WARWICK

ECONOMICS

CRETA

Centre for Research in Economic Theory and its Applications

Discussion Paper Series

Network Comparative Statics

Andrew Harkins

(This paper also appears in
Warwick Economics Research Papers series No: 1306)

October 2020 No: 64

CRETA
Centre for Research in Economic Theory and its Applications

Department of Economics
University of Warwick, Coventry,
CV4 7AL, United Kingdom

Network Comparative Statics

Andrew Harkins*

October 1, 2020

Abstract

This paper develops a framework for analyzing the effect of arbitrary changes to network structure in linear-quadratic games on networks. Changes to network structure which increase total activity and total utility are studied for the case of strategic complements and strategic substitutes. Changes which are welfare increasing are found to depend on a new measure of centrality which counts the total length of walks from a node.

Two optimal network design problems are then considered. Total activity is found to be a convex function of the edge weights of the network, which allows for convex optimization techniques to be applied to minimize total activity as in the traditional 'key player' problem. Welfare maximizing network structures are also studied and previous results which associate optimal networks with nested split graphs are generalized.

1 Introduction

Networks play an important role in our lives. We send information along communication networks, travel along transportation networks, purchase goods through supplier networks, find jobs through our professional networks and influence friends through our personal networks. These networks determine what we hear, who we can reach, what we can do and ultimately how we act. Yet some positions in the network

*Department of Economics, University of Warwick, Coventry CV4 7AL, United Kingdom. Email: andrew.harkins@warwick.ac.uk

are more privileged than others. This idea, referred to as network *centrality*, has long been studied by sociologists and has more recently come to the attention of economists.

Economists have used the tools of game theory to develop new measures of network centrality and provide theoretical foundations for existing measures. The seminal contribution of Ballester et al. (2006) highlighted an equivalence between the equilibrium actions of agents in their model and a particular centrality measure known as Katz-Bonacich centrality. This connection between activity and Katz-Bonacich centrality has proven to be insightful and far-reaching. Applications of this insight have been made in settings such as production networks (Acemoglu et al., 2012), political lobbying (Battaglini and Patacchini, 2018), intra-firm communication (Calvó-Armengol and de Martí, 2009), peer effects (Calvó-Armengol et al., 2009), monopoly pricing (Candogan et al., 2012; Bloch and Quérou, 2013), information acquisition (Myatt and Wallace, 2019) and interaction in cities (Helsley and Zenou, 2014).

Despite extensive and varied applications of the concept, comparatively little work has been done to examine how arbitrary changes to network structure impact Katz-Bonacich centrality. This paper addresses this gap by providing an analysis of how changes to network structure affect individuals' centralities in the network. All the results presented will apply to games on networks with linear best replies where agents' actions are strategic complements. Several of the key results will also carry over to the more complex case of strategic substitutes as studied in Bramoullé et al. (2014).

The question of how to alter the network structure to achieve a particular policy goal was partially addressed by Ballester et al. (2006) themselves in their analysis of 'key players' and later 'key links' (Ballester et al., 2010). The key player in a network is the node who generates the largest reduction in effort following their removal from the network. An overview of this literature is provided by Zenou (2016), including some empirical applications. The question of how to optimally redesign the network so as to lower aggregate effort (e.g. in a criminal network) was not addressed but will be examined in this paper.

Efficient network design in games with strategic complementarity has been studied by Belhaj et al. (2016). They identify a particular network intervention which leads to increases in welfare and use it to generate optimal network structures under a variety of cost functions. Acemoglu et al. (2016) also compare the efficiency properties of network structures in a similar framework, however they focus on the impact of shocks to individual nodes and analyze the first and second order impact of those shocks for different networks.

A related area of the literature which has received recent attention is the design of targeted interventions in networks. A planner may wish to increase aggregate activity or welfare by providing targeted subsidies or taxes. Both Demange (2017) and Galeotti et al. (2020) examine optimal targeting problems where a planner is able to allocate resources to individuals in order to achieve a desired outcome. Their papers differ to this one as subsidies are targeted at nodes to alter their private marginal payoff from action, rather than to alter their neighbors in the network.

The focus in Section 2 is on discrete changes to network structure where edges of any weight can be added or removed from the network. Section 3 then presents some results on continuous (infinitesimal) changes to network structure and outlines a convex optimization approach to optimal network design. Section 4 examines a different network design problem and generalizes some previous results on the optimality of nested split graphs using techniques related to vector majorization.

2 Discrete Changes to Network Structure

A network $\mathbf{g}$ is a list of edges defined over a set of nodes (or vertices) which is denoted by $N = \{1, \ldots, n\}$. The network has an associated weighted adjacency matrix $\mathbf{G}$ where $g_{ij} > 0$ indicates the existence of an edge from node i to node j. The neighborhood of i is the set $N_i = \{j \in N : g_{ij} > 0\}$. Unless otherwise stated, $\mathbf{G}$ is a possibly non-symmetric real matrix with off-diagonal entries $g_{ij} \geq 0$ and diagonal entries $g_{ii} = 0$.

The degree of a node is given by $d_i = \sum_{j=1}^{n} g_{ij}$ and d_{max} denotes the maximum weighted degree in the network, i.e. $d_{max} \equiv \max_{i \in N} d_i$. A vector $\mathbf{x}$ is non-negative (written $\mathbf{x} \geq 0$) if $x_i \geq 0$ for all i and $x_{i'} > 0$ for some i' and positive (written $\mathbf{x} > 0$) if $x_i > 0$ for all i. A matrix $\mathbf{A}$ is similarly non-negative (written $\mathbf{A} \geq 0$) if $a_{ij} \geq 0$ for all ij and $a_{i'j'} > 0$ for some $i'j'$.

For any matrix $\mathbf{A}$ let $\mathbf{a}_i$ denote the i^{th} row of $\mathbf{A}$ and let $\lambda_i(\mathbf{A})$ denote the i^{th} eigenvalue of $\mathbf{A}$. Whenever the eigenvalues of $\mathbf{A}$ are real they are indexed such that $\lambda_1(\mathbf{A}) \geq \lambda_2(\mathbf{A}) \geq \cdots \geq \lambda_n(\mathbf{A})$. We also let $\rho(\mathbf{A}) = \max_i\{|\lambda_i|\}$ denote the spectral radius of a matrix $\mathbf{A}$.

Each node in the network represents an individual who is assumed to maximize the utility function

$$u_i(x_i) = \alpha_i x_i + \delta \sum_{j \in N} g_{ij} x_i x_j - \frac{c}{2} x_i^2. \tag{1}$$

The variable $x_i \geq 0$ represents a privately costly activity such as the amount of effort expended, investment made or information collected. The cost parameter c is by convention normalized to $c = 1$. Each individual may differ in their marginal private benefit from activity, this is denoted by α_i. The parameter δ measures the impact of the activity by i's neighbors. By differentiating u_i with respect to x_i we see that agent i's action is a strategic complement for agent $j's$ action if $\delta > 0$ and a strategic substitute if $\delta < 0$.[1]

A key concept in the analysis of games played on networks is Katz-Bonacich centrality. This measure assigns a centrality score to each node based on their connectedness in the network. The Katz-Bonacich centralities are given by the vector $\mathbf{b} \equiv \mathbf{M}\boldsymbol{\alpha}$ where $\mathbf{M} = (\mathbf{I} - \delta\mathbf{G})^{-1}$. If $\rho(\delta\mathbf{G}) < 1$ then $\mathbf{M}$ can be expressed as $\mathbf{M} = \sum_{k=0}^{\infty}(\delta\mathbf{G})^k$ where $\mathbf{G}^k$ counts all walks of length k. This means that entry m_{ij} of $\mathbf{M}$ measures the total weighted walks from i to j and $b_i = \sum_j m_{ij}\alpha_j$ measures the total walks out of i weighted by the α_j of the destination node j.

When $\mathbf{G}$ is not symmetric there is a distinction between the vector of 'out-centralities' $\mathbf{M}\boldsymbol{\alpha}$ and the vector of 'in-centralities' $\boldsymbol{\alpha}^T\mathbf{M}$. In order to avoid ambiguity when the network is not symmetric we will present results for 'out-centralities' only.

[1] It is assumed throughout that all parameters and the network structure itself is known to the players. A treatment of network games with incomplete information is given in (Galeotti et al., 2010).

Results for 'in-centralities' can then be easily obtained by replacing $\mathbf{G}$ with $\mathbf{G}^T$, $\mathbf{M}$ with $\mathbf{M}^T$, and so on.

Case 1 - Strategic Complements ($\delta > 0$)

The standard case to which all the following results will apply is where actions are local strategic complements. In this case the weights $g_{ij} \geq 0$ indicate the relative impact of agent $j's$ action on $i's$ incentive to increase their action x_i. Best responses are linear and weakly increasing in the actions of other players. The system of best responses can be written as $\mathbf{x} = \boldsymbol{\alpha} + \delta \mathbf{G}\mathbf{x}$. As shown by Ballester et al. (2006), the unique Nash equilibrium in the case where $0 < \delta < 1/\rho(\mathbf{G})$ and $\boldsymbol{\alpha} > 0$ is when all agents pick $x_i = b_i$.

Belhaj et al. (2014) show that Katz-Bonacich centralities do not always characterize equilibrium actions when $\delta \geq 1/\rho(\mathbf{G})$ and the action set is bounded from above. We will therefore maintain the assumptions from Ballester et al. (2006) that $\rho(\delta\mathbf{G}) < 1$, $x_i \in [0, \infty)$ and $\alpha_i > 0$ for all $i \in N$.

Examples of where this version of the model could be applied include peer effects (Calvó-Armengol et al., 2009), production networks (Acemoglu et al., 2012) and monopoly pricing (Bloch and Quérou, 2013). The results presented will extend to models which do not share the same payoff function as in (1) but still involve Katz-Bonacich centrality (or some variant), for example the information acquisition models of (Calvó-Armengol et al., 2015) and (Myatt and Wallace, 2019) and the limiting case of the Banerjee et al. (2013) model of information diffusion.

Case 2 - Strategic Substitutes ($\delta < 0$)

The second case to which several of the central results of this paper also apply is where actions are local strategic substitutes. In this case $\delta < 0$ and the weights $g_{ij} \geq 0$ now indicate the relative impact of agent $j's$ action on $i's$ incentive to decrease x_i. The best responses again satisfy $\mathbf{x} = \boldsymbol{\alpha} + \delta \mathbf{G}\mathbf{x}$ if individuals have the utility function in (1).

As discussed in Bramoullé et al. (2014), we may extend the analysis to games which have different payoffs to (1) but are best response equivalent. For example, if individuals have the utility function

$$u_i(x_i) = f(x_i + \delta \sum_{j \in N} g_{ij} x_j) - c x_i, \tag{2}$$

where f is a differentiable, concave and strictly increasing function and $c > 0$ is the constant marginal cost of effort. If y is defined implicitly as $f'(y) = c$, then best responses satisfy

$$\begin{array}{ll} x_i = y - \mid \delta \mid \sum_j g_{ij} x_j & \text{if } \mid \delta \mid \sum_j g_{ij} x_j < y \\ x_i = 0 & \text{if } \mid \delta \mid \sum_j g_{ij} x_j \geq y. \end{array}$$

Normalizing the cost parameter c so that $y = 1$ we see that the efforts of each agent at an interior equilibrium are given by $\mathbf{x} = \mathbf{1} - |\delta|\mathbf{G}\mathbf{x}$ and therefore $\mathbf{x} = (\mathbf{I} - \delta\mathbf{G})^{-1}\mathbf{1}$, where $\mathbf{1}$ is a vector (of appropriate length) with all entries equal to 1. It is then possible to apply some results concerning changes in $\mathbf{x}$ to games which are best response equivalent, although any conclusions about welfare would not necessarily apply.

2.1 Comparative Statics

To introduce the notion of comparative statics between different networks, let $\widetilde{\mathbf{G}} \equiv \mathbf{G} + \mathbf{P}$ where $\mathbf{P}$ can be thought of as a perturbation to network structure or a *policy* imposed by a planner. Here we assume that the policy matrix is taken from the set of zero diagonal $n \times n$ matrices $\mathcal{P}_{n \times n}$. The main concern is then to analyze the impact of $\mathbf{P}$ on the vector of actions as it changes from $(\mathbf{I} - \delta\mathbf{G})^{-1}\boldsymbol{\alpha} = \mathbf{x}$ to $(\mathbf{I} - \delta\widetilde{\mathbf{G}})^{-1}\boldsymbol{\alpha} = \tilde{\mathbf{x}}$. One obvious impact to analyze would be the change in total activity $\mathbf{1}^T(\tilde{\mathbf{x}} - \mathbf{x})$. Another candidate would be the change in total utility at the Nash equilibrium, given by $\sum_i (u_i(\tilde{\mathbf{x}}) - u_i(\mathbf{x})) = \frac{1}{2}\tilde{\mathbf{x}}^T\tilde{\mathbf{x}} - \frac{1}{2}\mathbf{x}^T\mathbf{x}$.

A source of concern following the change to network structure is that a new equilibrium vector $\tilde{\mathbf{x}} \geq 0$ may not exist. It may be that the iterative process of summing up walks via $\sum_{k=0}^{\infty}(\delta\widetilde{\mathbf{G}})^k$ does not converge to a finite value, so even if $\widetilde{\mathbf{M}} \equiv (\mathbf{I}-\delta\widetilde{\mathbf{G}})^{-1}$ exists it may have negative entries. To ensures a new equilibrium $\tilde{\mathbf{x}} \geq 0$ exists and is unique we will make the following assumption:

Assumption 1: If $\delta > 0$ then $\boldsymbol{\alpha} > \mathbf{0}$ and $\rho(\delta\widetilde{\mathbf{G}}) < 1$.

When $\delta < 0$ it was shown by Bramoullé et al. (2014) that there may be multiple equilibria, some of which contain at least one inactive agent with $x_i = 0$. This presents two issues, firstly, if there is a change to network structure which results in multiple equlilibria how do we know one will be reached? Secondly, since we wish to consider only interior solutions, we must rule out equilibria with inactive agents whose actions are not described by the solution to $\mathbf{x} = (\mathbf{I} - \delta\mathbf{G})^{-1}\boldsymbol{\alpha}$. To resolve both of these issues, stronger assumptions are made when $\delta < 0$:

Assumption 2: If $\delta < 0$ then $\boldsymbol{\alpha} = \mathbf{1}$, $\mid \delta d_{max}(\mathbf{G}) \mid< 1$ and $\mid \delta d_{max}(\widetilde{\mathbf{G}}) \mid< 1$.

The assumption that $\mid \delta d_{max}(\widetilde{\mathbf{G}}) \mid< 1$ is sufficient to guarantee that $\rho(\delta\widetilde{\mathbf{G}}) < 1$ and is actually necessary when all agents have equal degrees. The interpretation of this assumption is that neighborhood activity is an imperfect substitute for individual activity, since the marginal benefit of an increase in neighbors' total effort is less than the marginal benefit from an increase in individual effort.

These assumptions are stronger than those made for Proposition 2 of Bramoullé et al. (2014) which only requires that $|\lambda_{min}(\delta\mathbf{G})| < 1$, however this allows us to extend their analysis to directed graphs and to obtain the stronger result that all agents play a strictly positive effort in the unique equilibrium.[2]

The following lemma gives sufficient conditions for the existence of a new equilibrium vector $\tilde{\mathbf{x}} > 0$ after a given policy $\mathbf{P}$:

Lemma 1. *Suppose Assumptions 1 and 2 hold, then* $\mathbf{x} > \mathbf{0}$ *and* $(\mathbf{I} - \delta\widetilde{\mathbf{G}})\tilde{\mathbf{x}} = \boldsymbol{\alpha}$ *has unique solution* $\tilde{\mathbf{x}} > \mathbf{0}$.

[2]The assumption that $\boldsymbol{\alpha}$ is a constant vector in the case of substitutes is also made in Bramoullé et al. (2014)

Proof. All proofs are contained in the appendix. □

By making Assumptions 1 and 2 when the network is undirected, we can also guarantee convergence to this new equilibrium $\tilde{\mathbf{x}}$ under the Nash tâtonnement dynamic as it is globally asymptotically stable (Bramoullé et al., 2014).

The next lemma allows us to express the difference between $\tilde{\mathbf{x}}$ and $\mathbf{x}$ in terms of $\mathbf{P}$ and is used in many of the results which will follow:

Lemma 2. $\widetilde{\mathbf{M}} - \mathbf{M} = \widetilde{\mathbf{M}}\delta\mathbf{P}\mathbf{M} = \mathbf{M}\delta\mathbf{P}\widetilde{\mathbf{M}}$.

Corollary 1. *If $\delta > 0$ then $\delta\mathbf{P}\mathbf{x} \geq \mathbf{0}$ implies that $\tilde{\mathbf{x}} \geq \mathbf{x}$.*

An immediate implication of post-multiplying $\widetilde{\mathbf{M}}\delta\mathbf{P}\mathbf{M}$ by $\boldsymbol{\alpha}$ is that $\delta\mathbf{P}\mathbf{x} \geq \mathbf{0}$ guarantees that $\tilde{\mathbf{x}} \geq \mathbf{x}$ whenever $\delta > 0$ because $\widetilde{\mathbf{M}}$ is a non-negative matrix. If $\delta < 0$ then $\widetilde{\mathbf{M}}$ may have negative entries and therefore a similar implication does not hold for the case of substitutes.[3]

The intuition for this in the case of complements is straightforward, if we fix the action profile at $\mathbf{x}$ and alter the network such that every agent experiences a net increase in the total effort of their neighbors, then all agents will have the incentive to increase their effort. The new equilibrium action profile in $\widetilde{\mathbf{G}}$ must weakly increase for all i due to strategic complementarity in actions. We can extend this intuition with the following result:

Proposition 1. *Suppose $\delta > 0$, then $\tilde{\mathbf{x}} > \mathbf{x}$ if and only if $\sum_{k=0}^{\ell}(\delta\widetilde{\mathbf{G}})^k\delta\mathbf{P}\mathbf{x} > 0$ for some $\ell \geq 0$.*

One way to interpret Proposition 1 is to consider an iterative process of myopic best responses after a change to network structure. Suppose each agent plays a myopic best response at each time $t \in \mathbb{N}$ given by

$$x_i^{(t)} = \alpha_i + \sum_{j \in N} \delta \left(g_{ij} + p_{ij}\right) x_j^{(t-1)}. \tag{3}$$

[3] Counter examples can easily be found even when $n = 3$.

This process of best responses will always converge to a new equilibrium $\tilde{\mathbf{x}}$ for any initial condition $\mathbf{x}^{(0)} \geq 0$ (Milgrom and Roberts, 1990). If each individual begins this process at $x_i^{(0)} = x_i$ and updates according to (3) then

$$\mathbf{x}^{(1)} - \mathbf{x}^{(0)} = \delta \left(\mathbf{G} + \mathbf{P} \right) \mathbf{x}^{(0)} - \delta \mathbf{G} \mathbf{x}^{(0)} = \delta \mathbf{P} \mathbf{x}^{(0)} \tag{4}$$

is the change in actions after one period. For $t \geq 2$ the per period change is

$$\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)} = \delta \widetilde{\mathbf{G}} \left(\mathbf{x}^{(t-1)} - \mathbf{x}^{(t-2)} \right). \tag{5}$$

Using (4) and (5) we see that after an update in period 2 the total change in action is

$$\left(\mathbf{x}^{(2)} - \mathbf{x}^{(1)} \right) + \left(\mathbf{x}^{(1)} - \mathbf{x}^{(0)} \right) = (\mathbf{I} + \delta \widetilde{\mathbf{G}}) \delta \mathbf{P} \mathbf{x}.$$

The result states that if at some t the cumulative effect of this myopic best response process (summarized by $(\sum_{k=0}^{t} (\delta \widetilde{\mathbf{G}})^k) \mathbf{P} \mathbf{x}$) results in all agents playing a strictly higher action than their original x_i then they must continue to increase their actions as $\mathbf{x}^{(t)}$ converges toward $\tilde{\mathbf{x}}$.

2.2 Undirected Networks

The benchmark case in models of games played on networks is the undirected graph where agents interact symmetrically. In this subsection we will make the assumption that $g_{ij} = g_{ji}$ but maintain the previous assumption that edges can be weighted. For the first result in this section we will also restrict attention to the case where $\boldsymbol{\alpha}$ is constant. In this simplified benchmark case we can establish the following result:

Theorem 1. *Let* $\mathbf{g}$ *be undirected and* $\boldsymbol{\alpha} = \mathbf{1}$. *If* $\mathbf{x}^T \delta \mathbf{P} \mathbf{x} > 0$ *then* $\sum_i \tilde{x}_i > \sum_i x_i$.

Any policy which satisfies $\mathbf{x}^T \delta \mathbf{P} \mathbf{x} = \delta \sum_{ij} p_{ij} x_i x_j > 0$ will exhibit a net increase in the total activity, in the cases of both strategic complements and strategic substitutes.

This condition is sufficient to guarantee an increase in total activity but is not necessary. An example of a change which does not satisfy $\delta \mathbf{x}^T \mathbf{P} \mathbf{x} > 0$ but leads to

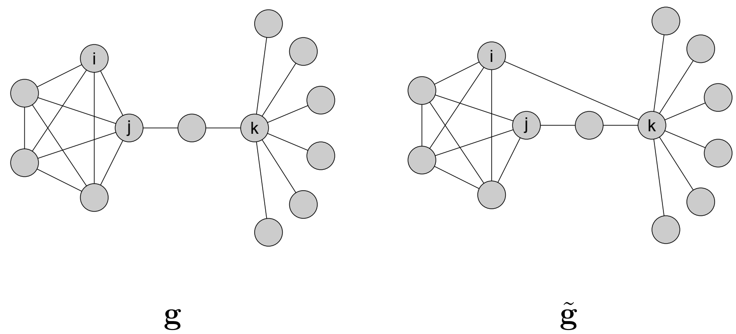

Figure 1: An Edge Switch

an increase in total activity levels is shown in Figure 1. For $\delta = 0.15$ the activity levels are $x_k = 2.51$ and $x_j = 2.82$ with the associated value of $\mathbf{x}^T\mathbf{P}\mathbf{x} = -0.24$, yet $\sum_i(\tilde{x}_i - x_i) = 0.10$.

Figure 1 highlights a policy which is of particular interest. We will call such a policy an *edge switch* by i from j to k, where an edge ij is removed and an edge ik is added. In an unweighted network an *edge switch* means that

$$\mathbf{x}^T\delta\mathbf{P}\mathbf{x} = \delta(2x_ix_k - 2x_ix_j) = \delta 2x_i(x_k - x_j).$$

This policy would increase total activity whenever $\delta(x_k - x_j) > 0$.

Corollary 2. *An edge switch by i from j to k where $\delta(x_k - x_j) > 0$ strictly increases total activity.*

If actions are strategic complements then we can increase total activity by switching edges to higher activity nodes. When applying this result to strategic substitutes the intuition changes slightly. If $x_k < x_j$ then the policy increases total activity if we switch an edge between nodes i and j for an edge between nodes i and k. Like the case of strategic complements, an edge switch which increases activity will typically see one prominent node gaining more edges. This is because nodes with *lower* activity tend to have more links when efforts are substitutes.

Application: Monopoly pricing with consumption externalities

One application of Theorem 1 is to the model of Bloch and Quérou (2013). In the main version of their model a group of consumers are located on an undirected network and decide whether to adopt a new technology provided by a monopolist seller. Consumers have a private valuation θ_i which is drawn from a uniform distribution on $[0, 1]$ but they also derive utility from adoption by their neighbors. The utility from i adopting is

$$u_i = \theta_i - p_i + \delta \sum_j g_{ij} \Pr[j \text{ adopts}].$$

The authors examine optimal discriminatory prices when costs are given by $c_i(q_i) = cq_i$ with $q_i = 1$ if i adopts and $q_i = 0$ if not. For linear costs, the monopolist charges a uniform price $p = \frac{1+c}{2}$ which is independent of the location of the consumer in the network. The vector of (expected) demands $\mathbf{q}(p)$ is shown to be $\mathbf{q}(p) = \frac{1-c}{2}(\mathbf{I} - \delta\mathbf{G})^{-1}\mathbf{1} = \frac{1-c}{2}\mathbf{x}$. Profits are therefore given by

$$\pi = \frac{(1-c)^2}{4}\mathbf{1}^T\mathbf{x},$$

which is an increasing function of the sum of Katz-Bonacich centrality scores.[4] Theorem 1 implies that if consumers in this market are encouraged to edge switch from low x_j consumers to higher x_k consumers then profit for the monopolist will increase.

2.3 Neighborhood Changes in Undirected Networks

The problem of characterizing the exact effects of changes to network structure on the vector of efforts $\mathbf{x}$ in full generality is challenging. Complex interactions between first order and higher order neighborhood changes are difficult to summarize in terms of a single sufficient statistic that is computationally easier than simply inverting $(\mathbf{I} - \widetilde{\mathbf{G}})$ and calculating the new $\tilde{\mathbf{x}}$ directly. A simplification which will make tackling this problem more straightforward is to only consider changes to a single node's

[4] Bloch and Quérou (2013) also provide a version of their model where the monopolist has convex costs. The implications of Theorem 1 also apply to that case. Details are provided in Appendix B.

neighborhood, rather than multiple simultaneous changes across the network.

The focus will now shift to the case where alterations are made to one individual's neighborhood while all other neighborhoods remain fixed. This restriction is realistic in some contexts and will allow us to strengthen some of the earlier results. Changes to network structure may occur over long periods of time and on a node-by-node basis. Actions then have time to adjust to the new network structure before another structural change takes place. Avoiding simultaneous changes to multiple neighborhoods eliminates interactions between these changes and simplifies the analysis to a significant degree.

Making a neighborhood change in an undirected network involves changing both the out-neighborhood and in-neighborhood simultaneously. If we let $\mathbf{p} = \tilde{\mathbf{g}}_i - \mathbf{g}_i$ denote the change to agent i's out-neighborhood, then $\mathbf{P}$ is simply $\mathbf{P} = \mathbf{e}_i\mathbf{p}^T + \mathbf{p}\mathbf{e}_i^T$ where $\mathbf{e}_i$ is a vector with $e_i = 1$ and $e_{j\neq i} = 0$. This leads to the following lemma:

Lemma 3. *Assume* $\mathbf{g}$ *is undirected and we make a symmetric change to the neighborhood of agent* i*. Then*

$$\widetilde{\mathbf{M}} - \mathbf{M} = \delta D^{-1}\left(c_1\mathbf{MPM} + \delta c_2\mathbf{m}_i\mathbf{m}_i^T + \delta m_{ii}\mathbf{M}\mathbf{p}\mathbf{p}^T\mathbf{M}\right), \tag{6}$$

where $c_1 = 1 - \delta\mathbf{p}^T\mathbf{m}_i > 0$*,* $c_2 = \mathbf{p}^T\mathbf{M}\mathbf{p} > 0$ *and* $D = c_1^2 - \delta^2 c_2 m_{ii} > 0$*.*

The main benefit of Equation (6) is in allowing for a deeper characterization of welfare improving structural changes. To examine this further, a new measure of network centrality is defined below.

Definition 1. The vector of *walk length centralities* is given by $\boldsymbol{\xi} \equiv \mathbf{M}^T\mathbf{M}\boldsymbol{\alpha}$.

As suggested by the name, the measure ξ_i has an interpretation in terms of summing up walks in the network and weighting them by a measure of walk length. In the original formulation of Katz-Bonacich centrality (Katz, 1953) the vector of centralities $\mathbf{x}_{KB}$ counts the total number of walks which leave each agent, discounting walks of length k by factor δ^k. This is given by

$$\mathbf{x}_{KB} = \delta\mathbf{G}\mathbf{x} = \left(\delta\mathbf{G} + \delta^2\mathbf{G}^2 + \delta^3\mathbf{G}^3 + \cdots\right)\mathbf{1}.$$

If $\mathbf{g}$ is symmetric then

$$\mathbf{M}^T\mathbf{M} = (\mathbf{I} + \delta\mathbf{G} + \delta^2\mathbf{G}^2 + \cdots)^2 = \sum_{k=1}^{\infty} k(\delta\mathbf{G})^{k-1},$$

provided $\rho(\delta\mathbf{G}) < 1$.[5] This implies for $\boldsymbol{\alpha} = \mathbf{1}$ that

$$\delta\mathbf{G}\boldsymbol{\xi} = \sum_{k=1}^{\infty} k(\delta\mathbf{G})^k\mathbf{1}. \tag{7}$$

For Katz-Bonacich centrality the i^{th} entry of $\delta\mathbf{G}\mathbf{x}$ gives a count of the total number of walks leaving i, discounting walks of length k by δ^k. By comparison, for walk length centrality the i^{th} entry of $\delta\mathbf{G}\boldsymbol{\xi}$ gives the total length of all discounted walks which leave i in terms of the number of edges traversed, where again walks of length k are also discounted by δ^k. If we instead consider $\boldsymbol{\xi} = \sum_{k=1}^{\infty} k(\delta\mathbf{G})^{k-1}\mathbf{1}$ we see that it provides a measure of total walk length in terms of nodes visited.

When $\mathbf{g}$ is not symmetric we can also interpret the walk length centrality measure as a Katz-Bonacich centrality measure in its own right. If we pre-multiply $\boldsymbol{\xi}$ by $\mathbf{I}-\delta\mathbf{G}^T$ we get $(\mathbf{I} - \delta\mathbf{G}^T)\boldsymbol{\xi} = \mathbf{x}$ and therefore

$$\boldsymbol{\xi} = \mathbf{x} + \delta\mathbf{G}^T\boldsymbol{\xi}. \tag{8}$$

Writing out this system of equations for each i we get $\xi_i = x_i + \delta\sum_{j=1}^{n} g_{ji}\xi_j$ and we can then interpret ξ_i as a Katz-Bonacich measure of centrality where $\alpha_i = x_i$.[6] This newly defined measure is now applied in Proposition 2 below.

Proposition 2. *Assume* $\mathbf{g}$ *is undirected. If we make a change to the neighborhood of agent* i *then:*

[5] Diagonalizing $\mathbf{G}$ gives $\mathbf{G} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^T$ where $\mathbf{Q}$ is orthogonal with i^{th} column equal to the i^{th} eigenvector of $\mathbf{G}$ and $\boldsymbol{\Lambda}$ is diagonal and contains the eigenvalues of $\mathbf{G}$. Differentiating $\sum_{k=0}^{\infty} a^k = (1-a)^{-1}$ with respect to a gives $\sum_{k=1}^{\infty} ka^{k-1} = (1-a)^{-2}$ when $|a| < 1$. Hence if $a = \delta\lambda_i$ then $\sum_{k=1}^{\infty} k(\delta\mathbf{G})^{k-1} = \mathbf{Q}(\mathbf{I} - \delta\boldsymbol{\Lambda})^{-2}\mathbf{Q}^T = \mathbf{M}^2$.

[6] Like Katz-Bonacich centrality, this measure can be negative when $\delta < 0$. If $\delta > 0$ then $\xi_i > 0$ for all i, as can be seen by iterating on equation (8).

- $\mathbf{p}^T\mathbf{x} \geq 0$ *and* $\mathbf{p}^T\boldsymbol{\xi} \geq 0$ *implies that* $\sum_i \tilde{u}_i > \sum_i u_i$ *whenever* $\delta > 0$.
- $\mathbf{p}^T\mathbf{x} \leq 0$ *and* $\mathbf{p}^T\boldsymbol{\xi} \leq 0$ *implies that* $\sum_i \tilde{u}_i > \sum_i u_i$ *whenever* $\delta < 0$ *and* $\xi_i > 0$.

To guarantee an increase in utility we need to consider how neighborhood changes interact with the two centrality measures $\mathbf{x}$ and $\boldsymbol{\xi}$. If we fix a profile of actions and change the neighborhood of i such that there is a net increase in the Katz-Bonacich and walk length centralites of i's neighbors then total utility will increase.

From Theorem 1 we know that deleting edge ij and adding edge ik would be guaranteed to increase total activity if $\delta(x_k - x_j) \geq 0$. However, in Belhaj et al. (2016) it was shown that $\delta(x_k - x_j) \geq 0$ is not sufficient to also guarantee an increase total utility. Using Proposition 2 we can obtain a new sufficient condition for an edge switch to increase total utility.

Corollary 3. *If the assumptions of Proposition 2 hold then an edge switch by i from j to k strictly increases total utility if* $\delta(x_k - x_j) \geq 0$ *and* $\delta(\xi_k - \xi_j) \geq 0$.

2.4 Neighborhood Changes in Directed Networks

We will now focus on the case where an agent can only change their neighborhood in one direction, altering either their out-neighbors or their in-neighbors. By doing so we make the assumption that the network is directed but maintain the previous assumption that edges are possibly weighted. As before, $\mathbf{x}$ is the out-centrality vector which results from post-multiplying $\mathbf{M}$ by $\boldsymbol{\alpha}$ and $\overleftarrow{\mathbf{x}}$ will be the in-centrality vector which results from pre-multiplying $\mathbf{M}$ by $\boldsymbol{\alpha}$.

2.4.1 Changes to Out-Neighborhoods

Network changes which are restricted to only the out-neighborhood of a given node are of particular interest and have a strikingly simple structure. If only the out-neighborhood of agent i is changed then we can write $\mathbf{P}$ as simply $\mathbf{P} = \mathbf{e}_i\mathbf{p}^T$. We can then view $\widetilde{\mathbf{M}} = (\mathbf{I} - \delta\mathbf{G} - \delta\mathbf{P})^{-1}$ as a rank one update of $\mathbf{M} = (\mathbf{I} - \delta\mathbf{G})^{-1}$ because the matrix $\mathbf{P}$ is rank one by construction. This allows us to apply the Sherman-Morrison formula for a rank one update of the inverse of a matrix to obtain

$$\widetilde{\mathbf{M}} = \mathbf{M} + \mathbf{M}\delta(\mathbf{e}_i\mathbf{p}^T)\mathbf{M}\left(\frac{1}{1-\delta\mathbf{p}^T\mathbf{M}\mathbf{e}_i}\right). \tag{9}$$

We then have the following result:

Theorem 2. *Assume* $\mathbf{g}$ *is directed and that* $\delta > 0$*. If we alter the* i^{th} *row of* $\mathbf{G}$ *from* $\mathbf{g}_i$ *to* $\tilde{\mathbf{g}}_i = \mathbf{g}_i + \mathbf{p}$ *then the following statements are equivalent:*

(i) $\mathbf{p}^T\mathbf{x} \geq 0$.

(ii) $\tilde{x}_i \geq x_i$.

(iii) $u_i(\tilde{\mathbf{x}}) \geq u_i(\mathbf{x})$.

(iv) $\tilde{x}_j \geq x_j$ *for all* $j \neq i$.

(v) $u_j(\tilde{\mathbf{x}}) \geq u_j(\mathbf{x})$ *for all* $j \neq i$.

(vi) $\sum_{k=1}^{n} \tilde{x}_k \geq \sum_{k=1}^{n} x_k$.

Fixing the action profile at $\mathbf{x}$ the condition that $\mathbf{p}^T\mathbf{x} \geq 0$ ensures that the total out-centrality of the neighbors of i increases. If $\delta > 0$ then it is intuitive that changes of this kind should then weakly increase the total number of out-walks in $\tilde{\mathbf{g}}$ since only the out-neighborhood of i has changed. Any directed walks which do not pass through i are unaffected, but walks which reach i at some point must increase in weight since the total walks out of i have increased. The central insight of Theorem 2 is firstly that this condition is necessary for such increases and it is also equivalent to a weak increase in the activity levels of all agents when $\delta > 0$.

It is also possible to state a similar result for strategic substitutes under the assumption that the maximum in-degree, denoted by $\overleftarrow{\mathbf{d}}_{max}$, is less than $1/ \mid \delta \mid$.

Theorem 3. *Assume* $\mathbf{g}$ *is directed,* $\delta < 0$ *and that* $\overleftarrow{\mathbf{d}}_{max} < 1/ \mid \delta \mid$*. If we alter the* i^{th} *row of* $\mathbf{G}$ *from* $\mathbf{g}_i$ *to* $\tilde{\mathbf{g}}_i = \mathbf{g}_i + \mathbf{p}$ *then the following statements are equivalent:*

(i) $\mathbf{p}^T\mathbf{x} \leq 0$.

(ii) $\tilde{x}_i \geq x_i$.

(iii) $u_i(\tilde{\mathbf{x}}) \geq u_i(\mathbf{x})$.

(iv) $\sum_{j=1}^{n} \tilde{x}_j \geq \sum_{j=1}^{n} x_j$.

Theorems 2 and 3 therefore have an important implication. They tell us that any decentralized process of network formation where agents act in sequence and have the ability to optimally alter their out-neighbors will naturally lead to increases in total activity. Each agent only needs to take account of the activity levels of their potential neighbors when switching links.

Furthermore, if $\delta > 0$ then any process where agents sequentially replace low activity neighbors which high activity neighbors will result in an increase in total utility. Consequently, a network is socially optimal in terms of maximizing aggregate utility if and only if it is privately optimal for each individual.[7]

2.4.2 Changes to In-Neighborhoods

Now consider the case where an agent best-responds to out-neighbor's actions but only control their in-links. As we shall see, the above link between private and social optimality under out-neighborhood changes no longer holds.

To remain consistent with the rest of the paper we will state results in this subsection in terms of how changes to in-neighborhoods affect out-centrality. However, we may also interpret these results in a setting where agents select their out-neighbors (who they follow) but pick actions based on the actions of their in-neighbors (who follows them). An example which fits the first interpretation is the inter-sectoral production model of Acemoglu et al. (2012) where firms/sectors pick inputs but they (and the planner) are typically concerned with levels of output.

Proposition 3. *Let* $\mathbf{g}$ *be directed and let* $\overleftarrow{\mathbf{g}}_i$ *denote the* i^{th} *row of* $\mathbf{G}^T$. *If we alter the in-neighborhood of agent* i *from* $\overleftarrow{\mathbf{g}}_i$ *to* $\tilde{\overleftarrow{\mathbf{g}}}_i = \overleftarrow{\mathbf{g}}_i + \mathbf{p}$ *then:*

[7] This of course ignores the fact that actions will still be below socially optimal levels.

- $\tilde{x}_i \geq x_i$ *if and only if* $\delta \mathbf{p}^T \mathbf{m}_i \geq 0$.
- $\sum_{j=1}^{n} \tilde{x}_j \geq \sum_{j=1}^{n} x_j$ *if and only if* $\delta \mathbf{p}^T \mathbf{M}^T \mathbf{1} \geq 0$.
- $\tilde{\mathbf{x}} \geq \mathbf{x}$ *if and only if* $\mathbf{M} \delta \mathbf{p} \geq 0$.

When making changes to the in-neighborhood of a node we lose the equivalence between changes which increase the activity of agent i and changes which increase total activity. It is no longer necessarily the case the linking to nodes with higher activity levels will lead to higher total activity.

Moreover, Proposition 3 shows that individual nodes cannot simply link to agents of higher activity to guarantee an increase in their utility. Increases in utility can however be guaranteed if they link to nodes with higher walk length centrality.

Proposition 4. *If we alter a single agent's in-neighborhood then* $\delta \mathbf{p}^T \boldsymbol{\xi} \geq 0$ *implies that* $\sum_{j=1}^{n} u_j(\tilde{x}_j) \geq \sum_{j=1}^{n} u_j(x_j)$.

When considering how to make a change to the out-neighborhood of agent i to increase total utility it was sufficient to use the vector of activities $\mathbf{x}$. When making changes to in-neighborhoods the vector of walk length centralities $\boldsymbol{\xi}$ should be used to inform this decision.

3 Continuous Changes to Network Structure

Except for relatively simple cases it is not always possible to summarize the impact of structural changes in terms of a single statistic or measure. The mathematical expressions for non-trivial changes to network structure are often unwieldy and require a planner to counterbalance the direct impact of a policy on the local neighborhood with its impact via higher order network effects.

The task of analyzing how arbitrary changes can impact activities can however be drastically simplified when considering continuous, rather than discrete, changes to network structure. In doing so we effectively take a linear approximation to the change in $\mathbf{x}$, which sidesteps the difficulties associated with summarizing the impact of higher order interactions. This simplification is appropriate in cases where the size of the perturbation is small in magnitude or in the case where a network is evolving continuously through time.

To modify the framework used so far, we assume that the strength of each edge changes continuously with a parameter t (e.g. time). These changes are represented by the matrix $\hat{\mathbf{P}}$ where for each entry $\hat{p}_{ij} = \frac{dg_{ij}}{dt}$. The change in activity x_i with respect to t can be stated as

$$\frac{dx_i}{dt} = \frac{d\left(\alpha_i + \delta \sum_j g_{ij} x_j\right)}{dt} = \sum_{j \in N} \delta \left(\frac{dg_{ij}}{dt} x_j + g_{ij} \frac{dx_j}{dt} \right). \tag{10}$$

The relationship in equation (10) holds for each node i, so writing this system of equations in matrix form we have $\frac{d\mathbf{x}}{dt} = \delta\hat{\mathbf{P}}\mathbf{x} + \delta\mathbf{G}\frac{d\mathbf{x}}{dt}$. We can then solve to get

$$\frac{d\mathbf{x}}{dt} = \mathbf{M}\delta\hat{\mathbf{P}}\mathbf{x}. \tag{11}$$

A number of results can be immediately derived from equation (11). Consider first the case of a directed graph where $\frac{dg_{ij}}{dt} = 1$ for a given edge ij and $\frac{dg_{k\ell}}{dt} = 0$ for all other edges, then using (11)

$$\frac{d\mathbf{x}}{dt} = \delta \begin{bmatrix} m_{1i}x_j \\ m_{2i}x_j \\ \vdots \\ m_{ni}x_j \end{bmatrix}.$$

If instead we have an undirected graph then $\frac{dg_{ij}}{dt} = 1$ implies that $\frac{dg_{ji}}{dt} = 1$ and therefore

$$\frac{d\mathbf{x}}{dt} = \delta \begin{bmatrix} m_{1i}x_j + m_{1j}x_i \\ m_{2i}x_j + m_{2j}x_i \\ \vdots \\ m_{ni}x_j + m_{nj}x_i \end{bmatrix}.$$

More complex changes to network structure can be easily constructed in this linear fashion by including other non-zero terms in $\hat{\mathbf{P}}$.

We can see from (11) that $d(\mathbf{1}^T\mathbf{x})/dt$ is given by $\mathbf{x}^T\delta\hat{\mathbf{P}}\mathbf{x}$ in the undirected case with constant $\boldsymbol{\alpha}$. It is also straightforward to calculate the rate of change for $\sum_i u_i = \sum_i x_i^2/2$ with respect to t. Differentiating for each i we get $du_i/dt = x_i\frac{dx_i}{dt}$ and summing up we get $d\left(\sum_i u_i\right)/dt = \mathbf{x}^T\frac{d\mathbf{x}}{dt} = \boldsymbol{\xi}^T\delta\hat{\mathbf{P}}\mathbf{x}$. Now walk length centrality plays a crucial role in determining whether a continuous change to network structure increases welfare.

3.1 Optimal Network Design

With explicit results for derivatives with respect to edge weight g_{ij} in hand, attention naturally turns to the optimal choice of edge weight with respect to some differentiable objective function. Using these results it is now possible to follow a general convex optimization approach to key player and optimal network design problems. The original 'key player' problem described in Ballester et al. (2006) involved a planner who was able to remove a single node from the network to optimally lower aggregate activity. It was shown in Ballester et al. (2006) that a planner would optimally target the node with highest intercentrality measure, defined as $ICM_i = x_i^2/m_{ii}$.

To extend and generalize the key player problem we will now consider an optimal network design problem using convex optimization techniques. A planner wishes to alter an undirected network $\mathbf{g}$ to minimize some undesirable activity which has local strategic complementarity, such as crime. Assuming that $\delta > 0$, all actions are weakly increasing in g_{ij} and hence the planner would optimally choose to disconnect all nodes in the network by reducing all g_{ij} to zero. However, the problem becomes non-trivial if we suppose that the planner also has a budget B for edge changes which cannot be exceeded.

Stating the problem formally for $\delta > 0$ the planner wishes to solve:

$$\begin{aligned} \underset{\widetilde{\mathbf{G}} \in \mathbb{R}^{\mathbf{n}\times\mathbf{n}}}{\text{minimize}} \quad & \textstyle\sum_i x_i \\ \text{subject to} \quad & \tilde{g}_{ij} \geq 0 \qquad (\forall i, j \neq i) \\ & \tilde{g}_{ii} = 0 \\ & \textstyle\sum_{ij} |p_{ij}| \leq B \end{aligned} \tag{12}$$

The first and second constraints ensure that the new matrix $\widetilde{\mathbf{G}}$ is non-negative and has a zero diagonal. The third constraint will always bind provided that $B \leq \mathbf{1}^T\mathbf{G}\mathbf{1}$ because reducing some $g_{ij} > 0$ will always lower total activity.

A further implicit constraint is that $\rho(\delta\widetilde{\mathbf{G}}) \leq 1$, however, this cannot bind at any optimum. This is true because the optimal policy cannot have any $p_{ij} > 0$ as reducing to $p_{ij} = 0$ will lower total activity and stay within the budget. Therefore $\mathbf{P} \leq 0$ at the optimum and consequently $\rho(\delta\widetilde{\mathbf{G}}) \leq \rho(\delta\mathbf{G}) < 1$.[8]

In order to solve this activity minimization problem we now prove that total activity, which is given by $f(\mathbf{G}) = \sum_i x_i$, is indeed a convex function. Let Γ be the set of symmetric non-negative $(n \times n)$ matrices $\widetilde{\mathbf{G}}$ with zero diagonal such that $\rho(\widetilde{\mathbf{G}}) < 1$.

Lemma 4. *If $\boldsymbol{\alpha} = c\mathbf{1}$ for scalar $c > 0$ and $\mathbf{G}$ is symmetric then $f(\mathbf{G}) = \sum_i x_i$ is strictly convex on Γ.*

Lemma 4 holds in the case of either strategic complements or strategic substitutes and now enables convex optimization techniques to be applied to this problem.[9] It is

[8]This follows from the Courant-Fischer theorem which states that $\lambda_1(\widetilde{\mathbf{G}}) = \max_{||\mathbf{v}||=1} \frac{\mathbf{v}^T\widetilde{\mathbf{G}}\mathbf{v}}{\mathbf{v}^T\mathbf{v}}$ which is attained at $\mathbf{v} = \mathbf{v}_1$ where $\mathbf{v}_1$ is the principal eigenvector of $\widetilde{\mathbf{G}}$. Since λ_1 and $\mathbf{v}_1$ can be chosen to be nonnegative for a nonnegative matrix (Meyer, 2000) we see that $\frac{\mathbf{v}_1^T\mathbf{G}\mathbf{v}_1}{\mathbf{v}_1^T\mathbf{v}_1} = \lambda_1(\widetilde{\mathbf{G}}) - \frac{\mathbf{v}_1^T\mathbf{P}^*\mathbf{v}_1}{\mathbf{v}_1^T\mathbf{v}_1} \geq \lambda_1(\widetilde{\mathbf{G}})$. Consequently, $\lambda_1(\mathbf{G}) = \max_{||\mathbf{v}||=1} \frac{\mathbf{v}^T\mathbf{G}\mathbf{v}}{\mathbf{v}_1^T\mathbf{v}_1} \geq \lambda_1(\widetilde{\mathbf{G}})$ and $\rho(\delta\widetilde{\mathbf{G}}) \leq \rho(\delta\mathbf{G}) < 1$.

[9]Lemma 4 also provides some further intuition for Theorem 1. The first order Taylor approximation of a convex function is always a global under-estimator of that function. Therefore $f(\delta\widetilde{\mathbf{G}}) \geq f(\delta\mathbf{G}) + vec(\delta\mathbf{P})^T(\mathbf{x} \otimes \mathbf{x})$ where $vec(\delta\mathbf{P})$ denotes the vectorization of $\delta\mathbf{P} = \delta\widetilde{\mathbf{G}} - \delta\mathbf{G}$ and $\mathbf{x} \otimes \mathbf{x}$ denotes the Kronecker product of $\mathbf{x}$ with itself, which is the gradient vector of f. It can then be verified that this is equivalent to $f(\widetilde{\mathbf{G}}) \geq f(\mathbf{G}) + \mathbf{x}^T\delta\mathbf{P}\mathbf{x}$ and therefore $f(\delta\widetilde{\mathbf{G}}) - f(\delta\mathbf{G}) \geq \mathbf{x}^T\delta\mathbf{P}\mathbf{x}$

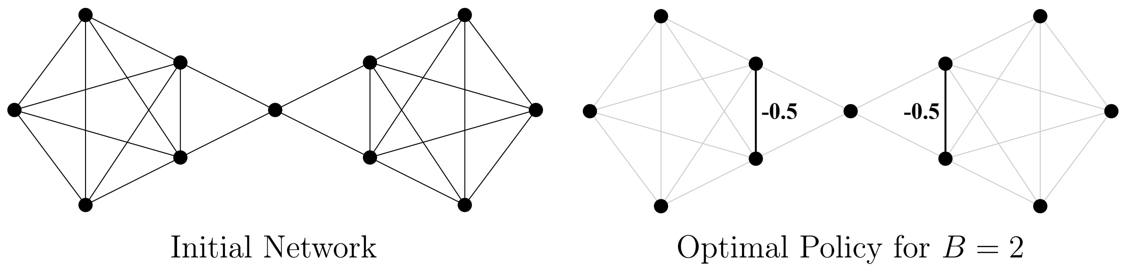

Figure 2: The Optimal Policy when $\delta = 0.2$ and $B = 2$

worth noting that convexity of $\sum_i x_i$ always holds when the network is undirected and agents have identical types α_i but may not hold otherwise. Numerical simulations reveal that x_i and $\sum_i x_i$ are not convex when agents can have heterogeneous types (i.e. $\alpha_i \neq \alpha_j$). Moreover, x_i, $\sum_i x_i$ and $\sum_i u_i$ are not convex if $\mathbf{G}$ can be non-symmetric.

The network in Figure 2 was used in Ballester et al. (2006) to illustrate the difference between key players and agents with highest activity. When $\delta = 0.2$ it is the individual who acts as a bridge between the two cliques who has the highest intercentrality measure, even though each of their 4 neighbors have the highest Katz-Bonacich centralities in the network.

The policy which minimizes total activity when $B = 2$ is also shown in Figure 2. Interestingly, the optimal policy does not alter any edges connected to the key player but instead lowers the weights on edges connecting the agents with highest Katz-Bonacich centralities.[10] When B is increased to 8 so that the removal of the key player from the network is feasible, the optimal intervention involves lowering the weight on all edges until the degree of each individual equals $\frac{48-B}{|N|} \approx 3.63$. The optimal network for $B = 8$ is shown in Figure 3.

As can be seen in this second example, removing the 'key player' from the network is not optimal when the planner has the option of varying all edge weights. In fact, it can never be optimal for a planner to completely remove any player from the network unless all players are being removed. Intuitively, this comes from the fact that $\sum_i x_i$ is

[10]Calculations given in this section were made using the semidefinite programming routines of the MATLAB package *cvx*.

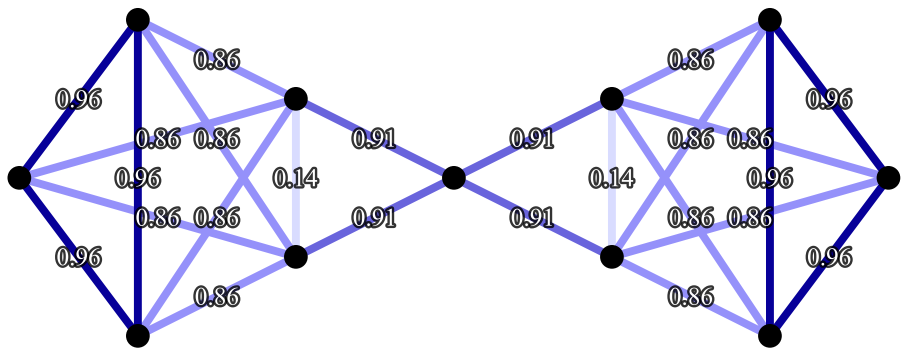

Figure 3: The Optimal Network when $\delta = 0.2$ and $B = 8$

convex in the entries of $\mathbf{G}$ but the 'cost' of lowering edge weight is constant, therefore there are decreasing returns to lowering weight along any given edge.

To show this formally we can examine the Kuhn-Tucker conditions which are both necessary and sufficient when the objective function and constraints are convex and differentiable. Although the constraint in (12) is non-differentiable, we can rewrite the constraint as $\sum_{ij} g_{ij} - \tilde{g}_{ij} \leq B$ and then add a constraint for $p_{ij} \leq 0$, since we know that $p_{ij} \leq 0$ at any optimum. Now writing the Lagragian we have

$$\mathcal{L} = \sum_{i \in N} \tilde{x}_i - \sum_{i<j} \lambda_{ij} \tilde{g}_{ij} + \sum_{i<j} \mu_{ij} (\tilde{g}_{ij} - g_{ij}) + \nu \sum_{i<j} (2(g_{ij} - \tilde{g}_{ij}) - B),$$

which leads to the following necessary and sufficient conditions for each edge ij with $i < j$:

$$\begin{aligned} \delta \tilde{x}_i \tilde{x}_j &= \nu && \text{if } 0 < \tilde{g}_{ij} < g_{ij} \\ \delta \tilde{x}_i \tilde{x}_j &= \nu + \lambda_{ij} && \text{if } 0 = \tilde{g}_{ij} < g_{ij} \\ \delta \tilde{x}_i \tilde{x}_j &= \nu - \mu_{ij} && \text{if } 0 < \tilde{g}_{ij} = g_{ij} \end{aligned}$$

If it was optimal to remove individual i, then for every agent j that had been connected to i it must hold that $\delta \tilde{x}_i \tilde{x}_j = \nu + \lambda_{ij} \geq \nu$. Now $\tilde{x}_i = 1$ yet for any $k \in N$ with $\tilde{g}_{jk} > 0$ it holds that $x_k > 1$ and consequently $\delta \tilde{x}_j \tilde{x}_k > \delta \tilde{x}_i \tilde{x}_j \geq \nu$, which contradicts optimality. This argument shows that if it is ever optimal to remove i it must be optimal to remove any j also previously linked to i, and so on throughout the network.[11]

To conclude this section a brief comment on the case of strategic substitutes is necessary. Although convexity of the objective function also applies when $\delta < 0$, the activity minimization problem requires further constraints. Because activity is a local public good when $\delta < 0$, the optimal policy always involves adding weight to $\mathbf{G}$, which in turn leads to an increase in $\rho(\delta\mathbf{G})$. With no restriction on row sums it may be possible at the optimum that $\mathbf{I} - \delta\widetilde{\mathbf{G}}$ is no longer positive definite or even invertible. However, restricting the budget B so that row sums of $\delta\widetilde{\mathbf{G}}$ are always less than 1 will ensure that $\tilde{x}_i > 0$ for each i and that $\mathbf{I} - \delta\widetilde{\mathbf{G}}$ is positive definite. Convex optimization techniques can then be applied.

4 Dominance and Nested Split Graphs

This section provides another application of the results from Section 2 to an optimal network design problem. We will now develop a framework for generating welfare improvements whilst holding the number of edges or total edge weight fixed. Past work has identified optimal structures for undirected networks in the unweighted case (Belhaj et al., 2016). Optimal weighted networks have been examined by Olaizola and Valenciano (2019) in the context of connections models such as Bala and Goyal (2000) and Bloch and Dutta (2009), however their framework does not apply to the linear-quadratic model we study here. Recent work by Li (2020) examines the optimal design of directed networks but does so assuming differentiability of the objective function. Results in this section will provide a more general framework which includes these previous results as special cases. Throughout this section it is assumed that $\delta > 0$.[12]

[11] This argument also applies if the network is not connected. If we delete one component we must also delete the other as $\delta\tilde{x}_i\tilde{x}_j = \delta = \nu$ for the deleted component but $\delta\tilde{x}_{i'}\tilde{x}_{j'} > \delta$ for i' and j' in the other component.

[12] The case of strategic substitutes is more complex. Welfare maximizing networks were studied in a similar context by Bramoullé and Kranton (2007). They show that optimal networks can contain

Research on the interaction between network formation and games on networks has frequently highlighted the special role of nested split graphs. Although several equivalent definitions of nested split graphs have been given, we will use the definition which is most common in the economics literature:

Definition 2. An undirected and unweighted network $\mathbf{g}$ is a nested split graph if $d_i \geq d_j \implies N_i \cup \{i\} \supseteq N_j$.

If $g_{ij} = 1$ in a nested split graph then $g_{ik} = 1$ for any agent k such that $d_k \geq d_j$. The efficiency properties of nested split graphs (NSGs) were demonstrated by Belhaj et al. (2016) for the case of undirected and unweighted networks, where it is shown that NSGs maximize total utility under linear linking costs.

When agents simultaneously make linking decisions and take complementary actions it has been shown by Hiller (2017) that NSGs are the only pairwise stable network structures. NSGs have also been shown by König et al. (2014) to arise as the stochastically stable outcome of a preferential attachment process of network formation with strategic complementary in actions.

The results of Section 2 and the framework developed below will now be used to add to these earlier results on NSGs. Throughout this section we will index the nodes according to their activity levels, i.e. $x_1 \geq x_2 \geq \cdots \geq x_n$. A vector $\mathbf{x} \in \mathbb{R}^n$ *dominates* a vector $\mathbf{y} \in \mathbb{R}^n$ (written $\mathbf{x} \succcurlyeq \mathbf{y}$) if $\sum_{i=1}^k x_i \geq \sum_{i=1}^k y_i$ for all $k \in \{1, \dots, n\}$. When $\mathbf{x}$ and $\mathbf{y}$ are probability vectors this is equivalent to familiar notion of stochastic dominance. If $\mathbf{x}^{\downarrow}$ denotes the vector $\mathbf{x}$ which has been rearranged in decreasing order (i.e. $x_1 \geq x_2 \geq ... \geq x_n$), then $\mathbf{x}$ *weakly submajorizes* $\mathbf{y}$ (written $\mathbf{x} \succ_w \mathbf{y}$) if $\mathbf{x}^{\downarrow} \succcurlyeq \mathbf{y}^{\downarrow}$.[13]

It is well-known that if $\mathbf{x} \succ_w \mathbf{y}$ then $f(\mathbf{x}) \geq f(\mathbf{y})$ for any increasing Schur-convex function f (Marshall et al., 2011). Several examples of Schur-convex functions are discussed in Marshall et al. (2011) but for our purposes it is important to note that $f(\mathbf{x})$ is Schur-convex if $f(\mathbf{x}) = \sum_i \phi(x_i)$ where ϕ is an increasing and convex function.

inactive agents who contribute no effort, a scenario not covered by the framework in this paper.

[13]If $\mathbf{x}^{\downarrow} \succcurlyeq \mathbf{y}^{\downarrow}$ and $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ then $\mathbf{x}$ *majorizes* $\mathbf{y}$. If in addition $\sum_{i=1}^k x_i^{\downarrow} \geq \sum_{i=1}^k y_i^{\downarrow}$ holds strictly for at least one $k \in \{1, \dots, n-1\}$ then $\mathbf{x}$ *strictly majorizes* $\mathbf{y}$.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

$\mathbf{L}$ $\qquad$ $\mathbf{L}^{-1}$

Figure 4: $\mathbf{L}$ and $\mathbf{L}^{-1}$

Therefore, both $f(\mathbf{x}) = \sum_{i=1}^{n} x_i$ and $f(\mathbf{x}) = \sum_{i=1}^{n} \frac{1}{2} x_i^2$ are Schur-convex functions.[14]

Let $\mathbf{L}$ denote the lower triangular matrix with $\mathbf{L}_{ij} = 1$ for $i \geq j$ and $\mathbf{L}_{ij} = 0$ for $i < j$ (see Figure 4). Pre-multiplying any vector $\mathbf{x}$ by $\mathbf{L}$ returns a running sum of that vector and hence $\mathbf{Lx} \geq \mathbf{Ly} \iff \mathbf{x} \succcurlyeq \mathbf{y}$.[15] Pre-multiplying any matrix $\mathbf{A}$ by $\mathbf{L}$ returns a matrix with columns equal to a running sum of the columns of $\mathbf{A}$. If we let $\mathbf{U} \equiv \mathbf{L}^T$ then post-multiplying any matrix $\mathbf{A}$ by $\mathbf{U}$ returns a matrix with rows equal to a running sum of the rows of $\mathbf{A}$.

The inverses of $\mathbf{L}$ and $\mathbf{U}$ also have useful properties. Consider first $\mathbf{L}^{-1}$ which is also shown in Figure 4. Pre-multiplying $\mathbf{L}^{-1}$ by any vector $\mathbf{x}$ gives $\mathbf{x}^T\mathbf{L}^{-1} \geq 0$ if $\mathbf{x}$ is decreasing and non-negative. Equivalently for $\mathbf{U}^{-1} = (\mathbf{L}^{-1})^T$, $\mathbf{U}^{-1}\mathbf{x} \geq 0$ if $\mathbf{x}$ is decreasing and non-negative.[16]

Using these matrices allows us to easily prove a number of facts relating dominance to increases in equilibrium actions or welfare, for example the following proposition:

Proposition 5. *Assume $\delta > 0$. If each row of $\widetilde{\mathbf{G}}$ dominates the corresponding row of $\mathbf{G}$ then $\tilde{x}_i \geq x_i$ for all $i \in N$.*

The focus in the remainder of this section will be on policies which are *rewirings* of the network, that is, policies which hold the total number of edges or edge weight fixed. These policies are natural in some contexts, for example when in-degree or out-

[14]Other examples of Schur-convex functions include the variance of $\mathbf{x}$ and the Gini coefficent of $\mathbf{x}$. A function is Schur-concave if $-f$ is Schur-convex, examples include the entropy of $\mathbf{x}$ and $\Pi_{i=1}^{n} x_i$.

[15]Note that $\mathbf{Lx} = \mathbf{Ly} \iff \mathbf{x} = \mathbf{y}$ because $\mathbf{L}$ is invertible.

[16]These matrices were originally used in the context of 'monotone' Markov chains, see Keilson and Kester (1977). Applications to economics were later made by Conlisk (1985).

$$\underbrace{\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}}_{\tilde{\mathbf{G}}} - \underbrace{\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}}_{\mathbf{G}} = \underbrace{\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & -1 \\ 1 & -1 & -1 & 0 \end{pmatrix}}_{\mathbf{P}}$$

$$\begin{array}{c}(\mathbf{LPU})_{ij} \text{ sums all elements} \\ \text{of } \mathbf{P} \text{ from } \mathbf{P}_{11} \text{ to } \mathbf{P}_{ij}.\end{array} \quad \rightarrow \quad \mathbf{LPU}: \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

Figure 5: A positive rewiring in a symmetric network for $n = 4$

degree is fixed for each agent (e.g. if each agent has a fixed amount of time to spend with others). This also applies when a planner is asked to construct a network from scratch with a given budget B of edges or edge weight. Alternatively, the planner may face a cost function which is increasing in the total edge weight used.

Within this class of network rewirings, a particular subset of policies called *positive rewirings* will be of interest.

Definition 3. A rewiring $\mathbf{P}$ is a *positive rewiring if* $\mathbf{LPU} \geq 0$.

The main intuition behind positive rewirings is that they are a net redistribution of edge weight upwards and/or leftwards in the adjacency matrix. This is seen by noting that each entry ij of the matrix $\mathbf{LPU}$ is given by $\mathbf{LPU}_{ij} = \sum_{i' \leq i, j' \leq j} p_{i'j'}$.

For any positive rewiring we see that the degree distribution of $\tilde{\mathbf{g}}$ dominates $\mathbf{g}$, at least weakly, since

$$\mathbf{L}(\mathbf{G} + \mathbf{P})\mathbf{1} \geq \mathbf{LG1} \iff \mathbf{LP1} \geq \mathbf{0},$$

which holds due to the fact that $\mathbf{LPU} \geq 0$ and $\mathbf{U}^{-1}\mathbf{1} \geq \mathbf{0}$.

Two examples of positive rewirings are worth highlighting. The first is the *edge switch* from j to k in the case where $x_k > x_j$ which, as was seen in Section 2, is guaranteed to increase total activity when $\mathbf{g}$ is symmetric. The second example is the *neighborhood switch* defined by Belhaj et al. (2016), where for two nodes j and k with $x_k > x_j$ the planner takes all edges where $g_{ij} = 1$ but $g_{ik} = 0$ for a given $i \in N$ and simultaneously switches them so that $g_{ij} = 0$ and $g_{ik} = 1$.

When the network is undirected and edges are unweighted there is a close relationship between positive rewirings and NSGs. We will now show that $\mathbf{g}$ is a nested split graph if and only if $\mathbf{g}$ admits no positive rewiring. In the process we will also show that this is equivalent to $\mathbf{G}$ being a 'stepwise' matrix. Following König et al. (2014), the definition of a stepwise matrix is given below:

Definition 4. A stepwise matrix $\mathbf{A}$ is a symmetric, binary $(n \times n)$ matrix with elements a_{ij} satisfying the following condition: if $i < j$ and $a_{ij} = 1$, then $a_{k\ell} = 1$ whenever $k < \ell \leq j$ and $k \leq i$.

Proposition 6. *Assume that $\delta > 0$. If $\mathbf{g}$ is symmetric and unweighted then the following statements are equivalent:*

(i) $\mathbf{g}$ *is a nested split graph.*

(ii) $\mathbf{G}$ *is a stepwise matrix.*

(iii) $\mathbf{g}$ *admits no positive rewiring.*

NSGs have stepwise adjacency matrices (König et al., 2014) and hence by continually moving edges upwards and leftwards in the adjacency matrix we can reach a nested split graph.

To now state the final result of this section, let $\mathbf{g}^*(E)$ be a network which maximizes $\Phi(\mathbf{x})$ where Φ is a Schur-convex function of $\mathbf{x}$ and E is a fixed total number of edges or edge weight. To ensure a well-defined maximum exists the following assumption is made:

Assumption 3: If $\mathbf{g}$ is unweighted then $\frac{1}{n-1} > \delta > 0$ and if $\mathbf{g}$ is weighted then $\frac{1}{E} > \delta > 0$.

This assumption guarantees that any permutation of edges or edge weight will always converge to a finite $\tilde{\mathbf{x}}$.

Theorem 4. *Assume $\boldsymbol{\alpha}$ is constant and that Assumption 3 holds. Then the optimal $\mathbf{g}^*(E)$ admits no positive rewiring.*

Belhaj et al. (2016) show that any network which is not an NSG admits a welfare improving neighborhood switch. The result above extends this intuition to a larger class of networks (including directed networks) and any welfare function which is Schur-convex. The reader will notice that all of the results in this section are based on techniques related to vector majorization and do not rely on differentiability of the welfare function Φ.

For undirected but weighted networks there is a unique network which admits no positive rewiring, that is where $g_{12} = g_{21} = E/2$ and $g_{ij} = 0$ for all other ij pairs. For networks which are both directed and weighted there is a potential infinity of networks which admit no positive rewiring. However, it can be shown that the network which maximizes utility when edges are weighted and directed is exactly the same as in the undirected case (i.e. $g_{12} = g_{21} = E/2$).[17]

Putting these results in the context of the literature it is clear that nested split graphs (and their generalization via the concept of positive rewirings) occupy a special status as both efficient and stable networks.

5 Conclusion

This paper has made two main contributions to the literature. The first is to develop a general framework for analyzing how changes to network structure affect equilibrium activity and welfare in the linear-quadratic setting. This complements recent work by Galeotti et al. (2020) on node level targeting by examining how changes to network structure impact welfare in games on networks. Neighborhood changes in directed networks prove to be the most tractable to analyze and an equivalence between policies which are privately and socially beneficial is highlighted.

The second contribution has been to develop and extend approaches to optimal network design problems. Convex optimization techniques can be applied to show how a planner with a fixed budget would optimally adjust edge weights to minimize

[17] Writing $x_i = 1 + \delta g_{ij}(1 + \delta g_{ji} x_i)$ and solving for $x_i = \frac{1+\delta g_{ij}}{1-\delta^2 g_{ij} g_{ji}}$ it is then straightforward to show that $g^*_{ij} = g^*_{ji} = E/2$ under the constraint $g_{ji} = E - g_{ij}$.

total activity, generalizing previous work on 'key player' policies. A new technique based on vector majorization can be used when the planner must reallocate edges in the network to maximize total activity or welfare. These results extend those of Belhaj et al. (2016) to a larger class of networks.

Further extensions and applications of the convex optimization approach outlined in Section 3 are possible. It would be of interest to extend this to the more general network models introduced in Acemoglu et al. (2016), which include non-linear aggregation functions. It may also be of interest to study policies which minimize the variance of activity in settings where coordination is important, for example a supply chain network or a communication network. Initial simulations for this objective function are suggestive of convexity in undirected networks but whether this actually holds is an open question.

Appendix A - Proofs

Proof of Lemma 1. For the case of complements see Ballester et al. (2006). For substitutes we will first show that if a Nash equilibrium $\mathbf{x}$ exists then $x_i \in (0, 1)$ for each $i \in N$. Taking the utility function in 2 and normalizing c so that $f'(1) = c$, the best responses satisfy

$$\begin{array}{ll} x_i = 1- \mid \delta \mid \sum_j g_{ij}x_j & \text{if } \mid \delta \mid \sum_j g_{ij}x_j < 1 \\ x_i = 0 & \text{if } \mid \delta \mid \sum_j g_{ij}x_j \geq 1. \end{array}$$

Since $\mid \delta d_{max}(\mathbf{G}) \mid < 1$ we infer that $x_i \leq 1$. This implies that $\delta \sum_j g_{ij}x_j \leq \delta \sum_j g_{ij} < 1$ therefore that $x_i \in (0, 1)$ and hence if a Nash equilibrium exists it is interior.

The interior Nash equilibrium is unique if and only if there is a unique solution to the system $\mathbf{x} = \mathbf{1} + \delta \mathbf{G}\mathbf{x}$, which occurs if and only if $\mathbf{I} - \delta \mathbf{G}$ is invertible. By Gerschgorin's circle theorem, if λ_i is the i^{th} eigenvalue of $\delta \mathbf{G}$ it holds that $|\lambda_i| \leq \mid \delta d_{max}(\mathbf{G}) \mid < 1$ for all i. The spectral norm therefore satisfies $\rho(\delta \mathbf{G}) < 1$ which is well-known to imply that $\lim_{k \to \infty} (\delta \mathbf{G})^k = 0$ and that $(\mathbf{I} - \delta \mathbf{G})^{-1}$ exists and is equal to $\sum_{k=0}^{\infty}(\delta\mathbf{G})^k$, see for example Meyer (2000). ☐

Proof of Lemma 2. $\widetilde{\mathbf{M}}$ satisfies $\widetilde{\mathbf{M}}(\mathbf{I}-\delta\mathbf{G}-\delta\mathbf{P})=\mathbf{I}$, or equivalently $\widetilde{\mathbf{M}}(\mathbf{I}-\delta\mathbf{G})=\mathbf{I}+\widetilde{\mathbf{M}}\delta\mathbf{P}$. Post-multiplying by $\mathbf{M}$ and rearranging gives $\widetilde{\mathbf{M}}-\mathbf{M}=\widetilde{\mathbf{M}}\delta\mathbf{P}\mathbf{M}=\mathbf{M}\delta\mathbf{P}\widetilde{\mathbf{M}}$ where the final equality follows from $(\mathbf{I}-\delta\mathbf{G}-\delta\mathbf{P})\widetilde{\mathbf{M}}=\mathbf{I}$ by an identical argument. ☐

Proof of Proposition 1. To show sufficiency begin by recursively substituting in to the right-hand side of $\tilde{\mathbf{x}}-\mathbf{x}=\delta\widetilde{\mathbf{G}}\,(\tilde{\mathbf{x}}-\mathbf{x})+\delta\mathbf{P}\mathbf{x}$ a total of ℓ times. This results in

$$\tilde{\mathbf{x}}-\mathbf{x}=\left(\delta\widetilde{\mathbf{G}}\right)^{\ell+1}(\tilde{\mathbf{x}}-\mathbf{x})+\left(\sum_{k=0}^{\ell}(\delta\widetilde{\mathbf{G}})^k\right)\delta\mathbf{P}\mathbf{x}. \tag{13}$$

Solving equation (13) for $\tilde{\mathbf{x}}-\mathbf{x}$ yields

$$\tilde{\mathbf{x}}-\mathbf{x}=\left(\mathbf{I}-(\delta\widetilde{\mathbf{G}})^{\ell+1}\right)^{-1}\left(\sum_{k=0}^{\ell}(\delta\widetilde{\mathbf{G}})^k\right)\delta\mathbf{P}\mathbf{x}$$

where invertibility of $\mathbf{I}-(\delta\widetilde{\mathbf{G}})^{\ell+1}$ follows from the fact that $\rho((\delta\widetilde{\mathbf{G}})^{\ell+1})=(\rho(\delta\widetilde{\mathbf{G}}))^{\ell+1}<1$ and therefore $\mathbf{I}-(\delta\widetilde{\mathbf{G}})^{\ell+1}$ is a non-singular M-matrix (Berman and Plemmons, 1994).

Non-singular M-matrices have non-negative inverses and hence $(\mathbf{I}-(\delta\widetilde{\mathbf{G}})^{\ell+1})^{-1}\geq 0$ for any $\ell\geq 0$. Therefore, if $(\sum_{k=0}^{\ell}(\delta\widetilde{\mathbf{G}})^k)\delta\mathbf{P}\mathbf{x}>0$ for some ℓ then $\tilde{\mathbf{x}}-\mathbf{x}>0$, since $\sum_{k=0}^{\infty}(\delta\widetilde{\mathbf{G}})^{(\ell+1)k'}$ has a positive diagonal and non-negative entries elsewhere.

To show necessity, rewrite (13) as

$$\left(\mathbf{I}-\left(\delta\widetilde{\mathbf{G}}\right)^{\ell+1}\right)(\tilde{\mathbf{x}}-\mathbf{x})=\left(\sum_{k=0}^{\ell}(\delta\widetilde{\mathbf{G}})^k\right)\delta\mathbf{P}\mathbf{x}.$$

Since $\rho(\delta\widetilde{\mathbf{G}})<1$ this implies that $\lim_{\ell\to\infty}(\delta\widetilde{\mathbf{G}})^{\ell+1}(\tilde{\mathbf{x}}-\mathbf{x})=\mathbf{0}$. Therefore if $\tilde{\mathbf{x}}-\mathbf{x}>0$ then there exists an ℓ sufficiently high such that $\tilde{\mathbf{x}}-\mathbf{x}>(\delta\widetilde{\mathbf{G}})^{\ell+1}(\tilde{\mathbf{x}}-\mathbf{x})$, proving that $\left(\sum_{k=0}^{\ell}(\delta\widetilde{\mathbf{G}})^k\right)\delta\mathbf{P}\mathbf{x}>0$ for some $\ell\geq 0$. ☐

Proof of Theorem 1. From Lemma 2 we have $\tilde{\mathbf{x}}-\mathbf{x}=\widetilde{\mathbf{M}}\delta\mathbf{P}\mathbf{x}=\mathbf{M}\delta\mathbf{P}\tilde{\mathbf{x}}$ and therefore $\mathbf{1}^T(\tilde{\mathbf{x}} - \mathbf{x}) = \mathbf{x}^T\mathbf{P}\tilde{\mathbf{x}}$. If $\mathbf{Px} = \mathbf{0}$ then $\mathbf{1}^T(\tilde{\mathbf{x}} - \mathbf{x}) = 0$ so assume instead that $\mathbf{Px} \neq \mathbf{0}$.

Under the assumption that $\delta\lambda_1(\mathbf{G}) < 1$, the symmetric matrix $\mathbf{I} - \delta\mathbf{G}$ has all positive eigenvalues. Similarly $\delta\lambda_1(\widetilde{\mathbf{G}}) < 1$ implies that $\widetilde{\mathbf{M}}$ is also positive definite and thus $\mathbf{M}\delta\mathbf{P}\widetilde{\mathbf{M}}\delta\mathbf{P}\mathbf{M} = \mathbf{M}\delta\mathbf{P}(\widetilde{\mathbf{M}} - \mathbf{M})$ is positive semi-definite. Hence $\mathbf{1}^T\mathbf{M}\delta\mathbf{P}(\widetilde{\mathbf{M}} - \mathbf{M})\mathbf{1} \geq 0$ and we obtain $\mathbf{x}^T\delta\mathbf{P}\tilde{\mathbf{x}} \geq \mathbf{x}^T\delta\mathbf{P}\mathbf{x}$. This shows that $\mathbf{x}^T\mathbf{P}\mathbf{x} > 0$ implies $\mathbf{1}^T(\tilde{\mathbf{x}} - \mathbf{x}) > 0$ for $\delta > 0$ and $\mathbf{x}^T\mathbf{P}\mathbf{x} < 0$ implies $\mathbf{1}^T(\tilde{\mathbf{x}} - \mathbf{x}) < 0$ for $\delta < 0$. □

Proof of Lemma 3. This lemma is an application of the Woodbury formula for a rank 2 update of an inverse (see Horn and Johnson (2013) pg.19). Let $\mathbf{V} = (\mathbf{e}_i, \mathbf{p})$ be an $n \times 2$ matrix with columns $\mathbf{e}_i$ and $\mathbf{p}$ and let $\mathbf{W}^T = \begin{pmatrix} \mathbf{p}^T \\ \mathbf{e}_i^T \end{pmatrix}$ which is a corresponding $2 \times n$ matrix. Notice that $\mathbf{V}\mathbf{W}^T = \mathbf{e}_i\mathbf{p}^T + \mathbf{p}\mathbf{e}_i^T = \mathbf{P}$.

The Woodbury formula then states that

$$\widetilde{\mathbf{M}} = \left(\mathbf{I} - \delta\mathbf{G} - \delta\mathbf{V}\mathbf{W}^T\right)^{-1} = \mathbf{M} + \delta\mathbf{M}\mathbf{V}\left(\mathbf{I} - \delta\mathbf{W}^T\mathbf{M}\mathbf{V}\right)^{-1}\mathbf{W}^T\mathbf{M}$$

To invert $\mathbf{I} - \delta\mathbf{W}^T\mathbf{M}\mathbf{V}$ we write it as

$$\mathbf{I} - \delta\mathbf{W}^T\mathbf{M}\mathbf{V} = \begin{pmatrix} 1 - \delta\mathbf{p}^T\mathbf{M}\mathbf{e}_i & -\delta\mathbf{p}^T\mathbf{M}\mathbf{p} \\ -\delta\mathbf{e}_i^T\mathbf{M}\mathbf{e}_i & 1 - \delta\mathbf{e}_i^T\mathbf{M}\mathbf{p} \end{pmatrix}.$$

Let $D := \det\left(\mathbf{I} - \delta\mathbf{W}^T\mathbf{M}\mathbf{V}\right)$, inverting then gives:

$$\left(\mathbf{I} - \delta\mathbf{W}^T\mathbf{M}\mathbf{V}\right)^{-1} = D^{-1}\begin{pmatrix} 1 - \delta\mathbf{e}_i^T\mathbf{M}\mathbf{p} & \delta\mathbf{p}^T\mathbf{M}\mathbf{p} \\ \delta\mathbf{e}_i^T\mathbf{M}\mathbf{e}_i & 1 - \delta\mathbf{p}^T\mathbf{M}\mathbf{e}_i \end{pmatrix}$$

Because $\mathbf{p}^T\mathbf{M}\mathbf{e}_i = \mathbf{e}_i^T\mathbf{M}\mathbf{p}$ we now have

$$\left(\mathbf{I} - \delta\mathbf{W}^T\mathbf{M}\mathbf{V}\right)^{-1} = D^{-1}\left((1 - \delta\mathbf{p}^T\mathbf{M}\mathbf{e}_i)\mathbf{I} + \begin{matrix} 0 & \delta\mathbf{p}^T\mathbf{M}\mathbf{p} \\ \delta\mathbf{e}_i^T\mathbf{M}\mathbf{e}_i & 0 \end{matrix}\right),$$

and so $\widetilde{\mathbf{M}} - \mathbf{M} =$

$$\delta D^{-1}\left[(1-\delta\mathbf{p}^T\mathbf{M}\mathbf{e}_i)\mathbf{M}\delta\mathbf{P}\mathbf{M} + \left(\delta\mathbf{e}^T\mathbf{M}\mathbf{e}\right)\mathbf{M}\mathbf{p}\mathbf{p}^T\mathbf{M} + \left(\delta\mathbf{p}_i^T\mathbf{M}\mathbf{p}_i\right)\mathbf{M}\mathbf{e}_i\mathbf{e}_i^T\mathbf{M}\right].$$

The terms $\mathbf{p}^T\mathbf{M}\mathbf{p}$ and $\mathbf{e}_i^T\mathbf{M}\mathbf{e}_i$ are positive because $\mathbf{M}$ is positive definite. It remains to prove that $1-\delta\mathbf{p}^T\mathbf{M}\mathbf{e}_i > 0$ and $D > 0$. Note that

$$\mathbf{I} - \delta\mathbf{G} - \delta\mathbf{V}\mathbf{W}^T = [\mathbf{I} - \delta\mathbf{G}]\left[\mathbf{I} - \mathbf{M}\delta\mathbf{V}\mathbf{W}^T\right],$$

and therefore

$$\det\left(\mathbf{I} - \delta\mathbf{G} - \delta\mathbf{V}\mathbf{W}^T\right) = \det\left(\mathbf{I} - \delta\mathbf{G}\right)\det\left(\mathbf{I} - \mathbf{M}\delta\mathbf{V}\mathbf{W}^T\right).$$

By Theorem 1.3.22 of Horn and Johnson (2013) the non-zero eigenvalues of $(\mathbf{M}\mathbf{V})\,\mathbf{W}^T$ are equal to the non-zero eigenvalues of $\mathbf{W}^T(\mathbf{M}\mathbf{V})$, and so

$$\Pi_{i=1}^n\left(1-\delta\lambda_i\left(\mathbf{M}\mathbf{V}\mathbf{W}^T\right)\right) = \Pi_{i=1}^2\left(1-\delta\lambda_i\left(\mathbf{W}^T\mathbf{M}\mathbf{V}\right)\right).$$

Thus using the relationship between determinants and eigenvalues we have that

$$\det\left(\mathbf{I}_n - \delta\mathbf{M}\mathbf{V}\mathbf{W}^T\right) = \det\left(\mathbf{I}_2 - \delta\mathbf{W}^T\mathbf{M}\mathbf{V}\right).$$

Therefore $\det\left(\mathbf{I}_2 - \delta\mathbf{W}^T\mathbf{M}\mathbf{V}\right) > 0$ if $\det(\mathbf{I} - \delta\widetilde{\mathbf{G}}) > 0$ and $\det(\mathbf{I} - \delta\mathbf{G}) > 0$, which holds whenever $\lambda_1(\widetilde{\mathbf{G}})$ and $\lambda_1(\mathbf{G})$ are less than $1/\delta$.

Finally, to show $1-\delta\mathbf{p}^T\mathbf{M}\mathbf{e}_i > 0$ assume $\delta\mathbf{p}^T\mathbf{M}\mathbf{e}_i > 0$. Starting with $D = 1-2\delta\mathbf{p}^T\mathbf{M}\mathbf{e}_i+\delta^2\left(\mathbf{p}^T\mathbf{M}\mathbf{e}_i\right)^2-\delta^2\mathbf{p}^T\mathbf{M}\mathbf{p}\mathbf{e}_i^T\mathbf{M}\mathbf{e}_i > 0$ and applying the Cauchy-Schwartz inequality to $\mathbf{p}^T\mathbf{M}^{1/2}\mathbf{M}^{1/2}\mathbf{e}_i$ we see that $1-2\delta\mathbf{p}^T\mathbf{M}\mathbf{e}_i > 0$ because $\left(\mathbf{p}^T\mathbf{M}\mathbf{e}_i\right)^2 \leq \mathbf{p}^T\mathbf{M}\mathbf{p}\mathbf{e}_i^T\mathbf{M}\mathbf{e}_i$. Therefore $1-\delta\mathbf{p}^T\mathbf{M}\mathbf{e}_i > 0$ as required. □

Proof of Proposition 2. The change in utility is $(1/2)\tilde{\mathbf{x}}^T\tilde{\mathbf{x}}-(1/2)\mathbf{x}^T\mathbf{x} = (1/2)(\tilde{\mathbf{x}}-\mathbf{x})^T(\tilde{\mathbf{x}}-\mathbf{x})+\mathbf{x}^T(\tilde{\mathbf{x}}-\mathbf{x})$. Using Lemma 3 and the definition of $\mathbf{P}$ as $\mathbf{P} = \mathbf{e}_i\mathbf{p}^T + \mathbf{p}\mathbf{e}_i^T$

we get

$$\mathbf{x}^T(\tilde{\mathbf{x}} - \mathbf{x}) = \delta D^{-1}\left(c_1 x_i \boldsymbol{\xi}^T \mathbf{p} + c_1 \xi_i \mathbf{p}^T \mathbf{x} + \delta c_2 \xi_i x_i + \delta c_3 \boldsymbol{\xi}^T \mathbf{p}\mathbf{p}^T \mathbf{x}\right).$$

Lemma 3 established that $c_1 > 0$,$c_2 > 0$ and $D > 0$, so therefore $\boldsymbol{\xi}^T \mathbf{p} > 0$ and $\mathbf{p}^T \mathbf{x} > 0$ guarantee that utility increases for $\delta > 0$. For the $\delta < 0$ case the extra assumption that $\xi_i > 0$ guarantees that the term inside the parentheses is negative whenever $\boldsymbol{\xi}^T \mathbf{p} < 0$ and $\mathbf{p}^T \mathbf{x} < 0$. □

Proof of Theorem 2. It can be seen from equation (9) that

$$\tilde{x}_i - x_i = \frac{m_{ii} \delta \mathbf{p}^T \mathbf{x}}{1 - \delta \mathbf{p}^T \mathbf{M} \mathbf{e}_i}.$$

As $m_{ii} > 0$, statements (i) and (ii) are equivalent provided $\delta \mathbf{p}^T \mathbf{M} \mathbf{e}_i < 1$ holds for any choice of $\mathbf{p}$ such that $\rho(\widetilde{\mathbf{G}}) < 1$. The formula for a rank one update of a determinant is stated below without proof (see Meyer (2000) Section 6.2):

Lemma 5 (**Rank one update of a determinant**). *If* $\mathbf{A}$ *is non-singular and* $\mathbf{c}$ *and* $\mathbf{d}$ *are column vectors of appropriate length then* $\det(\mathbf{A} + \mathbf{c}\mathbf{d}^T) = \det(\mathbf{A}) \det(1 + \mathbf{d}^T \mathbf{A}^{-1} \mathbf{c})$.

Letting $\mathbf{A} = \mathbf{I} - \delta \mathbf{G}$, $\mathbf{c} = \mathbf{e}_i$ and $\mathbf{d} = -\delta \mathbf{p}$, this implies that $\det(\mathbf{I} - \delta \mathbf{G} - \delta \mathbf{e}_i \mathbf{p}) = \det(\mathbf{I} - \delta \mathbf{G})(1 - \delta \mathbf{p}^T \mathbf{M} \mathbf{e}_i)$. If $\delta \mathbf{p}^T \mathbf{M} \mathbf{e}_i \geq 1$ then $\det(\mathbf{I} - \delta \widetilde{\mathbf{G}}) \leq 0$ and therefore $\det(\widetilde{\mathbf{M}}) \leq 0$, which is a contradiction. When $\delta > 0$, $\widetilde{\mathbf{M}}$ has a positive determinant due to being an inverse of an M-matrix (Berman and Plemmons, 1994). Hence if $\det(\widetilde{\mathbf{M}}) > 0$ then $\delta \mathbf{p}^T \mathbf{M} \mathbf{e}_i < 1$ and (i) $\iff$ (ii). To prove the other claims:

(ii) $\iff$ (iii): At equilibrium $u_i = (1/2) x_i^2$, which is strictly increasing in x_i.

(i) $\iff$ (iv): $\tilde{\mathbf{x}} - \mathbf{x} = \mathbf{M} \mathbf{e}_i \delta \mathbf{p}^T \mathbf{x} (1 - \delta \mathbf{p}^T \mathbf{M} \mathbf{e}_i)^{-1}$ where $\mathbf{M} \mathbf{e}_i$ is the i^{th} column of $\mathbf{M}$. Since $\mathbf{M}$ is a non-negative matrix $\tilde{\mathbf{x}} - \mathbf{x} \geq 0 \iff \delta \mathbf{p}^T \mathbf{x} \geq 0$.

(iv) $\iff$ (v): Identical to (ii) $\iff$ (iii).

(i) $\iff$ (vi): Since $\mathbf{1}^T\mathbf{M}\mathbf{e}_i > \mathbf{0}$ always holds, then $\mathbf{1}^T(\tilde{\mathbf{x}} - \mathbf{x}) = \mathbf{1}^T\mathbf{M}\mathbf{e}_i\delta\mathbf{p}^T\mathbf{x}(1 - \delta\mathbf{p}^T\mathbf{M}\mathbf{e}_i)^{-1} \geq 0$ holds if and only if $\delta\mathbf{p}^T\mathbf{x} \geq 0$. □

Proof of Theorem 3. The proof is similar to Theorem 2 but first we need to establish two basic facts about $\mathbf{M}$ when $\delta < 0$.

Lemma 6. *If $\delta < 0$ then $m_{ii} > 0$.*

Proof. Letting Assumption 2 hold and applying the Gershgorin Circle Theorem we have that $\mid 1 - \mu_i(\mathbf{I} - \delta\mathbf{G}) \mid \leq \sum_{j \neq i} \mid \delta g_{ij} \mid < 1$ where μ_i is the i^{th} eigenvalue of $\mathbf{I} - \delta\mathbf{G}$. This implies that every real eigenvalue of $\mathbf{I} - \delta\mathbf{G}$ is positive since it lies in the interval $(0, 2)$. Any complex eigenvalues come in conjugate pairs and therefore if $\mu_i = a + b\mathrm{i}$ is an eigenvalue then $\mu_j = a - b\mathrm{i}$ is also an eigenvalue. Therefore $\det(\mathbf{I} - \delta\mathbf{G}) = \prod_{i=1}^{n} \mu_i > 0$ due to the positivity of all real eigenvalues and the positivity of the product of any conjugate pairs of complex eigenvalues.

Note that the above argument also holds for any principal submatrix of $\mathbf{I} - \delta\mathbf{G}$. Now let $(\mathbf{I} - \delta\mathbf{G})_{-i}$ denote the principal submatrix of $\mathbf{I} - \delta\mathbf{G}$ where the i^{th} row and column have been removed. Apply the co-factor formula for an inverse of matrix to obtain

$$m_{ii} = \det((\mathbf{I} - \delta\mathbf{G})_{-i}) / \det(\mathbf{I} - \delta\mathbf{G}) > 0.$$

□

Lemma 7. *If $\delta < 0$ and $\overleftarrow{\mathbf{d}}_{max} < 1$ then $\mathbf{1}^T\mathbf{M}\mathbf{e}_i > 0$.*

Proof. $\mathbf{M} = (\mathbf{I} + \delta\mathbf{G} + (\delta\mathbf{G})^2 \cdots)$ which can be written as $\mathbf{M} = (\mathbf{I} + \delta\mathbf{G})(\mathbf{I} + (\delta\mathbf{G})^2 + (\delta\mathbf{G})^4 + \cdots) = (\mathbf{I} + \delta\mathbf{G})(\mathbf{I} - (\delta\mathbf{G})^2)^{-1}$. Observe that $(\delta\mathbf{G})^2 \geq 0$ and that $\rho((\delta\mathbf{G})^2) = \rho(\delta\mathbf{G})^2 < 1$ and consequently that $(\mathbf{I} - (\delta\mathbf{G})^2)^{-1} \geq 0$ since it is the inverse of an M-matrix. $\mathbf{1}^T(\mathbf{I} + \delta\mathbf{G}) > 0$ is guaranteed by the assumption on in-degree and the result then follows. □

Given Lemmas 6 and 7 the proof follows nearly identical steps to Theorem 2. The only modification when $\delta < 0$ is to show that $\det(\widetilde{\mathbf{M}}) > 0$ and therefore $\delta\mathbf{p}^T\mathbf{M}\mathbf{e}_i < 1$, which is now done by applying the argument used in proving Lemma 6. □

Proof of Proposition 3. Applying the Sherman-Morrison formula,

$$\tilde{\mathbf{x}} - \mathbf{x} = \mathbf{M}\delta\mathbf{p}\frac{x_i}{1-\delta\mathbf{e}_i^T\mathbf{M}\mathbf{p}}. \tag{14}$$

Since $\frac{1}{2}\tilde{\mathbf{x}}^T\tilde{\mathbf{x}} - \frac{1}{2}\mathbf{x}^T\mathbf{x} = \frac{1}{2}(\tilde{\mathbf{x}}-\mathbf{x})^T(\tilde{\mathbf{x}}-\mathbf{x}) + \mathbf{x}^T(\tilde{\mathbf{x}}-\mathbf{x})$ it suffices to show that $\mathbf{x}^T(\tilde{\mathbf{x}}-\mathbf{x}) \geq 0$. Again using equation (14) we can pre-multiply by $\mathbf{x}^T$ to get $\mathbf{x}^T(\tilde{\mathbf{x}}-\mathbf{x}) = \delta\boldsymbol{\xi}^T\mathbf{p}\frac{x_i}{1-\delta\mathbf{e}_i^T\mathbf{M}\mathbf{p}}$ .and the result follows provided that $1-\delta\mathbf{e}_i^T\mathbf{M}\mathbf{p} > 0$.

By using an identical argument to the proof of Theorem 2, $\det(\mathbf{I}-\delta\widetilde{\mathbf{G}}) = \det(\mathbf{I}-\delta\mathbf{G})(1-\delta\mathbf{e}_i^T\mathbf{M}\mathbf{p})$. By applying the argument used in the proof of Lemma 6 we know that both $\det(\mathbf{I}-\delta\mathbf{G}) > 0$ and $\det(\mathbf{I}-\delta\widetilde{\mathbf{G}}) > 0$, and hence that $1-\delta\mathbf{e}_i^T\mathbf{M}\mathbf{p} > 0$. □

Proof of Lemma 4. Let $c = 1$ without loss of generality. The function $f(\mathbf{G}) = \sum_i x_i = \mathbf{1}^T(\mathbf{I}-\delta\mathbf{G})^{-1}\mathbf{1}$ is strictly convex if for any $\mathbf{G} \in \Gamma$ and $\widetilde{\mathbf{G}} \in \Gamma$ where $\widetilde{\mathbf{G}} \neq \mathbf{G}$ we have that $\lambda f(\mathbf{G}) + (1-\lambda)f(\widetilde{\mathbf{G}}) > f(\lambda\mathbf{G} + (1-\lambda)\widetilde{\mathbf{G}})$ for $\lambda \in (0,1)$. We will show that the following inequality holds

$$\lambda\mathbf{1}^T\mathbf{M}\mathbf{1} + (1-\lambda)\,\mathbf{1}^T\widetilde{\mathbf{M}}\mathbf{1} > \mathbf{1}^T(\mathbf{I}-\delta\lambda\mathbf{G}-\delta(1-\lambda)\tilde{\mathbf{G}})^{-1}\mathbf{1}.$$

To do so we employ the following well known result about Schur complements (see e.g. Horn and Johnson (2013)). If a symmetric matrix $\mathbf{X}$ is partitioned as

$$\mathbf{X} = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B}^T \\ \hline \mathbf{B} & \mathbf{C} \end{array}\right),$$

then $\mathbf{X}$ is positive semidefinite if and only if $\mathbf{C}$ is positive definite and its Schur complement $\mathbf{A}-\mathbf{B}^T\mathbf{C}^{-1}\mathbf{B}$ is positive semidefinite. Now let $\mathbf{A} = \lambda\mathbf{v}^T\mathbf{M}\mathbf{v}$, $\mathbf{B} = \mathbf{v}$, $\mathbf{C} = \lambda(\mathbf{I}-\delta\mathbf{G})$ for $\mathbf{v} \in \mathbb{R}^n$ where $\mathbf{v} \neq 0$. Observe that $\mathbf{X}$ is positive semidefinite.

Let a matrix $\mathbf{X}'$ be partitioned similarly, with $\mathbf{A}' = (1-\lambda)\mathbf{v}^T\widetilde{\mathbf{M}}\mathbf{v}$, $\mathbf{B}' = \mathbf{v}$, $\mathbf{C}' = (1-\lambda)(\mathbf{I}-\delta\widetilde{\mathbf{G}})$ and again observe that $\mathbf{X}'$ is positive semidefinite. Therefore $\mathbf{Y} = \mathbf{X}+\mathbf{X}'$ must be positive semidefinite, from which we conclude that

$$\mathbf{w}^T(\mathbf{A}+\mathbf{A}'-4\mathbf{B}^T(\mathbf{C}+\mathbf{C}')^{-1}\mathbf{B})\mathbf{w} \geq 0$$

for any $\mathbf{w} \in \mathbb{R}^n$ where $\mathbf{w} \neq 0$. Since $\mathbf{A} + \mathbf{A}'$ and $4\mathbf{B}^T(\mathbf{C} + \mathbf{C}')^{-1}\mathbf{B}$ are scalars

$$\lambda \mathbf{v}^T \mathbf{M} \mathbf{v} + (1-\lambda)\, \mathbf{v}^T \widetilde{\mathbf{M}} \mathbf{v} \geq 4[\mathbf{v}^T(\lambda(\mathbf{I} - \delta \mathbf{G}) + (1-\lambda)(\mathbf{I} - \delta \widetilde{\mathbf{G}}))^{-1}\mathbf{v}].$$

The right hand side of this inequality is positive as $\mathbf{C} + \mathbf{C}'$ is the convex combination of two positive definite matrices. Now setting $\mathbf{v} = \mathbf{1}$ we can conclude that $\lambda \mathbf{1}^T \mathbf{M} \mathbf{1} + (1-\lambda)\, \mathbf{1}^T \widetilde{\mathbf{M}} \mathbf{1} > \mathbf{1}^T(\mathbf{I} - \delta\lambda \mathbf{G} - \delta(1-\lambda)\tilde{\mathbf{G}})^{-1}\mathbf{1}$ for $\lambda \in (0,1)$. □

Proof of Proposition 5. If $\widetilde{\mathbf{G}}$ row dominates $\mathbf{G}$ then $\widetilde{\mathbf{G}}\mathbf{U} \geq \mathbf{G}\mathbf{U}$. We then see that $\mathbf{P}\mathbf{U} \geq 0$ and therefore $\delta \mathbf{P}\mathbf{x} = \delta \mathbf{P}\mathbf{U}\mathbf{U}^{-1}\mathbf{x} \geq 0$. The result follows because $\tilde{\mathbf{x}} - \mathbf{x} = \widetilde{\mathbf{M}}\delta \mathbf{P}\mathbf{x}$ where $\widetilde{\mathbf{M}} \geq 0$. □

Proof of Proposition 6. The proof proceeds by showing that (i) $\Longrightarrow$ (iii) $\Longrightarrow$ (ii) $\Longrightarrow$ (i).

(i) $\Longrightarrow$ (iii): We will prove that the existence of a positive rewiring implies that $\mathbf{g}$ is not a nested split graph. To do so we use two results stated in Mahadev and Peled (1995) where they refer to nested split graphs as threshold graphs.

The degree sequence of a graph is the vector $\mathbf{d} = (d_1, \ldots, d_n)$. Every graph with a degree sequence equal to $\mathbf{d}$ is called a *realization* of $\mathbf{d}$. A degree sequence $\mathbf{d}$ is *unigraphic* if all of its realizations are isomorphic.

Theorem 5. *Let $\mathbf{d} = (d_1, \ldots, d_n)$ be the degree sequence of graph $\mathbf{g}$. Then $\mathbf{g}$ is a threshold graph if and only if $\mathbf{d}$ is unigraphic.*

Proof. See Theorem 3.2.1 in Mahadev and Peled (1995). □

Theorem 5 states that a degree sequence represents a unique graph (up to a relabeling of nodes) if and only if $\mathbf{g}$ is a nested split graph.

Theorem 6. *Let $\mathbf{d} = (d_1, \ldots, d_n)$ be the degree sequence of graph $\mathbf{g}$. Then $\mathbf{g}$ is a threshold graph if and only if there does not exist a graph $\mathbf{g}'$ with a degree sequence $\mathbf{d}'$ which strictly majorizes $\mathbf{d}$.*

Proof. See Theorem 3.2.2 in Mahadev and Peled (1995). □

To prove that existence of a positive rewiring implies that $\mathbf{g}$ is not a nested split graph, assume to the contrary that $\mathbf{g}$ is. As shown in the text, a positive rewiring guarantees that the degree sequence of $\tilde{\mathbf{g}}$ weakly dominates the degree sequence of $\mathbf{g}$. By Theorem 5, $\tilde{\mathbf{g}}$ cannot have the same degree sequence as $\mathbf{g}$ so it must strictly dominate it, i.e. $\sum_{i=1}^{k} d_i(\tilde{\mathbf{g}}) \geq \sum_{i=1}^{k} d_i(\mathbf{g})$ holds strictly for some $k \in \{1, \ldots, n\}$.

Hence the existence of a positive rewiring implies that $\mathbf{g}$ is not a nested split graph since its degree sequence is strictly dominated (and therefore strictly majorized) by the degree sequence of another graph, contradicting Theorem 6.

(iii) $\implies$ (ii): No possible positive rewiring implies that if $i < j$ then $g_{ij} = 1 \implies g_{k\ell} = 1$ for $k < \ell \leq j$ and $k \leq i$, otherwise we could switch the entries g_{ij} and $g_{k\ell}$ to generate a positive rewiring. This is precisely the definition of a stepwise matrix.

(ii) $\implies$ (i): If $\mathbf{G}$ is stepwise when $x_1 \geq \cdots \geq x_n$ then also $d_1 \geq \cdots \geq d_n$ because all edges have been shifted leftwards in the adjacency matrix. For all entries strictly above the main diagonal, $g_{ij} = 1 \implies g_{k\ell} = 1$ if $d_k \geq d_i$ and $d_\ell \geq d_j$, possibly after some permutation of nodes with identical neighborhoods. By setting $\ell = j$ then $d_k \geq d_i \implies N_k \cup \{k\} \supseteq N_i$ which is the nested neighborhood condition of an NSG. □

Proof of Theorem 4. The bounds on δ guarantee via Gershgorin's theorem that any configuration of edge weight in $\mathbf{G}$ will give a convergent $\sum_{k=0}^{\infty}(\delta\mathbf{G})^k$. We will show that for any network $\mathbf{g}$ which admits a positive rewiring there exists a network $\tilde{\mathbf{g}}$ which does not admit a positive rewiring and where $\tilde{\mathbf{x}} \succ_w \mathbf{x}$ holds strictly, implying a strict welfare improvement.

If $\mathbf{g}$ is undirected and unweighted then the result follows from Theorem 1 of Belhaj et al. (2016) via the equivalence demonstrated in Proposition 6. Consider two remaining cases:

<u>Case 1 (undirected networks)</u>

Assume that $\mathbf{g}$ is an undirected but weighted network. This case requires the following definition and lemma:

Definition 5. Let $\bar{\mathbf{g}}_i \equiv (g_{1i}, \ldots, g_{(i-1)i}, g_{(i+1)i}, \ldots, g_{ni})^T$ denote the vector of in-links for a given node i . A graph $\mathbf{g}$ is *in-dominant* if $\bar{\mathbf{g}}_i \succcurlyeq \bar{\mathbf{g}}_{i+1}$ for $i \leq n-1$ and $\bar{\mathbf{g}}_n \succcurlyeq \mathbf{0}$.

Lemma 8. *Assume that $\delta > 0$. If $\mathbf{P}$ is a positive rewiring of $\mathbf{g}$ and $\tilde{\mathbf{g}}$ is in-dominant then $\tilde{\mathbf{x}} \succ_w \mathbf{x}$.*

Proof. To proceed we first show that $\mathbf{L}\widetilde{\mathbf{M}}\mathbf{L}^{-1} \geq 0$. If we post-multiply $\delta\widetilde{\mathbf{G}}$ by $\mathbf{L}^{-1}$ we get

$$\delta\widetilde{\mathbf{G}}\mathbf{L}^{-1} = \delta \begin{pmatrix} \tilde{g}_{11} & \tilde{g}_{12} & \cdots & \tilde{g}_{1n} \\ \tilde{g}_{21} & \tilde{g}_{22} & \cdots & \tilde{g}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{g}_{n1} & \tilde{g}_{n2} & \cdots & \tilde{g}_{nn} \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & -1 & 1 \end{pmatrix}$$

$$= \delta \begin{pmatrix} \tilde{g}_{11} - \tilde{g}_{12} & \tilde{g}_{12} - \tilde{g}_{13} & \cdots & \tilde{g}_{1n} \\ \tilde{g}_{21} - \tilde{g}_{22} & \tilde{g}_{22} - \tilde{g}_{23} & \cdots & \tilde{g}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{g}_{n1} - \tilde{g}_{n2} & \tilde{g}_{n2} - \tilde{g}_{n3} & \cdots & \tilde{g}_{nn} \end{pmatrix}$$

Pre-multiplying by $\mathbf{L}$ we then get

$$\delta\mathbf{L}\widetilde{\mathbf{G}}\mathbf{L}^{-1} = \delta \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \tilde{g}_{11} - \tilde{g}_{12} & \tilde{g}_{12} - \tilde{g}_{13} & \cdots & \tilde{g}_{1n} \\ \tilde{g}_{21} - \tilde{g}_{22} & \tilde{g}_{22} - \tilde{g}_{23} & \cdots & \tilde{g}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{g}_{n1} - \tilde{g}_{n2} & \tilde{g}_{n2} - \tilde{g}_{n3} & \cdots & \tilde{g}_{nn} \end{pmatrix}$$

$$= \delta \begin{pmatrix} \tilde{g}_{11} - \tilde{g}_{12} & \tilde{g}_{12} - \tilde{g}_{13} & \cdots & \tilde{g}_{1n} \\ \tilde{g}_{11} + \tilde{g}_{21} - \tilde{g}_{12} - \tilde{g}_{22} & \tilde{g}_{12} + \tilde{g}_{22} - \tilde{g}_{13} - \tilde{g}_{23} & \cdots & \tilde{g}_{1n} + \tilde{g}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \overleftarrow{\tilde{d}}_1 - \overleftarrow{\tilde{d}}_2 & \overleftarrow{\tilde{d}}_2 - \overleftarrow{\tilde{d}}_3 & \cdots & \overleftarrow{\tilde{d}}_n \end{pmatrix}.$$

If $\tilde{g}$ is in-dominant then each off-diagonal element of $\delta(\mathbf{L}\widetilde{\mathbf{G}}\mathbf{L}^{-1})_{ij} = \delta(\sum_{k=1}^{i} \tilde{g}_{kj} - \sum_{k=1}^{i} \tilde{g}_{k(j+1)})$ is non-negative by construction.

Let $Z^{n\times n} = \{\mathbf{A} \in \mathbb{R}^{n\times n} : a_{ij} \leq 0, i \neq j\}$ and notice that $(\mathbf{I} - \mathbf{L}\delta\widetilde{\mathbf{G}}\mathbf{L}^{-1}) \in Z^{n\times n}$. We now make use of a well-known result in linear algebra due to Fiedler and Pták (1962) which states (Theorem 4.3) that a matrix $\mathbf{A} \in Z^{n\times n}$ is inverse-positive (i.e. $\mathbf{A}^{-1} \geq 0$) if and only if every real eigenvalue of $\mathbf{A}$ is positive.

Therefore, since every real eigenvalue of $\mathbf{I} - \delta\widetilde{\mathbf{G}}$ is positive by assumption and because $\mathbf{I} - \mathbf{L}\delta\widetilde{\mathbf{G}}\mathbf{L}^{-1} = \mathbf{L}(\mathbf{I} - \delta\widetilde{\mathbf{G}})\mathbf{L}^{-1}$ is a similarity transformation of $\mathbf{I} - \delta\widetilde{\mathbf{G}}$ then by using the fact that similarity transformations preserve eigenvalues (Horn and Johnson (2013), p.58) we can establish that every real eigenvalue of $\mathbf{L}(\mathbf{I} - \delta\widetilde{\mathbf{G}})\mathbf{L}^{-1}$ is positive. Thus

$$\left(\mathbf{I} - \mathbf{L}\delta\widetilde{\mathbf{G}}\mathbf{L}^{-1}\right)^{-1} = \left(\mathbf{L}\left(\mathbf{I} - \delta\widetilde{\mathbf{G}}\right)\mathbf{L}^{-1}\right)^{-1} = \mathbf{L}\widetilde{\mathbf{M}}\mathbf{L}^{-1} \geq 0.$$

Establishing $\mathbf{L}\widetilde{\mathbf{M}}\mathbf{L}^{-1} \geq 0$ implies $\mathbf{L}\widetilde{\mathbf{M}}\delta\mathbf{P}\mathbf{x} \geq 0$ when $\mathbf{P}$ is a positive rewiring. From $\widetilde{\mathbf{M}}\delta\mathbf{P}\mathbf{x} = \tilde{\mathbf{x}} - \mathbf{x}^{\downarrow}$ we see that $\mathbf{L}\tilde{\mathbf{x}} \geq \mathbf{L}\mathbf{x}^{\downarrow}$ and hence $\tilde{\mathbf{x}} \succcurlyeq \mathbf{x}^{\downarrow} \implies \tilde{\mathbf{x}}^{\downarrow} \succcurlyeq \mathbf{x}^{\downarrow} \iff \tilde{\mathbf{x}} \succ_w \mathbf{x}$. □

Proof of Case 1. Shifting all edge weight upwards (and leftwards) so that $\tilde{g}_{12} = \tilde{g}_{21} > 0$ and $\tilde{g}_{ij} = 0$ otherwise. Now $\tilde{\mathbf{g}}$ is in-dominant and since this policy is a positive rewiring $\tilde{\mathbf{x}} \succ_w \mathbf{x}$ due to 8. To prove this holds strictly (i.e. $\sum_{i=1}^{k} \tilde{x}_i^{\downarrow} \geq \sum_{i=1}^{k} x_i^{\downarrow}$ holds strictly for some $k \in \{1, \ldots, n\}$) it suffices to show that either $\tilde{x}_1 > x_1$ or $\tilde{x}_1 + \tilde{x}_2 > x_1 + x_2$. This is true because $\tilde{x}_k = 0 < x_k$ for at least one $k \in \{3, \ldots, n\}$ if $\mathbf{g} \neq \tilde{\mathbf{g}}$. Therefore either $\tilde{x}_1 > x_1$ or $\tilde{x}_1 + \tilde{x}_2 > x_1 + x_2$ otherwise $\sum_{i=1}^{k} \tilde{x}_i^{\downarrow} \geq \sum_{i=1}^{k} x_i^{\downarrow}$ for at least one $k \in \{3, \ldots, n\}$, which contradicts $\tilde{\mathbf{x}} \succ_w \mathbf{x}$.

Case 2 (directed networks)

Assume that $\mathbf{g}$ is a directed and possibly weighted network. Shifting edge weight leftwards in the adjacency matrix guarantees that $\mathbf{PU} \geq \mathbf{0}$ and $\tilde{\mathbf{x}} \geq \mathbf{x}$. Let $\mathbf{P}'$ denote the policy which shifts edge weight as far left as possible and note that $\mathbf{P}'\mathbf{U} \geq \mathbf{P}''\mathbf{U}$ for any $\mathbf{P}''$ which also shifts edge weight leftwards and consequently $\mathbf{x}' \geq \mathbf{x}''$. If the graph is directed and unweighted then it has improved on $\mathbf{g}$ and does not admit a positive rewiring because no edge weight can be shifted upwards (or leftwards).

If the graph is directed and weighted then

$$\mathbf{G}' = \begin{pmatrix} 0 & g'_{12} & 0 & 0 & \cdots & 0 \\ g'_{21} & 0 & 0 & 0 & \cdots & 0 \\ g'_{31} & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ g'_{n1} & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Now shift all edge weight in the first column upwards to g'_{21} and we see that $\tilde{\mathbf{x}} - \mathbf{x}' = \widetilde{\mathbf{M}}\delta\mathbf{P}\mathbf{x}'$ is given by

$$\begin{pmatrix} \tilde{m}_{11} & \tilde{m}_{12} & 0 & 0 & \cdots & 0 \\ \tilde{m}_{21} & \tilde{m}_{22} & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 0 \\ \sum_{k=3}^{n} g'_{k1} \\ -g'_{31} \\ \vdots \\ -g'_{n1} \end{pmatrix} \delta x'_1.$$

By noting that $\tilde{m}_{12} + \tilde{m}_{22} > 1$ we see that $\mathbf{L}(\tilde{\mathbf{x}} - \mathbf{x}') \geq 0$ and we can directly verify that $\tilde{\mathbf{x}} \succ_w \mathbf{x}$ holds strictly due to this final step. □

Appendix B - Additional Details

When costs of production are convex, optimal prices depend on the Katz-Bonacich centrality of agents in the network.

$$\mathbf{p} = \mathbf{1} - \frac{1}{2}(\mathbf{I} - \delta\mathbf{G})\left((1+c)\mathbf{I} - \delta\mathbf{G}\right)^{-1}\mathbf{1}$$

$$\mathbf{q} = \frac{1}{2}\left((1+c)\mathbf{I} - \delta\mathbf{G}\right)^{-1}\mathbf{1}$$

It is also possible to show that the monopolist's profits are given by

$$\pi = \mathbf{1}^T\mathbf{q} - \mathbf{q}^T\left((1+c)\mathbf{I} - \delta\mathbf{G}\right)\mathbf{q} = \frac{1}{2}\mathbf{1}^T\mathbf{q}$$

$$\pi = \frac{1}{4}\mathbf{1}^T \left[(1+c)\mathbf{I} - \delta\mathbf{G}\right]^{-1}\mathbf{1} = \frac{1}{4(1+c)}\mathbf{1}^T \mathbf{x}_{(1+c)\delta},$$

where $\mathbf{x}_{(1+c)\delta}$ indicates the Katz-Bonacich centrality vector for decay factor of value $\delta' = (1+c)\delta$. This is again an increasing function of the sum of Katz-Bonacich centrality scores, this time with a modified decay factor δ'.

References

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012): "The Network Origins of Aggregate Fluctuations," *Econometrica*, 80, 1977–2016.

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2016): "Networks, Shocks, and Systemic Risk," in *Oxford Handbook on the Economics of Networks*, 569–607.

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

——— (2010): "Delinquent Networks," *Journal of the European Economic Association*, 8, 34–61.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013): "The Diffusion of Microfinance," *Science*, 341.

Battaglini, M. and E. Patacchini (2018): "Influencing Connected Legislators," *Journal of Political Economy*, 126, 2277–2322.

Belhaj, M., S. Bervoets, and F. Deroïan (2016): "Efficient Networks in Games with Local Complementarities," *Theoretical Economics*, 11, 357–380.

Belhaj, M., Y. Bramoullé, and F. Deroïan (2014): "Network Games under Strategic Complementarities," *Games and Economic Behavior*, 88, 310–319.

Berman, A. and R. J. Plemmons (1994): *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, second ed.

Bloch, F. and B. Dutta (2009): "Communication Networks with Endogenous Link Strength," *Games and Economic Behavior*, 66, 39–56.

Bloch, F. and N. Quérou (2013): "Pricing in Social Networks," *Games and Economic Behavior*, 80, 243–261.

Bramoullé, Y. and R. Kranton (2007): "Public Goods in Networks," *Journal of Economic Theory*, 135, 478–494.

Bramoullé, Y., R. Kranton, and M. D'Amours (2014): "Strategic Interaction and Networks," *American Economic Review*, 104, 898–930.

Calvó-Armengol, A. and J. de Martí (2009): "Information Gathering in Organizations: Equilibrium, Welfare, and Optimal Network Structure," *Journal of the European Economic Association*, 7, 116–161.

Calvó-Armengol, A., J. de Martí, and A. Prat (2015): "Communication and influence," *Theoretical Economics*, 10, 649–690.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009): "Peer Effects and Social Networks in Education," *Review of Economic Studies*, 76, 1239–1267.

Candogan, O., K. Bimpikis, and A. Ozdaglar (2012): "Optimal Pricing in Networks with Externalities," *Operations Research*, 60, 883–905.

Conlisk, J. (1985): "Comparative Statics for Markov Chains," *Journal of Economic Dynamics and Control*, 9, 139–151.

Demange, G. (2017): "Optimal Targeting Strategies in a Network under Complementarities," *Games and Economic Behavior*, 105, 84–103.

Fiedler, M. and V. Pták (1962): "On Matrices with Non-positive Off-diagonal Elements and Positive Principal Minors," *Czechoslovak Mathematical Journal*, 12, 382–400.

Galeotti, A., B. Golub, and S. Goyal (2020): "Targeting Interventions in Networks," *Working Paper*.

Galeotti, A., S. Goyal, M. O. Jackson, F. Vega-Redondo, and L. Yariv (2010): "Network Games," *Review of Economic Studies*, 77, 218–244.

Helsley, R. W. and Y. Zenou (2014): "Social Networks and Interactions in Cities," *Journal of Economic Theory*, 150, 426–466.

Hiller, T. (2017): "Peer Effects in Endogenous Networks," *Games and Economic Behavior*, 105, 349–367.

Horn, R. A. and C. R. Johnson (2013): *Matrix Analysis*, Cambridge University Press, second ed.

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

Keilson, J. and A. Kester (1977): "Monotone Matrices and Monotone Markov Processes," *Stochastic Processes and their Applications*, 5, 231–241.

König, M. D., C. J. Tessone, and Y. Zenou (2014): "Nestedness in Networks: A Theoretical Model and some Applications," *Theoretical Economics*, 9, 695–752.

Li, X. (2020): "Designing Weighted and Directed Networks under Complementarities," *Available at SSRN: https://ssrn.com/abstract=3299331*.

Mahadev, N. and U. Peled (1995): *Threshold Graphs and Related Topics*, North Holland.

Marshall, A. W., I. Olkin, and B. C. Arnold (2011): *Inequalities: Theory of Majorization and Its Applications*, Springer, second ed.

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

Milgrom, P. and J. Roberts (1990): "Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58, 1255–1277.

Myatt, D. P. and C. Wallace (2019): "Information Acquisition and use by Networked Players," *Journal of Economic Theory*, 182, 360–401.

Olaizola, N. and F. Valenciano (2019): "Dominance of Weighted Nested Split Graph Networks in Connections Models," *International Journal of Game Theory*, 49, 75–96.

Zenou, Y. (2016): "Key Players," in *Oxford Handbook on the Economics of Networks*, 244–274.