A Network Approach to Public Goods[*]

Matthew Elliott[†]   Benjamin Golub[‡]
(Job Market Paper)

January 14, 2013

Abstract

We study settings where each agent can exert costly effort that creates nonrival, heterogeneous benefits for some of the others. For example, municipalities can forgo consumption to reduce pollution. How do the prospects for efficient cooperation depend on asymmetries in the effects of players' actions? We approach this question by analyzing a network that describes the marginal benefits agents can confer on one another. The first set of results explains how the largest eigenvalue of this network measures the marginal gains available from cooperating; as an application, we describe the players whose participation is essential to achieving any Pareto improvement on an inefficient status quo. Next, we examine mechanisms all of whose equilibria are Pareto efficient and individually rational; an outcome is called *robust* if it is an equilibrium outcome in every such mechanism. Robust outcomes exist and correspond to the Lindahl public goods solutions. The main result is a characterization of effort levels at these outcomes in terms of players' centralities in the benefits network. It entails that an outcome is robust if and only if agents contribute in proportion to how much they value the efforts of those who help them.

Keywords: collective action, externalities, public goods, network centrality, implementation theory, Lindahl equilibrium

1 Introduction

An economy can be thought of as a network in which the nodes are agents and links among them represent heterogeneous opportunities for exchange or cooperation. This paper argues that studying properties of such a network – how dense it is, how "central" various agents are in it – yields insights about issues such as the efficiency and fragility of an economic system, as well as its market outcomes.

We develop this conceptual point in a model of a public goods economy: one in which each agent can incur a private cost to take an action – e.g., reducing pollution – that creates nonrival but heterogeneous benefits for others. To fix ideas and demonstrate the network we will focus on, suppose there are three towns: X, Y and Z, located as shown in Figure 1a, each generating air and water pollution during production. The air pollution of a town affects only those east of it due to the prevailing wind. A river flows westward, so Z's water pollution affects X but not Y, which is located away from the river. Each town, i, can forgo $a_i \geq 0$ units of production at a net cost of a dollar per unit,[1] reducing its pollution and creating positive externalities for others affected by that pollution. Let $u_i(a_X, a_Y, a_Z)$ denote i's payoff.
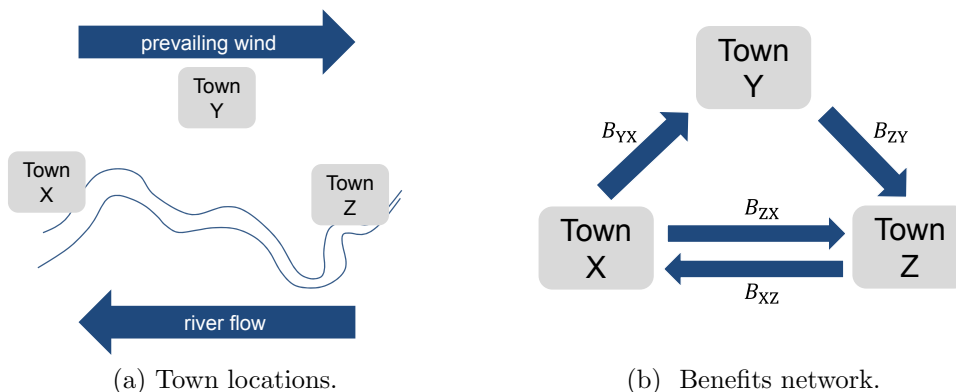

(a) Town locations. (b) Benefits network.

Figure 1: Town i benefits from j's pollution reduction when wind or water carries pollution from j to i; let $B_{ij} = \partial u_i / \partial a_j$ be the marginal benefit to i from j's reduction; these numbers may vary with the action profile, (a_X, a_Y, a_Z).

Suppose that the status quo at which everyone chooses $a_i = 0$ is not Pareto efficient (as will typically be the case when externalities are not internalized). To obtain a Pareto improvement, each town is willing to commit to an institution – a mechanism – in which it reduces pollution and others reciprocate.[2] Whose participation is essential to achieving a Pareto improvement on the status quo? Is the system fragile, in that the exclusion of one or a few agents destroys the potential for cooperation? Of the many Pareto efficient outcomes, can we consider some especially robust? If so, how do the robust solutions depend on the economic environment? We argue that insight into these questions can be gained by studying a *benefits network*, in which the links represent the marginal benefits players can confer on one another (Figure 1b).

---

[1]That is, the value of forgone production outweighs private environmental benefits. The net marginal cost is normalized to 1 in this example for convenience.

[2]We focus on favor-trading through the provision of public goods and abstract from side transfers of private goods. This keeps the modeling clean and is relevant for many practical negotiations.

Our first set of results focuses on a new measure of the marginal return to concerted investment in public goods: the largest eigenvalue of the benefits network.[3] These results have some immediate implications. First, cycles in the benefits network (e.g., X can help Y, who can help Z, who can help X) are critical to finding a Pareto improvement on a given outcome. The network's largest eigenvalue quantifies the marginal returns available from exploiting such cycles. For instance, agents can achieve a Pareto improvement on the status quo of zero effort if and only if the benefits matrix evaluated at that status quo has a largest eigenvalue exceeding 1. Second, there is a simple algorithm to find the players who are essential to a negotiation – in the sense that without their participation, there is no Pareto improvement on the status quo. They are the ones whose removal causes a sufficiently large disruption of cycles in the benefits network, as measured by the decrease in its largest eigenvalue.

When Pareto improvements on the status quo do exist, there is a question of how to implement them. We thus turn to the problem faced by a designer of a mechanism that the towns will use to address their collective action problem.[4] The designer wants a *reliable* mechanism for determining levels of effort – one such that, under whatever preferences the agents might later have, all Nash equilibria in the mechanism will be Pareto efficient and unanimously preferred to the status quo. We focus on the (nonempty) class of *robust* effort profiles: ones that are equilibrium outcomes of every reliable mechanism. These turn out to be precisely the *Lindahl outcomes*, which are analogues of Walrasian market allocations under individualized pricing of public goods.[5] There is an alternative way to look at these outcomes: if the designer wants to minimize the equilibrium selection problem and find a reliable mechanism with as few equilibria as possible (e.g., with a unique one if possible), she must choose a mechanism that yields precisely the Lindahl outcomes in equilibrium.[6]

The main result characterizes the Lindahl outcomes in terms of the benefits network. It states that a nonzero action profile is a Lindahl outcome if and only if, for all players i,

$$a_i = \sum_{j \neq i} B_{ij}\, a_j, \tag{1}$$

where B_{ij} measures the marginal value to i of j's effort at the outcome $\mathbf{a}$. Thus, players contribute in proportion to how much they value the efforts of those who help them. We will deduce two main practical consequences from this characterization. First, it is the benefits an agent *receives*, rather than those he can confer, that determine his level of effort at a robust outcome. Second, the players contributing the most are those who are most "central" in the benefits network, in the sense that they receive strong direct and indirect benefit flows from others.

---

[3]That is, the largest eigenvalue of the matrix with entries B_{ij} for $i \neq j$ and zeros on the diagonal.

[4]This designer could represent the towns themselves at an *ex ante* stage where they make the rules for their future interactions.

[5]Imagine that each agent must pay a personalized tax on each unit provided of public goods that he values; the tax is specific both to the good and to the agent. He also receives the taxes others pay for the public good he supplies. A Lindahl outcome is a profile of public goods provision that is optimal for each agent subject to budget balance at the given tax schedule. A formal definition appears in Section 4.2.

[6]Details appear in Section 4. For the formal results, we impose a continuity requirement on the reliable mechanisms and some conditions on the space of possible preferences.

The remainder of the introduction describes the key concepts and results in more detail, and also discusses connections to prior work.

The Benefits Network   Assume that $\frac{\partial u_i}{\partial a_i}(\mathbf{a}) < 0$ for all action profiles $\mathbf{a}$ and all agents i.[7] For $i \neq j$, define

$$B_{ij} = \frac{\partial u_i / \partial a_j}{-\partial u_i / \partial a_i}.$$

This quantity (which depends on the action profile) is i's marginal rate of substitution between decreasing own effort and receiving help from j. In other words, it is how much i values the help of j, measured in the number of units of effort that i would be willing to put forth in order to receive one unit of j's effort. Let $B_{ii} = 0$ for all i. We write $\mathbf{B}(\mathbf{a})$ for the matrix with these entries evaluated at $\mathbf{a}$. This matrix can also be represented as a directed, weighted graph, where the towns are the nodes and the links correspond to strictly positive entries of the matrix. For instance, the link directed to Y from X has weight $B_{\mathrm{YX}}(\mathbf{a})$, which measures the marginal benefit that accrues to Y when X increases its action at a profile $\mathbf{a}$. The network corresponding to our example is depicted in Figure 1b; links are drawn only for nonzero weights. There is a link to town i from town j if pollutants flow in that direction, allowing j to provide benefits to i by reducing pollution.

Pareto Efficient Outcomes and the Magnitude of Inefficiency   We begin with basic results on diagnosing and measuring inefficiency; these results do not depend on a selection of a particular Pareto improvement on the status quo, and do not involve Lindahl outcomes.

Proposition 1 in Section 3.1 shows that *an interior action profile $\mathbf{a}$ is Pareto efficient if and only if* 1 *is a largest eigenvalue of $\mathbf{B}(\mathbf{a})$*. The reason for this is as follows. The matrix $\mathbf{B}(\mathbf{a})$ is a linear system describing how investments translate into returns at the margin. Consider a particular sequence of investments: in Figure 1b, Y can increase his action slightly and provide a marginal benefit to Z. Then Z, in turn, can "pass forward" some of the resulting increase in his utility, investing effort to help X. Finally, X can similarly create benefits for Y, completing a cycle. If they can all receive back more than they invest in such an adjustment, then the starting point is not Pareto efficient. It is precisely in such cases that the linear system $\mathbf{B}(\mathbf{a})$ is "expansive": there is scope for everyone to get out more than they put in. And expansive systems are characterized by having a largest eigenvalue exceeding 1. If the largest eigenvalue of $\mathbf{B}(\mathbf{a})$ is less than 1, then everyone can be made better off by reducing investment. As a result, the interior Pareto efficient outcomes have a benefits matrix with a largest eigenvalue exactly equal to 1. Moreover, a Pareto improvement on the status quo outcome of $\mathbf{a} = \mathbf{0}$ exists if and only if $\mathbf{B}(\mathbf{0})$ has a largest eigenvalue exceeding 1. Section 3.1 derives these results. Section 3.4.2 formalizes the idea that cycles in the benefits network are crucial to Pareto improvements.

---

[7]That is, each agent has (more than) exhausted the net benefits available by unilaterally increasing his own action. This restricts the analysis to the interesting regime where there is a collective action problem.

A substantive implication of this result is that it allows us to analyze, in terms of the benefits network, which agents are essential to successful negotiations and to give a simple matrix criterion for finding them (Sections 3.4.4 and 3.4.5). These crucial agents are ones whose removal breaks important cycles in the benefits network, dropping the spectral radius of $\mathbf{B}(\mathbf{a})$ from above 1 to below 1. Section 8.1 applies these results in analyzing the temptation to free-ride.

Beyond characterizing Pareto efficiency, we would like to quantify the magnitude of inefficiency at the margin – defined loosely as the return on investment in public goods per unit of cost. If agents' utilities are denominated in a transferable numeraire (say, dollars), then it is easy to quantify this: just find the small change in contribution levels that maximizes the ratio (ρ) of overall marginal benefits to overall marginal costs in dollars. According to Coasian logic, all players agree that such an adjustment is the best one. There is an second way of looking at this measure of inefficiency via a hypothetical friction: suppose that to sacrifice \$1 for the sake of pollution reduction, a town must also waste an additional \$$t$ (e.g. due to polluting industries' lobbying). That makes the real cost per unit of effective contribution equal to $\tau = 1 + t$. If the collective return to a unit of effort, ρ, were equal to the coefficient τ then – given the frictions – the players' actions would be Pareto efficient.[8] Thus, we can equivalently define the returns on on investment ρ at an outcome as the tax coefficient τ that would be necessary to render that outcome Pareto efficient.

If agents' preferences are not quasilinear in any numeraire, utility is no longer transferable and the first approach to quantifying inefficiency no longer makes sense. We show, however, that the second way – the one based on a hypothetical friction – is portable easily to a world without such a numeraire (or without the possibility of transfers at all). Indeed, Section 3.2 derives that the measure of inefficiency defined via τ is equal to the largest eigenvalue of the benefits matrix. Therefore, that invariant is useful not only for diagnosing inefficiency but also for rigorously quantifying its magnitude. Section 3.3 gives another sense in which the largest eigenvalue of the benefits matrix measures inefficiency – one based on marginal improvements yielding the same returns on investment for everyone. Section 3.4.6 discusses implications when exact values of marginal costs and benefits are not known, but bounds are available.

Characterizing the Lindahl Outcomes through Network Centrality Theorem 1 in Section 4 formally establishes an equivalence between robust outcomes as defined in our discussion above and the Lindahl public goods solution. The main result of the paper characterizes the Lindahl outcomes in terms of eigenvector centrality in the benefits network. We now outline this result.

What is Eigenvector Centrality? In general, suppose we are given a network, represented by a matrix $\mathbf{M}$, where the nonnegative number M_{ij} captures the strength of the link ij (and is zero if that link is absent); this may differ from the strength of the

---

[8]In other words, whatever friction is inducing the inefficiency in the initial economy is comparable to a tax rate of t on contributions.

link ji. A *centrality measure* is an assignment of a number to each node measuring how important that node is based on the connections in the network.[9]

A widely studied measure is one that takes as a starting point the notion that an important node is one that is connected to other important nodes. A (right-hand) *eigenvector centrality* of the network $\mathbf{M}$ is simply a nonnegative, nonzero vector $\mathbf{e}$ whose entries sum to 1, with the property that for each i,

$$e_i = \gamma \sum_j M_{ij} e_j \tag{2}$$

for some proportionality constant $\gamma > 0$. If a vector $\mathbf{e}$ satisfying these conditions exists and is uniquely determined, then we say e_i is *the* eigenvector centrality of node i according to matrix $\mathbf{M}$. The equation says that the centrality of agent i is proportional to a weighted sum of the centralities of others, where e_j is weighted by M_{ij} – the strength of the connection ij. It can be shown that the most central agents are the ones with many paths leading to them in the network $\mathbf{M}$ (see Section 5.5 for a precise statement of this).

Perhaps the most famous application of this idea is the PageRank measure introduced as a part of Google's early algorithms to rank search results (Brin and Page, 1998). In that context, the nodes are web pages and M_{ij} is the number of hyperlinks that go to page i from page j, normalized by the total number of hyperlinks on page j. Then equation (2) says that a page is important if it is linked to by other important pages. Other applications include identifying those sectors in the macroeconomy that contribute the most to aggregate volatility via a network of intersectoral linkages (Acemoglu et al., 2012); the measurement of intellectual influence (Palacios-Huerta and Volij, 2004); and many others. We discuss the prior economic applications most closely related to our work – by Ballester, Calvó-Amengol, and Zenou (2006) and by Bramoullé, Kranton, and d'Amours (2011) – in the section on related literature below.

Centrality measures are intuitive and have useful mathematical properties, so it is worth knowing when they can be useful in economic analysis. In all prior work we are aware of, conditions such as (2) are obtained in economic models by positing a setting with particular parametric structure (linear-quadratic, Cobb-Douglas, etc.) on preferences or production. We find a nonparamteric connection between network centrality and an economic solution concept.

The Characterization Result. In our main result, we give a full characterization of the Lindahl outcomes in terms of eigenvector centrality.

An *eigenvector centrality action profile* is defined by the property that

$$a_i = \sum_j B_{ij}(\mathbf{a}) a_j. \tag{3}$$

for every i: player i's action is the weighted sum of actions of other players j, where

---

[9]There are very many ways to define such a measure, depending on the interpretation of the network and the notion of importance that are relevant in a given context. A survey of many measures with references is provided by Jackson (2008, Section 2.2.4).

the weight on a_j captures how much i cares about the effort of j at the margin. This condition states that the action profile itself is a right-hand eigenvector of the marginal benefits matrix with eigenvalue[10] equal to 1 – and this is intimately related to the efficiency characterization discussed above.

The main result of the paper, Theorem 2 in Section 5.2, asserts that an action profile is a Lindahl outcome if and only if it is an eigenvector centrality action profile. Intuitively, this means that the Lindahl outcomes are characterized by the condition that the agents contributing a lot to the public goods are the ones who are benefiting a lot, at the margin, from the efforts of others who are contributing.

Applications and Consequences. The conceptual import of the characterization result is that it provides a general connection between efficient, robust outcomes of public goods problems and the theory of network centrality. We now discuss some more concrete consequences.

The first application of the main theorem is solving for the Lindahl outcome in an environment with *global* public goods, where each agent's benefits depend only on the sum of others' actions (Section 5.3). Here we find a very simple characterization of Lindahl outcomes: an agent contributes a fraction of the overall public good quantity equal to his marginal rate of substitution between the public good and private consumption.

Because the condition (3) involves $\mathbf{B(a)}$, which varies as $\mathbf{a}$ changes, we also build intuition by considering (in Section 6) special cases where the fixed point defined by (3) can be expressed explicitly in terms of an *exogenous* network that does not depend on $\mathbf{a}$. Substantively, the theme of these results is that it is the intensity of benefits that an agent *receives*, and not of those he generates, that determine his effort level. However, it is not just i's *own* links in the benefits network that matter: chains of benefits of the form "k helps j who helps i" are crucial to determining i's centrality in the benefits network, and thus i's effort level at Lindahl outcomes. Proposition 5 in Section 5.5 relates such chains to Lindahl outcomes generally.

In studying these special cases, we also provide price-theoretic foundations for three of the most important centrality measures that have been studied in the networks literature.

Related Work   This paper is at the intersection of the theory of public goods provision and the theory of economic networks. A recent literature has studied Nash equilibria of one-shot games in networks when best responses are linear in others' actions and has related those equilibria to eigenvalue and centrality conditions. Key papers include Ballester, Calvó-Amengol, and Zenou (2006) on skill investment with externalities, and Bramoullé, Kranton, and d'Amours (2011) on local public goods; they offer more comprehensive surveys of the literature.[11]   A fundamental contrast is that our work focuses on institutions that implement Pareto efficient public goods

---

[10]In terms of equation (2), this definition of an eigenvector centrality action profile requires $\gamma = 1$.

[11]Most recently, Allouch (2012) has studied a network version of the setting introduced in the seminal paper of Bergstrom, Blume, and Varian (1986) on the voluntary (static Nash) private provision of public goods. Generalizing results of Bramoullé, Kranton, and d'Amours (2011), he derives comparative statics of public goods provision using network centrality tools.

outcomes, rather than on Nash behavior in the basic game without additional structure. From a technical perspective, these prior papers can be seen as addressing the question: "In which economic settings are eigenvalue and network centrality methods useful?" We provide new insights on this that are complementary to the prior work. Since the Lindahl solution is an analogue of a Walrasian equilibrium, the main theorem of the present paper can be seen as making a general link between network centrality and price equilibria.[12] A final contrast is that our main characterization result does not impose functional form assumptions on preferences.

There is a large literature in implementation theory on game forms that yield Walrasian and Lindahl outcomes. Samuelson (1954) doubted that the price-theoretic public goods solutions first contemplated by Wicksell (1896) and Lindahl (1919) could be implemented despite agents' incentives to manipulate a social institution. This inspired the early literature on implementation theory and mechanism design, with Groves and Ledyard (1977) showing that a game *could* be designed yielding Pareto efficient public goods provision in every Nash equilibrium. Key papers in the subsequent literature on implementing Walrasian and Lindahl outcomes include Hurwicz (1979a), Hurwicz (1979b), and Hurwicz, Maskin, and Postlewaite (1995). A broad survey is presented by Jackson (2001). The paper most important to us in this line of work is Hurwicz (1979a) who first suggested strategic foundations for Lindahl outcomes based on more primitive criteria – Pareto efficiency and individual rationality. We rely on the insights of this paper to establish what makes Lindahl outcomes special in our world.[13]

Complementary foundations for Lindahl outcomes based on equilibria of bargaining games are offered by Dávila, Eeckhout, and Martinelli (2009) and Penta (2011); these works are part of a broader literature on Walrasian bargaining – see Yildiz (2003) and Dávila and Eeckhout (2008).

The Plan of the Paper Section 2 lays out the basic assumptions and notation. Section 3 collects all results on efficiency and on measuring inefficiency. Section 4 formally presents the implementation theory framework; defines the Lindahl solution; and states the theorem on its unique robustness. Section 5 presents the main theorem, characterizing the Lindahl outcomes as the eigenvector centrality action profiles. Section 6 considers special cases of the characterization to show how network structure affects robust outcomes; this also yields new economic foundations for three important network centrality measures. Section 7 ties up the theory; its main purpose is to explain the proof of the characterization result. Finally, Section 8 discusses limitations and extensions.

---

[12]Du, Lehrer, and Pauzner (2012) connect market outcomes to a network centrality condition by showing that locations on an unweighted graph can be ranked using the equilibrium prices of an associated exchange economy in which agents have Cobb-Douglas utility functions.

[13]For technical reasons we must provide our own proofs; e.g., Hurwicz (1979a) assumes positive endowments of all private goods, an assumption that does not have a reasonable analogue in our setting.

2 Framework

2.1 The Environment

Each member i of a set $N = \{1, 2, \ldots, n\}$ of players (or agents) simultaneously chooses an effort level, or action[14], $a_i \in \mathbb{R}_+$. Taking a higher action should be interpreted as doing more of something that helps the other agents – for instance, mitigating pollution. Each player has a utility function $u_i : \mathbb{R}_+^n \to \mathbb{R}$; player i's payoff when the action profile $\mathbf{a}$ is played is $u_i(\mathbf{a})$.

2.2 Main Assumptions

Unless otherwise noted, the following assumptions are maintained throughout the paper.

Each $u_i : \mathbb{R}_+^n \to \mathbb{R}$ is concave and continuously differentiable. The main substantive assumptions about the payoffs are:

Assumption 1 (Costly Actions)**.** Each player finds it costly to invest effort, holding others' actions fixed: $\frac{\partial u_i}{\partial a_i}(\mathbf{a}) < 0$ for any $\mathbf{a} \in \mathbb{R}_+^n$ and $i \in N$.

Assumption 2 (Positive Externalities)**.** Increasing any player's action level weakly benefits all other players: $\frac{\partial u_i}{\partial a_j}(\mathbf{a}) \geq 0$ for any $\mathbf{a} \in \mathbb{R}_+^n$ whenever $j \neq i$.

Because the externalities are positive and nonrival, this is a public goods environment. Together, the two assumptions above make the setting a potential tragedy of the commons. The unique Nash equilibrium of a game in which players choose their actions a_i entails that everyone contributes nothing – $a_i = 0$ for each i – even though other outcomes may Pareto dominate this one.

Two additional technical assumptions are useful:

Assumption 3 (Connectedness of Benefits)**.** For all $\mathbf{a} \in \mathbb{R}_+^n$, if M is a nonempty proper subset of N, then there exist $i \in M$ and $j \notin M$ (which may depend on $\mathbf{a}$) such that $\frac{\partial u_i}{\partial a_j}(\mathbf{a}) > 0$.

This posits that it is not possible to find an outcome and partition society into two nonempty groups such that, at that outcome, one group does not care about the effort of the other at the margin.[15]

Finally, we assume that the set of points where everybody wants to scale up all effort levels is bounded. To state this, we introduce a few definitions. Under a preference profile $\mathbf{u}$, action profile $\mathbf{a}' \in \mathbb{R}_+^n$ *Pareto dominates* another profile $\mathbf{a} \in \mathbb{R}_+^n$ if $u_i(\mathbf{a}') \geq u_i(\mathbf{a})$ for all $i \in N$, and the inequality is strict for some i. We say $\mathbf{a}'$ *strictly Pareto dominates* $\mathbf{a}$ if $u_i(\mathbf{a}') > u_i(\mathbf{a})$ for all $i \in N$. Finally, $\mathbf{a}$ is *Pareto efficient* if no other action profile Pareto dominates it.

---

[14]We use $\mathbb{R}_+$ (respectively, $\mathbb{R}_{++}$) to denote to the set of nonnegative (respectively, positive) real numbers. We write $\mathbb{R}_+^n$ (respectively, $\mathbb{R}_{++}^n$) for the set of vectors $\mathbf{v}$ with n entries such that each entry v_i is in $\mathbb{R}_+$ (respectively, $\mathbb{R}_{++}$). When we write an inequality between vectors, e.g. $\mathbf{v} > \mathbf{w}$, that means the inequality holds coordinate by coordinate, i.e. $v_i > w_i$ for each $i \in N$.

[15]See Section 8.4 for a discussion of extending the analysis when this assumption does not hold.

Assumption 4 (Bounded Improvements)**.** The set

$$\{\mathbf{a} \in \mathbb{R}^n_+ : \text{there is an } s > 1 \text{ so that } s\mathbf{a} \text{ strictly Pareto dominates } \mathbf{a}\}$$

is bounded.[16]

This assumption is necessary to keep the problem well-behaved and ensure the existence of Lindahl equilibrium outcomes; see Section 7.1.

2.3   Key Notions

The Jacobian   We write $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ for a profile of utility functions. To keep track of the marginal costs and benefits of actions, we define the Jacobian, $\mathbf{J}(\mathbf{a}; \mathbf{u})$, to be the n-by-n matrix whose (i, j) entry is

$$J_{ij}(\mathbf{a}; \mathbf{u}) = \frac{\partial u_i}{\partial a_j}(\mathbf{a}).$$

The Benefits Matrix   A close relative of the Jacobian turns out to be useful. The *benefits matrix* $\mathbf{B}(\mathbf{a}; \mathbf{u})$ is defined as follows:

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

As discussed in the introduction, when $i \neq j$, the quantity $B_{ij}(\mathbf{a}; \mathbf{u})$ is i's marginal rate of substitution between decreasing own effort and receiving help from j. In other words, it is how much i values the help of j, measured in the number of units of effort that i would be willing to put forth in order to receive one unit of j's effort.

Suppose $\mathbf{u}$ satisfies the assumptions of Section 2.2. Since $J_{ii}(\mathbf{a}; \mathbf{u}) < 0$ by Assumption 1, this matrix is well-defined. By Assumption 2, it is entrywise nonnegative. Assumption 3 is equivalent to the statement that this matrix is irreducible[17] at every $\mathbf{a}$.

In discussing both the Jacobian and the benefits matrix, when there is no ambiguity about what $\mathbf{u}$ is, we suppress it.

The Spectral Radius (Largest Eigenvalue)   For any nonnegative matrix $\mathbf{M}$, we define $r(\mathbf{M})$ as the magnitude of a largest eigenvalue of $\mathbf{M}$, also called the *spectral radius*. That is,

$$r(\mathbf{M}) = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } \mathbf{M}\},$$

where $|\lambda|$ denotes the absolute value of the complex number λ. By the Perron-Frobenius Theorem (see Section 2.4 below for a formal statement), any such matrix has a *real, positive* eigenvalue equal to the maximum of all eigenvalues' magnitudes.

---

[16]This condition is substantially weaker than assuming that the set of Pareto efficient outcomes is bounded, which fails to be true in many reasonable environments.

[17]A matrix $\mathbf{M}$ is *irreducible* if it is not possible to find a nonempty subset S of indices so that $M_{ij} = 0$ for every $i \in S$ and $j \notin S$.

Thus, we may equivalently think of $r(\mathbf{M})$ as the largest eigenvalue of $\mathbf{M}$ on the real line.

This quantity can be thought of as a single measure of how expansive a matrix is as a linear operator – how much it can scale up vectors that it acts on. When applied to the matrix $\mathbf{B}$, it will quantify, in a sense we will make precise, the marginal gains that can be generated by cooperating.

2.4   The Perron-Frobenius Theorem

The key mathematical tool we use is the Perron-Frobenius Theorem, a fundamental result about nonnegative matrices which will be applied to analyze the benefits matrix $\mathbf{B}(\mathbf{a})$. The facts stated in this section are necessary only for proofs, and skipping this section (or any of the proofs in the paper) will not interrupt the flow. Nevertheless, since this powerful theorem underlies so much of the analysis, we state it here for easy reference.

Theorem (Perron-Frobenius[18])**.** Let $\mathbf{M}$ be an irreducible square matrix with no negative entries and spectral radius $r(\mathbf{M})$. Then:

  (i) The real number $r(\mathbf{M})$ is an eigenvalue of $\mathbf{M}$.

 (ii) There is a vector $\mathbf{p}$ (called a Perron vector) with only positive entries such that $\mathbf{M}\mathbf{p} = r(\mathbf{M})\mathbf{p}$.

(iii) If $\mathbf{v}$ is a nonzero vector with nonnegative entries such that $\mathbf{M}\mathbf{v} = q\mathbf{v}$ for some $q \in \mathbb{R}$, then $\mathbf{v}$ is a positive scalar multiple of $\mathbf{p}$, and $q = r(\mathbf{M})$.

Statement (i) says that there is always a real, positive eigenvalue of the matrix $\mathbf{M}$ that is at least as large as any other eigenvalue. Statement (ii) says that this eigenvalue, $r(\mathbf{M})$, is associated with a strictly positive eigenvector, the Perron vector. Statement (iii) says that *any* nonnegative eigenvector of the matrix $\mathbf{M}$ is a scalar multiple of the Perron vector, and is associated with the special eigenvalue $r(\mathbf{M})$.

REMARK **1.** Note that because a matrix has exactly the same eigenvalues as its transpose, all the same statements are true, with the same eigenvalue $r(\mathbf{M}) = r(\mathbf{M}^{\mathsf{T}})$, when we replace $\mathbf{M}$ by its transpose $\mathbf{M}^{\mathsf{T}}$. This observation yields a *left*-hand Perron eigenvector of $\mathbf{M}$, i.e. a row vector $\mathbf{w}$ such that $\mathbf{w}\mathbf{M} = r(\mathbf{M})\mathbf{w}$, which enjoys the analogue of property (iii) in the theorem.

3   Efficiency and the Spectral Radius

The thesis of this paper is that we can gain insight about the economics of the public goods problem by constructing, for any outcome $\mathbf{a}$ under consideration, a network in

---

[18]Meyer (2000, Section 8.3) has a comprehensive exposition of this theorem, its proof, and related results. Conventions vary regarding whether the Perron-Frobenius Theorem encompasses all these statements or just (i), but we will use the name as an umbrella term for all these facts; which one is relevant will be clear from context.

which the agents are nodes and the weighted links among them measure the marginal benefits available by increasing actions. The matrix corresponding to this network is $\mathbf{B}(\mathbf{a})$, defined in Section 2.3 above. (Throughout this section, we fix a profile $\mathbf{u}$ of utility functions and suppress it in the notation for the Jacobian and benefits matrix.)

This section offers support for our thesis by showing that an important statistic of this network – the size of the largest eigenvalue – can be used to diagnose whether an outcome is Pareto efficient (Section 3.1), and if not, used to measure the magnitude of the inefficiency. This measurement can be done in two ways: via a hypothetical friction (Section 3.2) or by considering egalitarian marginal improvements (Section 3.3). After presenting these general results, we discuss interpretations (especially in terms of cycles) and applications in Section 3.4.

3.1   A Characterization of Pareto Efficiency

PROPOSITION 1. Under the assumptions in Section 2.2, an interior action profile $\mathbf{a} \in \mathbb{R}^n_{++}$ is Pareto efficient if and only if the spectral radius of $\mathbf{B}(\mathbf{a})$ is 1.

Proof of Proposition 1: For any nonzero $\boldsymbol{\theta} \in \mathbb{R}^n_+$, define $\mathcal{P}(\boldsymbol{\theta})$, the Pareto problem with weights $\boldsymbol{\theta}$ as:

$$\text{maximize} \sum_{i \in N} \theta_i u_i(\mathbf{a}) \text{ subject to } \mathbf{a} \in \mathbb{R}^n_+.$$

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 as $\boldsymbol{\theta}$ ranges over all the nonnegative vectors not equal to $\mathbf{0}$.

Suppose that an interior action vector $\mathbf{a}^*$ is Pareto efficient, and therefore solves $\mathcal{P}(\boldsymbol{\theta})$ for a vector $\boldsymbol{\theta}$. Since $\mathbf{a}^*$ is interior, that solution satisfies the system of first-order conditions $\boldsymbol{\theta} \mathbf{J}(\mathbf{a}^*) = \mathbf{0}$. Let $\mathbf{D}(\mathbf{a})$ be a matrix with $D_{ii}(\mathbf{a}) = -J_{ii}(\mathbf{a})$ and zeros off the diagonal. The diagonal entries of this matrix are positive at all $\mathbf{a} \in \mathbb{R}^n_+$ by Assumption 1. By definition, $\mathbf{B}(\mathbf{a}) = \mathbf{D}(\mathbf{a})^{-1}\mathbf{J}(\mathbf{a}) + \mathbf{I}$, where $\mathbf{I}$ is the n-by-n identity matrix. Define $\boldsymbol{\gamma} = \boldsymbol{\theta}\mathbf{D}(\mathbf{a}^*)$ and note that the following statements are equivalent:

$$\boldsymbol{\theta}\mathbf{J}(\mathbf{a}^*) = \mathbf{0}$$
$$[\boldsymbol{\theta}\mathbf{D}(\mathbf{a}^*)][\mathbf{D}(\mathbf{a}^*)^{-1}\mathbf{J}(\mathbf{a}^*)] = \mathbf{0}$$
$$\boldsymbol{\gamma}[\mathbf{D}(\mathbf{a}^*)^{-1}\mathbf{J}(\mathbf{a}^*) + \mathbf{I}] = \boldsymbol{\gamma}$$
$$\boldsymbol{\gamma}\mathbf{B}(\mathbf{a}^*) = \boldsymbol{\gamma}.$$

Thus $\mathbf{B}(\mathbf{a}^*)$ has an eigenvalue of 1 with corresponding nonnegative left-hand eigenvector $\boldsymbol{\gamma}$. Since $\mathbf{B}(\mathbf{a}^*)$ is irreducible by Assumption 3 and has only nonnegative entries, the Perron-Frobenius Theorem applies to this matrix. That theorem (part (iii) of the statement in Section 2.4) says that the only eigenvalue of $\mathbf{B}(\mathbf{a}^*)$ that is associated with a nonnegative eigenvector is the spectral radius itself. Thus, the spectral radius of $\mathbf{B}(\mathbf{a}^*)$ must be 1.

Conversely, if $\mathbf{B}(\mathbf{a}^*)$ has a largest eigenvalue of 1, the Perron-Frobenius Theorem guarantees the existence of a nonnegative left-hand eigenvector $\boldsymbol{\gamma}$ such that $\boldsymbol{\gamma}\mathbf{B}(\mathbf{a}^*) =$

$\boldsymbol{\gamma}$. Consequently, the first-order conditions of the Pareto problem are satisfied for weights $\boldsymbol{\theta} = \boldsymbol{\gamma}\mathbf{D}(\mathbf{a}^*)^{-1}$ by the calculation above. By the assumption of concave utilities, it follows that $\mathbf{a}^*$ solves the Pareto problem for weights $\boldsymbol{\theta}$ (i.e., the first-order conditions are sufficient for optimality). ∎

An intuition for the result based on passing marginal benefits around a cycle and thinking of $\mathbf{B}(\mathbf{a})$ an expansive linear system was presented in the introduction.

REMARK 2. The condition that the spectral radius of $\mathbf{B}(\mathbf{a})$ is 1 is independent of how different players' cardinal utilities are measured – as, of course, it must be, since Pareto efficiency is an ordinal notion. To see how the benefits matrix changes under reparameterizations of cardinal utility, suppose we define, for each $i \in N$, new utility functions $\widehat{u}_i(\mathbf{a}) = f_i(u_i(\mathbf{a}))$ for some differentiable, strictly increasing functions f_i. If we let $\widehat{\mathbf{B}}$ be the benefits matrix obtained from these new utility functions, then $\mathbf{B}(\mathbf{a}) = \widehat{\mathbf{B}}(\mathbf{a})$; this follows by applying the chain rule to the numerator and denominator in the definition of the benefits matrix.

Proposition 1 focuses on interior outcomes. One particular boundary outcome – the profile $\mathbf{a} = \mathbf{0}$ in which nobody contributes – is particularly important. The following proposition characterizes when it is Pareto efficient. A proof appears in Section A.2.[19]

PROPOSITION 2. Under Assumptions 1 and 2, the outcome $\mathbf{0}$ is Pareto efficient if and only if $r(\mathbf{B}(\mathbf{0})) \leq 1$.

3.2 Quantifying Inefficiency: The Frictions Approach

The spectral radius of the benefits matrix does not merely diagnose inefficiency, but also measures its magnitude. In particular, we will show that to eliminate all Pareto improvements at a profile $\mathbf{a}$, effort must be taxed at a rate equal to the spectral radius of $\mathbf{B}(\mathbf{a})$ minus one.

To formalize this, this subsection focuses on a setting where costs and benefits are separable.[20] Suppose that for each i,

$$u_i(\mathbf{a}) = v_i(\mathbf{a}_{-i}) - c_i(a_i),$$

where $\mathbf{a}_{-i}$ is the vector of actions of players excluding i. Suppose also that the resulting utility function satisfies the assumptions Section 2.2. Here v_i is a function mapping others' contributions to i's benefits, and c_i is a cost function taking as an argument i's own effort. Construct a related economy in which frictions (e.g., due to political economy issues internal to each agent) make it more costly to contribute to public goods. More precisely, we posit that the cost function of each agent is

---

[19]Note that this proposition does not require the assumption that $\mathbf{B}(\mathbf{a})$ is irreducible.

[20]This is not essential to the result, but without separability, the counterfactual economy with taxes becomes more complicated to define in ways that do not contribute insight; details are available from the authors upon request.

multiplied by τ, for some $\tau \geq 0$. The utility functions under this assumption are given, for each i, by

$$u_i^{(\tau)}(\mathbf{a}) = v_i(\mathbf{a}_{-i}) - \tau c_i(a_i).$$

The quantity $t = \tau - 1$ may be thought of as a tax rate on effort.[21] Then we have the following corollary of Proposition 1.

COROLLARY 1. Fix a positive real number τ. The interior action profile $\mathbf{a}$ is a Pareto efficient outcome under $\mathbf{u}^{(\tau)}$ if and only if τ is the spectral radius of $\mathbf{B}(\mathbf{a}; \mathbf{u})$.

Proof This follows from Proposition 1 by observing that, for all $\mathbf{a} \in \mathbb{R}_+^n$, it holds that $\mathbf{B}(\mathbf{a}; \mathbf{u}^{(\tau)}) = \tau^{-1}\mathbf{B}(\mathbf{a}; \mathbf{u})$. ∎

Thus, to rationalize an outcome $\mathbf{a}$ as being Pareto efficient, a tax rate of $t = r(\mathbf{B}(\mathbf{a})) - 1$ on effort is required; effort must be taxed at a rate equal to the spectral radius of $\mathbf{B}(\mathbf{a})$ minus one. Correspondingly, if $r(\mathbf{B}(\mathbf{a})) < 1$, then effort must be subsidized at the margin to render the outcome $\mathbf{a}$ efficient. This is a sense in which the size of the spectral radius of $\mathbf{B}(\mathbf{a})$ is a quantitative measure of inefficiency.

REMARK 3. The spectral radius is a continuous function of the matrix entries (see, e.g., Wilkinson, 1965). It is reassuring that our measure of inefficiency varies continuously in the data of the problem.

3.3 Quantifying Inefficiency: The Value of Egalitarian Improvements

There is another way to think of the spectral radius of $\mathbf{B}(\mathbf{a})$ as measuring inefficiency. In brief: there is a way in which actions can be changed at $\mathbf{a}$ to yield marginal benefits per unit of marginal cost equal to the spectral radius of $\mathbf{B}(\mathbf{a})$ for *all* agents simultaneously.

Let Δ_n denote the simplex in $\mathbb{R}_+^n$ defined by $\Delta_n = \{\mathbf{a} \in \mathbb{R}_+^n : \sum_i a_i = 1\}$.

DEFINITION 1. The *bang for the buck* vector $\mathbf{b}(\mathbf{a}, \mathbf{d})$ at an action profile $\mathbf{a}$ along a direction $\mathbf{d} \in \Delta_n$ is defined by

$$b_i(\mathbf{a}, \mathbf{d}) = \frac{\sum_{j:j \neq i} J_{ij}(\mathbf{a})d_j}{-J_{ii}(\mathbf{a})d_i}.$$

The thought experiment behind this definition is as follows: suppose that we start at point $\mathbf{a}$ and consider increasing each agent i's effort by ϵd_i, for some small $\epsilon > 0$. Then $b_i(\mathbf{a}, \mathbf{d})$ measures how much agent i receives in benefits (from others' increases in effort) per unit of cost incurred in increasing his own action[22]. In other words, it

---

[21]Formally, since τ can be between 0 and 1, the setup also allows for effort being subsidized by an outside benefactor; in that case, $t < 0$.

[22]Formally, i's cost is defined as the difference in i's payoff between (i) leaving his own action unchanged at a_i (ii) increasing his own action by ϵd_i (with a_j unchanged for every $j \neq i$). Similarly, i's benefit is defined as the difference in his payoff between (iii) all other j increasing actions by ϵd_j while i holds his own action fixed at a_i and (iv) nobody increasing actions.

is equal to the ratio

$$\frac{i\text{'s marginal benefit}}{i\text{'s marginal cost}}$$

evaluated at $\mathbf{a}$ with respect to this hypothetical deviation from $\mathbf{a}$. An agent gains from everyone's changing actions in the direction $\mathbf{d}$ when this ratio exceeds 1, and loses when this ratio is less than 1.

We say a direction $\mathbf{d} \in \Delta_n$ is *egalitarian* if every entry of $b_i(\mathbf{a}, \mathbf{d})$ is the same. That is, any agent's cost of increasing his action (relative to free-riding on this marginal improvement by leaving his own effort unchanged) converts into benefits (received by that agent) at the same rate. This can also be viewed as a Rawlsian direction: it maximizes the minimum return any agent obtains from increasing effort.[23]

PROPOSITION 3. At any $\mathbf{a}$, there is a unique egalitarian direction $\mathbf{d}^{\text{eg}}(\mathbf{a})$. Every entry of $\mathbf{b}(\mathbf{a}, \mathbf{d}^{\text{eg}}(\mathbf{a}))$ is equal to the spectral radius of $\mathbf{G}(\mathbf{a})$.

Proof of Proposition 3: Fix $\mathbf{a}$ and denote by r the spectral radius of $\mathbf{B}(\mathbf{a})$. Since $\mathbf{B}(\mathbf{a})$ is nonnegative and irreducible, the Perron-Frobenius Theorem guarantees that it has a right-hand eigenvector $\mathbf{d}$ such that

$$\mathbf{B}(\mathbf{a})\mathbf{d} = r\mathbf{d}. \tag{4}$$

This is equivalent to $\mathbf{b}(\mathbf{a}, \mathbf{d}) = r\mathbf{1}$, where $\mathbf{1}$ is the column vector of all ones. Therefore, there is an egalitarian direction that generates a bang-for-the-buck of r (the spectral radius of $\mathbf{B}(\mathbf{a})$) for everyone.

Now suppose $\widetilde{\mathbf{d}} \in \Delta_n$ is any egalitarian direction, i.e. for some b we have

$$\mathbf{b}(\mathbf{a}, \widetilde{\mathbf{d}}) = b\mathbf{1}.$$

This implies

$$\mathbf{B}(\mathbf{a})\widetilde{\mathbf{d}} = b\widetilde{\mathbf{d}}. \tag{5}$$

By the Perron-Frobenius Theorem (statement (iii)), the only real number b and vector $\widetilde{\mathbf{d}} \in \Delta_n$ satisfying (5) are $b = r$ and $\widetilde{\mathbf{d}} = \mathbf{d}$.

Thus, $\mathbf{d}^{\text{eg}}(\mathbf{a}) = \mathbf{d}$ has all the properties claimed in the proposition's statement. ∎

When the spectral radius of $\mathbf{B}(\mathbf{a})$ exceeds 1, then by increasing actions in the direction given by $\mathbf{d}^{\text{eg}}(\mathbf{a})$, an egalitarian Pareto improvement can be achieved, and the benefit-to-cost ratio of that change is equal to $r(\mathbf{B}(\mathbf{a}))$ for every agent. When the spectral radius of $\mathbf{B}(\mathbf{a})$ is less than 1, then by decreasing actions in the direction $-\mathbf{d}^{\text{eg}}(\mathbf{a})$, each agent foregoes $r(\mathbf{B}(\mathbf{a}))$ units of benefit per unit of cost saved, and that yields a Pareto improvement as well.

If agents negotiate over small improvements and fairness or equality concerns are important, egalitarian directions may be a relevant benchmark because agents intrinsically care about them. However, and perhaps surprisingly, egalitarian directions also arise naturally in the context of solutions to public goods problems with more fundamental strategic foundations (see Remark 6 in Section 5.2).

---

[23]For a proof, see Section A.1

3.4 Interpretations and Applications

While the spectral radius of the benefits matrix has the economic interpretations laid out in the previous subsections, it is useful to give some intuition about what this quantity measures in the structure of a network.

Section 3.4.1 gives a very simple example in which the benefits matrix is a cycle and the spectral radius measures the magnitude of the benefits that are generated along that cycle. Section 3.4.2 lays out the general facts about how cycles in the benefits network relate to its spectral radius, fleshing out the claims in the introduction about the importance of cycles to enabling Pareto improvements.

Section 3.4.3 shows a monotonicity result: when links in the benefits network get weaker or disappear, its spectral radius decreases. Sections 3.4.4 and 3.4.5 explore an important special case, discussing the decrease in a group's potential for cooperation when some of its members do not contribute. These sections focus particularly on the question of which agents are *essential*, in the sense that their absence eliminates the potential for any Pareto improvements on the status quo. Finally, Section 3.4.6 applies the monotonicity result to bounding the gains from cooperation under imperfect measurement of marginal costs and benefits.

3.4.1 A Simple Example

Suppose $N = \{1, 2, 3\}$ and

$$\mathbf{B}(\mathbf{0}) = \begin{bmatrix} 0 & 0 & 7 \\ 5 & 0 & 0 \\ 0 & 6 & 0 \end{bmatrix}.$$

This network is depicted in Figure 2. It can be computed that $r(\mathbf{B}(\mathbf{0})) = (5 \cdot 6 \cdot 7)^{1/3} \approx 5.94$. Thus, there is substantial potential for cooperation in this network. The spectral radius of $\mathbf{B}(\mathbf{0})$ is the geometric mean of weights in the benefits network along the cycle – the product of these weights to the power $1/\ell$, where ℓ is the cycle's length. We state without proof that this is always the case when there is just one cycle in the network. Facts 1 below is a general bound that builds on this observation.
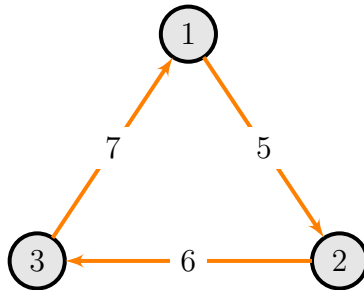

Figure 2: The network of the example. An arrow from from i to j indicates that $B_{ji} > 0$ (i.e. benefits flow from i to j) and the link is labeled by the weight B_{ji}.

3.4.2 Interpreting the Spectral Radius: Cycles of Mutual Benefit

As noted in the introduction, if we can find a cycle of players such that each can help the next, then that creates scope for cooperation. The larger the benefits that can be passed along a cycle, the greater that scope. This section formalizes the idea that the spectral radius of a matrix can be interpreted generally as measuring the intensity of such cycles.

Definitions   A (directed) *cycle of length ℓ in the matrix* $\mathbf{M}$ is a sequence

$$(c(1), c(2), \ldots, c(\ell))$$

of elements of N (player indices), where $c(1) = c(\ell)$ – i.e., the cycle starts and ends at the same node such that $M_{c(t)c(t+1)} > 0$ for each $t \in \{1, \ldots, \ell - 1\}$.[24] Let $\mathcal{C}(\ell; \mathbf{M})$ be the set of all cycles of length ℓ in matrix $\mathbf{M}$. For any nonnegative matrix $\mathbf{M}$, define the value of a cycle $c \in \mathcal{C}(\ell; \mathbf{M})$ as

$$v(c; \mathbf{M}) = \prod_{t=1}^{\ell-1} M_{c(t)c(t+1)}. \tag{6}$$

This is the number obtained by taking the product of all weights along the cycle c in the weighted directed graph defined by $\mathbf{M}$. Define

$$V(\ell; \mathbf{M}) = \sum_{c \in \mathcal{C}(\ell; \mathbf{M})} v(c; \mathbf{M}).$$

This is the sum of the values of all cycles of length ℓ in matrix $\mathbf{M}$.

Having made these definitions, we can show how the spectral radius of the benefits matrix is related to the total value of such cycles.

A Note on the Directions of Links and Cycles   Before this, it is worth remarking on a potentially confusing wrinkle of the notation. Since the thing that is naturally thought of as traveling among players is the positive externalities they confer on each other, we think of the strength of the link directed <u>from</u> i <u>to</u> j (setting own marginal cost of effort to 1) as $B_{ji} = \partial u_j / \partial a_i$ – i.e., how much i can help j (recall Figure 1 or Figure 2, for example). However, when a matrix $\mathbf{B}$ is represented as a weighted directed graph, the typical convention is to label an arrow from i to j by B_{ij}. These two conventions conflict. This problem goes away if we instead focus on $\mathbf{B}^{\mathsf{T}}$, the transpose of $\mathbf{B}$, when we draw the benefits network or discuss cycles in it. For example, if we consider the cycle $c = (1, 2, 3)$, then $v(c; \mathbf{B}^{\mathsf{T}}) = B_{21}B_{32}B_{13}$ multiplies three factors: how much 1 helps 2, how much 2 helps 3, and how much 3 helps 1.

---

[24]Nodes other than the first and last may also be repeated in this sequence. The convention on length – that the cycle i, j, i is said to be of length 3 even though it involves only two "steps" – is slightly awkward but useful for consistency with the discussion in Section 5.5.

The Results   First, any cycle provides a simple lower bound on the spectral radius of a matrix. A proof of this fact is in Section A.9.

FACT **1.** For any nonnegative matrix $\mathbf{M}$ and any cycle c of length ℓ in $\mathbf{M}$,

$$r(\mathbf{M}) \geq v(c; \mathbf{M})^{1/\ell}.$$

Moreover, the total value of long cycles provides an asymptotically exact estimate to the spectral radius in the following sense:

FACT **2.** For any nonnegative matrix $\mathbf{M}$,

$$r(\mathbf{M}) = \limsup_{\ell \to \infty} V(\ell; \mathbf{M})^{1/\ell}.$$

Proof of Fact 2: Note that $V(\ell; \mathbf{M}) = \text{trace}\left(\mathbf{M}^{\ell}\right)$. With this replacement, the fact is standard – see, e.g. Milnor (2001). ∎

From these formulas we can conclude that $V(\ell; \mathbf{B}(\mathbf{a})^{\mathsf{T}})$, the total magnitude of benefit cycles[25] at action profile $\mathbf{a}$, grows roughly as $e^{\ell \cdot \log r(\mathbf{B}(\mathbf{a}))}$ the sense that $V(\ell; \mathbf{B}(\mathbf{a})^{\mathsf{T}})$ gets that high infinitely often (and never gets higher). Thus, the logarithm of the spectral radius of the benefits matrix can be thought of as the growth rate of the cycle values as we increase the cycles' length.

Intuitively, benefit networks with an imbalanced structure, in which it is rare for the beneficiaries of one agent's effort to be able to directly or indirectly "give back", will have a lower spectral radius and, in the senses we have articulated earlier in this section, less scope for cooperation. Section 3.4.5 gives a concrete example where this insight is key to determining which agents are essential.

3.4.3   Comparing Marginal Returns

We recall a standard fact from linear algebra (see, e.g., Debreu and Hernstein 1953, Theorem I*).

FACT **3.** If $\mathbf{B}$ and $\widehat{\mathbf{B}}$ are two nonnegative matrices such that $\mathbf{B} \geq \widehat{\mathbf{B}}$, then $r(\mathbf{B}) \geq r(\widehat{\mathbf{B}})$.

A key economic consequence of this is a version of diminishing marginal returns: the spectral radius of the benefits matrix decreases as actions increase.

PROPOSITION **4.** If $\mathbf{a}' \geq \mathbf{a}$, then $r(\mathbf{B}(\mathbf{a}')) \leq r(\mathbf{B}(\mathbf{a}))$.

---

[25]Here we use that the spectral radius does not change if we take the transpose. Of course, it is perfectly legitimate to think of directed cycles in $\mathbf{B}$, without transposing; this way of looking at things simply leads to a different interpretation. In that case, the cycle $(1, 2, 3)$ is a cycle of dependence – 1 depends on the help of 2, who depends on the help of 3.

Proof of Proposition 4: By the concavity of the u_i, the inequality $\mathbf{B}(\mathbf{a}') \leq \mathbf{B}(\mathbf{a})$ holds (where the inequality between matrices means entrywise inequality).[26] Then the fact above implies the result. ∎

To formalize the interpretation of this as diminishing marginal returns, recall that according to Section 3.3, the spectral radius measures the returns to egalitarian investments at the margin.

3.4.4   How Essential is a Player?

The efficiency results provide a simple way of quantifying how essential any given player is to the negotiations. Suppose for a moment that a given player exogenously may or may not be able to participate in an institution to negotiate an outcome that Pareto dominates the status quo. If he is not able, then his action is set to the status quo level of $a_i = 0$. How much does such an exclusion hurt the prospects for cooperation by the rest of the society?

Without player i, the benefits matrix at the status quo of $\mathbf{0}$ is equal to the original $\mathbf{B}(\mathbf{0})$ without row and column i; equivalently, each entry in that row and column may be set to 0. Call a matrix constructed that way $\mathbf{B}^{[-i]}(\mathbf{0})$. By Proposition 3, the spectral radius of $\mathbf{B}^{[-i]}(\mathbf{0})$ is smaller than that of $\mathbf{B}(\mathbf{0})$. The most dramatic case is one in which the spectral radius of $\mathbf{B}(\mathbf{0})$ exceeds 1, but the spectral radius of $\mathbf{B}^{[-i]}(\mathbf{0})$ is less than 1. Then by Proposition 3 on egalitarian improvements, a Pareto improvement on $\mathbf{0}$ exists when i is present. But by Proposition 2, none exists when $\mathbf{0}$ is absent.

This argument shows that *player i's participation is essential to achieving any Pareto improvement on the status quo precisely when his removal changes the spectral radius of the benefits matrix at the status quo from being greater than* 1 *to being less than* 1.

More generally, the difference $r(\mathbf{B}(\mathbf{0})) - r(\mathbf{B}^{[-i]}(\mathbf{0}))$ measures how much the potential for cooperation at the margin is damaged when player i does not participate. The strategic implications of this are discussed in Section 8.1.

3.4.5   A More Elaborate Example: Who is Essential?

We now build on the example of Section 3.4.1 to illustrate what it means for a player to be essential. Suppose $N = \{1, 2, 3, 4\}$ and

$$\mathbf{B}(\mathbf{0}) = \begin{bmatrix} 0 & 0 & 7 & 0.5 \\ 5 & 0 & 6 & 0.5 \\ 0 & 0 & 0 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0 \end{bmatrix}. \tag{7}$$

---

[26]This is established as follows. Fix any i. Concavity of u_i implies that the Hessian of u_i in $\mathbf{a}$ is negative semidefinite. Therefore, each diagonal entry of this Hessian is nonpositive, and the principal minor of the Hessian corresponding to leaving only two indices i and j is nonpositive. Then it follows that $-J_{ii}(\mathbf{a})$ is increasing in a_i and $J_{ij}(\mathbf{a})$ is decreasing in a_j. But i and j were arbitrary, and we can get from $\mathbf{a}$ to $\mathbf{a}'$ by increasing actions one by one. That implies the needed inequality.

See Figure 3 for a graphical depiction of this benefits matrix (panel b) and a comparison with the example of Section 3.4.1.

(a) The example of Section 3.4.1.　　　　(b) The benefits network of equation (7).
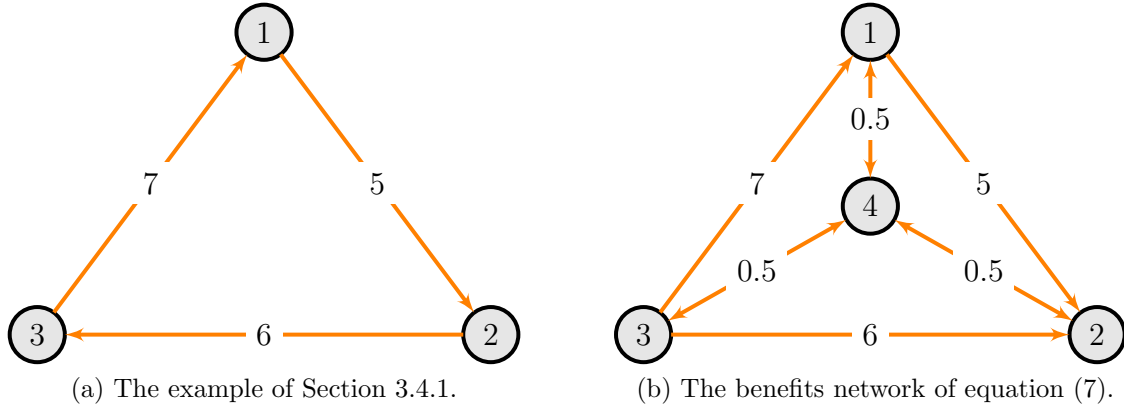
Figure 3: The benefit flows in (b) differ from those of (a) in two ways. First, the arrow from 2 to 3 has been flipped, destroying the directed cycle $1 \to 2 \to 3$. Second, a new player (#4) has been added, with bilateral links of weight 0.5 in the benefits matrix to all other players.

The import of the example is that player #4, even though he confers and receives the smallest benefits, is the only essential player. Without him, there are no cycles at all and the spectral radius of the corresponding benefits matrix $\mathbf{B}^{[-4]}(\mathbf{0})$ is 0 (by Fact 2). On the other hand, when he is present but any one other player $i \neq 4$ is absent, then there is always a cycle whose edges multiply to more than 1, and the spectral radius of $\mathbf{B}^{[-i]}(\mathbf{0})$ exceeds 1 (by Fact 1). Thus, the participation of a seemingly "small" player in negotiations can make an essential difference to the ability to improve on the status quo when that player completes cycles in the benefits network.

3.4.6　Bounding the Marginal Gains from Cooperation

In practical situations, the entries of the matrix $\mathbf{B}(\mathbf{a})$ may be known imperfectly, and some of them may not be known at all. Can an analyst give useful advice about how inefficient an outcome is in this circumstance?

We now explain the sense in which such advice under imperfect information is possible. If we can find an estimator $\widehat{\mathbf{B}}(\mathbf{a})$ such that $\widehat{B}_{ij}(\mathbf{a}) \leq B_{ij}(\mathbf{a})$, then $r(\widehat{\mathbf{B}}(\mathbf{a}))$ provides a lower bound on $r(\mathbf{B}(\mathbf{a}))$ by the monotonicity of the spectral radius (recall Section 3.4.3 above). In particular, if we have good theoretical reasons to believe that externalities are nonnegative for all pairs of players i and j but have no way to estimate some of the marginal benefits, then we can simply set $\widehat{B}_{ij} = 0$ for those pairs.

By the results of Sections 3.2 and 3.3, a lower bound on $r(\mathbf{B}(\mathbf{a}))$ has practical meaning: if we have such an estimate $\widehat{r}$ that exceeds 1, then we know that a tax on contributions of *at least* $\widehat{r} - 1$ is required to rationalize that outcome as Pareto efficient. We know that there is an egalitarian vector that yields a bang for the buck of *at least* $\widehat{r}$ at the current action profile. Thus, even partial knowledge of the network of benefit flows can be useful. However, the results in Section 3.4.2 do tell us that for

these estimates to yield a conclusion that the status quo is inefficient, there must be some cycles in the estimated benefits matrix.

Since in actual estimation, $\widehat{\mathbf{B}}(\mathbf{a})$ is a random variable, there is a question of formulating a probabilistic version of the monotonicity result. This is a direction for future work.

4 Lindahl Outcomes: Foundations

This section briefly reviews the framework and notation of implementation theory (Section 4.1) and formally defines Lindahl outcomes (Section 4.2). The purpose of the section is to state Theorem 1 in Section 4.3, which characterizes the Lindahl outcomes as the robustly implementable ones, in the sense discussed in the introduction.

4.1 The Implementation Theory Framework

We consider a designer of an institution who can leave future participants with a game to play. She does not know what preferences they will have. She believes they will play an equilibrium, but does not know which equilibrium. She would like to design the game so that all its equilibria have some desirable properties. What can she do? Which outcomes can she achieve?

Let $\mathcal{U}_A$ be the set of all functions $u : \mathbb{R}_+^n \to \mathbb{R}$. We denote by $\succeq_u$ and $\succ_u$ the weak and strict preference orderings, respectively, induced by $u \in \mathcal{U}_A$. The domain of possible preference profiles[27] is a set $\mathcal{U} \subseteq \mathcal{U}_A^n$; we will state specific assumptions on it in our results.

A *game form* is a tuple $H = (\Sigma_1, \dots, \Sigma_n, g)$ where:

- Σ_i is a set of *strategies* that agent i can play; we write $\Sigma = \prod_{i \in N} \Sigma_i$;

- $g : \Sigma \to \mathbb{R}_+^n$ is the *outcome function* that maps strategy profiles to action profiles.

DEFINITION 2. In a game form $H = (\Sigma_1, \dots, \Sigma_n, g)$, a strategy profile $\boldsymbol{\sigma} \in \Sigma$ is a *Nash equilibrium* for preference profile $\mathbf{u} \in \mathcal{U}$ if for any $i \in N$ and any $\widetilde{\sigma}_i \in \Sigma_i$, it holds that $g(\boldsymbol{\sigma}) \succeq_{u_i} g(\widetilde{\sigma}_i, \boldsymbol{\sigma}_{-i})$. We define $\Sigma^*(H, \mathbf{u})$ to be the set of all such $\boldsymbol{\sigma}$.

A *social choice correspondence* $F : \mathcal{U} \rightrightarrows \mathbb{R}_+^n$ maps each preference profile to a nonempty set of outcomes. Any game form for which equilibrium existence is guaranteed[28] naturally induces a social choice correspondence: its Nash equilibrium outcome correspondence $F_H(\mathbf{u}) = g(\Sigma^*(H, \mathbf{u}))$. The set $F_H(\mathbf{u})$ describes all the outcomes the participants with preferences $\mathbf{u}$ can end up with if they are left with a game form H

---

[27] The standard approach (e.g. Maskin, 1999) is to work with preference relations. We use sets of utility functions to avoid carrying around two parallel notations.

[28] Otherwise, we can still talk about the correspondence, but it will not be a social choice correspondence, which is required to be nonempty-valued.

and they play some Nash equilibrium. We say that F_H is the social choice correspondence that the game form H *implements*[29]. A social choice correspondence is said to be *implementable* if there is some game form H that implements it.

There are two basic normative criteria we impose on such correspondences.

DEFINITION 3. A social choice correspondence F is *Pareto efficient* if, for any $\mathbf{u} \in \mathcal{U}$ and $\mathbf{a} \in F(\mathbf{u})$, the profile $\mathbf{a}$ is Pareto efficient under $\mathbf{u}$.

DEFINITION 4. A social choice correspondence F is *individually rational* if, for any $\mathbf{u} \in \mathcal{U}$ and $\mathbf{a} \in F(\mathbf{u})$, it holds that $\mathbf{a} \succeq_{u_i} \mathbf{0}$.

An individually rational social choice correspondence is one that leaves every player no worse off than the status quo. For more on how to interpret this condition, see Section 8.1.

There is also a technical condition – upper hemicontinuity.

DEFINITION 5. A social choice correspondence F is *upper hemicontinuous* if, for any sequence $\left(\mathbf{u}^{(k)}\right)$ of preference profiles converging compactly[30] to $\mathbf{u}$, and any sequence of outcomes $\left(\mathbf{a}^{(k)}\right)$ such that $\mathbf{a}^{(k)} \in F\left(\mathbf{u}^{(k)}\right)$ for every k and $\mathbf{a}^{(k)} \to \mathbf{a}$, we have $\mathbf{a} \in F(\mathbf{u})$.

This condition has some normative appeal in that a social choice correspondence not satisfying upper hemicontinuity is sensitive to arbitrarily small changes in preferences that may be difficult for the agents themselves to detect.

4.2 Definition of Lindahl Outcomes

Informally, a Lindahl outcome is an analogue, in a public goods setting, of a Walrasian equilibrium allocation. Instead of standard prices for private goods, there are personalized taxes and subsidies: each player pays a tax for every public good he enjoys (in proportion to how much of that public good is produced), and receives a personalized subsidy (financed by others' taxes) per unit of effort he invests in the public good he provides. More formally:

DEFINITION 6. An action profile $\mathbf{a}^*$ is a *Lindahl outcome* for a preference profile $\mathbf{u}$ if there is an n-by-n matrix $\mathbf{P}$ so that the following conditions hold for every i:

(i) The inequality

$$\sum_{j:j \neq i} P_{ij} a_j \leq a_i \sum_{j:j \neq i} P_{ji} \qquad (\text{BB}_i(\mathbf{P}))$$

is satisfied when $\mathbf{a} = \mathbf{a}^*$;

(ii) for any $\mathbf{a}$ such that $\text{BB}_i(\mathbf{P})$ is satisfied, we have $\mathbf{a}^* \succeq_{u_i} \mathbf{a}$.

---

[29]To be more precise, this is full Nash implementation. Since we only consider this kind of implementation, we drop the adjectives.

[30]That is, converging uniformly on every compact set.

Note that there need not be any transferrable private commodity in which these prices are denominated. We can think of each player having access to artificial tokens, facing prices for the public goods denominated in these tokens, and being able to choose any outcome subject to not using more tokens than he receives from others. If all agents are choosing optimally at the given prices subject to this constraint, then we are at a Lindahl outcome.

The *Lindahl correspondence* $L : \mathcal{U} \rightrightarrows \mathbb{R}^n_+$ is defined by

$$L(\mathbf{u}) = \{\mathbf{a} \in \mathbb{R}^n_+ : \mathbf{a} \text{ is a Lindahl outcome for } \mathbf{u}\}.$$

In terms of interpretation, P_{ij} is the price i pays j for the effort of j, and $\sum_{j:j\neq i} P_{ji}$ is the "wage" or subsidy player i receives per unit of effort. Thus, $\mathrm{BB}_i(\mathbf{P})$ is simply a budget balance condition: expenditures are no greater than income.

The definition says that an action profile is a Lindahl outcome if there are some prices so that the amounts of externalities produced and consumed are optimal for each player, subject to budget balance at those prices. For any $\mathbf{u}$ satisfying the assumptions of Section 2.2, the set $L(\mathbf{u})$ is nonempty – see Section 7.1.

This definition, though simple, does not relate in an obvious way to realistic markets; Samuelson (1954) elaborates on this. It is the implementation-theoretic rationale that we are about to formally state that motivates our study of the Lindahl outcomes.

4.3   The Hurwicz Rationale for Lindahl Outcomes

Fix $\mathcal{U}$. Let $\mathcal{F}$ be the set of implementable social choice correspondences $F : \mathcal{U} \rightrightarrows \mathbb{R}^n_+$ that are Pareto efficient, individually rational, and upper hemicontinuous. For any $\mathbf{u} \in \mathcal{U}$, define the set of outcomes prescribed at $\mathbf{u}$ by *every* such correspondence:

$$R(\mathbf{u}) = \bigcap_{F \in \mathcal{F}} F(\mathbf{u}). \tag{8}$$

This defines a correspondence $R : \mathcal{U} \rightrightarrows \mathbb{R}^n_+$. We call this the *robustly attainable* correspondence.

If the set of possible preferences is rich enough, then the robustly attainable correspondence is precisely the Lindahl correspondence.

THEOREM 1. Suppose $\mathcal{U}$ is the set of all preference profiles satisfying the assumptions of Section 2.2, and the number of players n is at least 3. Then the robustly attainable correspondence is equal to the Lindahl correspondence: $R = L$.

As we note in Section 7.1 below, L is nonempty. Thus L is a minimal Pareto efficient, individually rational, upper hemicontinuous social choice correspondence: the outcomes it prescribes are a subset of those prescribed by any other social choice correspondence with these properties. In particular, if the designer wants a mechanism with a minimal set of equilibria (e.g., a unique equilibrium if possible), then she must choose a game form H such that its equilibrium outcomes are precisely the ones selected by the Lindahl solution.

A proof of Theorem 1 can be found in Section 7.2. As we discuss formally in Section A.4, it is not necessary to assume that $\mathcal{U}$, the set of all preference profiles the designer considers possible, is equal to the set of all preference profiles satisfying the assumptions of Section 2.2. It suffices, for example, to assume that $\mathcal{U}$ contains a neighborhood of every linear utility function. Actually, $\mathcal{U}$ can be assumed to be even smaller than this.

The above theorem is an analogue of Theorem 3 of Hurwicz (1979a). Because the environment studied in that paper (with assumptions such as nonzero endowments of all private goods) is not readily adapted to ours, we prove the result separately, using Hurwicz's insights combined with Maskin's Theorem.

5 The Main Result: A Network Centrality Characterization of the Lindahl Outcomes

Having explained what makes Lindahl outcomes special from an implementation perspective, we present our main results on their characterization. Section 5.1 states two equivalent definitions of eigenvector centrality action profiles. Section 5.2 is devoted to Theorem 2, which asserts, in essence, that the Lindahl outcomes are the same as the eigenvector centrality action profiles. Section 5.3 gives a simple first application: characterizing the outcome when there is a global public good whose production is additive and symmetric in the players' effort levels. Section 5.4 derives a general inequality on contribution levels at Lindahl outcomes: agents who enjoy all public goods more at the margin also provide more themselves.

5.1 Eigenvector Centrality Action Profiles

5.1.1 Definition

Fix a preference profile $\mathbf{u}$ throughout this subsection and the next. The main theorem characterizes Lindahl outcomes as ones satisfying a network centrality condition.

DEFINITION 7. An action profile $\mathbf{a} \in \mathbb{R}_+^n$ is an *eigenvector centrality action profile* if $\mathbf{a} \neq \mathbf{0}$ and $\mathbf{B}(\mathbf{a})\mathbf{a} = \mathbf{a}$.

The name comes from the fact that $\mathbf{a}$ is, according to the condition, a right-hand eigenvector of $\mathbf{B}(\mathbf{a})$ with eigenvalue 1. In other words, for each $i \in N$,

$$a_i = \sum_j B_{ij} a_j. \tag{9}$$

Equation (9) asserts that each player's contribution is a weighted sum of the other players' contributions, where the weight on a_j is proportional to the marginal benefits that j confers on i.

5.1.2 An Alternative Definition: Scaling-Indifference

An eigenvector centrality action profile is equivalently characterized by a convenient property called scaling-indifference, which we define and then explain.

DEFINITION 8. An action profile $\mathbf{a} \in \mathbb{R}^n_+$ satisfies *scaling-indifference* (or *is scaling-indifferent*) if $\mathbf{a} \neq \mathbf{0}$ and $\mathbf{J}(\mathbf{a})\mathbf{a} = \mathbf{0}$.

The vector $\mathbf{J}(\mathbf{a})\mathbf{v}$ gives the derivatives of utilities in ϵ when actions are changed from $\mathbf{a}$ to $\mathbf{a} + \epsilon\mathbf{v}$ for some vector $\mathbf{v} \in \mathbb{R}^n$. That is, to a first-order approximation, $\mathbf{u}(\mathbf{a} + \epsilon\mathbf{v}) \approx \mathbf{u}(\mathbf{a}) + \epsilon\mathbf{J}(\mathbf{a})\mathbf{v}$. Suppose now that actions are scaled by $1 + \epsilon$, for some small real number ϵ; this corresponds to setting $\mathbf{v} = \mathbf{a}$. If $\mathbf{J}(\mathbf{a})\mathbf{a} = \mathbf{0}$, then all players are indifferent, at the margin, to this small proportional perturbation in everyone's actions.

It is immediate to check that the vector $\mathbf{a}$ is an eigenvector centrality action profile if and only if it satisfies scaling-indifference.

5.2 The Characterization

We now present our main theorem, characterizing Lindahl outcomes.

THEOREM 2. Under the assumptions in Section 2.2, the following are equivalent for a nonzero $\mathbf{a} \in \mathbb{R}^n_+$:

 (i) $\mathbf{a}$ is a Lindahl outcome;

 (ii) $\mathbf{J}(\mathbf{a})\mathbf{a} = \mathbf{0}$ – i.e., $\mathbf{a}$ is scaling-indifferent;

 (iii) $\mathbf{B}(\mathbf{a})\mathbf{a} = \mathbf{a}$ – i.e., $\mathbf{a}$ is an eigenvector centrality action profile.

The key reasons behind this characterization are discussed in Section 7.3. The formal proof appears in Section A.5.

REMARK 4. The issue of when the profile $\mathbf{0}$ is a Lindahl outcome is treated in Proposition 8 in Section A.6; the most important fact is that $\mathbf{0}$ is a Lindahl outcome if and only if it is a Pareto efficient outcome.

The main theorem implies some simple consequences, tying up loose ends from the earlier exposition.

REMARK 5. By Theorem 2, at any nonzero Lindahl outcome $\mathbf{a}$, the matrix $\mathbf{B}(\mathbf{a})$ has a nonnegative right eigenvector $\mathbf{a}$ with eigenvalue 1, and therefore, by the Perron-Frobenius Theorem (part (iii)), a spectral radius of 1. Proposition 1 then implies the Pareto efficiency of $\mathbf{a}$ (we use Lemma 1 in Section A.3 to obtain that $\mathbf{a}$ is interior). Of course, the standard proof of the First Welfare Theorem also goes through: see, e.g., Foley (1970).

REMARK 6. The scaling-indifference condition is equivalent to the assertion that $\mathbf{d} = \mathbf{a}/\sum_i a_i$ is an egalitarian direction at $\mathbf{a}$ with corresponding bang for the buck $b(\mathbf{a}, \mathbf{d}) = 1$ (recall Section 3.3). Therefore, we can alternatively characterize Lindahl outcomes as Pareto efficient ones under which moving in the direction given by $\mathbf{a}$ itself is egalitarian.

REMARK 7. Note that the condition $\mathbf{J}(\mathbf{a})\mathbf{a} = \mathbf{0}$ does not change if we form new utility functions according to the transformation discussed in Remark 2 in Section 3.1. The Jacobian of the "new" utilities $\widehat{\mathbf{u}}$ is, by the chain rule, simply $\mathbf{F}(\mathbf{u}(\mathbf{a}))\mathbf{J}(\mathbf{a})$, where $\mathbf{F}(\mathbf{u})$ is a diagonal matrix with $F_{ii}(u_i) = f'_i(u_i)$. Since $\mathbf{F}(\mathbf{u}(\mathbf{a}))$ is nonsingular for the transformation being considered, it follows that $\mathbf{J}(\mathbf{a})\mathbf{a} = \mathbf{0}$ if and only if $\mathbf{F}(\mathbf{u}(\mathbf{a}))\mathbf{J}(\mathbf{a})\mathbf{a} = \mathbf{0}$.

REMARK 8. The condition $\mathbf{J}(\mathbf{a})\mathbf{a} = \mathbf{0}$ is a system of n equations in n unknowns (the coordinates of $\mathbf{a}$). By a standard Sard's Theorem argument, this entails that for generic utility functions satisfying our assumptions, the set of solutions will be of dimension zero in $\mathbb{R}^n_+$. Therefore, the set of Lindahl outcomes is typically small, as is usually the case with sets of market equilibria.

5.3   Example: A Global Public Good

The purpose of the general theory is, of course, to study environments with arbitrary heterogeneity in marginal costs and benefits across pairs. But to get a feeling for the mechanics of the characterization, it is helpful to consider a simple special case in which the public good is global.

Suppose each player's preferences are separable in costs and benefits, and the benefits depend only on the sum of all effort levels:

$$u_i(\mathbf{a}) = v_i\left(\sum_{j \in N} a_j\right) - c_i(a_i),$$

where each v_i is continuously differentiable, increasing, and concave, while c_i is continuously differentiable, increasing, and convex. To ensure that Assumption 1 holds, we assume $v'_i(0) - c'_i(0) < 0$. We can think of the v_i term as being the value of air quality and the c_i term as being the costs of foregone consumption.

To characterize the nonzero Lindahl outcomes, we use Theorem 2(ii). First, observe that:

$$J_{ij}(\mathbf{a}) = \begin{cases} v'_i\left(\sum_{k \in N} a_k\right) & \text{if } i \neq j \\ -c'_i(a_i) & \text{if } i = j. \end{cases}$$

Then the condition $\mathbf{J}(\mathbf{a})\mathbf{a} = \mathbf{0}$ can be rewritten:

$$\frac{a_i}{\sum_{j \in N} a_j} = \frac{v'_i\left(\sum_{j \in N} a_j\right)}{c'_i(a_i)} \tag{10}$$

Thus, each agent contributes a share[31] *of the total public good provision equal to his marginal rate of substitution between clean air and consumption.*

---

[31] The right-hand side is less than 1 because we assumed that $v'_i(0) - c'_i(0) < 0$, and our assumptions about v_i and c_i entail that $v_i - c_i$ is a concave function.

5.4 A Simple Inequality

The formula of Theorem 2(iii) equates an agent's contribution at a Lindahl outcome to quantities involving the marginal value of *incoming* benefits. This can be used to derive a general result on relative contributions: if, at every outcome, agent i's values the help of any other agent k more than j does (in the sense of marginal rates of substitution: i.e., $B_{ik}(\mathbf{a}) \geq B_{jk}(\mathbf{a})$ for all $\mathbf{a} \in \mathbb{R}^n_+$ and all k) then i always takes a weakly greater action than j at any Lindahl outcome $\mathbf{a}^*$ – that is, $a_i^* \geq a_j^*$. This generalizes (10) above.

5.5 An Interpretation of Eigenvector Centrality via Walks

In Section 3.4.2, we saw that the spectral radius of the benefits matrix could be interpreted through the values of long cycles. A related interpretation applies to eigenvector centrality action profiles. A *walk of length ℓ in the matrix* $\mathbf{M}$ is a sequence $(w(1), w(2), \ldots, w(\ell))$ of elements of N (player indices) such that $M_{w(t)w(t+1)} > 0$ for each $t \in \{1, 2, \ldots, \ell - 1\}$.[32] Let $\mathcal{W}_i^{\downarrow}(\ell; \mathbf{M})$ be the set of all walks of length ℓ in $\mathbf{M}$ such that $w(\ell) = i$ – that is, the set of walks ending at i. For a matrix $\mathbf{M}$, define the value of a walk w of length ℓ as the product of all weights along the walk:

$$v(w; \mathbf{M}) = \prod_{t=1}^{\ell-1} M_{w(t)w(t+1)}.$$

Note that such walks can repeat nodes – for example, they may cover the same cycle many times. Then we have the following:

PROPOSITION 5. Let $\mathbf{M} = \mathbf{B}(\mathbf{a})^{\mathsf{T}}$ and assume this matrix is aperiodic.[33] Then $\mathbf{a}$ is an eigenvector centrality action profile if and only if, for every i and j,

$$\frac{a_i}{a_j} = \lim_{\ell \to \infty} \frac{\displaystyle\sum_{w \in \mathcal{W}_i^{\downarrow}(\ell; \mathbf{M})} v(w; \mathbf{M})}{\displaystyle\sum_{w \in \mathcal{W}_j^{\downarrow}(\ell; \mathbf{M})} v(w; \mathbf{M})}.$$

A walk in $\mathbf{B}(\mathbf{a})^{\mathsf{T}}$ ending at i can be thought of as a chain of benefit flow: e.g., k helps j, who helps i. The weight of such a walk is the product of the marginal benefits along its links. According to Proposition 5, a player at an eigenvector centrality action profile (and hence a Lindahl outcome) contributes in proportion to the total strength of such benefit chains that he terminates.[34]

---

[32] As with cycles, defined in Section 3.4.2, nodes can be repeated in this sequence. Note also that a cycle of length ℓ is a walk of length ℓ.

[33] A *simple cycle* is one that has no repeated nodes except the initial/final one. A matrix is said to be *aperiodic* if the greatest common divisor of the lengths of all simple cycles in that matrix is 1.

[34] The formula of the proposition would also hold if had we defined $\mathbf{M} = \mathbf{B}(\mathbf{a})$ and replaced $\mathcal{W}_i^{\downarrow}(\ell; \mathbf{M})$ by $\mathcal{W}_i^{\uparrow}(\ell; \mathbf{M})$ – the set of walks of length ℓ in $\mathbf{M}$ that *start* at i. The convention we use above is in keeping with thinking of a walk in $\mathbf{B}(\mathbf{a})^{\mathsf{T}}$ capturing the way benefits flow; recall the discussion in Section 3.4.2.

6 Explicit Formulas for Lindahl Outcomes

The general characterization of Lindahl outcomes in terms of network centrality makes no parametric assumptions on preferences. The cost of this generality is that the characterization is implicit: an eigenvector centrality action profile $\mathbf{a}$ is defined with respect to a matrix, $\mathbf{B}(\mathbf{a})$, that depends on $\mathbf{a}$ itself. To build intuition, it is useful to study cases in which the Lindahl outcomes can be characterized explicitly in terms of exogenous parameters.

To this end, the present section examines a parametric family of preferences in which explicit formulas for the Lindahl outcomes are available. We study three different regimes within this parametric family, and in each of them there is a connection between Lindahl outcomes and an important centrality measure that has been studied independently in the networks literature.[35] In the first regime, actions are equal to degree centralities. In the second, actions are equal to Bonacich centralities. And in the third regime, actions approximate eigenvector centralities. In each case, the centralities are defined relative to an exogenous network, and all the results are special cases of Theorem 2.

The propositions of this section may also be viewed as microfoundations for the network centrality measures in terms of price equilibria, because the Lindahl outcomes are defined in terms of prices (recall Definition 6 in Section 4.2). Each result described below says that for particular preferences, the "market" levels of public good provision defined by Lindahl are equal to centralities according to a particular measure. Of course, in our framework the Lindahl outcomes are motivated primarily by their implementation-theoretic foundations, but the connection to prices is also worth noting.

Substantively, the import of this section is that in each case, what matters for the level of i's public good provision at a Lindahl outcome is the extent to which i is the *recipient* of benefits. In the simplest case – Section 6.2 – i's effort level is simply the sum of coefficients that determine how much each of his neighbors can help him. In Sections 6.3 and 6.4, what matters is not only links carrying benefits from other agents to i, but *chains* or *walks* of favor-giving links, involving multiple agents, that terminate at i.

6.1 A Parametric Family of Preferences and an Explicit Formula for Lindahl Outcomes

Let $\mathbf{G}$ and $\mathbf{H}$ be nonnegative matrices with zeros on the diagonal and $r(\mathbf{G}) < 1$. For each $i \in N$ let[36]

$$u_i(\mathbf{a}) = -a_i + \sum_j \left[G_{ij} a_j + H_{ij} \log a_j \right]. \tag{11}$$

---

[35]We will introduce each measure without assuming background in network topics.

[36]These should be viewed as functions $u_i : \mathbb{R}_+^n \to \mathbb{R} \cup \{-\infty\}$, with $0 \cdot \log 0$ understood as 0. In other words, preferences should be completed by continuity to the extended range. No result in the paper is affected by this slight departure from the framework of Section 2.

Then the benefits matrix $\mathbf{B}(\mathbf{a})$ is given by:

$$B_{ij}(\mathbf{a}) = \begin{cases} G_{ij} + \frac{H_{ij}}{a_j} & \text{if } i \neq j \\ 0 & \text{otherwise.} \end{cases}$$

Defining $h_i = \sum_j H_{ij}$, the condition for $\mathbf{a}$ to be an eigenvector centrality action profile – i.e., $\mathbf{a} = \mathbf{B}(\mathbf{a})\mathbf{a}$ – boils down to

$$\mathbf{a} = \mathbf{h} + \mathbf{G}\mathbf{a}$$

or

$$\mathbf{a} = (\mathbf{I} - \mathbf{G})^{-1}\mathbf{h}. \tag{12}$$

The vector $\mathbf{a}$ is well-defined and nonnegative[37] by the assumption that $r(\mathbf{G}) < 1$. This formula will yield all of the results in this section.

6.2 Degree Centrality

The most basic measure of connectedness is the total weight of links that a node has.

DEFINITION 9. For a nonnegative matrix $\mathbf{M}$ the degree centrality in $\mathbf{M}$ is defined by the following equation for each i:

$$\delta_i(\mathbf{M}) = \sum_j M_{ij}$$

Suppose $\mathbf{G}$ is the zero matrix and $\mathbf{H} = \mathbf{M}$ for a matrix $\mathbf{M}$ with zeros on the diagonal. Then it follows immediately from (12) that, under the preferences of (11) with these parameter values, we have

$$\mathbf{a} = \delta_i(\mathbf{M}).$$

In other words, when costs are linear in one's own action and benefits are logarithmic in others' actions, then an agent i's contribution is determined by how much he benefits from everyone else's effort at the margin: the sum of coefficients H_{ij} as j ranges across the other agents. The agents who are particularly dependent on the rest are the ones who are contributing the most. Note that how much others depend on i does not figure at all in this formula. If i's help suddenly becomes much more valuable to others, then his action does not change, though his share of the overall contribution goes down (because others' action levels go up).

Thus, even in the simplest case, this analysis already yields some stark comparative statics in levels of provision.

6.3 Bonacich Centrality

The second centrality measure we consider is due to Bonacich (1987).[38]

---

[37]See Ballester, Calvó-Armengol, and Zenou (2006, Section 3).
[38]An important antecedent was discussed by Katz (1953).

DEFINITION **10.** For a nonnegative matrix $\mathbf{M}$ and a constant $\alpha < 1/\rho(\mathbf{M})$, the Bonacich centrality of parameter α in $\mathbf{M}$ is defined by

$$\boldsymbol{\beta}(\mathbf{M}, \alpha) = [\mathbf{I} - \alpha\mathbf{M}]^{-1}\mathbf{1},$$

where $\mathbf{1}$ is the column vector of ones.

Dropping the arguments, the defining equation says that for every i, we have:

$$\beta_i = 1 + \alpha \sum_j M_{ij}\beta_j.$$

Thus, every node gets a baseline level of centrality (one unit) and then additional centrality in proportion to the centrality of those it is linked to.

Now, returning to the preferences defined by equation (11), let $\mathbf{G} = \alpha\mathbf{M}$ for any $\mathbf{M}$ with zeros on the diagonal and $\alpha < 1/r(\mathbf{M})$. If we fix any $\mathbf{H}$ such that each row sums to 1, then (12) implies

$$\mathbf{a} = \boldsymbol{\beta}(\mathbf{M}, \alpha). \tag{13}$$

An Interpretation via Walks To shed further light on this result, recall the definitions and notation related to walks from Section 5.5, and let

$$V_i(\ell; \mathbf{M}) = \sum_{w \in \mathcal{W}_i^{\downarrow}(\ell; \mathbf{M})} v(w; \mathbf{M}).$$

This is the sum of the values of all walks of length ℓ in $\mathbf{M}$ ending at i. Then we have:

FACT **4.** $\beta_i(\mathbf{M}, \alpha) = 1 + \sum_{\ell=1}^{\infty} \alpha^\ell V_i(\ell; \mathbf{M}^\mathsf{T})$.

Fact 4 is established, e.g., in Ballester, Calvó-Armengol, and Zenou (2006, Section 3). Thus, the Bonacich centrality is equal to one plus a weighted sum of values of all walks in $\mathbf{M}^\mathsf{T}$ terminating at i, with longer walks downweighted exponentially.

In contrast to the case of degree centrality treated in the previous section, it is not only how much i benefits from his immediate neighborhood that matters in determining his contribution, but also how much i's neighbors benefit from *their* neighbors, etc.

6.4   Eigenvector Centrality

Eigenvector centrality was defined in the introduction, and is a key notion throughout the paper. Indeed, one may ask why this section is necessary – after all, doesn't Theorem 2 establish a general connection between eigenvector centrality and Lindahl outcomes?

The answer is that Theorem 2 characterizes $\mathbf{a}$ through an *endogenous* eigenvector centrality condition – a condition that depends on $\mathbf{B}(\mathbf{a})$. In this section, we study the special case in which action levels approximate eigenvector centralities defined according to an *exogenous* network.

Fix any nonnegative, irreducible matrix $\mathbf{M}$. By the Perron-Frobenius Theorem, $\mathbf{M}$ has a unique eigenvector centrality vector $\mathbf{e}$, with entries summing to 1, so that $\mathbf{e} = \gamma \cdot \mathbf{M}\mathbf{e}$, where $\gamma = 1/r(\mathbf{M})$.

In the specification (11), let $\mathbf{H}$ be any matrix such that each row of $\mathbf{H}$ sums to 1. For any $z < 1$, let $\mathbf{G} = \frac{z}{r(\mathbf{M})}\mathbf{M}$.

By the result just established in Section 6.3,

$$\mathbf{a} = \boldsymbol{\beta}\left(\mathbf{M}, \frac{z}{r(\mathbf{M})}\right).$$

The key fact that allows us to connect outcomes to *eigenvector* centrality as opposed to Bonacich centrality is this: as $z \uparrow 1$, it holds that $a_i/a_j \to e_i/e_j$ for every i, j. That is, each player's share of the total of all actions converges to his eigenvector centrality according to $\mathbf{M}$. The essential reason for this convergence is presented in the proof of Theorem 3 of Golub and Lever (2010); see also Bonacich (1991).

7 Tying up the Theory

This section focuses on the proofs of the two most important results – Theorem 1 and Theorem 2 – as well as on the existence of Lindahl outcomes.

First, Section 7.1 states the existence of outcomes satisfying the three equivalent conditions of Theorem 2, and in particular the existence of Lindahl outcomes. Section 7.2 uses Theorem 2 to prove Theorem 1 (the theorem on what makes Lindahl outcomes special from the implementation perspective). Section 7.3 sketches the proof of our main result, Theorem 2; the full proof appears in the appendix.

7.1 Existence of Lindahl Outcomes

For completeness, we state a result on the existence of Lindahl outcomes in our setting.[39]

PROPOSITION **6.** If $\mathbf{u}$ satisfies the assumptions of Section 2.2, then $L(\mathbf{u})$ is nonempty, i.e., there is a Lindahl outcome for this preference profile. Moreover, either $\mathbf{0} \in L(\mathbf{u})$ or there is a Lindahl outcome in which all actions are strictly positive.

The proof appears in Section A.8, and works by directly showing the existence of scaling-indifferent outcomes; existence of Lindahl outcomes then follows by Theorem 2. The existence of scaling-indifferent outcomes may be of some independent interest, and is established without reference to prices, since the condition $\mathbf{J}(\mathbf{a})\mathbf{a} = \mathbf{0}$ does not involve prices.

---

[39]See Foley (1970) for a general existence proof. Note, however, that adapting Foley's environment to ours – in which, for example, there are no *ex ante* upper bounds on public good provision – is not entirely trivial.

7.2 Proof of Theorem 1 on the Hurwicz Rationale for Lindahl Outcomes

In this section, we give a proof of Theorem 1 in Section 4.3, which asserts that the only robustly attainable outcomes are the Lindahl outcomes.

To do this, we briefly recall Maskin's Theorem. Assuming that the number of agents n is at least 3 and and that a social choice correspondence F satisfies *no veto power*[40] (a condition that is vacuously satisfied in our setting), then F is implementable if and only if it satisfies *Maskin monotonicity*.

DEFINITION 11. A social choice correspondence $F : \mathcal{U} \rightrightarrows \mathbb{R}_+^n$ satisfies *Maskin monotonicity* if, for all $\mathbf{a}^* \in \mathbb{R}_+^n$ and all $\widehat{\mathbf{u}}, \mathbf{u} \in \mathcal{U}$, if $\mathbf{a}^* \in F(\widehat{\mathbf{u}})$ and

$$\forall i \in N, \forall \mathbf{a} \in \mathbb{R}_+^n, \quad \mathbf{a}^* \succeq_{\widehat{u}_i} \mathbf{a} \ \Rightarrow \ \mathbf{a}^* \succeq_{u_i} \mathbf{a} \tag{14}$$

then $\mathbf{a}^* \in F(\mathbf{u})$.[41]

Proof of Theorem 1: First we show that[42] $R \subseteq L$. By the definition that $R(\mathbf{u}) = \bigcap_{F \in \mathcal{F}} F(\mathbf{u})$, it suffices to show that $L \in \mathcal{F}$; i.e. that L is an implementable, individually rational, Pareto efficient, and upper hemicontinuous social choice correspondence. First, a social choice correspondence must be nonempty-valued; Proposition 6 in Section 7.1 guarantees that L complies. By Assumption 3, the no veto power condition is vacuous in our setting. It is verified immediately from Definition 6 that L satisfies Maskin monotonicity.[43] Thus, L is implementable by Maskin's Theorem. Also, L is individually rational since, by definition of a Lindahl outcome, each agent prefers a Lindahl outcome to $\mathbf{0}$, which is always feasible. By the standard proof of the First Welfare Theorem, L is Pareto efficient (see, e.g., Foley, 1970). Similarly, the standard argument for the upper hemicontinuity of equilibria in preferences transfers to our setting.

Now assume F is implementable, Pareto efficient, individually rational, and upper hemicontinuous. Fix $\mathbf{u} \in \mathcal{U}$ and $\mathbf{a}^* \in L(\mathbf{u})$. We will show $\mathbf{a}^* \in F(\mathbf{u})$. Define

$$\widehat{\mathbf{u}}(\mathbf{a}) = \mathbf{J}(\mathbf{a}^*; \mathbf{u})\mathbf{a}.$$

Lemma 2 in Section A.4 states that since F is individually rational, Pareto efficient, and upper hemicontinuous, it follows that $\mathbf{a}^* \in F(\widehat{\mathbf{u}})$.[44] Note that for all $\mathbf{a} \in \mathbb{R}_+^n$, we

---

[40] A social choice correspondence $F : \mathcal{U} \rightrightarrows \mathbb{R}_+^n$ satisfies no veto power if, for every $\mathbf{u}$, whenever there is an $\mathbf{a}$ such that $\mathbf{a} \succeq_{u_i} \mathbf{a}'$ for all $i \neq i'$ and all $\mathbf{a}' \in \mathbb{R}_+^n$, then $\mathbf{a} \in F(\mathbf{u})$.

[41] In words, if an alternative $\mathbf{a}$ was selected by F under $\mathbf{u}$ and then we change those preferences so that $\mathbf{a}$, in each agent's preference, defeats all the same alternatives as under $\mathbf{u}$ and perhaps some others, then $\mathbf{a}$ is still selected under $\widehat{\mathbf{u}}$.

[42] For two correspondences $F, F^\ddagger : \mathcal{U} \to \mathbb{R}_+^n$, we write $F \subseteq F^\ddagger$ if for every $\mathbf{u} \in \mathcal{U}$, it holds that $F(\mathbf{u}) \subseteq F^\ddagger(\mathbf{u})$. In this case, we say that F is a *sub-correspondence* of $F^\ddagger$.

[43] If $\widehat{\mathbf{u}}$ and $\mathbf{u}$ are as in the above definition of Maskin monotonicity and $\mathbf{a}$ is a Lindahl outcome under preferences $\widehat{\mathbf{u}}$, then using the same price matrix $\mathbf{P}$, the outcome $\mathbf{a}$ still satisfies condition (ii) in Definition 6.

[44] The proof of that lemma constructs a sequence of preference profiles $(\widehat{\mathbf{u}}^{(k)})$ converging to $\widehat{\mathbf{u}}$ such that individual rationality and Pareto efficiency alone force the set $F(\widehat{\mathbf{u}}^{(k)})$ to converge to $\mathbf{a}^*$. Then by upper hemicontinuity of F, it follows that $F(\widehat{\mathbf{u}})$ contains $\mathbf{a}^*$.

have
$$\widehat{\mathbf{u}}(\mathbf{a}^*) - \widehat{\mathbf{u}}(\mathbf{a}) = \mathbf{J}(\mathbf{a}^*; \mathbf{u})(\mathbf{a}^* - \mathbf{a}) \leq \mathbf{u}(\mathbf{a}^*) - \mathbf{u}(\mathbf{a})$$

by concavity of $\mathbf{u}$, so (14) holds. Since F is implementable, it satisfies Maskin monotonicity, so we conclude that $\mathbf{a}^* \in F(\mathbf{u})$. ∎

7.3 Centrality Characterization of Lindahl Outcomes: Proof Outline

It will be convenient to use in this section an equivalent definition of Lindahl outcomes:

DEFINITION **12.** An action profile $\mathbf{a}^*$ is a *Lindahl outcome* for a preference profile $\mathbf{u}$ if there exists an n-by-n matrix $\mathbf{P}$ with each column summing to zero so that the following conditions hold for every i:

(i) The inequality

$$\sum_{j \in N} P_{ij} a_j \leq 0 \qquad\qquad (\widehat{\mathrm{BB}}_i(\mathbf{P}))$$

is satisfied when $\mathbf{a} = \mathbf{a}^*$;

(ii) for any $\mathbf{a}$ such that $\widehat{\mathrm{BB}}_i(\mathbf{P})$ is satisfied, we have $\mathbf{a}^* \succeq_{u_i} \mathbf{a}$.

Given a Lindahl outcome defined as in Definition 6, set $P_{ii} = -\sum_{j:j\neq i} P_{ji}$ to find prices satisfying the new definition.[45] Conversely, the prices of Definition 12 work in Definition 6 without modification, since the original definition does not involve the diagonal terms of $\mathbf{P}$ at all.

The difficult part of Theorem 2 is showing that, for a nonzero $\mathbf{a}^*$, the condition $\mathbf{J}(\mathbf{a}^*)\mathbf{a}^* = \mathbf{0}$ implies that $\mathbf{a}^*$ is a Lindahl outcome. (The converse implication is straightforward, as we will note at the end of this section.) The key to the proof is to observe that the system of equations $\mathbf{J}(\mathbf{a}^*)\mathbf{a}^* = \mathbf{0}$ actually encodes all of the equilibrium conditions. It allows us to extract Pareto weights that support the outcome $\mathbf{a}^*$ as efficient, and using those Pareto weights and the Jacobian, we can construct prices that support $\mathbf{a}^*$ as a Lindahl outcome.

Now in more detail: suppose $\mathbf{J}(\mathbf{a}^*)\mathbf{a}^* = \mathbf{0}$. We will assume that $\mathbf{a}^*$ is interior, which is legitimate by Lemma 1 in Section A.3 (which shows that nonzero Lindahl outcomes are interior). Recalling Remark 5, the matrix $\mathbf{B}(\mathbf{a}^*)$ has a spectral radius of 1 and thus, by Proposition 1, the profile $\mathbf{a}^*$ is a Pareto efficient outcome. So there are Pareto weights $\boldsymbol{\theta}$ such that $\mathbf{a}^*$ solves $\mathcal{P}(\boldsymbol{\theta})$. (This is the point where the Perron-Frobenius Theorem plays a key role – recalling the proof of Proposition 1 in Section 3.1.) That implies that the first-order conditions of the Pareto problem $\mathcal{P}(\boldsymbol{\theta})$ defined in the proof of Proposition 1 are satisfied: for each j, we have

$$\sum_{i \in N} \theta_j J_{ij}(\mathbf{a}^*) = 0.$$

---

[45] In essence, $-P_{ii}$ is the wage of agent i, equal to the sum of taxes paid for his effort.

Let us guess prices $P_{ij} = \theta_j J_{ij}(\mathbf{a}^*)$. Then $\sum_i P_{ij} = 0$ for each j. Thus, the price matrix is a legitimate candidate for Definition 12. Moreover, the budget balance condition $\widehat{\mathrm{BB}}_i(\mathbf{P})$ is satisfied at these prices, since, for each i,

$$\sum_{j \in N} P_{ij} a_j = \theta_i \sum_{j \in N} J_{ij}(\mathbf{a}^*) a_j^* = 0.$$

The first equality is by construction of the prices, and the second is by the assumption that $\mathbf{J}(\mathbf{a}^*)\mathbf{a}^* = \mathbf{0}$. Thus, each agent is exhausting his budget. It only remains to see that each agent is optimizing at prices $\mathbf{P}$. The essential reason for this is that price ratios are equal to marginal rates of substitution by construction:

$$\frac{P_{ij}}{P_{ik}} = \frac{\theta_i J_{ij}(\mathbf{a}^*)}{\theta_i J_{ik}(\mathbf{a}^*)} = \frac{J_{ij}(\mathbf{a}^*)}{J_{ik}(\mathbf{a}^*)}. \tag{15}$$

The choice of prices $P_{ij} = \theta_i J_{ij}(\mathbf{a}^*)$ is intuitive. The price that i pays for j's effort is proportional to the marginal value to i of j's effort, and also proportional to i's Pareto weight θ_i. The reason that Pareto weights enter this way can be seen as follows. In the the constrained optimization problem of maximizing $u_i(\mathbf{a})$ subject to $\widehat{\mathrm{BB}}_i(\mathbf{P})$, the first-order conditions boil down to $P_{ij} = \mu_i^{-1} J_{ij}(\mathbf{a})$, where μ_i is the Lagrange multiplier on the constraint $\widehat{\mathrm{BB}}_i(\mathbf{P})$. Since μ_i is the marginal utility of relaxing the constraint $\widehat{\mathrm{BB}}_i(\mathbf{P})$ – i.e. the marginal utility of income – it follows that the Pareto weight θ_i plays the role of *inverse marginal utility of income*. Thus, agents with low marginal utilities of income are charged relatively higher prices.

The converse implication – that if $\mathbf{a}^*$ is a nonzero Lindahl outcome, then $\mathbf{J}(\mathbf{a}^*)\mathbf{a}^* = \mathbf{0}$ – is simple. As before, by Lemma 1 in Section A.3, we know a nonzero Lindahl outcome $\mathbf{a}^*$ is interior. Since agents are optimizing given prices, we have

$$\frac{P_{ij}}{P_{ik}} = \frac{J_{ij}(\mathbf{a}^*)}{J_{ik}(\mathbf{a}^*)},$$

which echoes (15) above. In other words, each row of $\mathbf{P}$ is a scaling of the same row of $\mathbf{J}(\mathbf{a}^*)$. Therefore, the condition that each agent is (exactly) exhausting his budget – which can be briefly written as $\mathbf{P}\mathbf{a}^* = \mathbf{0}$ – implies that $\mathbf{J}(\mathbf{a}^*)\mathbf{a}^* = \mathbf{0}$.

8 Concluding Discussion

8.1 Free-Riding

Throughout, we have studied the implementation problem under the requirements of Pareto efficiency and individual rationality. The latter is defined in terms of leaving everyone no worse off than at the status quo. In mechanism design and implementation problems, the individual rationality requirement is often motivated by the notion that participation is voluntary. For example, if a player may end up worse off after an auction than before he entered, then he may not wish to participate, paying nothing

and having no chance of winning the item. The value of this outcome can be normalized to 0 because it is independent of what happens in the mechanism without the agent.

In our context, the notion of opting out is more complicated. Since the goods in question are *public* goods, a player who "opts out" of a mechanism to provide them often cannot be excluded from enjoying the benefits of an arrangement reached in his absence – for example, one that leads to cleaner air or water. Thus, a player may wish to "opt out" in hopes of enjoying public goods provided by others without investing any effort himself.

In the analysis so far, we have implicitly assumed that players can commit to participate as part of their general power to commit to the rules of an institution – the function g in the language of Section 4.1. To explore the extent to which this is a binding assumption, we enrich the environment, giving each player an *inalienable* decision over whether to participate – one that he cannot make commitments about or delegate.[46] As in Section 3.4.4, agent i not participating means setting $a_i = 0$, i.e. making it it impossible to take positive action. If an agent chooses to participate, then he *can* commit to the rules of the subsequent game that sets the action profile **a**, as in Section 4.1.

Formally, we assume that the designer of Section 4.1 announces a family of game forms, $(H_P)_{P \subseteq N}$, one for every possible set P of participants. Players first simultaneously and unilaterally choose whether to participate, and the set P of those who chose to participate play the game H_P, just as in Section 4.1.

The idea of this section is that we can provide a sufficient condition under which no player wishes to use his power to opt out. Recall Section 3.4.4, in which we showed that player i's participation is essential to achieving any Pareto improvement on the status quo precisely when his removal changes the spectral radius of the benefits matrix at the status quo from being greater than 1 to being less than 1. If each player has this property, then his decision not to participate leaves the remaining players with no Pareto improvement on the status quo. Thus, assuming each H_P always implements Pareto efficient outcomes that are at least as good as **0** for each of the remaining participants, the only outcome a player can hope for after opting out is **0**.

In short, we have argued that *if every player is essential, then each player expects that opting out will lead to the outcome* **0***. Therefore, individual rationality as we defined it in our main analysis is equivalent to voluntary participation.*

There is a simple matrix characterization of what it means for every player to be essential: the spectral radius of $\mathbf{B}(\mathbf{0})$ exceeds 1, and the spectral radius of every submatrix obtained by removing one row and one column is less than 1.

This argument yields only a crude sufficient condition and suggests avenues for future work, of which we mention one. What if opting out decisions are made in equilibrium? An interesting framework is one in which they are sequential and players behave farsightedly, anticipating the effects of their opting out on others' incentives to participate in providing the public goods. How can we compute, given any $\mathbf{B}(\mathbf{0})$, the

---

[46]See Kartik and Tercieux (2012) for a discussion of inalienability in the context of implementation with evidence.

largest stable coalition not at risk of being unraveled by rational sequences of decisions to opt out? (See Genicot and Ray [2003] for a related analysis in a risk-sharing game.)

8.2   Lindahl Outcomes and Coalitional Deviations: The Core

One of the definitions of the core in games with externalities considers an outcome to be *outside* the core if a coalition can, by choosing a different action profile, improve the welfare of every one of its members no matter what the others do. Recall that in our setting, externalities across players are always nonnegative by Assumption 2, so in this case the uniformly worst punishment by the outsiders is to play 0. The outcomes that do not admit such a deviation are the core outcomes. Formally:

DEFINITION **13.** An action profile $\mathbf{a}$ is in the β-*core* under a preference profile $\mathbf{u}$ if there is no nonempty coalition $M \subseteq N$ and no other action profile $\mathbf{a}'$ of this game so that:

(i) $a_i' = 0$ for all $i \notin M$;

(ii) for each $i \in M$ we have $\mathbf{a}' \succ_i \mathbf{a}$.

Under nonnegative externalities, the Lindahl outcomes are in the β-core (Shapley and Shubik, 1969; Foley, 1970).

Leaving no coalition worse off than they could be without outsiders' help may be a desirable normative property. It is also important, intuitively speaking, for making sure that coalitions do not have incentives to secede. However, developing a full non-cooperative theory of play under the possibility of coalitional deviations is a very challenging and important direction for future work. Rubinstein (1980), Bernheim, Peleg, and Whinston (1987), and Horniaček (1996) make progress in different directions.

8.3   The Informational Assumptions of the Implementation Theory Framework

In the setup of Section 4.1, we follow the standard approach of complete information implementation in focusing on the Nash equilibria of the game form H given agents' true preferences. This implicitly assumes that agents know enough about each other's strategies to play a Nash equilibrium of the game.[47] On the other hand, the institutions in which they interact – the rules of the game – are long-run, and are designed to be robust to many realizations of preferences.

We do not view the assumption that agents play Nash equilibria of the complete information game as very restrictive or inappropriate when the negotiators represent municipalities or countries.[48] In those cases, decisions about strategy are made by institutions (e.g. governments) whose agents communicate internally, and consequently

---

[47]Viewed as elements of the strategy spaces Σ_i, not as maps from privately known types to such strategies.

[48]Roberts (2008) briefly critiques the informational assumptions of complete information Nash implementation when the agents in question are many citizens in a public finance setting.

the intentions of each government are difficult to conceal from counterparties. Relatedly, the parties often base their valuations on a more or less common pool of information. For instance, when the public good is environmental, then estimates of the economic sacrifices needed for abatement and the harms from pollution – which jointly determine the payoffs – are made based on information and models regarding the environment and the economy that are essentially public.[49]

Nevertheless, a setting with some incomplete information is certainly more realistic than the stark complete information benchmark, and further study of it is a direction for future work.[50]

8.4 Irreducibility of the Benefits Matrix

In Assumption 3, we posited that $\mathbf{B}(\mathbf{a})$ is irreducible – i.e., that it is not possible to find an outcome and partition society into two nonempty groups such that, at that outcome, one group does not care about the effort of the other at the margin.

How restrictive is this assumption? We now discuss how our analysis extends beyond it. Suppose that whether $B_{ij}(\mathbf{a})$ is positive does not depend on $\mathbf{a}$, so that the directed graph describing whose effort matters to whom is constant. Let $\mathbf{G}$ be a matrix defined by

$$G_{ij} = \begin{cases} 1 & \text{if } i \neq j \text{ and } B_{ij}(\mathbf{a}) > 0 \text{ for all } \mathbf{a} \\ 0 & \text{otherwise.} \end{cases}$$

We say a subset $S \subseteq N$ is *closed* if $G_{ij} = 0$ for every $i \in S$ and $j \notin S$. We say S is *irreducible* if $\mathbf{G}$ is irreducible when restricted to S.

We can always partition N into some closed, irreducible subsets

$$S^{(1)}, S^{(2)}, \ldots, S^{(m)}$$

and a remaining class T of agents who are in no closed, irreducible subset. The utility of any agent in a set $S^{(k)}$ is independent of the choices of anyone outside the set (and these are the minimal sets with that property). So it seems reasonable to consider negotiations restricted to each such set; that is, to take the set of players to be $S^{(k)}$. All our analysis then goes through without modification on each such subset.

When entries $B_{ij}(\mathbf{a})$ change from positive to zero depending on $\mathbf{a}$, then the analysis becomes substantially more complicated, and we leave it for future work.

---

[49]We thank Joe Shapiro for suggesting this point. Of course, the information and the models may themselves involve a great deal of uncertainty. This alone need not create a problem for complete information implementation because it is uncertainty common to the agents, and therefore does not create informational asymmetries – as long as the agents know counterparties' preferences and beliefs.

[50]See Tian (1996) for a study of implementing a rational expectations Lindahl equilibrium and Jackson (2001) for a survey of incomplete-information implementation theory.

References

ACEMOGLU, D., V. M. CARVALHO, A. OZDAGLAR, AND A. TAHBAZ-SALEHI
(2012): "The Network Origins of Aggregate Fluctuations," *Econometrica*, 80, 1977–2016.

ALLOUCH, N. (2012): "On the Private Provision of Public Goods on Networks,"
Tech. rep., Queen Mary, University of London, School of Economics and Finance.

BALLESTER, C., A. CALVÓ-ARMENGOL, AND Y. ZENOU (2006): "Who's Who in
Networks. Wanted: The Key Player," *Econometrica*, 74, 1403–1417.

BERGSTROM, T., L. BLUME, AND H. VARIAN (1986): "On the Private Provision
of Public Goods," *Journal of Public Economics*, 29, 25–49.

BERNHEIM, B., B. PELEG, AND M. WHINSTON (1987): "Coalition-Proof Nash
Equilibria: I. Concepts," *Journal of Economic Theory*, 42, 1–12.

BONACHICH, P. B. (1987): "Power and Centrality: A Family of Measures," *American Journal of Sociology*, 92, 1170–1182.

BONACICH, P. B. (1991): "Simultaneous Group and Individual Centralities," *Social Networks*, 13, 155–168.

BRAMOULLÉ, Y., R. KRANTON, AND D'AMOURS M. (2011): "Strategic Interaction
and Networks," Working paper, available at econ.duke.edu/people/kranton.

BRIN, S. AND L. PAGE (1998): "The Anatomy of a Large-Scale Hypertextual Web
Search Engine," in *Proceedings of the Seventh International World-Wide Web Conference*, Amsterdam: Elsevier Science Publishers B. V., WWW7, 107–117.

DÁVILA, J. AND J. EECKHOUT (2008): "Competitive Bargaining Equilibrium,"
Journal of Economic Theory, 139, 269–294.

DÁVILA, J., J. EECKHOUT, AND C. MARTINELLI (2009): "Bargaining Over Public
Goods," *Journal of Public Economic Theory*, 11, 927–945.

DEBREU, G. AND I. N. HERSTEIN (1953): "Nonnegative Square Matrices," *Econometrica*, 21, 597–607.

DU, Y., E. LEHRER, AND A. PAUZNER (2012): "Competitive Economy as a Ranking
Device over Networks," Mimeo.

FOLEY, D. K. (1970): "Lindahl's Solution and the Core of an Economy with Public
Goods," *Econometrica*, 38, 66–72.

GENICOT, G. AND D. RAY (2003): "Group Formation in Risk-Sharing Arrangements," *Review of Economic Studies*, 70, 87–113.

GOLUB, B. AND C. LEVER (2010): "The Leverage of Weak Ties: How Linking Groups Affects Inequality," Mimeo., available at http://www.mit.edu/~bgolub/papers/intergroup.pdf.

GROVES, T. AND J. LEDYARD (1977): "Optimal Allocation of Public Goods: A Solution to the 'Free Rider' Problem," *Econometrica*, 45, 783–809.

HORNIAČEK, M. (1996): "The Approximation of a Strong Perfect Equilibrium in a Discounted Supergame," *Journal of Mathematical Economics*, 25, 85–107.

HURWICZ, L. (1979a): "On Allocations Attainable through Nash Equilibria," *Journal of Economic Theory*, 21, 140–165.

——— (1979b): "Outcome Functions Yielding Walrasian and Lindahl Allocations at Nash Equilibrium Points," *Review of Economic Studies*, 46, 217–225.

HURWICZ, L., E. MASKIN, AND A. POSTLEWAITE (1995): "Feasible Nash Implementation of Social Choice Correspondences when the Designer Does Not Know Endowments or Production Sets," in *The Economics of Information Decentralization: Complexity, Efficiency, and Stability*, ed. by J. Ledyard, Amsterdam: Kluwer.

JACKSON, M. O. (2001): "A Crash Course in Implementation Theory," *Social Choice and Welfare*, 18, 655–708.

——— (2008): *Social and Economic Networks*, Princeton, NJ: Princeton University Press.

KARTIK, N. AND O. TERCIEUX (2012): "Implementation with Evidence," *Theoretical Economics*, 7, 323–355.

KATZ, L. (1953): "A New Status Index Derived from Sociometric Analysis," *Psychometrika*, 18, 39–43.

LINDAHL, E. (1919): "Just Taxation – A Positive Solution," in *Classics in the Theory of Public Finance*, ed. by R. A. Musgrave and A. T. Peacock, Macmillan, London, 98–123, 1958.

MASKIN, E. (1999): "Nash Equilibrium and Welfare Optimality," *Review of Economic Studies*, 66, 23–38.

MEYER, C. D. (2000): *Matrix Analysis and Applied Linear Algebra*, Philadelphia: SIAM.

MILNOR, J. W. (2001): "Matrix Estimates and the Perron-Frobenius Theorem," Notes, available at http://www.math.sunysb.edu/~jack/DYNOTES/dnA.pdf. Accessed November 3, 2012.

PALACIOS-HUERTA, I. AND O. VOLIJ (2004): "The Measurement of Intellectual Influence," *Econometrica*, 72, 963–977.

PENTA, A. (2011): "Multilateral Bargaining and Walrasian Equilibrium," *Journal of Mathematical Economics*, 47, 417–424.

PERKINS, P. (1961): "A Theorem on Regular Matrices," *Pacific Journal of Mathematics*, 11, 1529–1533.

ROBERTS, J. (2008): "Lindahl Equilibrium," in *The New Palgrave Dictionary of Economics, Second Edition*, Palgrave Macmillan.

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

SAMUELSON, P. A. (1954): "The Pure Theory of Public Expenditure," *The Review of Economics and Statistics*, 36, 387–389.

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

TIAN, G. (1996): "Continuous and Feasible Implementation of Rational-Expectations Lindahl Allocations," *Games and Economic Behavior*, 16, 135–151.

WICKSELL, K. (1896): "A New Principle of Just Taxation," in *Classics in the Theory of Public Finance*, ed. by R. A. Musgrave and A. T. Peacock, Macmillan, London, 1958.

WILKINSON, J. H. (1965): *The Algebraic Eigenvalue Problem*, Oxford: Oxford University Press.

YILDIZ, M. (2003): "Walrasian Bargaining," *Games and Economic Behavior*, 45, 465–487.

A Omitted Proofs and Additional Results

A.1 Rawlsian Interpretation of Bang for the Buck

This section gives a formal statement and proof of the Rawlsian interpretation of an egalitarian direction: that it maximizes the minimum return on investment.

PROPOSITION **7.** The egalitarian direction $\mathbf{d}^{\mathrm{eg}}(\mathbf{a})$ maximizes $\min_i b_i(\mathbf{a}, \mathbf{d})$ among all $\mathbf{d} \in \Delta_n$.

Proof of Proposition 7: The Collatz-Wielandt formula (Meyer, 2000, equation 8.3.3) says that $r(\mathbf{B}(\mathbf{a}))$ is the maximum of

$$\min_i \frac{[\mathbf{B}(\mathbf{a})\mathbf{d}]_i}{d_i}$$

as $\mathbf{v}$ ranges over nonzero, nonnegative vectors $\mathbf{d} \in \Delta_n$. Thus, the minimum bang for the buck is never greater than $r(\mathbf{B}(\mathbf{a}))$. The proof of Proposition 3 in Section 3.3 shows that the bang for the buck in the egalitarian direction $\mathbf{d}^{\mathrm{eg}}(\mathbf{a})$ is equal to $r(\mathbf{B}(\mathbf{a}))$ for everyone. This completes the proof. ∎

A.2 The Zero Action Profile as a Pareto Efficient Outcome

Proof of Proposition 2 in Section 3.1: First we prove the result in the case when $\mathbf{B}(\mathbf{0})$ is irreducible.

If $r(\mathbf{B}(\mathbf{0})) > 1$, then Proposition 3 yields an egalitarian direction at $\mathbf{0}$ with bang for the buck exceeding 1; this is a Pareto improvement at $\mathbf{0}$.

If $\mathbf{0}$ is not Pareto efficient, there is an $\mathbf{a}' \in \mathbb{R}_+$ such that $u_i(\mathbf{a}') \geq u_i(\mathbf{0})$ for each i, with strict inequality for some i. Using Assumption 3, namely the irreducibility of $\mathbf{B}(\mathbf{a}')$, as well as the continuity of the u_i, we can find[51] an $\mathbf{a}''$ with all positive entries so that $u_i(\mathbf{a}'') > u_i(\mathbf{0})$ for all i. Let $\mathbf{v}$ denote the derivative of $\mathbf{u}(\zeta \mathbf{a}'')$ in ζ evaluated at $\zeta = 0$. This derivative is strictly positive in every entry, since (by convexity of the u_i) the entry v_i must exceed $[u_i(\mathbf{a}'') - u_i(\mathbf{0})]/a_i''$. By the chain rule, $\mathbf{v} = \mathbf{J}(\mathbf{0})\mathbf{a}''$. From the fact that $\mathbf{v}$ is positive, it is immediate to deduce that there is a positive vector $\mathbf{w}$ so that $\mathbf{B}(\mathbf{0})\mathbf{w} > \mathbf{w}$. And from this it follows by the Collatz-Wielandt formula (Meyer, 2000, equation 8.3.3) that the spectral radius of $\mathbf{B}(\mathbf{0})$ exceeds 1.

Now assume $\mathbf{B}(\mathbf{0})$ is reducible.

If $r(\mathbf{B}(\mathbf{0})) > 1$, then the same is true when $\mathbf{B}(\mathbf{0})$ is replaced by one of its irreducible blocks, and in that case a Pareto improvement on $\mathbf{0}$ (involving only the agents in the irreducible block taking positive effort) is found as above.

---

[51]Suppose otherwise and let $\mathbf{a}''$ be chosen so that $\mathbf{u}(\mathbf{a}'') - \mathbf{u}(\mathbf{0}) \geq \mathbf{0}$ (note this is possible, since $\mathbf{a}'' = \mathbf{a}'$ satisfies this inequality) and so that the number of 0 entries in $\mathbf{u}(\mathbf{a}'') - \mathbf{u}(\mathbf{0})$ is as small as possible. Let S be the set of i for which $u_i(\mathbf{a}'') - u_i(\mathbf{a}) > 0$. Then by irreducibility of benefits, we can find $j \in S$ and $k \notin S$ such that $J_{kj}(\mathbf{0}) > 0$. Define $a_j''' = a_j'' + \epsilon$ and $a_i''' = a_i''$ for all $i \neq j$. If $\epsilon > 0$ is chosen small enough, then by continuity of the u_i we have $u_i(\mathbf{a}''') - u_i(\mathbf{a}) > 0$ for all $i \in S$, but also $u_k(\mathbf{a}''') - u_k(\mathbf{a}) > 0$, contradicting the choice of $\mathbf{a}''$.

Now suppose $\mathbf{0}$ is not Pareto efficient. There is an $\mathbf{a}' \in \mathbb{R}_+$ such that $u_i(\mathbf{a}') \geq u_i(\mathbf{0})$ for each i, with strict inequality for some i. Let $P = \{i : a_i' > 0\}$ – the set of agents taking positive actions at $\mathbf{a}'$. And let $\widehat{\mathbf{B}}(\mathbf{0})$ be obtained by restricting $\mathbf{B}(\mathbf{0})$ to P (i.e. by throwing away rows and columns not corresponding to indices in P). For each $i \in P$, there is a $j \in P$ such that $\widehat{B}_{ij}(\mathbf{0}) > 0$; otherwise, then i would be worse off than at $\mathbf{0}$. Therefore, each $i \in P$ is on a cycle[52] in $\widehat{\mathbf{B}}(\mathbf{0})$. And it follows that for each $i \in P$ there is a set $P_i \subseteq P$ such that $\widehat{\mathbf{B}}(\mathbf{0})$ is irreducible when restricted to P_i. Next, applying the argument of footnote 51 to each such P_i separately, we can find $\mathbf{a}''$ such that $u_i(\mathbf{a}'') > u_i(\mathbf{0})$ for each $i \in P$. From this point we can argue as above[53] to conclude that $r(\widehat{\mathbf{B}}(\mathbf{0})) > 1$. Since $\widehat{\mathbf{B}}(\mathbf{0})$ is a submatrix of $\mathbf{B}(\mathbf{0})$, Fact 3 on the monotonicity of the spectral norm implies that $r(\mathbf{B}(\mathbf{0})) > 1$. ∎

A.3 Nonzero Lindahl Outcomes are Interior

The following simple lemma underlies several important steps in the ensuing arguments.

LEMMA 1. If $\mathbf{a}^* \neq \mathbf{0}$ is a Lindahl outcome for preference profile $\mathbf{u}$, then $\mathbf{a}^* \in \mathbb{R}_{++}^n$.

Proof of Lemma 1: Assume, toward a contradiction, that $\mathbf{a}^*$ has some entries equal to 0. Let $\mathbf{P}$ be the matrix of prices given in Definition 6 supporting $\mathbf{a}^*$ as a Lindahl outcome. Let S be the set of i so that $a_i^* = 0$, which is a proper subset of N since $\mathbf{a}^* \neq \mathbf{0}$. By Assumption 3 (irreducibility of benefits), there is an $i \in S$ and a $j \notin S$ so that $J_{ij}(\mathbf{a}^*) > 0$. We will argue that this implies

$$P_{ij} > 0.$$

If this were not true, then an $\mathbf{a} \neq \mathbf{a}^*$ in which only j increases his action slightly relative to $\mathbf{a}^*$ would satisfy $\text{BB}_i(\mathbf{P})$ in Definition 6 and be better for i than $\mathbf{a}^*$, contradicting the definition of a Lindahl outcome.

Now consider $\text{BB}_i(\mathbf{P})$, the budget balance condition of agent i, at the outcome $\mathbf{a}^*$:

$$\sum_{k:k \neq i} P_{ik} a_k^* \leq a_i^* \sum_{k:k \neq i} P_{ki}.$$

Since $a_i^* = 0$, the right-hand side of this is zero. Since $P_{ij} > 0$, and $a_j^* > 0$ (since $j \notin S$), the left-hand side is positive. That is a contradiction. ∎

A.4 The Deferred Lemma in the Proof of Theorem 1, and a Stronger Version of the Theorem

As noted in Section 4.3, the Hurwicz rationale for the Lindahl outcomes is more general than is stated there. In this section, we formalize and prove this, also tying up the proof of Theorem 1. Let $\mathcal{A}$ be the set of preference profiles $\mathbf{u}$ satisfying the assumptions of Section 2.2.

---

[52]Recall the definition in Section 3.4.2.

[53]The Collatz-Wielandt formula does not assume irreducibility.

Fix any probability measure μ on $\mathbb{R}_+^n$ that is absolutely continuous with respect to Lebesgue measure – in particular, one assigning positive measure to every open set. We define a metric d on $\mathcal{A}$ by $d(u, \widehat{u}) = \int_{\mathbb{R}_+^n} |u - \widehat{u}| d\mu$. That metric induces a topology on the function space $\mathcal{A}$.

DEFINITION 14. A set of preferences $\mathcal{U} \subseteq \mathcal{A}$ is called *rich* if, for every $\mathbf{u} \in \mathcal{U}$ and $\mathbf{a}^* \in \mathbb{R}_+^n$, there is a (linear) preference profile $\widehat{\mathbf{u}} \in \mathcal{U}$ defined by[54]

$$\widehat{\mathbf{u}}(\mathbf{a}) = \mathbf{J}(\mathbf{a}^*; \mathbf{u})\mathbf{a} \tag{16}$$

and a neighborhood of $\widehat{\mathbf{u}}$ relative to $\mathcal{A}$ is contained in $\mathcal{U}$.

Richness of $\mathcal{U}$ requires that for every preference profile $\mathbf{u} \in \mathcal{U}$ and every $\mathbf{a}^* \in \mathbb{R}_+^n$, there are preferences in $\mathcal{U}$ that are linear over outcomes and have the same *marginal* tradeoffs as $\mathbf{u}$ does at $\mathbf{a}^*$, as well as a neighborhood of these preferences. To take a simple example, $\mathcal{A}$ itself is rich.

THEOREM 3. Suppose $\mathcal{U}$ is rich and the number of players, n, is at least 3. Then the robustly attainable correspondence is equal to the Lindahl correspondence: $R = L$.

The proof is exactly as in Section 7.2. The only thing that remains to do is to establish the following lemma used in that proof under the hypothesis that $\mathcal{U}$ is rich (the result needed in Section 7.2 is then a special case).

LEMMA 2. Fix $\mathbf{u}$ satisfying the assumptions of Section 2.2 and an $\mathbf{a}^* \in L(\mathbf{u})$. Define $\widehat{\mathbf{u}}$ as in (16). Suppose $F : \mathcal{U} \rightrightarrows \mathbb{R}_+^n$ is a Pareto efficient, individually rational, and upper hemicontinuous social choice correspondence. If $\mathcal{U}$ is rich, then $\mathbf{a}^* \in F(\widehat{\mathbf{u}})$.

Proof of Lemma 2: First, assume that $\mathbf{a}^* = \mathbf{0}$ is a Lindahl outcome. In that case, by Proposition 8 in Section A.6 below (or simply the First Welfare Theorem), the outcome $\mathbf{0}$ is Pareto efficient. In particular, $\mathbf{J}(\mathbf{0}; \mathbf{u})\mathbf{a}$ has a negative entry for every nonzero $\mathbf{a} \in \mathbb{R}_+^n$; otherwise, we would be able to find a (nearby) Pareto improvement on $\mathbf{0}$ under $\mathbf{u}$. It follows that $\mathbf{0}$ is the only individually rational and Pareto efficient outcome under $\widehat{\mathbf{u}}$. So $\mathbf{a}^* \in F(\widehat{\mathbf{u}})$.

Now assume $\mathbf{a}^* \neq \mathbf{0}$. By Lemma 1 in Section A.3, $\mathbf{a}^*$ is interior – all its entries are positive.

For $\gamma > 0$, and $i \in N$, define $\widehat{u}_i^{[\gamma]} : \mathbb{R}_+^n \to \mathbb{R}$ by

$$\widehat{u}_i^{[\gamma]}(\mathbf{a}) = J_{ii}(\mathbf{a}^*; \mathbf{u}) (\gamma + a_i)^{1+\gamma} + \sum_{j \neq i} J_{ij}(\mathbf{a}^*; \mathbf{u})a_j.$$

This is just an adjustment obtained from $\widehat{\mathbf{u}} = \widehat{\mathbf{u}}^{[0]}$ by building some convexity into the costs. Note that for all γ close enough to 0, the profile $\widehat{\mathbf{u}}$ is in $\mathcal{U}$ by the richness assumption.

Choose $\mathbf{a}^{(k)} \in F(\widehat{\mathbf{u}}^{[1/k]})$; this is legitimate since F is a social choice correspondence, and hence nonempty-valued. We will show that by the properties of F, a subsequence

---

[54]Of course, it suffices that there be some $\widehat{\widehat{\mathbf{u}}} \in \mathcal{U}$ that generates the same preferences for each agent.

of the sequence $(\mathbf{a}^{(k)})_k$ converges to $\mathbf{a}^*$. Then by upper hemicontinuity of F, it will follow that $\mathbf{a}^* \in F(\widehat{\mathbf{u}}^{[0]})$, as desired. The trickiest part of the argument is showing that the $\mathbf{a}^{(k)}$ lie in some compact set, so we can extract a convergent subsequence; it will then be fairly easy to show that the limit point of this sequence is $\mathbf{a}^*$.

Let $\mathrm{IR}^{[\gamma]}$ be the set of individually rational points under $\widehat{\mathbf{u}}^{[\gamma]}$ and let $\mathrm{PE}^{[\gamma]}$ be the set of Pareto efficient points under $\widehat{\mathbf{u}}^{[\gamma]}$. Let $a^*_{\max} = \max_i a^*_i$ and define the box $K = [0, 2a^*_{\max}]^n$.

CLAIM **1.** For all k, the point $\mathbf{a}^{(k)}$ is either in K or on the ray

$$Z = \{\mathbf{a} \in \mathbb{R}^n_+ : \mathbf{J}(\mathbf{a}^*; \mathbf{u})\mathbf{a} = \mathbf{0}\}.$$

To show the claim, we first establish that

$$\mathrm{IR}^{[0]} = Z.$$

The proof is as follows. First note that $\widehat{\mathbf{u}}^{[0]}(\mathbf{a}) = \mathbf{J}(\mathbf{a}^*; \mathbf{u})\mathbf{a}$. There cannot be an $\mathbf{a}$ such that $\mathbf{J}(\mathbf{a}^*; \mathbf{u})\mathbf{a}$ is nonnegative in all entries and positive in some entries.[55] Thus, if $\mathbf{J}(\mathbf{a}^*; \mathbf{u})\mathbf{a}$ is nonzero, it must have some negative entries; i.e. $\widehat{u}^{[0]}_i(\mathbf{a}) < 0$ for some i, and then $\mathbf{a} \notin \mathrm{IR}^{[0]}$, contradicting the fact that F is individually rational.

Note that for $\mathbf{a}$ outside the box K, we have for small enough γ

$$\widehat{\mathbf{u}}^{[\gamma]}(\mathbf{a}) \leq \widehat{\mathbf{u}}^{[0]}(\mathbf{a}).$$

From this and the fact that $\widehat{\mathbf{u}}^{[\gamma]}(\mathbf{0}) = \mathbf{0}$ for all γ, we have the relation

$$\mathrm{IR}^{[\gamma]} \cap K^c \subseteq \mathrm{IR}^{[0]} \cap K^c.$$

The claim follows.

It is easily checked[56] that if $\mathbf{a} \in Z$ and $\mathbf{a} > \mathbf{a}^*$, then for $\gamma > 0$ we have $r(\mathbf{B}(\mathbf{a}; \widehat{\mathbf{u}}^{[\gamma]})) < r(\mathbf{B}(\mathbf{a}^*; \mathbf{u})) = 1$, where the latter equality holds by Remark 5. Therefore, by Proposition 1, no point on the ray Z outside K is Pareto efficient for $\gamma > 0$. This combined with Claim 1 shows that $\mathrm{IR}^{[\gamma]} \cap \mathrm{PE}^{[\gamma]} \subseteq K$, and therefore (since F is Pareto efficient and individually rational) it follows that $\mathbf{a}^{(k)} \in K$ for all k.

As a result we can find a sequence $(j(k))_k$ such that the sequence $(\mathbf{a}^{(j(k))})_k$ converges to some $\bar{\mathbf{a}}$. Define $\bar{\mathbf{a}}^{(k)} = \mathbf{a}^{(j(k))}$ and set $\widehat{\mathbf{u}}^{(k)} = \widehat{\mathbf{u}}^{[1/j(k)]}$. Note that the $\widehat{\mathbf{u}}^{(k)}$ converge uniformly to $\widehat{\mathbf{u}}^{[0]}$ on K and, indeed, on any compact set (thus, they converge compactly to $\widehat{\mathbf{u}}^{[0]}$). By upper hemicontinuity of F, it follows that $\bar{\mathbf{a}} \in F(\widehat{\mathbf{u}}^{[0]})$. It remains only to show that $\bar{\mathbf{a}} = \mathbf{a}^*$, which we now do.

_____

[55]Otherwise, $\mathbf{a}^*$ would not have been Pareto efficient under $\mathbf{u}$: moving in the direction $\mathbf{a}$ would have yielded a Pareto improvement. But $\mathbf{a}^*$ *is* Pareto efficient – see Remark 5.

[56]We do a very similar calculation below.

If $\overline{\mathbf{a}} \notin Z$, then it is easy to see that for large enough k, we would have $\widehat{u}_i^{(k)}(\overline{\mathbf{a}}^{(k)}) < 0$ for some i. This would contradict the hypothesis that F is individually rational. Thus, $\overline{\mathbf{a}}^{(k)} \to \zeta \mathbf{a}^*$ for some $\zeta \geq 0$. If $\zeta = 0$ then inspection of the definition of $\widehat{\mathbf{u}}^{[\gamma]}$ shows that eventually $\overline{\mathbf{a}}^{(k)}$ is not Pareto efficient for preferences $\mathbf{u}^{(k)}$. But that contradicts the Pareto efficiency of F. So assume $\zeta > 0$. In that case we would have:

$$J_{ij}(\overline{\mathbf{a}}^{(k)}; \mathbf{u}^{(k)}) \to \begin{cases} \zeta J_{ij}(\mathbf{a}^*; \mathbf{u}) & \text{if } j = i \\ J_{ij}(\mathbf{a}^*; \mathbf{u}) & \text{otherwise.} \end{cases}$$

Thus,

$$\mathbf{B}(\overline{\mathbf{a}}^{(k)}; \mathbf{u}^{(k)}) \to \zeta \mathbf{B}(\mathbf{a}^*; \mathbf{u}).$$

Recall from Remark 5 that $\mathbf{B}(\mathbf{a}^*; \mathbf{u}) = 1$. Since the spectral radius is linear in scaling the matrix and continuous (recall Remark 3), it follows that

$$r(\mathbf{B}(\overline{\mathbf{a}}^{(k)}; \mathbf{u}^{(k)})) \to \zeta,$$

By the Pareto efficiency of F, we know that $r(\mathbf{B}(\overline{\mathbf{a}}^{(k)}; \mathbf{u}^{(k)})) = 1$ whenever $\overline{\mathbf{a}}^{(k)}$ is interior[57], which holds for all large enough k since $\zeta \neq 0$. Thus $\zeta = 1$. It follows that $\overline{\mathbf{a}} = \mathbf{a}^*$ and the proof is complete.

A.5   Proof of the Centrality Characterization of Lindahl Outcomes

Proof of Theorem 2 in Section 5.2: First, we show (i) implies (ii). Suppose $\mathbf{a}^* \in \mathbb{R}_+^n$ is a nonzero Lindahl outcome. Lemma 1 in Section A.3 above implies that $\mathbf{a}^* \in \mathbb{R}_{++}^n$ – that is, $\mathbf{a}^*$ has only positive entries. Let $\mathbf{P}$ be the matrix of prices in the modified definition of a Lindahl outcome (Definition 12). Consider the following program for each of $i \in M$, denoted by $\Pi_i(\mathbf{P})$:

maximize $u_i(\mathbf{a})$ subject to $\mathbf{a} \in \mathbb{R}_+^n$ and $\widehat{\mathrm{BB}}_i(\mathbf{P})$.

By definition of a Lindahl outcome, $\mathbf{a}^*$ solves $\Pi_i(\mathbf{P})$. By Assumption 3, which says there is always some a_j whose increase makes i better off, the budget balance constraint $\widehat{\mathrm{BB}}_i(\mathbf{P})$ is satisfied with equality, so that $\mathbf{P}\mathbf{a}^* = \mathbf{0}$. Because $\mathbf{a}^*$ is interior, the gradient of the maximand u_i must be orthogonal to the constraint set given by $\widehat{\mathrm{BB}}_i(\mathbf{P})$. In other words, row i of $\mathbf{J}(\mathbf{a}^*)$ is parallel to row i of $\mathbf{P}$. These facts together imply $\mathbf{J}(\mathbf{a}^*)\mathbf{a}^* = \mathbf{0}$.

The equivalence of (ii) and (iii) follows immediately from the definitions.

Finally, we claim that (iii) implies (i). Since $\mathbf{a}^*$ is a nonnegative right-hand eigenvector of $\mathbf{B}(\mathbf{a}^*)$, the Perron-Frobenius Theorem guarantees that 1 is a largest eigenvalue of $\mathbf{B}(\mathbf{a}^*)$. Arguing as in the proof of Proposition 1, we deduce that there is a nonzero vector $\boldsymbol{\theta}$ for which $\boldsymbol{\theta}\mathbf{J}(\mathbf{a}^*) = \mathbf{0}$. We need to find prices supporting $\mathbf{a}^*$ as a

---

[57]Recall Remark 5 again.

Lindahl outcome. Define the matrix $\mathbf{P}$ by $P_{ij} = \theta_i J_{ij}(\mathbf{a}^*)$ and note that for all $j \in N$ we have

$$\sum_{i \in N} P_{ij} = \sum_{i \in N} \theta_i J_{ij}(\mathbf{a}^*) = [\boldsymbol{\theta} \mathbf{J}(\mathbf{a}^*)]_j = 0, \tag{17}$$

where $[\boldsymbol{\theta} \mathbf{J}(\mathbf{a}^*)]_j$ refers to entry j of the vector $\boldsymbol{\theta} \mathbf{J}(\mathbf{a}^*)$.

Note that $\mathbf{B}(\mathbf{a}^*)\mathbf{a}^* = \mathbf{a}^*$ implies $\mathbf{J}(\mathbf{a}^*)\mathbf{a}^* = \mathbf{0}$ and each row of $\mathbf{P}$ is just a scaling of the corresponding row of $\mathbf{J}(\mathbf{a}^*)$. We therefore have:

$$\mathbf{Pa}^* = \mathbf{0}. \tag{18}$$

and these prices satisfy budget balance.

We claim that, for each i, the vector $\mathbf{a}^*$ solves $\Pi_i(\mathbf{P})$. This is because the gradient of u_i at $\mathbf{a}^*$, which is row i of $\mathbf{J}(\mathbf{a}^*)$, is normal to the constraint set by construction of $\mathbf{P}$ and by (18) above, $\mathbf{a}^*$ satisfies the constraint $\widehat{\mathrm{BB}}_i(\mathbf{P})$. The claim then follows by the concavity of u_i. ∎

A.6   The Zero Action Profile as a Lindahl Outcome

PROPOSITION 8. The following are equivalent:

(i) $r(\mathbf{B}(\mathbf{0})) \leq 1$;

(ii) $\mathbf{0}$ is a Pareto efficient action profile;

(iii) $\mathbf{0}$ is a Lindahl outcome.

Proof of Proposition 8: Proposition 2 establishes the equivalence between (i) and (ii).

(ii) $\Rightarrow$ (iii): The construction of prices is exactly analogous to the proof of Theorem 2 in Section A.5; the only difference is that rather than the Pareto weights, we use Pareto weights adjusted by the Lagrange multipliers on the binding constraints $a_i \geq 0$.

(iii) $\Rightarrow$ (ii): The standard proof of the First Welfare Theorem goes through without modification; see, e.g. Foley (1970). ∎

A.7   Proof of Proposition 5: Characterization of Eigenvector Centrality Action Profiles via Walks

Let $\mathcal{W}_i^\uparrow(\ell; \mathbf{M})$ be the set of all walks of length ℓ in a matrix $\mathbf{M}$ such that $w(1) = i$ – that is, the set of walks starting at i. The proof follows immediately from the following observation.

FACT 5. For any irreducible, nonnegative matrix $\mathbf{Q}$, and any i, j

$$\frac{p_i}{p_j} = \lim_{\ell \to \infty} \frac{\sum_{w \in \mathcal{W}_i^\uparrow(\ell; \mathbf{Q})} v(w; \mathbf{Q})}{\sum_{w \in \mathcal{W}_j^\uparrow(\ell; \mathbf{Q})} v(w; \mathbf{Q})},$$

where $\mathbf{p}$ is any nonnegative right-hand eigenvector of $\mathbf{Q}$ (i.e. a right-hand Perron vector in the terminology of Section 2.4).

To prove the proposition from this fact, set $\mathbf{Q} = \mathbf{B}(\mathbf{a}) = \mathbf{M}^\mathsf{T}$ and note that then the right-hand side of the equation in Fact 5 is equal to the right-hand side of the equation in Proposition 5. The statement that $\mathbf{a}$ is an eigenvector centrality action profile is equivalent to the statement that $\mathbf{a}$ is a right-hand Perron eigenvector of $\mathbf{Q} = \mathbf{B}(\mathbf{a})$.

Proof of Fact 5. Note that the formula above is equivalent to

$$\frac{p_i}{p_j} = \lim_{\ell \to \infty} \frac{\sum_k \left[\mathbf{Q}^\ell\right]_{ik}}{\sum_k \left[\mathbf{Q}^\ell\right]_{jk}}, \tag{19}$$

where $\left[\mathbf{Q}^\ell\right]_{ik}$ denotes the entry in the (i, k) position of the matrix $\mathbf{Q}^\ell$. To prove (19), let $\rho = r(\mathbf{Q})$ and note that

$$\lim_{\ell \to \infty}(\mathbf{Q}/\rho)^\ell = \mathbf{w}^\mathsf{T}\mathbf{p}, \tag{20}$$

where $\mathbf{w}$ is a left-hand Perron vector of $\mathbf{Q}$, and $\mathbf{p}$ is a right-hand Perron vector (recall Setion 2.4). This is statement (8.3.13) in Meyer (2000); the hypothesis that $\mathbf{Q}$ is primitive in that statement follows from the assumed aperiodicity of $\mathbf{Q}$ (see Theorems 1 and 2 of Perkins (1961)). To conclude, observe that (20) directly implies (19). □

A.8 Existence of a Lindahl Outcome

Proof of Proposition 6 in Section 7.1: We will use the Kakutani Fixed Point Theorem to find a Lindahl outcome. Define $Y = \{\mathbf{a} \in \mathbb{R}^n_+ : \min_i[\mathbf{J}(\mathbf{a})\mathbf{a}]_i > 0\}$, the set of action profiles $\mathbf{a}$ at which everyone has positive gains from scaling. It is easy to check that Y is convex.[58] Also, Y is bounded by Assumption 4. Thus, $\overline{Y}$, the closure of Y, is compact.[59]

Define the correspondence $F : \overline{Y} \setminus \{\mathbf{0}\} \rightrightarrows \overline{Y}$ by

$$F(\mathbf{a}) = \{z\mathbf{a} \in \overline{Y} : z \geq 0 \text{ and } \min_i[\mathbf{J}(z\mathbf{a})\mathbf{a}]_i \leq 0\}.$$

Finally, recalling the definition of $\mathbf{d}^{\text{eg}}(\mathbf{a})$ from Section 3.3, define the correspondence $G : \overline{Y} \rightrightarrows \overline{Y}$ by

$$G(\mathbf{a}) = F(\mathbf{d}^{\text{eg}}(\mathbf{a})).$$

Note that $\mathbf{d}^{\text{eg}}(\mathbf{a})$ is always nonzero, so that the argument of F is in its domain.[60] The function $\mathbf{d}^{\text{eg}}$ is continuous (Wilkinson, 1965, pp. 66–67), and F is clearly upper

---

[58]Given $\mathbf{a}, \mathbf{a}' \in Y$ and $\lambda \in [0, 1]$, consider $\mathbf{a}'' = \lambda\mathbf{a} + (1 - \lambda)\mathbf{a}'$. Note that for all $i \in N$ and $\epsilon \geq -1$

$$u_i((1 + \epsilon)\mathbf{a}'') \geq \lambda u_i((1 + \epsilon)\mathbf{a}) + (1 - \lambda)u_i((1 + \epsilon)\mathbf{a}')$$

by concavity of the u_i. Differentiating in ϵ at $\epsilon = 0$ yields the result.

[59]It is tempting to define $Y = \{\mathbf{a} \in \mathbb{R}^n_+ : \min_i[\mathbf{J}(\mathbf{a})\mathbf{a}]_i \geq 0\}$ instead and avoid having to take closures; but this set can be unbounded even when $\overline{Y}$ as we defined it above *is* bounded. For example, our assumptions do not exclude the existence of an (infinite) ray along which $\min_i[\mathbf{J}(\mathbf{a})\mathbf{a}]_i = 0$.

[60]Note that even though the domain of F is not a compact set, G is a correspondence from a compact set into itself.

hemicontinuous, so it follows that G is upper hemicontinuous. Finally, from the definitions of Y and F it follows that F is nonempty-valued.[61] Since $\overline{Y}$ is a compact and convex set, the Kakutani Fixed Point Theorem implies that there is an $\mathbf{a} \in \overline{Y}$ such that $\mathbf{a} \in F(\mathbf{d}^{\mathrm{eg}}(\mathbf{a}))$. Writing $\widehat{\mathbf{a}} = \mathbf{d}^{\mathrm{eg}}(\mathbf{a})$, this means that there is some $z \geq 0$ such that $\min_i[\mathbf{J}(z\widehat{\mathbf{a}})\widehat{\mathbf{a}}]_i \leq 0$. Let $\mathbf{a}^* = z\widehat{\mathbf{a}}$. We will argue that $\mathbf{a}^*$ is a Lindahl outcome.

Suppose $\mathbf{a}^* \neq \mathbf{0}$. Then by continuity of the function $z \mapsto \mathbf{J}(z\widehat{\mathbf{a}})\widehat{\mathbf{a}}$, there is some i for which we have $[\mathbf{J}(\mathbf{a}^*)\widehat{\mathbf{a}}]_i = 0$ – i.e., some player's marginal benefit to scaling is equal to marginal cost. Since $\widehat{\mathbf{a}}$ is an egalitarian direction at the action profile $\mathbf{a}^*$, the equation $[\mathbf{J}(\mathbf{a}^*)\widehat{\mathbf{a}}]_i = 0$ must hold for *all* i, and therefore $\mathbf{J}(\mathbf{a}^*)\widehat{\mathbf{a}} = \mathbf{0}$. Since $\widehat{\mathbf{a}}$ and $\mathbf{a}^*$ are parallel, we deduce $\mathbf{J}(\mathbf{a}^*)\mathbf{a}^* = \mathbf{0}$ and Theorem 2 shows that $\mathbf{a}^*$ is a Lindahl outcome. The condition $\mathbf{J}(\mathbf{a}^*)\mathbf{a}^* = \mathbf{0}$ along with Assumption 3 – irreducibility – imply that $\mathbf{a}^* \in \mathbb{R}_{++}^n$.

If $\mathbf{a}^* = \mathbf{0}$, then the bang for the buck $b(\mathbf{0}, \widehat{\mathbf{a}})$ of starting at $\mathbf{0}$ and moving in the egalitarian direction $\widehat{\mathbf{a}}$ is no greater than 1 – otherwise, $F(\widehat{\mathbf{a}})$ would not contain $\mathbf{a}^* = \mathbf{0}$. By Proposition 3, it follows that $r(\mathbf{B}(\mathbf{0})) \leq 1$. Then $\mathbf{0}$ is a Lindahl outcome by Proposition 8. ∎

A.9   Proofs of Minor Facts

Proof of Fact 1 in Section 3.4.2: Consider the submatrix $\mathbf{L}$ obtained by leaving only indices in the cycle c. Note $\mathbf{L}^\ell$ has diagonal entries, and hence row sums, bounded below by $v(c; \mathbf{M})$. Then by Meyer (2000, p. 497) it follows that $\rho(\mathbf{L}^\ell) \geq v(c; \mathbf{M})$, and by Fact 3 in Section 3.4.3 that $\rho(\mathbf{M}^\ell) \geq v(c; \mathbf{M})$. The desired result then follows because $\rho(\mathbf{M}^\ell) = \rho(\mathbf{M})^\ell$ for any nonnegative matrix. (By Perron-Frobenius there is always a real and positive eigenvalue equal to the spectral radius; moreover, λ is an eigenvalue of $\mathbf{M}$ if and only if λ^ℓ is an eigenvalue of $\mathbf{M}^\ell$.) ∎

---

[61] Toward a contradiction, take a nonzero $\mathbf{a}$ such that $F(\mathbf{a})$ is empty. Let $\overline{z}$ be the maximum z such that $z\mathbf{a} \in \overline{Y}$; such a $\overline{z}$ exists because $\mathbf{a}$ is nonzero and $\overline{Y}$ is compact. Since $\mathbf{J}(\overline{z}\mathbf{a})\mathbf{a} > \mathbf{0}$ it follows that for all i,

$$\frac{du_i((1+\epsilon)\overline{z}\mathbf{a})}{d\epsilon}\bigg|_{\epsilon=0} > 0,$$

from which it follows that $(\overline{z}+\delta)\mathbf{a} \in Y$ for small enough δ. This contradicts the choice of $\overline{z}$ (recalling the definition of Y).