = \boldsymbol{v}(\boldsymbol{a})$.

5.2.1. *Step 1.* Recall the definition of sustainable actions.

Definition. An action vector $\boldsymbol{a}$ of the stage game Γ is *sustainable* if for high enough $\delta < 1$, the infinite repetition of $\boldsymbol{a}$ occurs on the path of play of some strong perfect Nash equilibrium $\boldsymbol{\sigma}$ of the repeated game $\Gamma^*(\delta)$.

We define a weaker concept by dropping the word "perfect" from this definition.

Definition. An action vector $\boldsymbol{a}$ of the stage game Γ is *weakly sustainable* if for high enough $\delta < 1$, the infinite repetition of $\boldsymbol{a}$ occurs on the path of play of some strong Nash equilibrium $\boldsymbol{\sigma}$ of the repeated game $\Gamma^*(\delta)$.

Weakly sustainable action profiles may be enforced by the threat of painful grim-trigger punishments – ones taking everyone to autarky – that are not subgame-perfect, in the sense that the grand coalition of all players could find an improvement that everyone would strictly prefer to carrying out the punishments. Thus, the set of *weakly* sustainable outcomes is a superset of the sustainable outcomes. Nevertheless, a remarkable theorem of Rubinstein shows that this seemingly greater range of punishments does not actually add to the set of outcomes that can be sustained. (A key part of this equivalence is that δ is taken very close to 1 in the definition of sustainability.)

Lemma 1. The set of sustainable action profiles is the same as the set of weakly sustainable action profiles.

The proof of this lemma is a straightforward adaptation of the arguments in Rubinstein (1980). The main differences are that he works in utility space rather than action space, and that he focuses on average payoffs per action rather than discounted payoffs as δ tends to 1. However, these differences do not change the nature of the arguments in our setting.

5.2.2. *Step 2.*

Definition. An action profile $\boldsymbol{a}$ is in the β-core of the stage game Γ if there is no coalition $M \subseteq N$ and no other action profile $\boldsymbol{a}'$ of this game so that:

(i) $a'_i = 0$ for all $i \notin M$;
(ii) each $i \in M$ strictly prefers $\boldsymbol{\sigma}'$ to $\boldsymbol{\sigma}$.

Note that this defines the β-core in action space rather than, as is usual, in utility space. The formulation here is more convenient for our purposes than the standard one but involves no essential conceptual adjustments.

More generally the β-core is defined by players outside a deviating coalition punishing the deviating coalition by seeking to minimize their payoffs. In our environment players most severely punish a deviating coalition by choosing an action level of 0.[4]

Lemma 2. An action profile $\boldsymbol{a}$ is weakly sustainable if and only if it is in the β-core of Γ.

The proof of this lemma is a straightforward adaptation of the arguments in Aumann (1959). The main differences are that he works in utility space rather than action space, and that he focuses on average payoffs per action rather than discounted utility as δ tends to 1. However, these changes do not affect the main point of the argument, which is simply this: in the β-core concept, we envision society punishing deviations by taking action 0. An action profile is in the β-core if and only if a potential deviating coalition is deterred by this punishment. In a repeated game, society would implement the punishment by playing 0 in future periods. (Note that this does not have to be subgame-perfect, as we are now discussing weakly sustainable action vectors.) As $\delta \to 1$, the consequences of this punishment, as evaluated by a potential deviating coalition, become equivalent.

[4]In our environment the β-core also coincides with the α-core. The difference between the two is whether the punishment is implemented by non-deviating players in advance of the deviators' choice or after observing it. To define the α-core, we imagine the deviating coalition first choosing their actions to maximize their payoffs and then the other players choosing actions to minimize the payoffs of the deviating coalition given what has happened; to define the β-core, we imagine the non-deviating players first choosing their actions to minimize the deviating coalition's payoffs, and then the deviators choosing actions to maximize their payoffs given that. As actions levels of zero always minimize the payoffs of all members of a deviating coalition in Γ, the order of the moves does not matter.

5.2.3. *Step 3.* In this step, we construct an artificial competitive economy and show that its core is the same as the β-core of Γ. Define the economy E as follows:

- Each agent i is endowed with a large supply of labor, called good i.
- The (publicly available) production technology allows the following, for each $j \in N$:
  - put in one unit of good j;
  - get out one unit each of goods $1j, 2j, 3j, \ldots, nj$.
  - As will become clear in a moment, good ij is valued only by agent i.
- These output goods are separately tradeable.
- In an outcome of this economy:
  - set $\hat{a}_i$ equal to the total amount of labor i puts in;
  - set $\hat{a}_{ij}$ equal to the total amount of j's output good ij that i buys and consumes;
  - set utility of i to

$$\nu_i = u_i(\hat{a}_{i1}, \hat{a}_{i2}, \ldots, \hat{a}_i, \ldots, \hat{a}_{in}).$$

Here, $\hat{a}_i$ appears between $\hat{a}_{i[i-1]}$ and $\hat{a}_{i[i+1]}$.

Thus each agent chooses how much labor to supply and then exchanges the goods he produces with the other agents in a market economy to maximize his utility.

Lemma 3. The core of the economy E is the same as the β-core of the game with externalities Γ.

The proof of this lemma is a straightforward adaptation of the arguments in Shapley and Shubik 1969.

The main difference between the economy E and the game Γ is that the positive externalities present in Γ have been turned into goods and made tradable. Thus, in the economy E agents can withhold the "positive externalities" they produce but in the game Γ they cannot. It is therefore trivial that if there is a deviating coalition in the game Γ that this deviating coalition can be replicated in the economy E. Going the other way, the argument is as follows. Without loss of generality any coalition that deviates in E might as well allocate all the goods they produce in a Pareto-efficient way. Within the coalition this can only increase the profitability of the deviation and as the other goods produced are valued by no one inside the coalition, there is nothing lost from freely distributing them to agents outside of the coalition. Thus, whenever there is a profitable deviating coalition there is also a profitable deviating coalition that distributes the produced goods efficiently. However, any such coalitional deviation can be replicated by a deviating coalition in the game Γ.

5.2.4. *Step 4.* Assign prices $\boldsymbol{P}$ to all produced goods ij and to labor. Specifically, let $p_{ij} \geq 0$ be the price i pays to j for the good ij (when $i \neq j$) and let $p_{ii} \leq 0$ be minus the wage i is paid per unit of work.

A *competitive equilibrium* (CE) of the economy E is defined by prices $\boldsymbol{P}$, labor supplies $\{\hat{a}_i\}_{i=1}^{N}$, and an allocation of produced goods ij such that certain conditions hold. Before stating these, it is helpful to recall that equilibrium, the allocation of produced goods is efficient (by the First Welfare Theorem), which means that $\hat{a}_j = \hat{a}_{ij}$ for every $i \neq j$. Thus, we may denote all these quantities by a_j in equilibrium. (Having constructed such a vector $\boldsymbol{a}$, we will say that the competitive equilibrium *corresponds to* $\boldsymbol{a}$.) With this simplification of notation, an equilibrium is defined by the following conditions:

(1) Budget balance: $\boldsymbol{P}\,\boldsymbol{a} = \boldsymbol{0}$.
(2) Utility maximization/the marginal rates of substitution (MRS) condition: for all i, j, k we have $\frac{p_{ij}}{p_{ik}} = \frac{J_{ij}(\boldsymbol{a})}{J_{ik}(\boldsymbol{a})}$.
(3) Market wage: for all i, we have $p_{ii} = -\sum_{j \neq i} p_{ji}$.

The market wage condition asserts that the wage each agent receives for his labor is equal to the sum of prices paid for his output. Note that as each good is valued by just one agent, utility maximization and the market wage conditions are sufficient for ensuring the market clears and all produced goods ij are allocated to i.

The setup of economy E is such that a standard equilibrium existence theorem can be applied.[5] From the core property of competitive equilibria we know that all competitive equilibria are in the core and we can then apply Step 3 to show that the β-core of Γ is non-empty. Step 2 then implies that there exists a weak sustainable outcome and by Step 1 there exists a sustainable outcome. Reducing our repeated game to economy E is sufficient for guaranteeing existence. We now constructively find a class of sustainable outcomes by characterizing the competitive equilibria of the economy E. This is the key lemma of the paper.

Lemma 4. The action vector $\boldsymbol{a}$ corresponds to a competitive equilibrium of the economy E if and only if $\boldsymbol{G}(\boldsymbol{a})\,\boldsymbol{v}(\boldsymbol{a}) = \boldsymbol{v}(\boldsymbol{a})$.

Proof. Note that the condition $\boldsymbol{G}(\boldsymbol{a})\,\boldsymbol{v}(\boldsymbol{a}) = \boldsymbol{v}(\boldsymbol{a})$ is equivalent to the condition $\boldsymbol{J}(\boldsymbol{a})\,\boldsymbol{a} = \boldsymbol{0}$, so we will use them interchangeably.

First we show the *only if* direction: if $\boldsymbol{a}$ corresponds to to a competitive equilibrium allocation, then $\boldsymbol{J}(\boldsymbol{a})\,\boldsymbol{a} = \boldsymbol{0}$.

Suppose we are given a competitive equilibrium (and the associated prices and allocations). We will fix these quantities and drop the argument $\boldsymbol{a}$ on the Jacobian and other quantities that depend on on $\boldsymbol{a}$. Define $\boldsymbol{F}$ as the diagonal matrix with $F_{ii} = \frac{J_{ii}}{p_{ii}}$. From budget balance, we have:

$$\begin{aligned} \boldsymbol{P}\,\boldsymbol{a} &= \boldsymbol{0} \\ \Leftrightarrow \quad \boldsymbol{F}\boldsymbol{P}\,\boldsymbol{a} &= \boldsymbol{0} \end{aligned}$$

Using the MRS condition, we have that $\frac{p_{ij}}{p_{ik}} = \frac{J_{ij}}{J_{ik}}$, so the above is equivalent to:

$$\boldsymbol{J}\,\boldsymbol{a} = \boldsymbol{0}.$$

Now we show the *if* direction: if $\boldsymbol{J}\,\boldsymbol{a} = \boldsymbol{0}$, then $\boldsymbol{a}$ corresponds to to a competitive equilibrium allocation. Fix such an $\boldsymbol{a}$, and again drop all the arguments.

We seek non-negative prices such that competitive equilibrium conditions hold. Let us guess: $p_{ij} = \gamma_i J_{ij}$, for some parameters $\gamma_i > 0$ to be determined later. Equivalently $\boldsymbol{P} = \boldsymbol{D}\,\boldsymbol{J}$, where $\boldsymbol{D}$ is a diagonal matrix with $D_{ii} = \gamma_i$.

These prices $\boldsymbol{P}$ automatically satisfy the MRS condition: $\frac{p_{ij}}{p_{ik}} = \frac{J_{ij}}{J_{ik}}$.

[5]See, for example, Mas-Colell, Whinston and Green (1995). The concavity of the utility functions is essential here.

It can also be shown that these prices $\boldsymbol{P}$ satisfy the budget balance condition $\boldsymbol{P}\,\boldsymbol{a} = \boldsymbol{0}$, as follows:

$$\begin{aligned} & \boldsymbol{J}\,\boldsymbol{a} &= \boldsymbol{0} \\ \Leftrightarrow \quad & \boldsymbol{D}\boldsymbol{J}\,\boldsymbol{a} &= \boldsymbol{0} \\ \Leftrightarrow \quad & \boldsymbol{P}\,\boldsymbol{a} &= \boldsymbol{0} \end{aligned}$$

Finally, we show that price $\boldsymbol{P}$ satisfy the market wage condition. Since we have that $\boldsymbol{G}\boldsymbol{v} = \boldsymbol{v}$, that $\boldsymbol{v}$ is nonzero, and that $\boldsymbol{G}$ is irreducible, we can apply the Perron-Frobenius Theorem to obtain a positive vector $\boldsymbol{b}$ such that $\boldsymbol{b}\,\boldsymbol{G} = \boldsymbol{b}$. From this it follows immediately that there is a row vector $\boldsymbol{\gamma}$ such that $\boldsymbol{\gamma}\boldsymbol{J} = \boldsymbol{0}$.

Recall that $\boldsymbol{D}$ is the diagonal matrix with $D_{ii} = \gamma_i$, so that $\boldsymbol{J} = \boldsymbol{D}^{-1}\,\boldsymbol{P}$. Thus:

$$\gamma\,\boldsymbol{D}^{-1}\,\boldsymbol{P} \;\; = \;\; \boldsymbol{0}$$

Let $\boldsymbol{1}$ be the vector of ones, so that $\boldsymbol{\gamma}\,\boldsymbol{D}^{-1} = \boldsymbol{1}$. Thus:

$$\boldsymbol{1}\,\boldsymbol{P}(\boldsymbol{a}) = \boldsymbol{0} \tag{2}$$

This is the matrix version of the market wage condition. We have shown that when $\boldsymbol{J}\,\boldsymbol{a} = \boldsymbol{0}$, there exist prices such that utility is maximized, the market wage condition is satisfied and budgets are balanced. Thus $\boldsymbol{J}\,\boldsymbol{a} = \boldsymbol{0}$ implies a competitive equilibrium of the economy E. □

6. Discussion

We study repeated games in which agents can take costly actions to create benefits for other agents – we assume there are positive externalities. In order to do so we introduce two important concepts. The gift matrix describes the marginal benefits an agents can provide to another agent per unit of marginal costs he incurs. An agent's virtual cost is his marginal cost from increasing his action multiplied by his action level. We show that the Pareto frontier is characterized by the spectral radius of the gift matrix – the actions chosen by the players must induce a largest eigenvalue of exactly one in the gift matrix. When the associated right eigenvector of the gift matrix also corresponds to agents' virtual costs, then there are no profitable coalitional deviations and these actions are sustainable. Furthermore, actions always exist for which agents' virtual costs are a right eigenvector of the gift matrix with an associated eigenvalue of 1. To the best of our knowledge no other paper has used centrality conditions to describe equilibria robust to coalitional deviations or to characterize a Pareto frontier. We are able to do so without making parametric restrictions. Taking a similar approach to analyzing matrices derived from other first order conditions in other settings may prove fruitful.

Our condition only directly identifies a measure zero set of the actions that are sustainable. However, the point we identify is interior. To find our sustainable outcome we mapped our problem into a standard general equilibrium problem and then looked for competitive equilibria. This involved constructing (artificial) prices. It is the linearity of these prices that restrict us to identifying only a zero measure set of sustainable outcomes. By introducing (small) non-linearities to prices we can find additional sustainable outcomes. Formally, we could consider reparameterizing the action space of each agent, constructing new actions $\tilde{a}_i$ equal to some nonlinear function of the old actions, and then implement all the same analysis. In changing what "units" of action mean, we can allow agents to charge nonlinear prices in the old unit system. The only thing we must be careful of is to ensure that this transformation of action spaces preserves the concavity in our problem. Assuming the initial utility functions u_i are strictly concave, there exist reparametrizations that succeed on this score, and that implement any local movement of the identified sustainable point.

A second issue that may be of concern is computation of our sustainable outcomes. Without our results, computing a stable outcome by brute force would require checking $2^n - 1$ possible deviations, and there is evidence that similar computations are generally intractable (Deng and Papadimitriou 1994). However, using the scaling indifference condition which characterizes our sustainable outcome ($\boldsymbol{J}(\boldsymbol{a})\,\boldsymbol{a} = \boldsymbol{0}$) we have that $\boldsymbol{a}^T \boldsymbol{J}(\boldsymbol{a})^T \boldsymbol{J}(\boldsymbol{a})\,\boldsymbol{a} = \boldsymbol{0}$. This quadratic form can be minimized using a gradient descent algorithm. Under some convexity assumptions on the summands that constitute this quadratic form, we can even ensure that there is a unique global minimum of this function, thus guaranteeing that a single instance of gradient descent will find the solution efficiently.

A final issue is how a sustainable outcome might be implemented in practice if agents are wary of large discontinuous increases in their actions. For example, tariff reductions have in practice been implemented only in a piecemeal way through successive rounds of negotiations. Can actions be increased gradually to a sustainable outcome such that each small increase benefits all agents? The answer turns out to be positive under our assumptions. The path to a sustainable outcome consists of gradually scaling actions up proportionally to reach the sustainable (and thus Pareto efficient) point, making all agents better off with each increase. Formally, fix a sustainable outcome $\boldsymbol{a}$ and imagine a process taking place during times $t \in [0, 1]$; let us decree that actions at time t should be $t\boldsymbol{a}$. To see why all agents always prefer increasing actions as prescribed at time t, assuming everyone else does, consider again the scaling indifference condition $\boldsymbol{J}(\boldsymbol{a})\boldsymbol{a} = \boldsymbol{0}$. Concavity readily implies that for $t \in [0, 1)$, $\boldsymbol{J}(t\boldsymbol{a})t\boldsymbol{a} > \boldsymbol{0}$, since all entries of the Jacobian of a concave function are monotonic in t (recall that the entries of $\boldsymbol{a}$ are all positive).

References

Aumann, R. (1959): "Acceptable Points in General Cooperative n-Person Games," *Contributions to the Theory of Games IV, Annals of Mathematics Study*, 287–324.

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006): "Who's Who in Networks. Wanted: The Key Player," *Econometrica*, 74, 1403–1417.

Bramoullé, Y., R. Kranton, and d'Amours M. (2010): "Public Goods in Networks," .

Deng, X. and C. Papadimitriou (1994): "On the Complexity of Cooperative Solution Concepts," *Mathematics of Operations Research*, 19, 257–266.

Elliott, M. and B. Golub (2011): "Market Design for Multilateral Tariff Reduction Negotiations," .

Jackson, M. (2008): *Social and Economic Networks*, Princeton University Press.

Mas-Colell, A., M. Whinston, and J. Green (1995): *Microeconomic Theory*, vol. 1, Oxford University Press New York.

Rubinstein, A. (1980): "Strong Perfect Equilibrium in Supergames," *International Journal of Game Theory*, 9, 1–12.

Shapley, L. and M. Shubik (1969): "On the Core of an Economic System with Externalities," *The American Economic Review*, 59, 678–684.