Inequality and Network Structure[*]

Willemien Kets[†]   Garud Iyengar[‡]   Rajiv Sethi[§]   Samuel Bowles[¶]

December 17, 2010

Abstract

We explore the manner in which the structure of a social network constrains the level of inequality that can be sustained among its members, based on the following considerations: (i) any distribution of value must be stable with respect to coalitional deviations, and (ii) the network structure itself determines the coalitions that may form. We show that if players can jointly deviate only if they form a clique in the network, then the degree of inequality that can be sustained depends on the cardinality of the maximum independent set. For bipartite networks, the size of the maximum independent set fully characterizes the degree of inequality that can be sustained. This result extends partially to general networks and to the case in which a group of players can deviate jointly if they are all sufficiently close to each other in the network.

[†]Corresponding author: U.C. Irvine and Santa Fe Institute. E-mail: willemien.kets@santafe.edu.

[‡]Department of Industrial Engineering and Operations Research, Columbia University. E-mail: garud@ieor.columbia.edu.

[§]Department of Economics, Barnard College, Columbia University and the Santa Fe Institute. E-mail: rs328@columbia.edu.

[¶]Santa Fe Institute and University of Siena. E-mail: bowles@santafe.edu.

1 Introduction

In 494 BCE, the *plebs* of the Roman Republic, seeking relief from judicial harassment, indebtedness and poverty, left Rome *en masse* and threatened to settle permanently outside its walls, as a result extracting major concessions from the Roman patricians (Livy, 1960). Plantation owners in Hawaii a century ago expressly hired workers who spoke different native languages to ensure that communication among them would be limited, thus discouraging labor action (Takaki, 1983). U.S. employer efforts in the 1930s to build firm loyalty by sponsoring social activities led to stronger bonds between workers that they could use to mobilize their collective power and form effective unions (Estlund, 2003). And the extraordinary longevity of the Ottoman Empire (1300-1918) and its remarkable integration and taxation of diverse ethnic and religious communities was based on a network structure that "made peripheral elites dependent on the center, communicating only with the center rather than with one another" (Barkey, 2008).

A recurrent theme in these examples is the central role of coalitional deviations in determining the distribution of income, with the structure of social relations being a central determinant of the coalitions that form. This motivates us to explore formally the manner in which the structure of a social network constrains the level of inequality that can be sustained among its members. We develop a model of inequality on networks based on the following considerations: (i) any distribution of value must be stable with respect to coalitional deviations, and (ii) the set of feasible coalitions is itself constrained by the requirement that only groups of players that are mutually connected can jointly deviate. That is, we allow for deviations only by groups of individuals who form a *clique* in the network. A payoff distribution is said to be stable if there is no clique that can profitably deviate. The main research question is then the following: What is the relationship between the structure of the network and the maximum level of stable inequality?

To compare payoff distributions in terms of their level of inequality, we adopt the standard criterion of Lorenz dominance and define a value distribution to be *extremal* if it is stable with respect to clique deviations and does not Lorenz dominate any other stable distribution. Since Lorenz dominance provides only a partial ordering of value distributions, the extremal distribution for any given network may not be unique, and extremal distributions for different networks may be incomparable.

Our main contribution is to establish a connection between extremal inequality on a network and a natural measure of the sparseness of a network, the size of its maximum independent set.[1] This connection is especially strong in the case of bipartite networks, which have unique

---

[1]An independent set in a network is a set of vertices such that no pair of vertices in the set are connected to each other. An independent set is maximum if there is no independent set of greater size.

extremal distributions and can be completely ordered; we show that bipartite networks with larger maximum independent sets can sustain greater levels of extremal inequality. For general networks with arbitrary clique sizes, a weaker result holds: for any two networks, extremal inequality cannot be greater in the one with the smaller maximum independent set.

Our framework can be extended to include the case in which players can jointly deviate if they are all within distance k of each other (the case of clique deviations corresponds to $k = 1$). We explore the manner in which extremal inequality changes as k is varied. Although inequality (weakly) declines as k increases, it can do so at different rates in different networks. As a result, the ranking of networks by the extent of extremal inequality is not invariant in k.

A number of recent papers have explored the determinants of inequality in equilibrium networks (see Section 2 for details). In these papers, an agent's central position confers the ability to gain larger shares of the surplus, the intuition being that essential intermediaries can extract rents through their control of flows between players that are not otherwise connected (e.g., Goyal and Vega-Redondo, 2007; Hojman and Szeidl, 2008).[2] These "middleman" models are implicitly based on the idea that competition reduces inequality, and monopoly increases it.

While this intuition is undoubtedly correct in many settings, our model stresses another dimension that influences inequality: the ability of players to form viable coalitions. Intuitively, if the network is dense, inequality will be hard to sustain as disadvantaged players can communicate and coordinate on joint actions. Conversely, if the network is sparse, peripheral players can more readily be exploited. Hence, while in the so-called "middleman" models, a player can secure a large share of the surplus if he is well connected, under our approach this is the case if the *other* players are isolated. We show by example that our model gives rise to different predictions in a number of cases, underlining the importance of considering alternative approaches to the relationship between income distribution and social structure.

2   Related literature

The idea that network structure influences the allocation of value was initially proposed in a seminal paper by Myerson (1977), who assumed that a coalition of individuals could generate value if and only if they were all connected to each other along some path that did not involve anyone outside the coalition. Myerson's work motivated a significant literature on communication games (see Slikker and van den Nouweland, 2001, for a survey) and more generally on games on combinatorial structures (Bilbao, 2000); see e.g. Demange (2004) for

---

[2]Another important difference with our work is that these papers employ an exogenously given profile of payoff functions that determines for each network the allocation of value between players.

3

an important application to economic questions.

Our approach differs in two important respects from this line of work. First, while the aim of much of the literature cited above is to give a characterization of different solution concepts, to investigate their relation with each other, and to provide conditions for the existence of solutions in general classes of games, our focus is on the maximum degree of inequality that can be sustained in a restricted set of games where existence of stable distributions is guaranteed. Second, our setting naturally leads us to consider coalitional deviations that do not generate a combinatorial structure. For example, two feasible coalitions may overlap in our setting, without there being a feasible coalition (other than the complete network) that contains both, in contrast with the settings considered by Myerson (1977) and Demange (2004), for example. This means that there is no natural order in which the value can be allocated to the cliques; see Bilbao (2000) for a discussion.

Finally, Bloch, Genicot, and Ray (2008) study the stability of insurance networks for different levels of communication. As in the current paper, information transmission across network links (over limited distances) plays a crucial role in this work, and the sparseness of the network is an important determinant of the viability of various allocations. However, while Bloch et al. study the stability of insurance norms in different networks, we focus on sustainable levels of inequality. Furthermore, the notion of sparseness differs: while sparseness in our setting is determined by the size of independent sets, in the context of Bloch et al. sparseness is captured by the minimal length of cycles among triples of agents.

3   Distributions on networks

3.1   Networks

Players are located on a network. A *network* is a pair (N, g), where $N = \{1, \ldots, n\}$ is a set of *vertices* and g is an $n \times n$ matrix, with $g_{ij} = 1$ denoting that there is a *link* or *edge* between two vertices i and j, and $g_{ij} = 0$ meaning that there is no link between i and j. A link between i and j is denoted by $\{i, j\}$. We focus on undirected networks, so $g_{ij} = g_{ji}$. Moreover, we set $g_{ii} = 0$ for all i. In the following, we fix the vertex set N and denote a network by the matrix g. If $g_{ij} = 1$, that is, if there is a link between i and j, we say that i and j are *neighbors* or, alternatively, that they are *adjacent* in g. A *clique* is a set of pairwise adjacent vertices. Hence, an edge is a clique, and so is a triangle, where a *triangle* is a set $\{i, j, k\}$ of three distinct vertices that are all connected. It will be convenient to view a single vertex as a (singleton) clique. An important subclass of networks is the set of bipartite networks. A network is *bipartite* if the vertices can be partitioned into two classes, and there are no links within each class. Bipartite networks thus provide a natural model of trading relations

between buyers and sellers. Also, bipartite networks are of interest because they contain the class of minimally connected networks (trees), which often form the stable outcomes of strategic network formation models.

An *independent set* in a network is a set of vertices that are pairwise nonadjacent. A set of vertices forms a *maximum independent set* in g if it is an independent set and there is no independent set in g of a strictly greater size. Note that while a network may have multiple (maximum) independent sets, the size of a maximum independent set is unique.

3.2   Value generation

Consider a set of players $N = \{1, \ldots, n\}$, and a network g with vertex set N, so that each player is associated with a vertex. Following Jackson and Wolinsky (1996), we assume that the value generated by a group of players S is given by $v(g|_S)$, where $g|_S$ is the subgraph of g induced by S, i.e., the network obtained by removing the players not in S; for simplicity, we write $v(g)$ for $v(g|_N)$. As a normalization, assume that the value of the empty network g^E is $v(g^E) = 0$, where the empty network is the network without any vertices or edges.

We assume that the value function is *anonymous*, in the sense that the value generated by a group of players only depends on the number of players in the group and the way in which they are connected, but not on their identities. That is, if S and S' are subsets of players, then for any bijection π from S to S', it holds that

$$v(g|_S) = v(g^\pi|_{S'}),$$

where g^π is the network that has the same architecture as g, but with the players in S relabeled according to π, i.e., $g^\pi_{i,j} = 1$ if and only if $g_{\pi^{-1}(i), \pi^{-1}(j)} = 1$. This anonymity assumption allows us to abstract from the effects of productivity differences across players in order to isolate the role of network structure in determining inequality. The assumption implies in particular that any two cliques C_k, C'_k each having k players generate the same value, as they share the same network architecture:

$$v(g|_{C_k}) = v(g|_{C'_k}).$$

A special case of an anonymous value function is one in which the value generated by a group of players S depends only on their number, that is, $v(g|_S)$ equals $f(|S|)$ for any $S \subseteq N$, where f is some arbitrary function on the natural numbers.

3.3   Stability

The surplus generated by the group must be divided among its members subject to the constraint that no clique can deviate profitably. Formally, an *allocation* is any vector $x =$

$(x_1, \ldots, x_n) \in \mathbb{R}^N$. An allocation x is *stable* on g if no clique can gain by deviating: for each clique C in g,

$$\sum_{i \in C} x_i \geq v(g|_C). \tag{3.1}$$

That is, for an allocation to be stable, the members of each clique have to get at least as much collectively under the allocation as they would if they were to deviate collectively and form their own network.

A second natural constraint is that players cannot divide more than they produce:

$$\sum_{i \in N} x_i = v(g). \tag{3.2}$$

An allocation x is *feasible* if (3.2) is satisfied.

The definition of the set of feasible and stable allocations is reminiscent of the definition of the core in transferable-utility games (TU-games) for the special case that the value generated by a coalition does not depend on the way the players in the coalition are connected. The difference is that while inequality (3.1) needs to hold for *all* coalitions for x to be in the core, we only require the inequality to hold for subsets of players that are sufficiently close in the network. It follows that if we have two networks, g and g', and g is a connected subgraph of g' with the same number of players, then the set of feasible and stable allocations for g' is contained in the set of feasible and stable allocations for g (when the value function does not depend on the network architecture). In particular, the set of feasible and stable allocations for an arbitrary network is a superset of the set of feasible and stable allocations for the complete network, which coincides with the core of an appropriately defined TU-game, as illustrated in Example 4.1 below.

3.4 Lorenz dominance

We wish to compare allocations in terms of the inequality they generate. Corresponding to any allocation x is a *distribution* $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)$. The distribution $\bar{x}$ is simply a permutation of the elements of x that places them in (weakly) increasing order: $\bar{x}_1 \leq \bar{x}_2 \leq \cdots \leq \bar{x}_n$. We say that a distribution $\bar{x}$ is feasible and stable on a network g if there exists a corresponding allocation that is feasible and stable on g.

To compare distributions in terms of the level of inequality, we consider the widely-used criterion of Lorenz dominance. Consider two distributions $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)$, $\bar{y} = (\bar{y}_1, \ldots, \bar{y}_n) \in \mathbb{R}_+^n$ such that
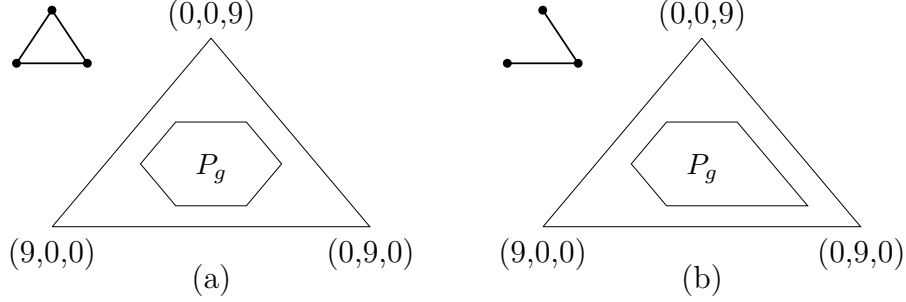
$$\sum_{i \in N} \bar{x}_i = \sum_{i \in N} \bar{y}_i.$$

Figure 4.1: (a) The set of feasible and stable allocations, denoted P_g, for the triangle; (b) The set of feasible and stable allocations, denoted P_g, for the line.

Then, we say that $\bar{x}$ *Lorenz dominates* $\bar{y}$ if, for each $m = 1, \ldots, n$,

$$\sum_{i=1}^{m} \bar{x}_i \geq \sum_{i=1}^{m} \bar{y}_i,$$

with strict inequality for some m. If $\bar{x}$ Lorenz dominates $\bar{y}$, we say that $\bar{x}$ is a *more equal* distribution than $\bar{y}$. If $\bar{x}$ does not Lorenz dominate $\bar{y}$ and $\bar{y}$ does not Lorenz dominate $\bar{x}$, we say that $\bar{x}$ and $\bar{y}$ are *incomparable*. We call a stable distribution $\bar{x}$ on g which is feasible *extremal* if there is no distribution $\bar{y}$ that is stable and feasible such that $\bar{x}$ Lorenz dominates $\bar{y}$. Since the Lorenz dominance criterion only provides a partial order on the set of feasible and stable distributions, there may be multiple extremal distributions for a given network. We say that a network g has a *unique extremal distribution* if the set of extremal distributions on g is a singleton.

4 Examples

The examples in this section illustrate the concepts of stability and inequality, and show that the model can give different predictions relative to other models of inequality on networks. Throughout this section, we focus on the special case in which the value generated by a group of players does not depend on the manner in which they are connected, but only on the number of players in the group. That is, there exists some function f such that $v(g|_S) = f(|S|)$ for any group of players S and any network g. Examples for the more general case can readily be constructed.

The first example illustrates how the structure of a network can affect the set of feasible and stable allocations.

Example 4.1 Consider two connected networks with three players, the triangle (illustrated in the upper-left corner of Figure 4.1(a)) and the line (illustrated in the upper-left corner of
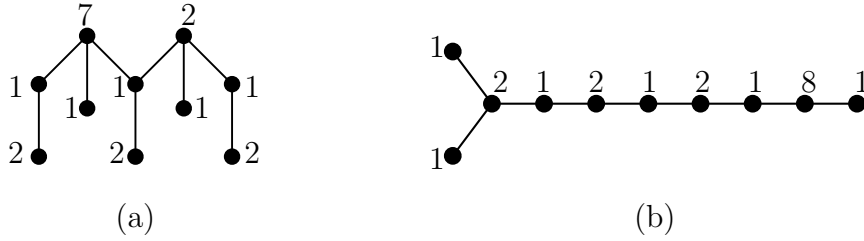
7

Figure 4.2: (a) The network h of Example 4.2. (b) The network h' of Example 4.2. The numbers represent one of the allocations consistent with the unique extremal distribution in each case.

Figure 4.1(b)), with player 2 as the center or hub. Assume that the value generated by a group of players of size m is $f(m) = n^2$. In both cases, the set of feasible and stable allocations is a subset of the simplex defined by $x_1 + x_2 + x_3 = f(n) = 9$. For the triangle, the set of extreme points consists of all permutations of $(1, 3, 5)$. For the line, the extreme points of the set of feasible and stable allocations are the allocations $(1, 3, 5), (5, 3, 1), (1, 5, 3), (3, 5, 1)$ and $(1, 7, 1)$ when player 2 is the center of the line. Hence, the set of feasible and stable allocations for the line is a superset of the set of feasible and stable allocations for the triangle; see panels (a) and (b) in Figure 4.1. Intuitively, there are fewer feasible coalitions in the line than in the triangle: the peripheral players (viz., player 1 and 3) in the line cannot jointly deviate, while each pair of players forms a feasible coalition in the triangle. That means that in the triangle, player 1 and 3 need to receive at least $f(2) = 4$ (as does any other pair of players), while in the line, they only need to get $f(1) + f(1) = 2$. For both networks, there is a unique extremal distribution, given by $\bar{x} = (1, 1, 7)$ for the line, and $\bar{x}' = (1, 3, 5)$ for the triangle. It is easily verified that the latter Lorenz dominates the former. ◁

In Example 4.1, what properties of network g allow it to support a more unequal distribution than g'? One possibility is the fact that the distribution of the number of neighbors that each player has in g is itself more unequal than that in g', i.e., an unequal distribution of value can be explained by inequality in the *degree distribution*. The following example shows that this is not the case.

Example 4.2 Suppose $f(1) = 1$, $f(2) = 3$, and $f(10) = 20$. Consider the networks h and h' in Figure 4.2(a) and (b), respectively. In both cases, the value generated by the network is equal to 20. The stability conditions require that each individual be assigned at least $f(1) = 1$, and each pair of neighbors be assigned at least $f(2) = 3$. Both networks have a unique extremal distribution, given by

$$\bar{x} = (1, 1, 1, 1, 1, 2, 2, 2, 2, 7), \quad \text{and} \quad \bar{x}' = (1, 1, 1, 1, 1, 1, 2, 2, 2, 8).$$
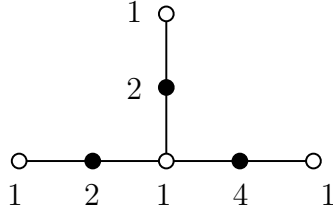
8

Figure 4.3: The network of Example 4.3. The player with the greatest degree and betweenness gets the lowest payoff in any extremal allocation.

Hence, $\bar{x}$ Lorenz dominates $\bar{x}'$. This is the opposite of what one would predict based on inequality in the degree distributions of h and h', which are given by:

$$\bar{d} = (1, 1, 1, 1, 1, 2, 2, 3, 3, 3), \qquad \text{and} \qquad \bar{d}' = (1, 1, 1, 2, 2, 2, 2, 2, 2, 3).$$

Clearly $\bar{d}'$ Lorenz dominates $\bar{d}$, even though $\bar{x}$ Lorenz dominates $\bar{x}'$. ◁

Like a player's degree, his betweenness is often taken as a measure of a player's prominence and as a determinant of a player's payoffs. The betweenness of a player i in a network is the number of shortest paths between v and w player i belongs to over the total number of all shortest paths between v and w, averaged over all v and w (see, for example, Jackson, 2008). However, inequality in betweenness fares no better in explaining extremal inequality, as the next example demonstrates.

Example 4.3 Suppose $f(1) = 1$, $f(2) = 3$, and $f(7) = 12$. Consider the network in Figure 4.3. It can be verified that the network has a unique extremal distribution, given by

$$\bar{x} = (1, 1, 1, 1, 2, 2, 4).$$

This distribution is consistent with different allocations to the players, but in any such allocation, each player represented by an open circle (∘) in Figure 4.3 is assigned $f(1) = 1$. This includes the player with the highest degree. This player also has the highest betweenness (0.43), more than double than that of his neighbors, both of whom receive higher payoffs. ◁

Taken together Examples 4.2 and 4.3 demonstrate that a focus on inequality in the degree or betweenness in attempting to understand the extent of inequality in social networks is misleading in two respects. First, networks with more equal degree or betweenness distributions may be capable of sustaining greater inequality than those with more unequal distributions. And second, by either measure, well-connected players can do substantially worse than less well-connected players in a given network.[3] This suggests that these measures, which are

---

[3]It can be shown by example that another important centrality measure, closeness, also fails to predict high payoffs, where the closeness of a player in the network is the average length of the shortest paths to other players (Jackson, 2008).

motivated by what we called "middleman" models in Section 1, fail to capture inequality caused by the differential ability of players to form deviating coalition, which lies at the heart of our model.

In the following section, we show that rather than the degree or betweenness distribution, it is the size of the largest independent sets that determines the degree of inequality that can be sustained.

5   Extremal inequality

5.1   Bipartite networks

In this section, we first show that any bipartite network has a unique extremal distribution when some conditions on the value functions are satisfied. We then investigate how the unique extremal distribution changes for bipartite networks when the network structure is varied. We make the following assumptions on the value function v: For any network g on player set N,

A1: $v(g|_{C_2}) \geq 2v(g|_{C_1})$;

A2: $2v(g) \geq n\, v(g|_{C_2})$,

where $n = |N|$ is the number of players, C_2 is a clique of size 2 in g (i.e., a pair of neighbors), and C_1 is a clique of size 1 (i.e., a single player).[4] By our anonymity assumption, the value of a clique of a given size does not depend on the identity of the players or on the wider network structure, so that **A1** and **A2** do not depend on which cliques C_1 and C_2 are being considered.

If **A1** is not satisfied, no allocation exists that is both feasible and stable for a nonempty network. If **A2** is not satisfied, an allocation that is stable and feasible potentially exists, but our results below will not hold for all bipartite networks. As only cliques of sizes 1 and 2 can deviate in a bipartite network, the egalitarian distribution, which gives an equal amount $v(g)/n$ to each player, is stable under these assumptions.

We first ask whether the extremal distribution is unique for this class of networks. Let A be a maximum independent set in g, and let $\ell \in N \setminus A$ be an arbitrary player not belonging to A. Define the allocation x^* by

$$
x_i^* = \begin{cases}
v(g|_{C_1}) & \text{if } i \in A, \\
v(g|_{C_2}) - v(g|_{C_1}) & \text{if } i \in N \setminus (A \cup \{\ell\}), \\
\begin{aligned}&v(g) - |A|v(g|_{C_1}) - \\ &\quad (n - |A| - 1)(v(g|_{C_2}) - v(g|_{C_1}))\end{aligned} & \text{if } i = \ell.
\end{cases}
\tag{5.1}
$$

---

[4]In the special case that the value generated by a group only depends on the size of the group, **A1** reduces to $f(2) \geq 2f(1)$, and **A2** becomes $2f(n) \geq nf(2)$.

By Assumptions **A1** and **A2**, $x_\ell^* \geq x_i^*$ for any $i \in N$. The corresponding distribution is denoted by $\bar{x}^*$.

The following result characterizes the extremal distribution (and establishes its uniqueness) for the case of bipartite networks.

Theorem 5.1 *Suppose assumptions **A1** and **A2** are satisfied. If g is a bipartite network, then $\bar{x}^*$ is its unique extremal distribution.*

As a corollary of Theorem 5.1, we find that bipartite networks can be ranked in terms of extremal inequality by the size of their maximum independent sets whenever they generate the same total value.[5] Hence, even though the Lorenz dominance relation is not a complete order, we obtain a complete order on the set of bipartite networks.

Corollary 5.2 *Suppose assumptions **A1** and **A2** are satisfied. Consider any two bipartite networks g, g' with vertex set N such that $v(g) = v(g')$. Let A and A' be any maximum independent sets in g and g', and let $\bar{x}$ and $\bar{x}'$ be their unique extremal distributions. Then, $\bar{x} = \bar{x}'$ if and only if $|A| = |A'|$. If $|A| \neq |A'|$, then $\bar{x}$ Lorenz dominates $\bar{x}'$ if and only if $|A| < |A'|$.*

Important for our results is that the size of a deviating coalition is at most 2 in a bipartite network. When we impose limits on the size of the deviating coalitions, the results extend to general networks. The next section shows that the size of the maximum independent set is still an important determinant of extremal inequality in general networks when coalitions of arbitrary size are allowed.

5.2   General networks

Our results on extremal inequality do not easily extend to general networks. There are two issues to consider: the uniqueness of the extremal distribution for a given network, and the ordering of networks with respect to their extremal distributions.

First, it can be shown by example that two networks with the *same* cardinality of their maximum independent sets can be unambiguously ranked with respect to their extremal distributions. Also, two networks that *differ* in the cardinality of their maximum independent set may have extremal distributions that cannot be ranked with respect to their extremal distributions.[6] Indeed, a companion paper (Iyengar, Kets, and Sethi, 2010) shows that the extreme points of the set of feasible and stable allocations for more general networks involves not only

---

[5]Note that Lorenz dominance is only defined for distributions $\bar{x}, \bar{x}^*$ such that the total value is equal, i.e., $\sum_i \bar{x}_i = \sum_i \bar{x}_i'$.

[6]See the earlier version of this paper (Kets et al., 2009).

the independent sets, but also network structures that consist of both edges and triangles. As any extremal distribution—and indeed any distribution at which an inequality measure such as the Gini index is maximized—must be consistent with an allocation at an extreme point of the set of feasible and stable allocations, these results suggest that features of a network other than the cardinality of maximum independent sets will be important for characterizing extremal inequality in general.[7]

However, it is possible to obtain a somewhat weaker result. To state this, we focus on the special case where the value generated by a group of players does not depend on the way they are connected, i.e., there exists a function f such that $v(g_S) = f(|S|)$ for every subset S of players.[8] We make the following assumption on the value function f:

An For all $k, \ell = 0, \ldots, n - 1$ such that $k > \ell$,

$$f(k + 1) - f(k) \geq f(\ell + 1) - f(\ell).$$

This is a strengthening of Assumptions **A1** and **A2** to ensure that a feasible and stable allocation always exists.

The following result provides a partial characterization of extremal inequality in general networks.

Theorem 5.3 *Suppose f satisfies **An**, and consider two networks g and g'. Let A and A' denote maximum independent sets in g and g', respectively. If $|A| < |A'|$, then there exists an extremal distribution $\bar{x}'$ for g' such that no extremal distribution $\bar{x}$ in g is Lorenz dominated by $\bar{x}'$.*

Note that Theorem 5.3 does not require extremal distributions to be unique. It also leaves open the possibility that the extremal distributions for two different networks are incomparable, or that extremal inequality does not change when the cardinality of the maximum independent set changes. We cannot rule out these possibilities because there may be multiple extremal distributions for a given network and, moreover, even if all extremal distributions are unique, the set of feasible and stable allocations may change with network structure in a nontrivial and unexpected manner (cf. Kalai et al., 1978). What Theorem 5.3 does rule out is that extremal inequality in a network with a smaller maximum independent set is greater than in a network with a larger one.

---

[7]The reason that inequality is maximized at the extreme points of the (convex) set of feasible and stable allocation is that inequality measures are (generally) convex functions.

[8]It is possible to generalize Theorem 5.3 to the case where the value function depends on network structure. However, additional conditions are needed to ensure that a feasible and stable allocation exists for an arbitrary network, and it is well known that such conditions can be very restrictive (Kaneko and Wooders, 1982).

We now turn to a natural application of Theorem 5.3, allowing players to coordinate a deviation over larger distances.

5.3   Broader coalitions

To this point we have assumed that players can coordinate on a deviation only if they form a clique. We now consider the possibility of deviations by coalitions of players that are all within some distance k of each other in the network.

Given a network, define a *k-coalition* to be a set of players that are all within distance k of each other. As in the previous section, assume that the value that a k-coalition C can obtain on its own is $f(|C|)$, i.e., it does not depend on how the players are connected. We say that an allocation x is *k-stable* on g if, for each k-coalition C in g,

$$\sum_{i \in C} x_i \geq f(|C|).$$

Hence, no k-coalition can profitably deviate from a k-stable allocation. Stability, as defined in Section 3, corresponds to k-stability for $k = 1$. A k-stable distribution $\bar{x}$ on g which is feasible is called *k-extremal* if there is no distribution $\bar{y}$ that is k-stable and feasible such that $\bar{x}$ Lorenz dominates $\bar{y}$.

An immediate observation is that the degree of inequality that can be sustained in a network weakly decreases when we increase k:

Observation 5.4 *For any network g and k, k' such that $k' > k$, if $\bar{x}', \bar{x}$ are extremal distributions in g for k and k', respectively, then either $\bar{x}' = \bar{x}$, or $\bar{x}'$ Lorenz dominates $\bar{x}$, or $\bar{x}$ and $\bar{x}'$ cannot be compared with respect to Lorenz dominance.*

Intuitively, a group of players that forms a k-coalition in a network g is a k'-coalition in g for $k' > k$, so that increasing k limits the degree of inequality that can be sustained. But while the degree of inequality that can be sustained in a network weakly decreases for any network if k increases, this decrease occurs at very different rates for different networks. The following example illustrates this.

Example 5.5 Consider the star network g_{star} and the line network g_{line} depicted in Figure 5.1(a) and (c), respectively, and suppose $f(m) = m^2$ for $m = 0, 1, \ldots n$. Corollary 5.2 shows that the unique extremal distribution $\bar{x}^1_{line}$ for the line Lorenz dominates the unique extremal distribution $\bar{x}^1_{star}$ for the star.

However, when $k = 2$, the situation is reversed. In the case of the star, all players can now form deviating coalitions, while for the line, the two players at the end of the star can still not coordinate a joint deviation. This has implications for the degree of inequality that can be
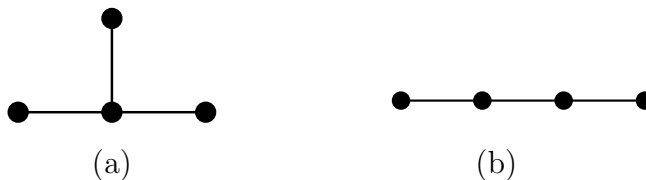
(a)   (b)

Figure 5.1: (a) The star network g_{star} of Example 5.5. (b) The network g_{line} of Example 5.5.

sustained. Also for $k = 2$, the extremal distributions for the line and star are unique; however, the unique extremal distribution $\bar{x}^2_{star}$ for the star now Lorenz dominates the unique extremal distribution for the line $\bar{x}^2_{star}$. ◁

6   Conclusions

We have studied how the degree of inequality that can be sustained on a network depends on its structure. The starting point of our analysis is the intuitive idea that players can jointly deviate only if they are sufficiently close to each other in terms of network distance. The key network property that determines inequality in our analysis is the cardinality of the maximum independent set.

Returning to the examples with which we began, the size of the maximum independent set provides a framework for understanding distributional conflicts on these networks. Factory employment, a common language and company-sponsored social events among workers have the opposite effect, reducing the cardinality of the maximum independent set and with it, the firm's feasible claims on the surplus of the network.

There are numerous avenues for further research. An immediate extension is to allow for deviations by larger cliques. The analysis of stable inequality in general networks by Iyengar, Kets, and Sethi (2010) suggests that a characterization of extremal inequality in terms of intuitive network properties is not possible for arbitrary clique sizes. However, it might be worth exploring this question for particular subclasses of networks. In a similar vein, it would be worth exploring how the result change under alternative assumptions on the coalitions that can form.

Finally, we have taken the social network as given. In our motivating examples the network that allows individuals to coordinate on possible deviations is typically formed for nonstrategic reasons, independent of the value-generating process. However, a recurrent theme in the network literature is that individuals typically create links to improve their position vis-à-vis others (e.g., Goyal and Vega-Redondo, 2007), which can lead to inefficiencies (Jackson, 2008). It would therefore be interesting to study the endogenous formation of networks in the current

14

setting. We leave to future research these and other unresolved issues concerning the subtle and interesting relationship between inequality and network structure.

Appendix A Proofs

A.1 Proof of Theorem 5.1

We first derive some preliminary results. Lemma A.1 shows that the set of vertices of any network can be partitioned into a maximum independent set and a set of vertices that are connected to at least one vertex in the maximum independent set.

Lemma A.1 *Consider a network g with at least two vertices, and let A be a maximum independent set in g. Define*

$$B := \{j \in N \mid \exists i \in A \text{ such that } g_{ij} = 1\}$$

to be the set of vertices that have at least one neighbor in A. Then the sets A and B form a partition of the vertex set N.

Proof. First we show that $A \cap B = \emptyset$. Suppose that there is a vertex $i \in A \cap B$. As $i \in A$ and since A is an independent set, there is no $j \in A$ such that $g_{ij} = 1$. However, we also have $i \in B$. By the definition of B, there exists $m \in A$ such that $g_{im} = 1$, a contradiction.

We now establish that $N = A \cup B$. Suppose there exists $i \in N$ that does not belong to $A \cup B$. Then, by the definition of B, there exists no $j \in A$ such that $g_{ij} = 1$. But then $A \cup \{i\}$ is an independent set, contradicting that A is a maximum independent set. $\qquad \square$

Lemma A.2 is a technical result on bipartite networks, which allows us to derive Corollary A.3, which will be an important ingredient of our characterization.

Before we can derive these results, we need some more definitions. The *endpoints* of an edge $e = \{i, j\}$ are the vertices i and j. A vertex is *incident* to an edge if it is one of the endpoints of that edge. A vertex without any neighbors is called an *isolated vertex*. An *edge cover* of a network with no isolated vertices is a set of edges L such that every vertex of the network is incident to some edge of L. A *minimum edge cover* of a network without isolated vertices is an edge cover of the network such that there is no edge cover with strictly smaller cardinality, see Figure A.1. Note that while a network can have multiple (minimum) edge covers, the cardinality of a minimum edge cover is well defined. A *subgraph* of a network (N, g) is a network (N', g') such that

(i) the vertex set of $(N'g')$ is a subset of that of (N, g), that is, $N' \subseteq N$;
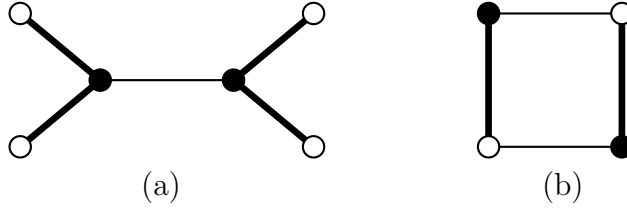
Figure A.1: Two bipartite networks; in each network, a minimum edge cover is indicated with bold lines, and vertices belonging to one of the maximum independent sets are marked by white circles (∘). Note that while the minimum edge cover and the maximum independent set are unique in the network in (a), there are two maximum independent sets and two minimum edge covers for the network in (b).

(ii) the edge set of (N', g') is a subset of (N, g), that is, $g'_{ij} = 1$ implies $g_{ij} = 1$ for all vertices i and j.

An *induced subgraph* is a subgraph obtained by deleting a set of vertices. A *component* of a network (N, g) is a maximal connected subgraph, that is, a subgraph $(N'g')$ that is connected and is not contained in another connected subgraph of (N, g). Given a network (N, g), the subgraph induced by the set non-isolated vertices is referred to as the *core subgraph* of (N, g).[9] Finally, a *star* is a tree consisting of one vertex adjacent to all other vertices. We refer to this vertex as the *center* of the star.

Lemma A.2 *Let (M, h) be a bipartite network, and let (M', h') be an induced subgraph of (M, h). For any maximum independent set A of the core subgraph (N, g) of (M', h'), there exists a minimum edge cover $L = \{\{i_1, j_1\}, \ldots, \{i_m, j_m\}\}$ of (N, g) such that*

$$\{i_1, \ldots, i_m\} = A, \qquad \{j_1, \ldots, j_m\} = N \setminus A,$$

and there exists no $j_\ell, j_k, j_\ell \neq j_k$ such that $i_\ell = i_k$.

Proof. First note that every induced subgraph of a bipartite network is again a bipartite network (that is, the class of bipartite networks is hereditary). Therefore, we can prove the statement in the lemma by proving that for any bipartite network (N, g) and any maximum independent set A of the core subgraph of (N, g), there exists a minimum edge cover $L = \{\{i_1, j_1\}, \ldots, \{i_m, j_m\}\}$ of the core subgraph with the desired properties (cf. West, 2001, Remark 5.3.20). Without loss of generality, we can restrict attention to a bipartite network (N, g) without isolated vertices. As before, we fix the vertex set N and denote the network (N, g) by g.

---

[9]Of course, if a network does not have isolated vertices, the core subgraph is just the network itself.

16

Let A be a maximum independent set in g. We will construct a minimum edge cover $L = \{\{i_1, j_1\}, \ldots, \{i_m, j_m\}\}$ with the desired properties. First note that for any minimum edge cover L' of g, for any vertex i belonging to A, there exists an edge e in L' such that i is an endpoint of e, as otherwise L' would not cover all vertices. Moreover, as A is an independent set, there is no edge in L' with two vertices from A as its endpoints. Hence, without loss of generality, we can take $L = \{\{i_1, j_1\}, \ldots, \{i_m, j_m\}\}$, with

$$\{i_1, \ldots, i_m\} \supseteq A.$$

By the Kőnig-Rado edge covering theorem (e.g. Schrijver, 2003, p. 317), the cardinality of a maximum independent set is equal to the cardinality of a minimum edge cover, so that

$$\{i_1, \ldots, i_m\} = A.$$

Since $\{i_1, \ldots, i_m\} = A$, for the vertices of $N \setminus A$ to be covered by L, we need

$$\{j_1, \ldots, j_m\} \supseteq N \setminus A.$$

As A is an independent set, we have

$$\{j_1, \ldots, j_m\} = N \setminus A.$$

Finally, suppose that there exist distinct j_ℓ, j_k such that $i_\ell = i_k =: i$. First note that for any minimum edge cover Λ the following holds. If both endpoints of an edge e belong to edges in Λ other than e, then $e \notin \Lambda$, because otherwise $\Lambda \setminus \{e\}$ would also be an edge cover of the network, contradicting that Λ is a minimum edge cover. Hence, each component formed by edges of L has at most one vertex with more than one neighbor and is a star. By assumption, j_ℓ and j_k belong to the same component in L; the center of this component is i. Since each vertex in A is associated with at least one edge in L, this means that $|L| > |A|$, which cannot hold by the Kőnig-Rado edge covering theorem. $\qquad \square$

Lemma A.2 shows that for each maximum independent set of a bipartite network, there exists a minimum edge cover such that each vertex i in the network not belonging to the maximum independent set is matched to a vertex j in the maximum independent set to which it is connected in the network, and there is no other vertex i' that is matched to j. Note that vertices not belonging to the maximum independent set will typically be connected to multiple vertices in the maximum independent set, see e.g. the network in Figure A.1(a).

Corollary A.3 states that for bipartite networks, there exists an injective mapping from vertices not belonging to a maximum independent set to the vertices in the maximum independent set, in such a way that the vertices that are matched in this way are neighbors in the network.

Corollary A.3 *Let (M, h) be a bipartite network, and let (M', h') be an induced subgraph of (M, h). For any maximum independent set A of (M', h'), there exists an injective mapping π from $M' \setminus A$ to A such that $h'_{i\pi(i)} = 1$ for all $i \in M' \setminus A$.*

Proof. Denote the set of isolated vertices in (M', h') by B. By Lemma A.2, there exists a minimum edge cover $L = \{\{i_1, j_1\}, \dots, \{i_m, j_m\}\}$ for the core subgraph (N, g) of (M', h') such that

$$\{i_1, \dots, i_m\} = A \setminus B, \qquad \{j_1, \dots, j_m\} = M' \setminus (A \cup B),$$

and there exists no $j_m, j_k, j_m \neq j_k$ such that $i_m = i_k$. Moreover, $B \subseteq A$. Hence, the mapping $\pi : \{j_1, \dots, j_m\} \to \{i_1, \dots, i_m\} \cup B$ defined by

$$\pi(j_t) = i_t$$

for $t = 1, \dots, m$ satisfies the desired properties. $\qquad \square$

Finally, Lemma A.4 establishes that the allocation x^* (Eq. 5.1) is feasible and stable for a bipartite network.

Lemma A.4 *Suppose assumptions **A1** and **A2** are satisfied. Consider a bipartite network g with at least two vertices. Let A be a maximum independent set in g, and let ℓ be an arbitrary player in $N \setminus A$. Then, the allocation x^* is feasible and stable.*

Proof. The allocation x^* is feasible by definition: Condition (i) is satisfied by construction:

$$\sum_{i \in N} x_i^* = v(g).$$

To show that the allocation x is stable, we need to establish the following:

(i) For each $i \in N$, it holds that $x_i \geq v(g|_{C_1})$.

(ii) For each pair $i, j \in N$ such that $g_{ij} = 1$, it holds that $x_i + x_j \geq v(g|_{C_2})$.

Condition (i) is satisfied, since $v(g|_{C_1}) \leq v(g|_{C_2}) - v(g|_{C_1}) \leq x_\ell$ for any cliques C_1, C_2 in g of size 1 and 2, respectively, where the first inequality follows from **A1**, and the second from **A2**. To see that (ii) holds, note that by **A1** and Lemma A.1, each pair of neighbors $i, j \in N \setminus \{\ell\}$ gets at least $v(g|_{C_2}) - v(g|_{C_1}) + v(g|_{C_1}) = v(g|_{C_2})$. It then follows from **A2** that each pair of neighbors k, m receives at least $v(g|_{C_2})$. $\qquad \square$

We are now ready to prove Theorem 5.1. Consider a bipartite network (N, g). As before, we fix N and denote the network by g. When there is one player, i.e., $n = 1$, it is easy to see

18

that the set of feasible and stable allocations is the singleton $\{x^*\}$, so that trivially $\bar{x}^*$ is the unique extremal distribution.

Hence, consider the case $n \geq 2$. Let A be a maximum independent set of N, and for each t, define

$$Y_t := \sum_{i=1}^{t} \bar{x}_i^*$$

to be the sum of the t smallest assignments under $\bar{x}^*$, and note that

$$Y_t^* = \begin{cases} t\, v(g|_{C_1}) & \text{if } t \leq |A|; \\ |A|\, v(g|_{C_1}) + (t - |A|)\, (v(g|_{C_2}) - v(g|_{C_1})) & \text{if } |A| < t \leq n - 1; \\ v(g) & \text{if } t = n; \end{cases} \quad \text{(A.1)}$$

where C_1 and C_2 are arbitrary cliques in g of size 1 and 2, respectively. By Lemma A.4, x^* is stable and feasible. It remains to show that for any distribution $\bar{y}$ on g that is stable and feasible, either $\bar{y} = \bar{x}^*$ or $\bar{y}$ Lorenz dominates $\bar{x}^*$. Suppose not. Then there exists t such that

$$Y_t^* > Y_t,$$

where we have defined $Y_t := \sum_{i=1}^{t} \bar{y}_i$ to be the sum of the t smallest assignments under $\bar{y}$. Let Q_t be any subset of vertices of cardinality t such that

$$\sum_{i \in Q_t} y_i = Y_t,$$

and let $A_t \subseteq Q_t$ be a maximum independent set in the subgraph induced by Q_t. Clearly, $|A_t| \leq |A|$.

By Lemma A.1, the set Q_t can be partitioned into A_t and the set B_t of vertices that have at least one neighbor in A_t. By Corollary A.3, there is an injective mapping π from B_t to A_t such that for each $i \in B_t$, $\{i, \pi(i)\}$ is an edge in the subgraph induced by Q_t. Define

$$U_t := \{i \in A_t \mid i = \pi(j) \text{ for some } j \in B_t\}$$

to be the set of players in A_t that are matched with a player in B_t by the mapping π.

In a bipartite network, only singleton coalitions or coalitions consisting of pairs of neighbors can form. Hence, by stability of $\bar{y}$, each individual player needs to be assigned at least $v(g|_{C_1})$ under $\bar{y}$. By **A1**, it holds that $2v(g|_{C_1}) \leq v(g|_{C_2})$. Hence, under a stable allocation, two neighboring players cannot both be assigned $v(g|_{C_1})$, except when $2v(g|_{C_1}) = v(g|_{C_2})$. In the latter case, giving each player other than ℓ his "autarky value" $v(g|_{C_1})$, and the remainder $v(g) - (n - 1)v(g|_{C_1})$ to ℓ clearly gives the extremal distribution, and this distribution equals $\bar{x}^*$.

So suppose $2v(g|_{C_1}) > v(g|_{C_2})$. Combining our earlier results gives

$$
\begin{aligned}
Y_t &= \sum_{i \in Q_t} y_i \\
&= \sum_{i \in B_t} (y_i + y_{\pi(i)}) + \sum_{i \in A_t \setminus U_t} y_i \\
&\geq \sum_{i \in B_t} v(g|_{C_2}) + \sum_{i \in A_t \setminus U_t} v(g|_{C_1}) \\
&= (t - |A_t|) v(g|_{C_2}) + (|A_t| - (t - |A_t|)) v(g|_{C_1}) \\
&= t \left(v(g|_{C_2}) - v(g|_{C_1}) \right) + |A_t| \left(2v(g|_{C_1}) - v(g|_{C_2}) \right) \\
&\geq t \left(v(g|_{C_2}) - v(g|_{C_1}) \right) + |A| \left(2v(g|_{C_1}) - v(g|_{C_2}) \right), \tag{A.2}
\end{aligned}
$$

where C_1 and C_2 are arbitrary cliques in g of size 1 and 2, respectively, as before. The last inequality follows from $|A_t| \leq |A|$ and $2v(g|_{C_1}) - v(g|_{C_2}) < 0$, which holds by assumption.

We need to consider three cases. First, if $t \leq |A|$, then $Y_t^* = t\, v(g|_{C_1})$. Since by stability, $y_i \geq v(g|_{C_1})$ for all $i \in N$, it follows that $Y_t^* \leq Y_t$. Second, suppose $|A| < t \leq n - 1$. Then it follows from (A.1) and (A.2) that

$$
Y_t^* = t \left(v(g|_{C_2}) - v(g|_{C_1}) \right) + |A| \left(2v(g|_{C_1}) - v(g|_{C_2}) \right) \leq Y_t.
$$

Finally, if $t = n$, then $Y_t^* = Y_t = v(g)$. Hence, for all t, $Y_t^* \leq Y_t$, a contradiction. $\qquad\square$

A.2 Proof of Theorem 5.3

We first construct an allocation that is feasible and stable in g' and gives $f(1)$ to all players in A'. Define the allocation y' by

$$
y_i' = \begin{cases} f(1) & \text{if } i \in A', \\ \frac{f(n) - |A'| f(1))}{n - |A'|} & \text{otherwise.} \end{cases}
$$

This allocation satisfies the requirement that $y_i' = f(1)$ for all $i \in A'$. Note that by Assumption **An**, $y_i' \geq f(1)$ for all i.

It can easily be checked that y' is feasible. We now show that it is stable in g'. Let $C \subseteq N$ be a clique in g', and note that either $C \cap A' = \emptyset$ or $|C \cap A'| = 1$. Feasibility ensures that the allocation is stable when $|C| = n$, so suppose $|C| = 1, 2, \ldots, n - 1$. Then,

$$
\sum_{i \in C} y_i' \geq f(1) + \frac{(c - 1)(f(n) - af(1)))}{n - a},
$$

where $c := |C|$ and $a := |A'|$. It is therefore sufficient to show that

$$
(c - 1)[f(n) - af(1))] \geq (n - a)\big(f(c) - f(1))\big). \tag{A.3}
$$

This is clearly satisfied with equality when the clique size is $c = 1$. If the condition is satisfied for $c = n-1$, it then follows from Assumption $\mathbf{An}$ that the inequality holds for all $c = 1, \ldots, n-1$. It remains to show that the condition holds for $c = n-1$. If $n = 2$, then this follows immediately from the fact that the condition holds for $c = 1$. So suppose $n > 2$. If $c = n - 1$, then it is not hard to see that the size of the maximum independent set must be $a = 2$. Substituting these values into (A.3) and using that $n > 2$ gives the condition

$$f(n) - 2f(1) \geq f(n-1) - f(1).$$

Rearranging terms gives

$$f(n) - f(n-1) \geq f(1).$$

But this holds by Assumption $\mathbf{An}$ and the normalization $f(0) = 0$. It follows that the allocation y' is stable. Denote the corresponding distribution by $\bar{y}'$.

Suppose $\bar{z}'$ is an extremal distribution for g'. Then, since $\bar{z}'$ is stable, $\bar{z}'_i \geq f(1)$ for all i. If $\bar{z}'_i > f(1)$ for all i, $\bar{y}'$ is also extremal. Otherwise, $\bar{z}' = f(1)$ for all $i = 1, \ldots, |A'|$. In that case, there exists an extremal distribution $\bar{x}'$ in g' such that $\bar{x}'_i = f(1)$ for $i = 1, \ldots, |A'|$.

For any extremal distribution $\bar{x}$ in g, $\bar{x}_i \geq f(1)$ for $i = 1, \ldots, n$. Since A is a maximum independent set in g, any set $S \subseteq N$ with $|S| > |A|$ must contain at least two adjacent vertices. Hence, we cannot have $\bar{x}_i = f(1)$ for some $i > |A|$, so that either $\bar{x}$ is more equal than $\bar{x}'$, or $\bar{x}$ and $\bar{x}'$ are incomparable. $\qquad \square$

References

Barkey, K. (2008). *Empire of Difference: The Ottomans in Comparative Perspective.* Cambridge, UK: Cambridge University Press.

Bilbao, J. M. (2000). *Cooperative Games on Combinatorial Structures*, Volume 26 of *Theory and Decision Library C.* Berlin: Springer.

Bloch, F., G. Genicot, and D. Ray (2008). Informal insurance in social networks. *Journal of Economic Theory 143*, 36–58.

Demange, G. (2004). On group stability in hierarchies and networks. *Journal of Political Economy 112*, 754 – 778.

Estlund, C. (2003). *Working Together.* Oxford: Oxford University Press.

Goyal, S. and F. Vega-Redondo (2007). Structural holes in social networks. *Journal of Economic Theory 137*, 460–492.

Hojman, D. and A. Szeidl (2008). Core and periphery in networks. *Journal of Economic Theory 139*, 295–309.

Iyengar, G., W. Kets, and R. Sethi (2010). Stable inequality and coalitional deviations in networks. Mimeo.

Jackson, M. and A. Wolinsky (1996). A strategic model of social and economic networks. *Journal of Economic Theory 71*, 44 – 74.

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

Kalai, E., A. Postlewaite, and J. Roberts (1978). Barriers to trade and disadvantageous middlemen: Nonmonotonicity of the core. *Journal of Economic Theory 19*, 200–209.

Kaneko, M. and M. Wooders (1982). Cores of partitioning games. *Mathematical Social Sciences 3*, 312–327.

Kets, W., G. Iyengar, R. Sethi, and S. Bowles (2009). Inequality and network structure. mimeo.

Livy, T. (1960). *The History of Rome from its Foundation*. London: Penguin.

Myerson, R. B. (1977). Graphs and cooperation in games. *Mathematics of Operations Research 2*, 225–229.

Schrijver, A. (2003). *Combinatorial Optimization: Polyhedra and Efficiency*, Volume 24 of *Algorithms and Combinatorics*. Berlin: Springer-Verlag.

Slikker, M. and A. van den Nouweland (2001). *Social and Economic Networks in Cooperative Game Theory*. Boston, MA: Kluwer Academic.

Takaki, R. T. (1983). *Pau Hana: Plantation Life and Labor in Hawaii 1835-1920*. Honolulu: University of Hawaii Press.

West, D. B. (2001). *Introduction to Graph Theory* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.