TARGETING INTERVENTIONS IN NETWORKS

ANDREA GALEOTTI, BENJAMIN GOLUB, AND SANJEEV GOYAL

Abstract. We study the design of optimal interventions in network games, where individuals' incentives to act are affected by their network neighbors' actions. A planner shapes individuals' incentives, seeking to maximize the group's welfare. We characterize how the planner's intervention depends on the network structure. A key tool is the decomposition of any possible intervention into *principal components*, which are determined by diagonalizing the adjacency matrix of interactions. There is a close connection between the strategic structure of the game and the emphasis of the optimal intervention on various principal components: In games of strategic complements (substitutes), interventions place more weight on the top (bottom) principal components. For large budgets, optimal interventions are *simple* – targeting a single principal component.

Date Printed. March 15, 2019.
We are grateful to the co-editor, Dirk Bergemann, and four anonymous referees for helpful comments. We thank Drew Fudenberg, Eric Maskin, John Urschel, Xavier Vives, Rakesh Vohra, Alex Wolitzky, and Glen Weyl for helpful conversations. Joerg Kalbfuss, Gustavo Paez, and Eduard Talamás provided exceptional research assistance. We thank Sihua Ding, Joerg Kalbfuss, Fakhteh Saadatniaki, Alan Walsh, and Yves Zenou for detailed readings of earlier drafts. Andrea Galeotti is grateful to the European Research Council for support through the ERC-consolidator grant (award no. 724356) and to the European University Institute for support through the Internal Research Grant. Galeotti: Department of Economics, London Business School, agaleotti@london.edu. Golub: Department of Economics, Harvard University and University of Pennsylvania, bgolub@fas.harvard.edu; Goyal: Faculty of Economics and Christ's College, University of Cambridge, sg472@cam.ac.uk.

1. Introduction

We consider an environment where individuals' actions have spillover effects, which can be strategic (altering the incentives of others to act) or non-strategic (taking the form of pure externalities). A utilitarian planner with limited resources can intervene to change individuals' incentives for taking the action. Our goal is to understand how the planner can use her knowledge of the spillovers to best target the intervention.

We now lay out the main elements of the model.

Individuals play a simultaneous-move game in which everyone chooses an action. This action confers *standalone* benefits on the individual, independent of anyone else's action, but it also creates spillovers. The intensity of these spillovers is described by a network, with the strength of a link between two individuals reflecting how strongly the action of one affects the marginal benefits experienced by the other. The effects may take the form of strategic complements or strategic substitutes. In addition, there may be positive and negative externalities imposed by network neighbors on each other, separate from strategic effects. This framework encompasses a number of well-known economic examples from the literature: spillovers in educational/criminal effort (Ballester, Calvó-Armengol, and Zenou, 2006), research collaboration among firms (Goyal and Moraga-Gonzalez, 2001), local public goods (Bramoulle and Kranton, 2007), investment games and beauty contests (Angeletos and Pavan, 2007; Morris and Shin, 2002), and peer effects in smoking (Jackson et al., 2017).

Before the individuals play this simultaneous-move game, the planner can target some individuals and alter their standalone marginal benefits. The cost of the intervention is increasing in the magnitude of the change and is separable across individuals. The planner seeks to maximize the utilitarian welfare under equilibrium play of the game, subject to a budget constraint. Our results characterize the optimal intervention policy.

A key observation is that an intervention on one individual potentially has direct and indirect effects on the incentives of others. These effects are mediated by the network in a way that depends on the nature of the game. For example, suppose the planner targets a given individual and increases his standalone marginal benefits to effort: this induces more effort by the targeted individual. If actions are strategic complements, this will push up the incentives of the targeted individual's neighbors. That will in turn increase the efforts of the neighbors of these neighbors, and so forth. In contrast, under strategic substitutes an intervention that encourages an individual to exert more effort will discourage the individual's neighbors. This in turn may well push up the efforts of the neighbors of these neighbors. At the heart of our approach is a particular way to organize these spillover effects in terms of the *principal components* of the matrix of interactions. We now spell out our approach and summarize our main results.

An intervention is a change in the vector of standalone (marginal) returns from individual actions. A crucial step is to express this vector in a new basis in which the strategic spillovers are simple to analyze. If interactions are symmetric,[1] there is an orthonormal basis for the space of all possible interventions consisting of eigenvectors of the adjacency matrix of the network. (This basis comes from diagonalizing the adjacency matrix.) These can be viewed as *principal components* for interventions.[2] Consider now an intervention that changes incentives in proportion to one of these components. It turns out that the change this intervention causes in equilibrium actions is confined to the same component; in other words, it is proportional to the original intervention. Moreover, its magnitude is equal to the product of the magnitude of the intervention and an amplification factor characteristic of that principal component. Thus the principal component approach gives us a set of orthogonal basis vectors in which the effect of interventions is simple to describe.

Our main result, Theorem 1, builds on these observations to characterize the optimal intervention in terms of how *similar* it is to various principal components (in the sense of cosine similarity, a standard notion of how close two vectors are in terms of their directions). This characterization identifies the role of the four primitives of the model: (i) the initial standalone marginal returns to individuals; (ii) the matrix of interaction between individuals; (iii) the nature of the strategic interaction; (iv) the budget of the planner. Equipped with this characterization, we examine more closely the properties of the optimal intervention.

Corollary 1 shows that the relationship between the optimal intervention and the principal components is decisively shaped by the nature of the strategic interaction. We order the principal components of the network (recall that these are eigenvectors of the network) by their associated eigenvalues (from high to low). In games of strategic complements, the optimal intervention is most similar to the first principal component, which is the familiar *eigenvector centrality*, and progressively less similar as we move down the principal components. In games of strategic substitutes, by contrast, the optimal intervention is most similar to the *last* principal component. The "top" principal components capture the more global structure of the network, which is important for taking advantage of strategic complementarities. The "later" or "bottom" principal components capture the local structure of the network: they help the planner to target the intervention so that it does not cause crowding out between adjacent neighbors, which is an important concern when actions are strategic substitutes.

We next examine the circumstances under which the optimal intervention is *simple* in the sense that the relative emphasis of the intervention on different individuals depends only on their network positions, and not on the details of the initial profile of incentives or the exact size of the budget. Propositions 1 and 2 show that for large enough budgets the optimal intervention is (approximately) simple in this sense. Moreover, the network structure

[1] That is, the strategic effect of individual i on individual j is the same as the effect of individual j on i.
[2] For an illustration of different principal components in a simple circle network, see Figure 1.

determines how large the budget must be for optimal interventions to be simple. In games with strategic complements, the difference between the largest and second-largest eigenvalues of the network is important: when that difference is large, even at moderate budgets the intervention is invariant to all primitives, except the network structure. In the case of strategic substitutes, the relevant spectral statistic is the difference between the smallest and second-smallest eigenvalues. We develop examples that highlight these connections between optimal interventions and spectral graph theory.

To illustrate how our approach applies beyond our benchmark intervention problem, we propose and solve two related problems. The first is in a setting where the planner does not know the standalone marginal returns to individuals but knows their distribution. In this setting, the intervention will target the mean and the variance–covariance matrix of the standalone marginal returns. We demonstrate that our approach can be used to characterize the optimal intervention in this setting of incomplete information, and that the key insights about the ordering of principal components extend. The second extension takes up a setting in which a planner provides monetary payments to individuals that alter their incentives to choose different courses of action. We show that this optimal intervention problem has the same mathematical structure as the one we study in our basic model. Going beyond symmetric adjacency matrices, our approach to finding optimal interventions builds on the *singular value decomposition* (SVD) of the adjacency matrix. The SVD is the appropriate generalization in our setting of the diagonalization, allowing us to describe the relationship between equilibrium actions and the aggregate equilibrium utility using suitable orthogonal decompositions.

We now place the paper in the context of the literature. At a basic level, the intervention problem concerns optimal policy in the presence of externalities. In this sense, it takes up a theme central to public economics. Research over the past two decades has deepened our understanding of the empirical structure of networks and the theory of the role networks play in strategic behavior (see, for example, Goyal, Moraga, and van der Leij (2006), Ballester, Calvó-Armengol, and Zenou (2006), Bramoullé, Kranton, and d'Amours (2014), and Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv (2010)). Our paper builds on this body of research to contribute to the study of optimal policy design.

Network interventions are currently an active subject not only in economics but also in related disciplines such as computer science, sociology, and public health. For a general introduction to the subject, see Rogers (1983), Kempe, Kleinberg, and Tardos (2003), Borgatti (2006), and Valente (2012). Within economics, a prominent early contribution is Ballester, Calvó-Armengol, and Zenou (2006); recent contributions include Banerjee, Chandrasekhar, Duflo, and Jackson (2013), Belhaj and Deroian (2017), Bloch and Querou (2013), Candogan, Bimpikis, and Ozdaglar (2012), Demange (2017), Fainmesser and Galeotti (2017), Galeotti and Goyal (2009), Galeotti and Rogers (2013), Leduc, Jackson, and Johari (2017), and

Akbarpour, Malladi, and Saberi (2017). Our paper is most closely related to the strand of work in which individuals engage in strategic interaction; for a survey of this work, see Zenou (2016). The main contribution of this paper lies in (i) using the principal components approach to decompose the effect of an intervention on social welfare and (ii) using the structure afforded by this decomposition to characterize optimal interventions. Specifically, we show that there is a mapping between the strategic structure (complements or substitutes) and the appropriate principal component to target.[3]

The rest of the paper is organized as follows. Section 2 presents the optimal intervention problem. Section 3 sets out notation and basic facts about the diagonalization of the adjacency matrix of interactions into principal components and presents its application to the network game. In Section 4 we characterize optimal interventions and study their properties. Section 5 discusses how our approach and methods can also be applied when we relax assumptions relating to the nature of externalities, the adjacency matrix, and the costs of intervention. Section 6 takes up intervention in games where the planner has incomplete information about the standalone marginal returns to the individuals. Section 7 relates our work to existing literature on other network measures. Section 8 concludes. Appendix A contains the proofs of all the main results of the paper. The Online Appendix takes up extensions and related technical issues.

2. The model

We consider a simultaneous-move game among individuals in the set $\mathcal{N} = \{1, \ldots, n\}$ with $n \geq 2$. Individual i chooses an action, $a_i \in \mathbb{R}$. The vector of actions is denoted by $\boldsymbol{a} \in \mathbb{R}^n$. The payoff to individual i depends on this vector, $\boldsymbol{a}$, the *network* with adjacency matrix $\boldsymbol{G}$, and other parameters, described below:

$$U_i(\boldsymbol{a}, \boldsymbol{G}) = \underbrace{a_i \left(b_i + \beta \sum_j g_{ij} a_j \right)}_{\text{returns from own action}} - \underbrace{\frac{1}{2} a_i^2}_{\substack{\text{private costs} \\ \text{of own action}}} + \underbrace{P_i(\boldsymbol{a}_{-i}, \boldsymbol{G}, \boldsymbol{b})}_{\text{pure externalities}}. \tag{1}$$

The private marginal returns from increasing the action a_i depend both on i's own action, a_i, and on others' actions. The coefficient $b_i \in \mathbb{R}$ corresponds to the part of i's marginal return that is independent of others' actions, and is thus called i's *standalone marginal return*. The contribution of others' actions to i's marginal return is given by the term $\beta \sum_j g_{ij} a_j$. Here $g_{ij} \geq 0$ is a measure of the strength of the interaction between i and j. The parameter β captures the overall magnitude and sign of strategic interdependencies. If $\beta > 0$, then actions are strategic complements; if $\beta < 0$, then actions are strategic substitutes. The function $P_i(\boldsymbol{a}_{-i}, \boldsymbol{G}, \boldsymbol{b})$ captures *pure externalities*, that is, spillovers that do not affect the

[3]Section 7 presents a discussion of the relationship between principal components and other network measures in the literature.

best response. The first-order condition for individual i's action to be optimal is:

$$a_i = b_i + \beta \sum g_{ij} a_j.$$

Any Nash equilibrium action profile $\boldsymbol{a}^*$ of this game satisfies

$$[\boldsymbol{I} - \beta \boldsymbol{G}]\boldsymbol{a}^* = \boldsymbol{b}. \tag{2}$$

If the matrix is invertible, the unique Nash equilibrium of the game can be characterized by

$$\boldsymbol{a}^* = [\boldsymbol{I} - \beta \boldsymbol{G}]^{-1}\boldsymbol{b}. \tag{3}$$

We now make two assumptions about the network and the strength of strategic spillovers.

Assumption 1. The adjacency matrix $\boldsymbol{G}$ is symmetric.

We extend our analysis to more general $\boldsymbol{G}$ in Section 5.2.

For our next assumption, recall that the spectral radius of a matrix is the maximum of its eigenvalues' absolute values.

Assumption 2. The spectral radius of $\beta \boldsymbol{G}$ is less than 1,[4] and all eigenvalues of $\boldsymbol{G}$ are distinct (the latter condition holds generically).

Assumption 2 ensures the existence of the inverse in (3), and also the uniqueness and stability of the Nash equilibrium; see Ballester et al. (2006) and Bramoullé et al. (2014) for detailed discussions of this assumption and the interpretation of the solution given by (3). It also makes a technical distinctness assumption on the eigenvalues that sharpens some statements but is not essential to the method.

The vector of equilibrium actions is denoted by $\boldsymbol{a}^*$. The utilitarian social welfare at equilibrium is given by the sum of the equilibrium utilities:

$$W(\boldsymbol{b}, \boldsymbol{G}) = \sum_i U_i(\boldsymbol{a}^*, \boldsymbol{G}).$$

We now introduce the planner. The planner wishes to maximize the aggregate equilibrium utility and can modify, at a cost, the incentive of each individual by changing the standalone marginal returns to the individuals from the *status quo*, $\hat{\boldsymbol{b}}$, to new values, $\boldsymbol{b}$. The timing is as follows. The planner moves first and chooses her intervention, and then individuals simultaneously choose actions. The incentive-targeting (IT) problem is given by

$$\begin{aligned} &\max_{\boldsymbol{b}} W(\boldsymbol{b}, \boldsymbol{G}) \qquad \text{(IT)} \\ &\text{s.t.: } \boldsymbol{a}^* = [\boldsymbol{I} - \beta \boldsymbol{G}]^{-1}\boldsymbol{b}, \\ K(\boldsymbol{b}, \hat{\boldsymbol{b}}) = &\sum_{i \in \mathcal{N}} \left(b_i - \hat{b}_i\right)^2 \leq C, \end{aligned}$$

[4]An equivalent condition is for β to be less than the reciprocal of the spectral radius of $\boldsymbol{G}$.

where C is a given budget. The marginal costs of altering the b_i are separable across individuals, and increasing in the magnitude of the change for each individual. For discussions and extensions on more general cost functions, see Sections 5.3 and 5.4.

We next present three economic applications to illustrate the scope of our model.

Example 1 (The investment game)**.** Individual i makes an investment a_i at a cost $\frac{1}{2}a_i^2$. The private marginal return on that investment is $b_i + \beta \sum_j g_{ij} a_j$, where b_i is individual i's standalone marginal return and $\sum_j g_{ij} a_j$ is the aggregate local effort. The utility of i is

$$U_i(\boldsymbol{a}, \boldsymbol{G}) = a_i \left(b_i + \beta \sum_j g_{ij} a_j \right) - \frac{1}{2} a_i^2.$$

The case with $\beta > 0$ is the canonical case of investment complementarities as in Ballester et al. (2006). Here, an individual's marginal returns are enhanced when his neighbors work harder; this creates both strategic complementarities and positive externalities. The case of $\beta < 0$ corresponds to strategic substitutes and negative externalities; this can be microfounded via a model of competition in a market after the investment decisions have been made, as in Goyal and Moraga-Gonzalez (2001).

It can be verified that the Nash equilibrium action $\boldsymbol{a}^*$ satisfies condition (3), and that the equilibrium utilities, $U_i(\boldsymbol{a}^*, \boldsymbol{G})$, and the utilitarian social welfare at equilibrium, $W(\boldsymbol{b}, \boldsymbol{G})$, are as follows:

$$U_i(\boldsymbol{a}^*, \boldsymbol{G}) = \frac{1}{2}(a_i^*)^2 \text{ and } W(\boldsymbol{b}, \boldsymbol{G}) = \frac{1}{2} (\boldsymbol{a}^*)^{\mathsf{T}} \boldsymbol{a}^*.$$

Example 2 (Investment game with coordination concerns)**.** This example is inspired by Morris and Shin (2002) and Angeletos and Pavan (2007). Individuals trade off the returns from effort against the costs, as in the first example, but also care about coordinating with others. These considerations are captured in the following payoff:

$$U_i(\boldsymbol{a}, \boldsymbol{G}) = a_i \left(\tilde{b}_i + \tilde{\beta} \sum_j g_{ij} a_j \right) - \frac{1}{2} a_i^2 - \frac{\gamma}{2} \sum_j g_{ij} [a_j - a_i]^2,$$

where we assume that $\tilde{\beta} > 0$ and $\gamma > 0$ and that $\sum_j g_{ij} = 1$ for all i, so the total interaction is the same for each individual. This formulation also relates to the theory of teams and organizational economics (see, for example, Dessein et al. (2016), Marschak and Radner (1972), and Calvó-Armengol et al. (2015)). We may interpret individuals as managers in different divisions within an organization. Each manager selects the action that maximizes the private returns for the division, but the manager also cares about coordinating with other divisions' actions.[5] This is a game of strategic complements; moreover, an increase in j's

[5]In the Online Appendix B.2 we study optimal interventions in a standard (local) beauty contest game in which $U_i(\boldsymbol{a}, \boldsymbol{G}) = -(a_i - \tilde{b}_i)^2 - \gamma \sum_j g_{ij} [a_j - a_i]^2$. Here, we focus on a modification of the standard beauty contest game that makes the mapping to our formulation easier to present.

action has a positive effect on individual i's utility if and only if $a_j < a_i$. It can be verified that the first-order condition for individual i is given by

$$a_i = \frac{\tilde{b}_i}{1+\gamma} + \frac{\tilde{b}_i + \gamma}{1+\gamma} \sum g_{ij} a_j.$$

By defining $\beta = \frac{\tilde{\beta}+\gamma}{1+\gamma}$ and $\boldsymbol{b} = \frac{1}{1+\gamma}\tilde{\boldsymbol{b}}$, we obtain a best-response structure exactly as in condition (2). Moreover, the aggregate equilibrium utility is $W(\boldsymbol{b}, \boldsymbol{g}) = \frac{1}{2}\left(\boldsymbol{a}^*\right)^\mathsf{T} \boldsymbol{a}^*$.

Example 3 (Local public good)**.** Following Bramoulle and Kranton (2007), Galeotti and Goyal (2010), and Allouch (2015, 2017), we consider a local public goods problem—for instance, collecting non-rival information. Without information-acquisition costs, the optimal amount of information to acquire would be τ.[6] Individual i has an amount $\tilde{b}_i < \tau$ of information to begin with. He can expend effort to personally acquire additional information, increasing his amount of information to $\tilde{b}_i + a_i$. If his neighbors acquire information, then he can also access $\tilde{\beta} \sum_j g_{ij} a_j$, with $\tilde{\beta} \in (0, 1]$ capturing a loss in the transmission of information. The total information that individual i has is

$$x_i = \tilde{b}_i + a_i + \tilde{\beta} \sum_j g_{ij} a_j.$$

The utility of i is

$$U_i(\boldsymbol{a}, \boldsymbol{G}) = -\frac{1}{2}(\tau - x_i)^2 - \frac{1}{2} a_i^2.$$

This is a game of strategic substitutes and positive externalities. Setting $\beta = -\tilde{\beta}/2$ and $b_i = [\tau - \tilde{b}_i]/2$ yields a best-response structure exactly as in condition (2). The aggregate equilibrium utility is $W(\boldsymbol{b}, \boldsymbol{G}) = -\left(\boldsymbol{a}^*\right)^\mathsf{T} \boldsymbol{a}^*$.

These three canonical examples of network games all have the technically convenient property that the aggregate equilibrium utility is proportional to the sum of the squares of the equilibrium actions:

Property A. The aggregate equilibrium utility is proportional to the sum of the squares of the equilibrium actions, that is, $W(\boldsymbol{b}, \boldsymbol{G}) = w\left(\boldsymbol{a}^*\right)^\mathsf{T} \boldsymbol{a}^*$ for some $w \in \mathbb{R}$, where $\boldsymbol{a}^*$ is the Nash equilibrium of the network game.

While this property is convenient for exposition, it is not essential. Section 5 presents an example where this assumption is violated, and discusses how our results can be extended to such settings.

[6] This can be taken to be the maximum amount of information available; equilibrium acquisitions will always be less than this.

3. Principal components

This section introduces basic concepts that we use to characterize optimal interventions. A key aspect of our analysis is the use of a convenient basis for the space of standalone marginal returns and actions. The basis is obtained by diagonalizing the interaction (adjacency) matrix $\boldsymbol{G}$. The advantage of this basis is that strategic effects take a very simple form; moreover, the objective function of the planner remains simple. To diagonalize $\boldsymbol{G}$, we rely on the assumption that $\boldsymbol{G}$ is symmetric, that is, Assumption 1. We generalize the analysis to arbitrary $\boldsymbol{G}$ in Section 5.2.

3.1. **Principal components: notation and definitions.** The following statement introduces our notation for the diagonalization of a matrix (Meyer, 2000).

Fact 1. If $\boldsymbol{G}$ satisfies Assumption 1, then $\boldsymbol{G} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathsf{T}}$, with the right-hand side satisfying the following conditions:

1. $\boldsymbol{\Lambda}$ is an $n \times n$ diagonal matrix whose diagonal entries $\Lambda_{ll} = \lambda_l$ are the eigenvalues of $\boldsymbol{G}$ (which are real numbers), ordered from greatest to least: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.
2. $\boldsymbol{U}$ is an orthogonal matrix. The ℓ^{th} column of $\boldsymbol{U}$, which we call $\boldsymbol{u}^\ell$, is a real eigenvector of $\boldsymbol{G}$, namely the eigenvector associated to the eigenvalue λ_ℓ, which is normalized in the Euclidean norm: $\|\boldsymbol{u}^\ell\| = 1$.

For generic $\boldsymbol{G}$, the decomposition is uniquely determined, except that any column of $\boldsymbol{U}$ is determined only up to multiplication by -1.

An important interpretation of this diagonalization is as a decomposition into *principal components*. We can think of the columns of $\boldsymbol{G}$ as n data points. The first principal component of $\boldsymbol{G}$ is defined as the n-dimensional vector that minimizes the sum of squares of the distances to the columns of $\boldsymbol{G}$. The first principal component can therefore be thought of as a fictitious column that "best summarizes" the dataset of all columns of $\boldsymbol{G}$. To characterize the next principal component, we orthogonally project all columns of $\boldsymbol{G}$ off this vector and repeat this procedure for the new columns. We continue in this way, projecting orthogonally off the (subspace generated by) vectors obtained to date, to find the next principal component. A well-known result is that the eigenvectors of $\boldsymbol{G}$ that diagonalize the matrix (i.e., the columns of $\boldsymbol{U}$) are indeed the principal components of $\boldsymbol{G}$ in this sense. Moreover, the eigenvalue corresponding to a given principal component quantifies the residual variation explained by that vector. When we refer to the ℓ^{th} principal component of $\boldsymbol{G}$, we mean the ℓ^{th} eigenvector of $\boldsymbol{G}$, which we denote by $\boldsymbol{u}^\ell(\boldsymbol{G})$. In Section 5.2, we discuss how the singular value decomposition generalizes this for non-symmetric $\boldsymbol{G}$.

Figure 1 illustrates some eigenvectors/principal components of a circle network with 14 nodes and with links all having equal weight given by 1. For each principal component, the color of a node indicates the sign of the entry of that node in that principal component (the

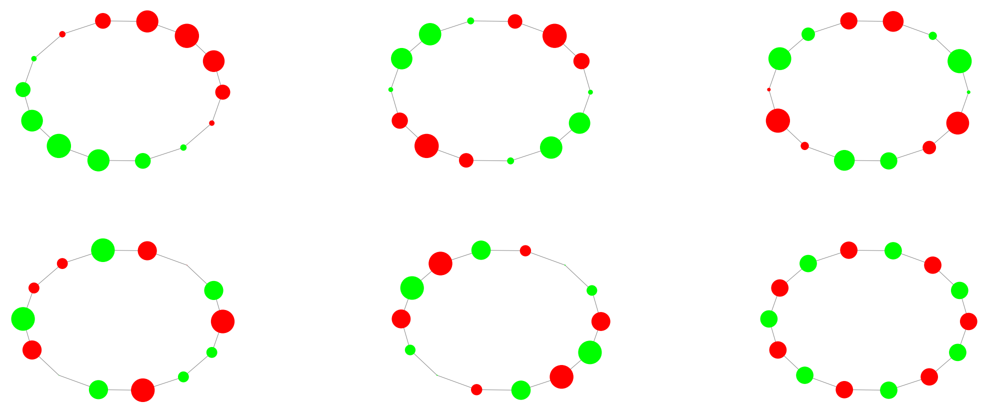


FIGURE 1. (Top) Eigenvectors 2, 4, 6. (Bottom) Eigenvectors 10, 12, 14.

color red means negative), while the size of a node indicates the absolute value of that entry. A general feature that is worth noting is that the weights for the top principal components (smaller values of ℓ) are clustered among neighboring nodes, while the weights for the bottom principal components (larger values of ℓ) tend to be negatively correlated among neighboring nodes.[7]

3.2. **Analysis of the game using principal components.** For any vector $\boldsymbol{z} \in \mathbb{R}^n$, let $\underline{\boldsymbol{z}} = \boldsymbol{U}^\mathsf{T}\boldsymbol{z}$. From now on, we will refer to $\underline{z}_\ell$ as the projection of $\boldsymbol{z}$ onto the ℓ^{th} principal component, or the magnitude of $\boldsymbol{z}$ in that component. Substituting the expression $\boldsymbol{G} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\mathsf{T}$ into equation (2), which characterizes equilibrium, we obtain

$$[\boldsymbol{I} - \beta \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\mathsf{T}]\boldsymbol{a}^* = \boldsymbol{b}.$$

Multiplying both sides of this equation by $\boldsymbol{U}^\mathsf{T}$ gives us an analogue of (3) characterizing the solution of the game:

$$[\boldsymbol{I} - \beta\boldsymbol{\Lambda}]\underline{\boldsymbol{a}}^* = \underline{\boldsymbol{b}} \qquad \Longleftrightarrow \qquad \underline{\boldsymbol{a}}^* = [\boldsymbol{I} - \beta\boldsymbol{\Lambda}]^{-1}\underline{\boldsymbol{b}}.$$

This system is diagonal, and we denote the ℓ^{th} diagonal entry of $[\boldsymbol{I} - \beta\boldsymbol{\Lambda}]^{-1}$ by $\frac{1}{1-\beta\lambda_\ell}$. Hence, for every $\ell \in \{1, 2, \ldots, n\}$,

$$\underline{a}^*_\ell = \frac{1}{1 - \beta\lambda_\ell}\underline{b}_\ell. \tag{4}$$

As stated earlier, the principal components of $\boldsymbol{G}$ constitute a basis in which strategic effects are easily described. *In each principal component, the equilibrium action is simply a scaling of the corresponding entry of the standalone marginal returns vector* $\underline{\boldsymbol{b}}$. Indeed, the equilibrium action $\underline{a}^*_\ell$ in the ℓ^{th} principal component of $\boldsymbol{G}$ is the product of an amplification factor

[7]As the network is perfectly symmetric, there is a degree of freedom with regard to the partitioning of nodes. The key element here is the pattern of clustering and negative correlation across neighboring nodes, as we move from the top to the bottom principal components.

(determined by the strategic parameter β and the eigenvalue λ_ℓ) and $\underline{b}_\ell$, which is simply the projection of $\boldsymbol{b}$ onto that principal component. Under Assumption 2, and for a generic $\boldsymbol{G}$, $1-\beta\lambda_\ell > 0$ for all ℓ (the spectral radius assumption implies that $\beta\boldsymbol{\Lambda}$ has no entries larger than 1). Moreover, when $\beta > 0$, the amplification factor is greater for principal components with greater eigenvalues, that is, it is decreasing in ℓ. When $\beta < 0$, the amplification factor is greater for principal components with lower eigenvalues, that is, it is increasing in ℓ.

Rewriting the equilibrium action in the original coordinates:

$$a_i^* = \sum_\ell \frac{1}{1-\beta\lambda_\ell} u_i^\ell \underline{b}_\ell.$$

Thus, individual i's action is proportional to the representations of i in the principal components (u_i^ℓ), the representations of the standalone marginal returns vector in the principal components ($\underline{b}_\ell$), and the magnifications from the corresponding factors, $\frac{1}{1-\beta\lambda_\ell}$.

3.3. **A notion of vector similarity.** An intervention can be described via a vector of changes to individuals' standalone marginal returns relative to the status quo. As we shall see, to describe the determinants of these changes in the optimal intervention, it is useful to define a measure that allows us to compare two vectors, in terms of similarity. A standard measure is cosine similarity.

Definition 1. The cosine similarity of two nonzero vectors $\boldsymbol{z}$ and $\boldsymbol{y}$ is

$$\rho(\boldsymbol{z},\boldsymbol{y}) = \frac{\boldsymbol{z}\cdot\boldsymbol{y}}{\|\boldsymbol{z}\|\|\boldsymbol{y}\|}$$

This is the cosine of the angle between the two vectors in the plane determined by $\boldsymbol{y}$ and $\boldsymbol{z}$. When $\rho(\boldsymbol{z},\boldsymbol{y}) = 1$, vector $\boldsymbol{z}$ is a positive scaling of $\boldsymbol{y}$. When $\rho(\boldsymbol{z},\boldsymbol{y}) = 0$, vectors $\boldsymbol{z}$ and $\boldsymbol{y}$ are orthogonal. When $\rho(\boldsymbol{z},\boldsymbol{y}) = -1$, vector $\boldsymbol{z}$ is a negative scaling of $\boldsymbol{y}$.

4. Optimal interventions

This section develops a characterization of optimal interventions and studies their properties. Recall that under Property A, the planner's payoff as a function of the equilibrium actions $\boldsymbol{a}^*$ is $W(\boldsymbol{b},\boldsymbol{G}) = w\left(\boldsymbol{a}^*\right)^\mathsf{T}\boldsymbol{a}^*$.

We begin by dealing with a straightforward case of the planner's problem. If $w < 0$, the planner wishes to minimize the sum of the squares of the equilibrium actions. In this case, when the budget is large, that is, $C \geq \|\hat{\boldsymbol{b}}\|$, the planner can allocate resources to ensure that individuals have a zero target action by setting $b_i^* = 0$ for all i. It follows from the best-response equations that all individuals choose action 0 in equilibrium, and so the planner achieves the first-best.[8] The next assumption implies that the planner's bliss point cannot be achieved, so that there is an interesting optimization problem:

[8]In the local public good application, recall Example 3, $w = -1$, and so when $C \geq \|\hat{\boldsymbol{b}}\|$, $b_i^* = 0$. Recalling our change of variables there ($b_i = [\tau - \tilde{b}_i]/2$), the optimal intervention in that case is to modify the endowment

Assumption 3. Either $w < 0$ and $C < \|\hat{\boldsymbol{b}}\|$, or $w > 0$.

Let $\boldsymbol{b}^*$ solve the incentive-targeting problem (IT), and let $\boldsymbol{y}^* = \boldsymbol{b}^* - \hat{\boldsymbol{b}}$ be the changes in individuals' standalone marginal benefits at the optimal intervention. Furthermore, let

$$\alpha_\ell = \frac{1}{(1 - \beta\lambda_\ell)^2}$$

and recall that $a^*_\ell = \sqrt{\alpha_\ell}\underline{b}_\ell$ is the equilibrium action in the ℓ^{th} principal component of $\boldsymbol{G}$ (see equation (4)).

Theorem 1. Suppose Assumptions 1–3 hold and the network game satisfies Property A. At the optimal intervention, the similarity between $\boldsymbol{y}^*$ and principal component $\boldsymbol{u}^\ell(\boldsymbol{G})$ satisfies the following proportionality:

$$\rho(\boldsymbol{y}^*, \boldsymbol{u}^\ell(\boldsymbol{G})) \quad \propto \quad \rho(\hat{\boldsymbol{b}}, \boldsymbol{u}^\ell(\boldsymbol{G}))\frac{w\alpha_\ell}{\mu - w\alpha_\ell}, \quad \ell = 1, 2, \ldots, n, \tag{5}$$

where μ, the shadow price of the planner's budget, is uniquely determined as the solution to

$$\sum_\ell \left(\frac{w\alpha_\ell}{\mu - w\alpha_\ell}\right)^2 \hat{\underline{b}}_\ell^2 = C \tag{6}$$

and satisfies $\mu > w\alpha_\ell$ for all ℓ, so that all denominators are positive.

We briefly sketch the main argument here, and interpret the quantities in the formula; the proof of this theorem and all the other results in this section are presented in the Appendix.

Define $x_\ell = (\underline{b}_\ell - \hat{\underline{b}}_\ell)/\hat{\underline{b}}_\ell$ as the relative change of $\underline{b}_\ell$ relative to $\hat{\underline{b}}_\ell$, the projection of $\hat{\boldsymbol{b}}$ onto the ℓ^{th} principal component. By rewriting the objective function of the intervention using the expression for the equilibrium action in terms of the principal components of $\boldsymbol{G}$ (expression (4)), we obtain

$$W(\boldsymbol{b}, \boldsymbol{G}) = \sum_\ell w\alpha_\ell(1 + x_\ell)^2.$$

In the same variables, the budget constraint of the planner is

$$\sum_\ell \hat{\underline{b}}_\ell^2 x_\ell^2 \leq C.$$

If the planner allocates a marginal unit of the budget to principal component ℓ, the marginal return and marginal cost are proportional, respectively, to

$$\underbrace{w\alpha_\ell(1 + x_\ell)}_{\text{marginal return}} \text{ and } \underbrace{\mu x_\ell}_{\text{marginal cost}}$$

of each individual so that everyone accesses the optimal level of the local public good without investing personally.

It follows that $\frac{w\alpha_\ell}{\mu - w\alpha_\ell}$ is exactly the value of x_ℓ at which the marginal return and the marginal cost are equalized.[9] Moreover, by the definitions of x_ℓ and cosine similarity, the optimal value satisfies:

$$\frac{w\alpha_\ell}{\mu - w\alpha_\ell} = x_\ell^* = \frac{\|\boldsymbol{y}^*\| \ \rho(\boldsymbol{y}^*, \boldsymbol{u}^\ell(\boldsymbol{G}))}{\|\hat{\boldsymbol{b}}\| \ \rho(\hat{\boldsymbol{b}}, \boldsymbol{u}^\ell(\boldsymbol{G}))}.$$

Rearranging this yields the proportionality expression (5) in the theorem. Now, up to a scaling, the similarities $\rho(\boldsymbol{y}^*, \boldsymbol{u}^\ell(\boldsymbol{G}))$ determine the optimal intervention $\boldsymbol{y}^*$ as a linear combination of the $\boldsymbol{u}^\ell(\boldsymbol{G})$. The scaling is easily computed by exhausting the budget, and this yields (6).

Thus Theorem 1 provides a full characterization of the optimal intervention. Next, we discuss the formula for the similarities given in expression (5). The similarity between $\boldsymbol{y}^*$ and $\boldsymbol{u}^\ell(\boldsymbol{G})$ measures the extent to which principal component $\boldsymbol{u}^\ell(\boldsymbol{G})$ is represented in the optimal intervention $\boldsymbol{y}^*$. Equation (5) tells us that this is proportional to two factors. The first factor, $\rho(\hat{\boldsymbol{b}}, \boldsymbol{u}^\ell(\boldsymbol{G}))$, corresponds to the similarity of the ℓ^{th} principal component with the status-quo vector $\hat{\boldsymbol{b}}$. Therefore, this factor summarizes how much the initial condition influences or biases the optimal intervention for a given budget.

The second factor, $\frac{w\alpha_\ell}{\mu - w\alpha_\ell}$, is determined by two quantities: the eigenvalue corresponding to $\boldsymbol{u}^\ell(\boldsymbol{G})$ (via $\alpha_\ell = \frac{1}{1-\beta\lambda_\ell}$), and the budget C (via the shadow price μ—recall expression (6)). To focus on this second factor $\frac{w\alpha_\ell}{\mu - w\alpha_\ell}$ we define the *similarity ratio* of $\boldsymbol{u}^\ell(\boldsymbol{G})$ to be the fraction

$$r_\ell^* = \frac{\rho(\boldsymbol{y}^*, \boldsymbol{u}^\ell(\boldsymbol{G}))}{\rho(\hat{\boldsymbol{b}}, \boldsymbol{u}^\ell(\boldsymbol{G}))}.$$

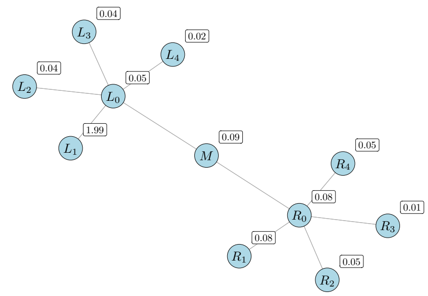
Theorem 1 shows that, as we vary ℓ, the similarity ratio r_ℓ^* is proportional to $\frac{w\alpha_\ell}{\mu - w\alpha_\ell}$. As λ_ℓ is decreasing in ℓ, it follows that the similarity ratio is greater, in absolute value, for the principal components ℓ with the greatest α_ℓ. Intuitively, those are the components where the intervention makes the largest change relative to the status quo profile of incentives. Which principal components these are is determined by the nature of the game, as follows:

Corollary 1. Suppose Assumptions 1–3 hold and the network game satisfies Property A. If the game has the strategic complements property ($\beta > 0$), then $|r_\ell^*|$ is decreasing in ℓ; if the game has the strategic substitutes property ($\beta < 0$), then $|r_\ell^*|$ is increasing in ℓ.

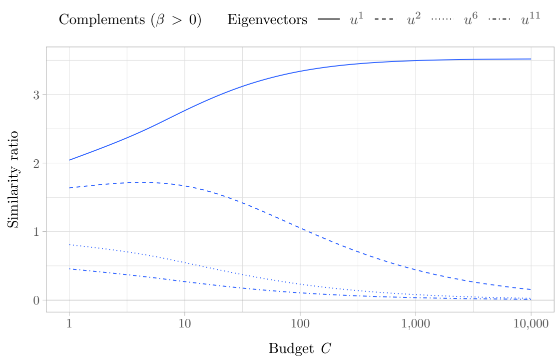
In some problems there may be a nonnegativity constraint on actions, in addition to the constraints in the statement of problem (IT). Note that as long as the status quo actions $\hat{\boldsymbol{b}}$ are positive, this constraint will be respected for all C less than some $\hat{C}$, and so our approach will give information about the relative effects on various components in this case as well.

We present numerical examples to illustrate the results in this section. They are based on payoffs taken from Example 1, and we assume undirected networks with binary links. For the

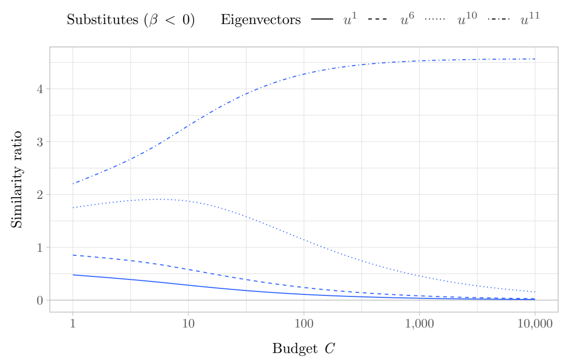
[9] It can be verified that the ratio for every $\ell \in \{1, \ldots, n-1\}$, $x_\ell / x_{\ell+1}$ is increasing (decreasing) in β for the case of strategic complements (substitutes): thus the intensity of the strategic interaction shapes the relative importance of different principal components.



(A) Simple network



(B) Similarity ratio for strategic complements



(C) Similarity ratio for strategic substitutes

FIGURE 2. Similarity ratios

case of strategic complements, we set $\beta = 0.1$, and for strategic substitutes we set $\beta = -0.1$. Assumptions 1 and 2 are satisfied and Property A holds.

To illustrate the content of the Corollary we consider a simple 11-node network containing two hubs, L_0 and R_0, that are connected by an intermediate node M; the network is depicted in Figure 2(A). The numbers next to the nodes are the initial standalone marginal returns. Figures 2(B) and 2(C) illustrate the monotonicity of the similarity ratio across network components. In the case of strategic complements, for every level of C the similarity ratio of the first principal component is higher than that of the other network components (Figure 2(B)). In the case of strategic substitutes, the monotonicity is reversed (Figure 2(C)).

The optimal intervention takes an especially simple form for the case of small and large budgets. From equation (6), we can deduce that the shadow price μ is decreasing in C. For $w > 0$, it follows then that an increase in C raises $\frac{w\alpha_\ell}{\mu - w\alpha_\ell}$ and that the principal components with larger α_ℓ become larger in relative terms as well; that is, if $w > 0$ and $\alpha_\ell > \alpha_{\ell'}$, then $r^*_\ell / r^*_{\ell'}$ is increasing in C.[10] For simplicity of exposition, we suppress the dependence of

[10]Analogously, when $w < 0$, $\frac{w\alpha_\ell}{\mu - w\alpha_\ell}$ and $r^*_\ell / r^*_{\ell'}$ are both decreasing in C.

outcomes on C in the following statement, but note that $\boldsymbol{y}^*$ and thus the r^*_ℓ are all functions of C.

Proposition 1. Suppose Assumptions 1–3 hold and the network game satisfies Property A. Then the following hold:

1. As $C \to 0$, in the optimal intervention, $\frac{r^*_\ell}{r^*_{\ell'}} \to \frac{\alpha_\ell}{\alpha_{\ell'}}$.
2. As $C \to \infty$, in the optimal intervention
   - 2a. If the game has the strategic complements property, $\beta > 0$, then the similarity of $\boldsymbol{y}^*$ and the first principal component of the network tends to 1, $\rho(\boldsymbol{y}^*, \boldsymbol{u}^1(\boldsymbol{G})) \to 1$.
   - 2b. If the game has the strategic substitutes property, $\beta < 0$, then the similarity of $\boldsymbol{y}^*$ and the last principal component of the network tends to 1, $\rho(\boldsymbol{y}^*, \boldsymbol{u}^n(\boldsymbol{G})) \to 1$.

This result can be understood by recalling equation (5) in Theorem 1. First, consider the case of small C. When the planner's budget becomes small, the shadow price μ tends to ∞.[11] Equation (5) then implies that the similarity ratio of the ℓ^{th} principal component becomes proportional to α_ℓ. Turning now to the case where C grows large, the shadow price converges to $w\alpha_1$ if $\beta > 0$, and to $w\alpha_n$ if $\beta < 0$ (by equation (6)). Plugging this into equation (5), we find that in the case of strategic complements, the optimal intervention shifts individuals' standalone marginal returns (very nearly) in proportion to the first principal component of $\boldsymbol{G}$, so that $\boldsymbol{y}^* \to \sqrt{C}\boldsymbol{u}^1(\boldsymbol{G})$. In the case of strategic substitutes, on the other hand the planner changes individuals' standalone marginal returns (very nearly) in proportion to the last principal component, namely $\boldsymbol{y}^* \to \sqrt{C}\boldsymbol{u}^n(\boldsymbol{G})$.[12]

Figure 3 presents optimal targets in the same simple network when the budget is large—in particular, for $C = 500$. The top-left figure illustrates the first eigenvector, and the top-right figure depicts optimal targets in a game with strategic complements. The bottom-left figure illustrates the last eigenvector, and the bottom-right figure depicts the optimal targets when the game has strategic substitutes. The node size represents the size of the intervention, $|b^*_i - \hat{b}_i|$; its color represents the sign of the intervention (with green signifying a positive intervention and red indicating a negative intervention).

In line with part 2 of Proposition 1 for large C, the optimal intervention is guided by the "main" component of the network (corresponding to the largest or smallest eigenvalue). Under strategic complements, this is the "first" eigenvector of the network, which corresponds to individuals' eigenvector centrality. Intuitively, by increasing the standalone marginal

[11] As costs are quadratic, small relaxation in the budget around zero can have a large impact on aggregate welfare.

[12] When individuals' initial standalone marginal returns are zero ($\hat{\boldsymbol{b}} = \boldsymbol{0}$), we can dispense with the approximations invoked for a large budget C. Assuming that $\boldsymbol{G}$ is generic, if $\hat{\boldsymbol{b}} = \boldsymbol{0}$ then, regardless of the level of C, the entire budget is spent either (i) on changing $\underline{b}_1$ (if $\beta > 0$) or (ii) on changing $\underline{b}_n$ (if $\beta < 0$). To see this, set $\hat{\underline{\boldsymbol{b}}} = \boldsymbol{0}$ in the maximization problem (IT-PC), the principal component version of (IT); note that if the allocation is not monotonic, then the effort can be reallocated profitably among the principal components without changing the cost.

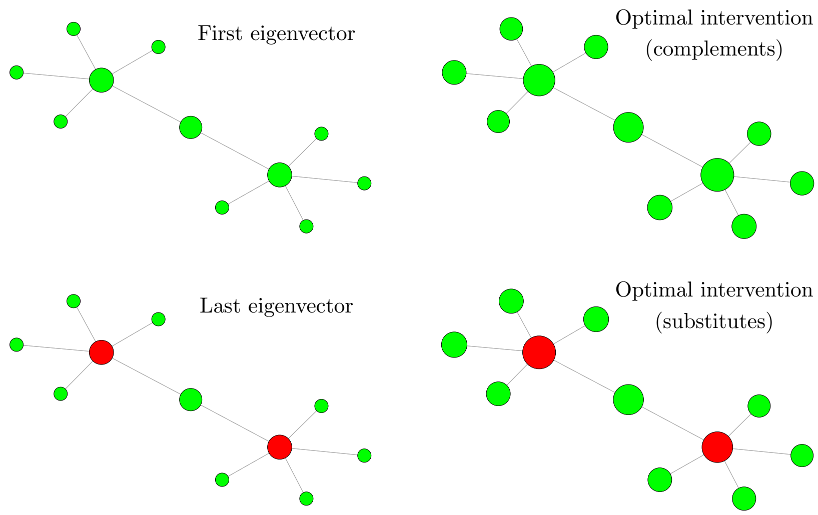


FIGURE 3. Optimal targets with large budgets

return of each individual in proportion to his eigenvector centrality, the planner targets the individuals in proportion to their global contributions to strategic feedbacks, and this is welfare maximizing.

Under strategic substitutes, optimal targeting is determined by the "last" eigenvector of the network, corresponding to its smallest eigenvalue. This network component contains information about the local structural properties of the network: it determines the way to partition the set of nodes into two sets so that most of the links are across individuals in different sets.[13] The optimal intervention increases the standalone marginal returns of all individuals in one set and decreases those of individuals in the other set. The planner wishes to target neighboring nodes asymmetrically, as this reduces possible crowding-out effects by the strategic substitutes property of individuals' best replies.

To see why this happens, it is instructive to examine the nature of best replies: an increase in b_i raises a_i, and, by the strategic substitutes property, this exerts a downward pressure on neighbor j's action, a_j. A smaller a_j in turn pushes a_i up further, and that lowers a_j even more, and so forth, until we reach a new equilibrium configuration. This process is amplified if, for adjacent nodes i and j, we simultaneously increase b_i and decrease b_j. On the other hand, if we were to raise b_i and b_j simultaneously, then the pressure toward a greater effort by both i and j would tend to cancel them against each other; that would be wasteful.

[13]The last eigenvector of a graph is useful in determining the bipartiteness of a graph and its chromatic number. Desai and Rao (1994) characterize the smallest eigenvalue of a graph and relate it to the degree of bipartiteness of a graph. Alon and Kahale (1997) demonstrate that the last eigenvector of a graph corresponds to a coloring of the underlying graph, that is, a labeling of nodes by a minimal set of integers such that no neighboring nodes share the same label.

4.1. **When are interventions simple?** We have just seen examples illustrating how, with large budgets, interventions are guided by the principal components corresponding to extreme eigenvalues. Our final result in this section gives the corresponding formal statement: For large budgets C, interventions that target players' incentives according to the two extreme principal components—*simple interventions*—are approximately optimal, that is, they generate most of the maximum achievable welfare.

Definition 2 (Simple interventions)**.** An intervention is *simple* if, for all $i \in \mathcal{N}$,

- $b_i - \hat{b}_i = \sqrt{C} u_i^1$ when the game has the strategic complements property ($\beta > 0$),
- $b_i - \hat{b}_i = \sqrt{C} u_i^n$ when the game has the strategic substitutes property ($\beta < 0$).

Let W^* be the aggregate utility under the optimal intervention, and let W^s be the aggregate utility under the simple intervention.

Proposition 2. Suppose $w > 0$, Assumptions 1 and 2 hold, and the network game satisfies Property A. Then we have the following:

1. If the game has the strategic complements property, $\beta > 0$, then for any $\epsilon > 0$, if $C > \frac{2\|\hat{\boldsymbol{b}}\|^2}{\epsilon}\left(\frac{\alpha_2}{\alpha_1 - \alpha_2}\right)^2$, then $W^*/W^s < 1 + \epsilon$ and $\rho(\boldsymbol{y}^*, \sqrt{C}\boldsymbol{u}^1) > \sqrt{1-\epsilon}$.
2. If the game has the strategic substitutes property, $\beta < 0$, then for any $\epsilon > 0$, if $C > \frac{2\|\hat{\boldsymbol{b}}\|^2}{\epsilon}\left(\frac{\alpha_{n-1}}{\alpha_n - \alpha_{n-1}}\right)^2$, then $W^*/W^s < 1 + \epsilon$ and $\rho(\boldsymbol{y}^*, \sqrt{C}\boldsymbol{u}^n) > \sqrt{1-\epsilon}$.

Proposition 2 gives a condition on the size of the budget beyond which (a) simple interventions achieve most of the optimal welfare and (b) the optimal intervention is very similar to the simple intervention. This bound depends on the status quo standalone marginal returns and the structure of the network.

- Consider the status quo benefits: Observe that the first term on the right-hand side of the inequality is proportional to $\|\hat{\boldsymbol{b}}\|$. This inequality is therefore easier to satisfy when the standalone status quo marginal returns are smaller, in the sense of having a smaller norm. The inequality is harder to satisfy when these marginal returns are large and/or heterogeneous.[14]
- The role of the network: Recall that $\alpha_\ell = (1 - \beta\lambda_\ell)^{-2}$; thus if $\beta > 0$, the term $\alpha_2/(\alpha_1 - \alpha_2)$ of the inequality is large when $\lambda_1 - \lambda_2$, the "spectral gap" of the graph, is small. Networks with a small spectral gap require large budgets before the optimal intervention approximates the first principal component. If $\beta < 0$, then the term $\alpha_{n-1}/(\alpha_{n-1} - \alpha_n)$ is small when the "bottom gap" of the graph, the difference $\lambda_{n-1} - \lambda_n$, is small. Networks with a small bottom gap require large budgets before the optimal intervention concentrates on the last principal component.

[14]Recall that $\|\frac{1}{n}\hat{\boldsymbol{b}}\|^2$ is equal to the sum of $\left(\frac{1}{n}\sum_i \hat{b}_i\right)^2$ (the squared mean of the entries of $\boldsymbol{b}$) and the sum of squared deviations of the entries of the vector $\hat{\boldsymbol{b}}$ from their mean.

Figure 4 illustrates the role of the network structure in shaping the rate (in terms of the size of the budget C) at which the optimal intervention converges to a simple intervention (Proposition 2). This example will highlight the idea that in some networks there are two different embedded sub-structures that offer a similar potential for amplifying the effect of interventions. In such networks, interventions will *not* be simple for reasonable budgets. Mathematically, this property of the network is captured by a small spectral or bottom gap.

Under strategic complements, the optimal intervention converges faster in a network that has a large spectral gap. Recall that the two largest eigenvalues can be expressed in terms of the corresponding eigenvectors as follows:

$$\lambda_1 = \max_{\boldsymbol{u}:\|\boldsymbol{u}\|=1} \sum_{ij} g_{ij} u_i u_j \qquad \lambda_2 = \max_{\substack{\boldsymbol{u}:\ \|\boldsymbol{u}\|=1 \\ \boldsymbol{u}\cdot\boldsymbol{u}^1=0}} \sum_{ij} g_{ij} u_i u_j .$$

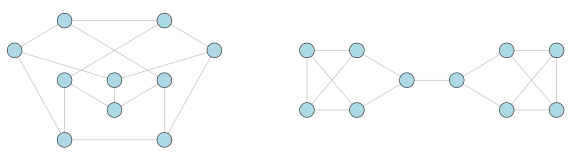
Eigenvector $\boldsymbol{u}^1 = \arg\max_{\boldsymbol{u}:\|\boldsymbol{u}\|=1} \sum_{ij} g_{ij} u_i u_j$ (corresponding to λ_1) assigns the same sign, (say) positive, to adjacent nodes in the network. Clearly, eigenvector $\boldsymbol{u}^2$ must assign negative values to some of the nodes (as it is orthogonal to $\boldsymbol{u}^1$). In the network on the right side of 4(A), by assigning positive signs to nodes in one community and negative signs to nodes in the other community, λ_2 will be almost as large as λ_1, because most of the adjacent nodes will have the same sign. This will yield a small spectral gap. In the network on the left side of Figure 4(A) this assignment is not possible; as a result, λ_2 will be much smaller than λ_1, leading to a large spectral gap. In words, the spectral gap measures the level of "cohesiveness" of the network, and it is this property that dictates fast convergence to simple interventions.[15]

Turning next to strategic substitutes, the convergence of the optimal intervention to the simple intervention is faster in networks with a larger bottom gap. Figure 4(D) illustrates the impact of the bottom gap on the rate at which the optimal intervention strategy converges to the simple intervention. Recall that the smallest two eigenvalues, λ_n and λ_{n-1}, can be written in terms of the corresponding eigenvectors as follows:

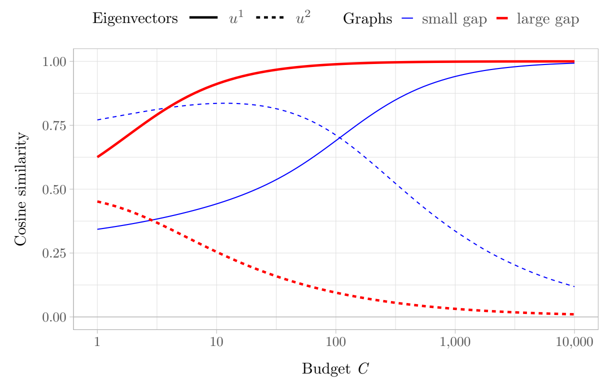
$$\lambda_n = \min_{\boldsymbol{u}:\|\boldsymbol{u}\|=1} \sum_{ij} g_{ij} u_i u_j \qquad \lambda_{n-1} = \min_{\substack{\boldsymbol{u}:\ \|\boldsymbol{u}\|=1 \\ \boldsymbol{u}\cdot\boldsymbol{u}^n=0}} \sum_{ij} g_{ij} u_i u_j . \tag{7}$$

This tells us that $|\lambda_n|$ is large when the eigenvector $\boldsymbol{u}^n = \arg\min_{\boldsymbol{u}:\|\boldsymbol{u}\|=1} \sum_{ij} g_{ij} u_i u_j$ (corresponding to λ_n) assigns opposite signs to most pairs of adjacent nodes, that is, if $g_{ij} = 1$ then $\text{sign}(u_i^n) = -\text{sign}(u_j^n)$. In other words, the last eigenvalue is large when nodes can be partitioned into two sets and most of the connections are across sets: thus $|\lambda_n|$ is maximized in a bipartite graph. The second-smallest eigenvalue of $\boldsymbol{G}$ reflects the extent to which the next-best eigenvector (orthogonal to $\boldsymbol{u}^n$) is good at solving the same minimization problem. Hence, the bottom gap of $\boldsymbol{G}$ is small when there are two orthogonal ways to partition the

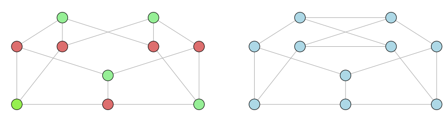
[15]See Hartfiel and Meyer (1998), Levin et al. (2009), and Golub and Jackson (2012a,b,c) for discussions and further citations to the literature on spectral gap.



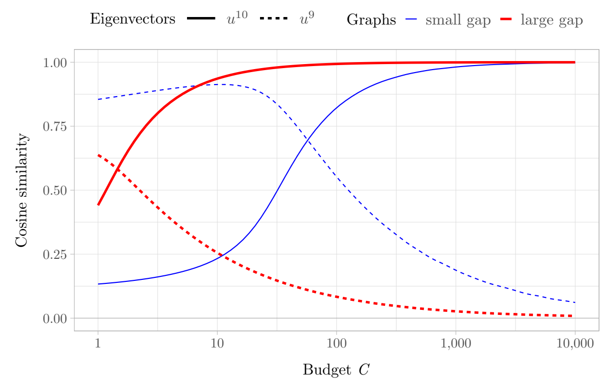
(A) Large (left) and small (right) spectral gap



(B) Cosine similarity and strategic complements



(C) Large (left) and small (right) bottom gap



(D) Cosine similarity and strategic substitutes

FIGURE 4. Spectral gap, bottom gap, and optimal interventions

network into two sets so that, either way, the "quality" of the bipartition, as measured by $\sum_{ij} g_{ij}u_iu_j$, is similar.

We illustrate this with a comparison of the two graphs in Figure 4(C). The left-hand graph is bipartite, so there is a normalized vector $\boldsymbol{u}$ for which $\sum_{ij} g_{ij}u_iu_j = -3$: set $u_i = \pm\frac{1}{\sqrt{3}}$, with the green nodes having a positive sign. This turns out to be an eigenvector $\boldsymbol{u}^n$ associated to $\lambda_n = -3$. The eigenvector $\boldsymbol{u}^{n-1}$ associated to λ_{n-1} is orthogonal to $\boldsymbol{u}^n$ and the uniform vector $\boldsymbol{u}^1$. This essentially forces half the green nodes (in a weighted sense) to get a positive sign in $\boldsymbol{u}^{n-1}$, and similarly for the red nodes.[16] This situation forces some adjacent nodes to have the same sign in $\boldsymbol{u}^{n-1}$. Hence, the second-best "bipartition" is much worse than the best (where all edges have opposite signs at their ends with extreme magnitudes) and this is reflected by the considerably less extreme value of $\lambda_{n-1} = -1.64$. For a contrast, consider the graph on the right of Figure 4(C): note that this is obtained by rewiring only two links from the left graph. This small change has real implications for targeting. First, this graph is not bipartite and so the bottom eigenvalue is not as extreme; it is only $\lambda_n = -2.62$ in this

[16]Moreover, as this $\boldsymbol{u}^{n-1}$ must have norm 1, the magnitudes of some of the entries must be considerable.

case. Second, there is another normalized $\boldsymbol{u}^{n-1}$, orthogonal to $\boldsymbol{u}^n$, achieving a similar value of $\sum_{ij} g_{ij} u_i u_j$; the second best is $\lambda_{n-1} = -2.30$.[17]

This gives some intuition for why the left-hand graph has a large bottom gap, while the right-hand graph has a small one. And we see the consequences reflected in targeting policy as we vary C in Figure 4(D): in the graph with large bottom gap, it does not take much growth in the budget for the intervention to stop putting mass on $\boldsymbol{u}^{n-1}$ eigenvector; but it takes much more growth for this to happen with small bottom gap.

We conclude by noting the influence of the initial vector of standalone marginal returns in shaping optimal interventions for small budgets. For a small budget C, the cosine similarity of the optimal intervention for non-main network components can be higher than the one for the main component. This is true when the initial vector $\hat{\boldsymbol{b}}$ is similar to some of the non-main network components. For small C, in the small spectral gap network the cosine similarity for the second principal component is larger than the one for the first principal component (see Figure 4(B)). Figure 4(D) shows that in both the small and large bottom gap networks the cosine similarity of the optimal intervention is higher for the next-to-last principal component. In all cases, the influence of the initial condition vanishes as the planner's budget becomes sufficiently large.

5. Discussion

This section extends our basic model to study settings where (a) Property A is not satisfied, (b) the matrix $\boldsymbol{G}$ is non-symmetric, (c) the exact quadratic cost specification does not hold, and (d) the interventions occur via monetary incentives for activity.

5.1. **General non-strategic externalities.** Section 4 characterizes optimal interventions for network games that satisfy Property A. We now relax this assumption. Recall that player i's utility for action profile $\boldsymbol{a}$ is

$$U_i(\boldsymbol{a}, \boldsymbol{G}) = \hat{U}_i(\boldsymbol{a}, \boldsymbol{G}) + P_i(\boldsymbol{a}_{-i}, \boldsymbol{G}, \boldsymbol{b}),$$

where $\hat{U}_i(\boldsymbol{a}, \boldsymbol{G}) = a_i(b_i + \sum_j g_{ij} a_j) - \frac{1}{2} a_i^2$.

At an equilibrium $\boldsymbol{a}^*$, it can be checked that $\sum_i \hat{U}_i(\boldsymbol{a}^*, \boldsymbol{G}) \propto (\boldsymbol{a}^*)^{\mathsf{T}} \boldsymbol{a}^*$. Therefore, a sufficient condition for Property A to be satisfied is that $\sum_i P_i(\boldsymbol{a}^*_{-i}, \boldsymbol{G}, \boldsymbol{b})$ is also proportional to $(\boldsymbol{a}^*)^{\mathsf{T}} \boldsymbol{a}^*$. Examples 1–3 satisfy this property. However, as the next example shows, there are natural environments in which it is violated.

Example 4 (Social interaction and peer effects)**.** Individual decisions on smoking and alcohol consumption are susceptible to peer effects (see Jackson et al. (2017) for references to the extensive literature on this subject). For example, an increase in smoking among an

[17] Intuitively, when we solve the minimization problem for λ_{n-1}, orthogonality to $\boldsymbol{u}^n$ does not force as many neighboring nodes to have positive $u_i u_j$ products, because $\boldsymbol{u}^n$ does not correspond to a perfect bipartition.

adolescent's friends increases her incentives to smoke and, at the same time, has negative effects on her welfare. These considerations are reflected in the following payoff function:

$$U_i(\boldsymbol{a}, \boldsymbol{G}) = \hat{U}_i(\boldsymbol{a}, \boldsymbol{G}) - \gamma \sum_{j \neq i} a_j,$$

where $\beta > 0$ and γ is positive and sufficienctly large. It can be checked that the aggregate equilibrium welfare is:

$$W(\boldsymbol{b}, \boldsymbol{G}) = \frac{1}{2} \left(\boldsymbol{a}^*\right)^{\mathsf{T}} \boldsymbol{a}^* - n\gamma \sum_i a_i^*, \tag{8}$$

with $\boldsymbol{a}^*$ given by expression (3).[18]

To extend the analysis beyond Property A, we allow the non-strategic externality term $P_i(\boldsymbol{a}_{-i}, \boldsymbol{G}, \boldsymbol{b})$ to take a form that allows for flexible externalities within the linear–quadratic family:[19]

$$P_i(\boldsymbol{a}_{-i}, \boldsymbol{G}) = m_1 \sum_j g_{ij} a_j + m_2 \sum_j g_{ij} a_j^2 + m_3 \sum_{j \neq i} a_j + m_4 \left(\sum_{j \neq i} a_j \right)^2 + m_5 \sum_{j \neq i} a_j^2.$$

We also make the following assumption on the matrix $\boldsymbol{G}$:

Assumption 4. The total interaction is constant across individuals, that is, $\sum_j g_{ij} = 1$ for all $i \in \mathcal{N}$.

Using equation (3) and Assumption 4, we can rewrite the expression for the aggregate equilibrium utility as follows:

$$W(\boldsymbol{b}, \boldsymbol{G}) = w_1 \left(\boldsymbol{a}^*\right)^{\mathsf{T}} \boldsymbol{a}^* + \frac{w_2}{n} \left(\sum_i a_i^* \right)^2 + \frac{w_3}{\sqrt{n}} \sum_i a_i^*,$$

where $w_1 = 1 + m_2 + m_5 + (n-1)m_4$, $w_2 = nm_5(n-2)$, and $w_3 = \sqrt{n}[m_1 + (n-1)m_3]$.

Observe that Property A clearly holds when $w_2 = w_3 = 0$. On the other hand, if (say) $w_1 = w_2 = 0$, then the planner's objective is to maximize the sum of the equilibrium actions, which is a fairly different type of objective.[20]

Under Assumption 4, the sum of the equilibrium actions is proportional to the sum of the standalone marginal returns. Because $\boldsymbol{u}^1$ is proportional to the all-ones vector $\mathbf{1}$, this sum in turn is equal to $\underline{b}_1$.

Together, these facts allow us to extend our earlier analysis to the case of general w_2 and w_3. First, we can still express the objective function simply in terms of the singular value

[18] In this specification the last (externality) term is a global term. We can easily accommodate local negative externalities by replacing that term with $\sum_j g_{ij} a_j$.

[19] We can also accommodate externalities that depend directly on the b_i, but we omit this for brevity.

[20] A characterization of the optimal intervention when the planner's objective is to maximize the sum of the equilibrium actions can be found in Corollary 2 of Online Appendix B.1.

decomposition; the only difference is that now $\underline{b}_1$ will enter both in a quadratic term and in a linear term. In view of this, we first solve the problem (exactly analogously to the previous solution) for a given value of $\underline{b}_1$, and then we optimize over $\underline{b}_1$. This derivation is presented in Online Appendix B.1.

5.2. **Beyond symmetric and non-negative $\boldsymbol{G}$.** In this subsection we relax the assumption that $\boldsymbol{G}$ is symmetric. Recall that equilibrium actions are determined by:

$$\boldsymbol{a}^* = [\boldsymbol{I} - \beta\boldsymbol{G}]^{-1}\boldsymbol{b}.$$

When $\boldsymbol{G}$ is not symmetric, we employ the *singular value decomposition* (SVD) of the matrix $\boldsymbol{M} = \boldsymbol{I} - \beta\boldsymbol{G}$. This allows to diagonalize the system of equilibrium actions and to obtain an orthogonal decomposition of welfare. An SVD of $\boldsymbol{M}$ is defined to be a tuple $(\boldsymbol{U}, \boldsymbol{S}, \boldsymbol{V})$ satisfying:

$$\boldsymbol{M} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^\mathsf{T}, \tag{9}$$

where:

(1) $\boldsymbol{U}$ is an orthogonal $n \times n$ matrix whose columns are eigenvectors of $\boldsymbol{M}\boldsymbol{M}^\mathsf{T}$;
(2) $\boldsymbol{V}$ is an orthogonal $n \times n$ matrix whose columns are eigenvectors of $\boldsymbol{M}^\mathsf{T}\boldsymbol{M}$;
(3) $\boldsymbol{S}$ is an $n \times n$ matrix with all off-diagonal entries equal to zero and nonnegative diagonal entries $S_{ll} = s_l$, which are called *singular values* of $\boldsymbol{M}$. As a convention, we order the singular values so that $s_\ell > s_{\ell+1}$.

It is a standard fact that an SVD exists.[21] For expositions of the SVD, see Golub and Van Loan (1996) and Horn and Johnson (2012). The ℓ^{th} left singular vector of $\boldsymbol{M}$ corresponds to the ℓ^{th} principal component of $\boldsymbol{M}$. When $\boldsymbol{G}$ is symmetric, the SVD of $\boldsymbol{M} = \boldsymbol{I} - \beta\boldsymbol{G}$ can be taken to have $\boldsymbol{U} = \boldsymbol{V}$, and the SVD basis is one in which $\boldsymbol{G}$ is diagonal.

Let $\underline{\boldsymbol{a}} = \boldsymbol{V}^\mathsf{T}\boldsymbol{a}$ and $\underline{\boldsymbol{b}} = \boldsymbol{U}^\mathsf{T}\boldsymbol{b}$; then the equilibrium condition implies that:

$$\underline{a}_\ell^* = \frac{1}{s_\ell}\underline{b}_\ell^2,$$

and therefore the objective function is:

$$W(\boldsymbol{b}, \boldsymbol{G}) = w\left(\boldsymbol{a}^*\right)^\mathsf{T}\boldsymbol{a}^* = w\underline{\boldsymbol{a}}^{*\mathsf{T}}\underline{\boldsymbol{a}}^*.$$

It is now apparent that the analysis of the optimal intervention can be carried out in the same way as in Section 4. Theorem 1 applies, with the only difference that now $\alpha_\ell = 1/s_\ell^2$. We can also extend Proposition 1 and Proposition 2. As the budget tends to 0, $r_\ell^*/r_{\ell'}^*$ tends to $\alpha_\ell/\alpha_{\ell'}$; on the other hand, when C is very large, the optimal intervention is proportional to the first principal component of $\boldsymbol{M}$, and a simple intervention that focuses on the first principal component performs (nearly) as well as the optimal intervention. When $\boldsymbol{G}$ is symmetric, the

[21] The decomposition is uniquely determined up to a permutation that (i) reorders the singular values of $\boldsymbol{M}$ and correspondingly reorders the columns of $\boldsymbol{U}$ and $\boldsymbol{V}$, and (ii) flips the sign of any column of $\boldsymbol{U}$ and $\boldsymbol{V}$.

strategic property of the game (determined by β) pins down the principal component that most amplifies an intervention. If $\boldsymbol{G}$ is non-symmetric, the singular values s_l of $\boldsymbol{M}$ are not equal to $1-\beta\lambda_l$, where λ_l are the eigenvalues of $\boldsymbol{G}$; the singular vectors of $\boldsymbol{M}$ are not the eigenvectors of $\boldsymbol{G}$; and the left and right singular vectors need not be the same.

5.3. **More general costs of intervention.** In Section 4 we solved the optimal intervention problem under a specific cost function. This section develops properties that a reasonable cost function must satisfy. We then show that our analysis of the optimal intervention extends to this general class of cost functions as long as the budget is small.

We begin by developing properties that a reasonable cost function $(\boldsymbol{b}, \hat{\boldsymbol{b}}) \mapsto K(\boldsymbol{b}; \hat{\boldsymbol{b}})$ must satisfy.

Assumption 5.

(1) *Translation-invariance:* For any $\boldsymbol{z} \in \mathbb{R}^n$, we have $K(\boldsymbol{b}+\boldsymbol{z}; \hat{\boldsymbol{b}}+\hat{\boldsymbol{z}}) = K(\boldsymbol{b}; \hat{\boldsymbol{b}})$, that is., there is a function $\kappa : \mathbb{R}^n \to \mathbb{R}$ such that $K(\boldsymbol{b}; \hat{\boldsymbol{b}}) = \kappa(\boldsymbol{b} - \hat{\boldsymbol{b}})$.
(2) *Symmetry:* For any permutation σ of $\{1, \dots, n\}$, it is true that $\kappa(y_{\sigma(1)}, y_{\sigma(2)}, \dots, y_{\sigma(n)}) = \kappa(y_1, y_2, \dots, y_n)$.
(3) *Nonnegativity:* κ is nonnegative, and $\kappa(\boldsymbol{0}) = 0$.
(4) *Local separability:* $\frac{\partial^2 \kappa(\boldsymbol{y})}{\partial y_i \partial y_j} = 0$ evaluated at $\boldsymbol{0}$.
(5) *Well-behaved second derivative at 0:* κ is twice differentiable with $\frac{\partial^2 \kappa}{\partial y_i^2}(\boldsymbol{0}) > 0$ for all i.

Translational invariance says that there is no dependence on the starting point. Symmetry across players implies that names don't matter for costs. Nonnegativity implies that the planner cannot extract money from the system: $\kappa(\boldsymbol{0}) = 0$ is the definition of the status quo $\hat{\boldsymbol{b}}$, and it does not cost anything to enact $\hat{\boldsymbol{b}}$. Local separability across individuals requires that there are no spillovers in the *costs* of interventions. This is reasonable, as it ensures that the complementarities we study come from the benefits side and not from the costs of interventions. Finally, the twice differentiability of the function is a technical assumption to facilitate the analysis while the positive value of the second derivative at 0 rules out cost functions such as $\kappa(\boldsymbol{y}) = \sum_i y_i^4$ in which the increase in marginal costs at 0 is too slow.

Consider a cost function that satisfies Assumption 5: $\kappa(\boldsymbol{y}) = \sum_i \tilde{\kappa}(y_i)$, where $\tilde{\kappa}(y) = y^2 + cy^3 e^y + c' y^4$, with c and c' being arbitrary constants. Our main result is that the structure of interventions identified in Section 3.2 carries over to such cost functions as long as the budget is small.

Proposition 3. Consider the intervention problem (IT) with the modification that the cost function satisfies Assumption 5. Suppose Assumptions 1 and 2 hold and the network game satisfies Property A. At the optimal intervention, if $C \to 0$ we have $\frac{r_\ell^*}{r_{\ell'}^*} \to \frac{\alpha_\ell}{\alpha_{\ell'}}$.

In Online Appendix B.3 we further clarify the relationship between our original cost formulation and the requirements of Assumption 5. Specifically, we show that adding one more restriction—a generalization of homogeneity—yields a cost function that is equivalent to our original formulation—that is, the one with quadratic costs.

We conclude by noting that the assumption that the cost of intervention is convex implies that the planner will distribute the budget across individuals. With a linear cost function, that is, $K(\boldsymbol{b}, \hat{\boldsymbol{b}}) = \sum_i |b_i - \hat{b}_i|$, the optimal intervention will target a single individual. Online Appendix B.4 presents the details, and there we also show that the methods for identifying this single individual build on the methods that we develop in this paper.

5.4. **Monetary incentives.** In the basic model presented in Section 2, an intervention alters incentives for individual action through a direct change in marginal benefits/marginal costs. The convexity in the cost of changing these marginal benefits plays a key role in the analysis. In this section we provide a demonstration of how our approach can be applied beyond this cost setting. We do this by using our methods to solve the problem of offering monetary incentives to individuals for choosing between two actions.

Let us reinterpret a node i as a population; thus $\mathcal{N} = \{1, 2, \ldots, n\}$ is the set of populations. Within population i, there is a continuum of individuals distributed uniformly in $\mathcal{I} = [0, \overline{\tau}]$. Each individual in population i chooses whether to take action 1 or to take action 0. A strategy of an individual in population i is a function $q_i : [0, \overline{\tau}] \to [0, 1]$ that describes the probability that an individual of type $\tau_i \in [0, \overline{\tau}]$ chooses action 1. Without loss of generality, we focus on the equilibrium in which all the players within a population have the same strategy.

The payoff to an individual who chooses action 0 is normalized to 0. If individual τ_i takes action 1, then he incurs a cost τ_i and gets a benefit that depends on his population's standalone marginal benefit of action 1, b_i, and the number of other individuals he meets who have also taken action 1. We assume that the interaction between populations takes the form of random matching, with the following specification: An individual τ_i in population i meets someone from population j with probability g_{ij}, and, within population j, τ_i meets an individual selected uniformly at random. Suppose τ_i meets type τ_j, and let q_j be the strategy of individuals in population j. Then individual τ_i's payoff for the interaction with the random partner τ_j is

$$\widetilde{\beta} q_j(\tau_j) + b_i - \tau_i.$$

In this expression, $\widetilde{\beta} q_j(\tau_j)$ represents the payoffs from interacting with peers that have also taken action 1.

In Online Appendix B.5 we develop the analysis of this model. First, we show that the conditions for an equilibrium are isomorphic to those of the games we studied in Section 3.2.

We then consider a planner who intervenes in the system. The planner has complete information about the type of each individual in each population and can subsidize types to take action 1 or to take action 0, in a perfectly targeted manner. In doing this, the planner effectively shifts the b_i of some individuals in some populations. The cheapest individuals to influence are those who are close to being indifferent between the two actions, so that they do not need to be paid very much to change their behavior. Indeed, the payment to an individual is proportional to his distance x from the marginal type in equilibrium: Integrating across all the individuals whose actions are changed gives $\int_0^{y_i} x\,dx$, a cost that is quadratic in the magnitude of the change. The intervention problem turns out to be mathematically equivalent to (IT), and so all our results apply.

Note that the specific payoff functions we have taken here make the problem isomorphic to the setting of Example 1, but by suitably modifying the payoffs, we could capture more general externalities, along the lines of Section 5.1.

We focus throughout on maximizing aggregate utility, but we note that the results have applications to other kinds of objectives, such as implementing Pareto improvements. In some cases, interventions will make everyone better off without modification, when positive externalities are strong enough to overcome any negative welfare impacts. However, even when this is not the case, the planner may be able to achieve Pareto improvements. For example, consider a planner who is able to make lump sum transfers—e.g., award or take away discretionary compensation—in addition to any targeted incentives or contingent payments. In such cases, if an improvement in aggregate utility is possible, then the planner can use such transfers to compensate individuals (for instance, those harmed by negative externalities), and achieve a Pareto improvement. In the setting discussed in this subsection, combining lump-sum and action-contingent transfers would then implement a range of Pareto improvements. Even beyond the monetary-incentives setting under consideration here, lump sum transfers may be available to the planner in addition to whatever incentive-targeting scheme is being used, and in such a setting our comments here would apply also.

6. Incomplete information

In the basic model, we assumed that the planner knows the key payoff parameter—the standalone marginal return b_i—of every individual. This section studies settings where the planner does *not* know this parameter and shows that our approach of characterizing interventions based on their principal components, and our main results, have analogues in this environment. For purposes of exposition, we focus on network games that satisfy Property A.

Formally, fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The planner's belief over states is given by $\mathbb{P}$. This represents the planner's uncertainty, given all her information. The planner has control over the random vector (r.v.) $\mathcal{B}$, that is, a function $\mathcal{B} : \Omega \to \mathbb{R}^n$. The choice of $\mathcal{B}$ determines the cost of intervention. A realization of the random variable is denoted by $\boldsymbol{b}$. This realization is common knowledge among individuals when they choose their actions.

Thus, the game individuals play is one of complete information.[22] We also define a function K that gives the cost $K(\mathcal{B})$ of implementing the random variable $\mathcal{B}$.[23]

We solve the following generalized incomplete-information intervention problem:

$$\begin{aligned} &\text{choose r.v. } \mathcal{B} \text{ to maximize } \mathbb{E}\left[W(\boldsymbol{b};\boldsymbol{G})\right] \\ &\qquad \text{s.t. } [\boldsymbol{I}-\beta\boldsymbol{G}]\boldsymbol{a}^* = \boldsymbol{b}, \\ &\qquad\quad K(\mathcal{B}) \leq C. \end{aligned} \tag{IT-G}$$

Note that the intervention problem (IT) under complete information is a special case of a degenerate r.v. $\mathcal{B}$: one in which the planner knows the vector of standalone marginal returns exactly and implements a deterministic adjustment relative to it.

To guide our modeling of the cost of intervention, we now review the features of the distribution of $\mathcal{B}$ that matter for aggregate welfare. For network games that satisfy Property A, we can write:

$$\mathbb{E}\left[W(\boldsymbol{b};\boldsymbol{G})\right] = w\mathbb{E}[(\boldsymbol{a}^*)^\mathsf{T}\boldsymbol{a}^*] = w\mathbb{E}[\underline{\boldsymbol{a}}^\mathsf{T}\underline{\boldsymbol{a}}] = w\sum_{\ell}\alpha_\ell\left(\mathbb{E}[\underline{b}_\ell]^2 + \mathrm{Var}[\underline{b}_\ell]\right). \tag{10}$$

Therefore, we focus on the mean and variance of the realized components $\underline{b}_\ell$; these in turn are determined by the first and second moments of the chosen random variable $\mathcal{B}$. In view of this, we will consider intervention problems that can modify the mean and the covariance matrix of $\mathcal{B}$, and the cost of intervention will depend only on these modifications.

6.1. **Mean shifts.** We consider an intervention where there is an arbitrarily distributed vector of standalone marginal returns and the planner's intervention shifts it in a deterministic way. Formally, fix a random variable $\hat{\mathcal{B}}$, called the status quo, with typical realization $\hat{\boldsymbol{b}}$ (we use notation analogous to that which we defined for $\mathcal{B}$ and $\boldsymbol{b}$). The planner's policy is given by $\boldsymbol{b} = \hat{\boldsymbol{b}} + \boldsymbol{y}$, where $\boldsymbol{y} \in \mathbb{R}^n$ is a deterministic vector. We denote the corresponding random variable by $\mathcal{B}_{\boldsymbol{y}}$. In terms of interpretation, note that implementing this policy does not require knowing $\hat{\boldsymbol{b}}$ as long as the planner has an instrument that shifts incentives.

Assumption 6. The cost of implementing r.v. $\mathcal{B}_{\boldsymbol{y}}$ is

$$K(\mathcal{B}_{\boldsymbol{y}}) = \sum_i y_i^2,$$

and $K(\mathcal{B})$ is ∞ for any other random variable.

[22] It is possible to go further and allow for incomplete information among the individuals about each other's b_i. We do not pursue this substantial generalization here; see Golub and Morris (2017) and Lambert et al. (2018) for analyses in this direction.
[23] The domain of this function is the set of all random vectors taking values in $\mathbb{R}^n$ defined on our probability space.

In contrast to the analysis of Theorem 1, the vector $\hat{\boldsymbol{b}}$ is a random variable. But we take the analogue of the cost function used there, noting that in the deterministic setting this formula held with $\boldsymbol{y} = \boldsymbol{b} - \hat{\boldsymbol{b}}$.

Proposition 4. Consider problem (IT-G) with the cost of intervention satisfying Assumption 6. Suppose Assumption 1 and 2 hold and the network game satisfies Property A. The optimal intervention policy $\mathcal{B}^*$ is equal to $\mathcal{B}_{\boldsymbol{y}^*}$, where $\boldsymbol{y}^*$ is the optimal intervention in the deterministic problem with $\bar{\boldsymbol{b}} = \mathbb{E}[\hat{\boldsymbol{b}}]$ as the status quo vector of standalone marginal returns.

6.2. **Intervention on variances.** We next consider the case where the planner faces a vector of means, fixed at $\bar{\boldsymbol{b}}$, and can choose any random variable $\mathcal{B}$ subject to that mean. In that case, the difference in the expected welfare for two different interventions $\mathcal{B}$ and $\widetilde{\mathcal{B}}$ depends only on the variance–covariance matrix of $\mathcal{B}$ and $\widetilde{\mathcal{B}}$. Thus, the planner effectively faces *the problem of intervening on variances.* We prove a result on optimal intervention for all cost functions satisfying certain symmetries.

Assumption 7. The cost function satisfies two properties: (a) $K(\mathcal{B}) = \infty$ if $\mathbb{E}\boldsymbol{b} \neq \bar{\boldsymbol{b}}$; (b) $K(\mathcal{B}) = K(\widetilde{\mathcal{B}})$ if $\widetilde{\boldsymbol{b}} - \bar{\boldsymbol{b}} = \boldsymbol{O}(\boldsymbol{b} - \bar{\boldsymbol{b}})$, where $\boldsymbol{O}$ is an orthogonal matrix.

Part (a) is a restriction on feasible interventions, namely a restriction to interventions that are mean neutral. Part (b) means that rotations of coordinates around the mean do not affect the cost of implementing a given distribution. This assumption gives the cost a directional neutrality, which ensures that our results are driven by the benefits side rather than by asymmetries operating through the costs. For example, let $\boldsymbol{\Sigma}_{\mathcal{B}}$ be the variance–covariance matrix of the random variable $\mathcal{B}$. That is, $\sigma_{ii}^{\mathcal{B}}$ is the variance of b_i. Suppose that the cost of implementing $\mathcal{B}$ with $\mathbb{E}\boldsymbol{b} = \bar{\boldsymbol{b}}$ is a function of the sum of the variances of the b_i:

$$K(\mathcal{B}) = \begin{cases} \phi\left(\sum_i \sigma_{ii}^{\mathcal{B}}\right) & \text{if } \mathbb{E}\boldsymbol{b} = \bar{\boldsymbol{b}} \\ \infty & \text{otherwise.} \end{cases} \tag{11}$$

The cost function (11) satisfies property (a) of Assumption 7. Moreover, it satisfies property (b) of Assumption 7 because $\sum_i \sigma_{ii}^{\mathcal{B}} = \operatorname{trace} \boldsymbol{\Sigma}_{\mathcal{B}}$; this trace is the sum of the eigenvalues of $\boldsymbol{\Sigma}_{\mathcal{B}}$, which is invariant to the transformation defined in (b).[24]

Proposition 5 (Variance control)**.** Consider problem (IT-G) with the cost of intervention satisfying Assumption 7. Suppose Assumptions 1 and 2 hold and the network game satisfies Property A. Let the optimal intervention be $\mathcal{B}^*$. We have the following:

1. Suppose the planner likes variance (i.e., $w > 0$). If the game has strategic complements ($\beta > 0$), then $\operatorname{Var}(\boldsymbol{u}^\ell(\boldsymbol{G}) \cdot \boldsymbol{b}^*)$ is weakly decreasing in ℓ; if the game has strategic substitutes ($\beta < 0$), then $\operatorname{Var}(\boldsymbol{u}^\ell(\boldsymbol{G}) \cdot \boldsymbol{b}^*)$ is weakly increasing in ℓ.

[24]When we look at the variance–covariance matrix of $\widetilde{\boldsymbol{b}}$ defined by $\widetilde{\boldsymbol{b}} - \bar{\boldsymbol{b}} = \boldsymbol{O}(\boldsymbol{b} - \bar{\boldsymbol{b}})$, the variance–covariance matrix becomes $\boldsymbol{O}\boldsymbol{\Sigma}\boldsymbol{O}^{\mathsf{T}}$, and this has the same eigenvalues and therefore the same trace.

2. Suppose the planner dislikes the variance (i.e., $w < 0$). If the game has strategic complements ($\beta > 0$), then $\mathrm{Var}(\boldsymbol{u}^\ell(\boldsymbol{G}) \cdot \boldsymbol{b}^*)$ is weakly increasing in ℓ; if the game has strategic substitutes ($\beta < 0$), then $\mathrm{Var}(\boldsymbol{u}^\ell(\boldsymbol{G}) \cdot \boldsymbol{b}^*)$ is weakly decreasing in ℓ.

We now provide an intuition for Proposition 5. Shocks to individuals' standalone marginal returns create variability in the players' equilibrium actions. The assumption that the intervention is mean neutral (part (a) of Assumption 7) leaves the planner to control only the variances and covariances of these marginal returns with her intervention. Hence, the solution to the intervention problem describes what the planner should do to induce volatilities in actions that maximize the ex-ante expected welfare.

Suppose first that investments are strategic complements. Then a perfectly correlated shock in individual standalone marginal returns is amplified by strategic interaction. In fact, the type of shock that is most amplifying (at a given size) is the one that is perfectly correlated across individuals, with the magnitude of a given individual's shock proportional to the first principal component (his eigenvector centrality). These shocks are exactly what $\underline{b}_1^* = \boldsymbol{u}^1(\boldsymbol{G}) \cdot \boldsymbol{b}^*$ captures. Hence, this is the dimension of volatility that the planner most wants to increase if she likes variability in actions ($w > 0$) and most wants to decrease if she dislikes variability in actions ($w < 0$).

If investments are strategic substitutes, then a perfectly correlated shock does not create a lot of variability in actions: The first-order response of all individuals to an increase in their standalone marginal returns is to increase investment, but that in turn makes all individuals to decrease their investment somewhat because of the strategic substitutability with their neighbors. Hence, highly positively correlated shocks do not translate into high volatility. The shock profiles that create most variability in actions are the ones in which neighbors have *negatively* correlated shocks. A planner that loves variability in actions will then prioritize such shocks. Because the last eigenvector of the system describes the local connections across nodes, this is exactly the type of volatility that is of greatest concern, and this is what the planner will focus on most.

Example 5 (Illustration in the case of the circle)**.** Figure 1 depicts six of the eigenvectors/principal components, beginning with eigenvector 2, of a circle network with 14 nodes. The first principal component is a positive vector and so $\mathcal{B}$ projected on $\boldsymbol{u}^1(\boldsymbol{G})$ captures positively correlated shocks across all players. The second principal component (top left panel of Figure 1) splits the graph into two sides, one with positive entries and the other with negative entries. Hence, $\mathcal{B}$ projected on $\boldsymbol{u}^2(\boldsymbol{G})$ captures shocks that are highly positively correlated on each side of the circle network, with the two opposite sides of the circle being anti-correlated. As we move along the sequence, we can see that $\mathcal{B}$ projected on the ℓ^{th} eigenvector represents shocks that are more and more local, with shocks to any given node being anti-correlated with shocks to nearby nodes. For example, the $\ell = 10$ or $\ell = 12$ component (bottom-right panel of Figure 1) depict volatility that is locally highly anti-correlated—the shocks of neighbors are usually opposite—but not quite as strongly as in the last component, $\ell = 14$, where shocks of connected individuals are perfectly anti-correlated.

With this interpretation, we can now appreciate that when the game has strategic complements the intervention will give priority to changing the variances of the top principal components, as these are the ones representing the shocks that, by the strategic nature of the game, amplify more and so create greater volatility. Analogously, when the game has strategic substitutes the intervention will give priority to changing the variance of the last principal components. Whether the planner wants to decrease or increase the likelihood of these shocks will depend on whether the planner likes variance ($w > 0$) or dislikes it ($w < 0$).

7. Principal components and other network measures

This section discusses the relation between the principal components and related network statistics.

First principal component and eigenvector centrality: For ease of exposition, let the network be connected, that is, let $\boldsymbol{G}$ be irreducible. By the Perron–Frobenius Theorem, $\boldsymbol{u}^1(\boldsymbol{G})$ is entry-wise positive; indeed, this vector is the Perron vector of the matrix, also known as the vector of individuals' eigenvector centralities. Thus, our results of Section 4 imply that, under strategic complementarities, interventions that aim to maximize the aggregate utility should change individuals' incentives in proportion to their eigenvector centralities.

It is worth comparing this result with results that highlight the importance of Bonacich centrality. Under strategic complements, equilibrium actions are proportional to the individuals' Bonacich centralities in the network (Ballester et al., 2006).[25] Within the Ballester et al. (2006) framework, it can easily be verified that if the objective of the planner is linear in the sum of actions, then under a quadratic cost function the planner will target individuals in proportion to their Bonacich centralities (see also Demange (2017)). Bonacich centrality converges to eigenvector centrality as the spectral radius of $\beta\boldsymbol{G}$ tends to 1; otherwise the two vectors can be quite different (see, for example, Calvó-Armengol et al. (2015) or Golub and Lever (2010)).

The substantive point is that the objective of our planner when solving the intervention problem (IT) is to maximize the aggregate equilibrium *utility*, not the sum of actions, and that explains the difference in the targeting strategy. Indeed, our planner's objective (under Property A) can be written as follows (introducing a different constant factor for convenience):

$$\sum_i u_i \propto \frac{1}{n}\sum_i a_i^2 \quad = \quad \bar{a}^2 + \sigma_{\boldsymbol{a}}^2,$$

[25] For a different economic context in which eigenvector centrality reflects equilibrium outcomes, see also Elliott and Golub (2018).

where $\sigma^2_{\boldsymbol{a}}$ is the variance of the action profile and $\overline{a}$ is the mean action. Thus, our planner cares about the sum of actions and also their diversity, simply as a mathematical consequence of her objective. This explains the reason why her policies differ from those that would be in effect if just the mean action were the focus. To reiterate this point, we finally note that if we consider problem (IT) but we assume that the cost of intervention is linear, that is, $K(\boldsymbol{b}, \hat{\boldsymbol{b}}) = \sum_i |b_i - \hat{b}_i|$, then the optimal intervention will target only one individual (see also the discussion in Section 5.3). The targeted individual is not necessarily the individual with the highest Bonacich centrality; the optimal intervention is characterized in Online Appendix B.4.

Last principal component: We have shown that in games with strategic substitutes, for large budgets interventions that aim to maximize the aggregate utility target individuals in proportion to the eigenvector of $\boldsymbol{G}$ associated to the smallest eigenvalue of $\boldsymbol{G}$, the last principal component.

There is a connection between this result and the work of Bramoullé et al. (2014). Bramoullé et al. (2014) study the set of equilibria of a network game with linear best replies and strategic substitutes. They observe that such a game is a potential game, and they derive the potential function explicitly. From this, they can deduce that the smallest eigenvalue of $\boldsymbol{G}$ is crucial for whether the equilibrium is unique, and it is also useful for analyzing the stability of a particular equilibrium.[26] The basic intuition is that the magnitude of the smallest eigenvalue determines how small changes in individuals' actions propagate, via strategic substitutes, in the network. When these amplifications are strong, multiple equilibria can emerge. Relatedly, when these amplifications are strong around an equilibrium, that equilibrium will be unstable.

Our study of the strategic substitutes case is driven by different questions, and delivers different sorts of characterizations. We assume that there is a stable equilibrium which is unique at least locally, and then we characterize optimal interventions in terms of the eigenvectors of $\boldsymbol{G}$. In general, all the eigenvectors—not just the one associated to the smallest eigenvalue—can matter. Interventions will focus *more* on the eigenvectors with smaller eigenvalues. When the budget is sufficiently large, the intervention will (in the setting of Section 4) focus on only the smallest-eigenvalue eigenvector. As discussed in Section 4, the network determinants of whether targeting is simple can be quite subtle. To the best of our knowledge, these considerations are all new in the study of network games.

Nevertheless, at an intuitive level there are important points of contact between our intuitions and those of Bramoullé et al. (2014). In our context, as discussed earlier, our planner likes to move the incentives of adjacent individuals in opposite directions. The eigenvector associated to the smallest eigenvalue emerges as the one identifying the best way to do this at a given cost, and the eigenvalue itself measures how intensely the strategic

[26]For stability of equilibrium, what is relevant is the magnitude of the smallest eigenvalue of an appropriately defined subgraph of $\boldsymbol{G}$.

effects amplify. This "amplification" property involves forces similar to those that make the smallest eigenvalue important to stability and uniqueness in Bramoullé et al. (2014).

Spectral approaches to variance control: Acemoglu et al. (2016) give a general analysis of which network statistics matter for volatility of network equilibria. Baqaee and Farhi (2017) develop a rich macroeconomic analysis relating network measures to aggregate volatility. Though both papers note the importance of eigenvector centrality in (their analogues of) the case of strategic complements, their main focus is on how the *curvature* of best responses changes the volatility of an aggregate outcome, and which "second order" (curvature-related) network statistics are important. We use the principal components of the network to understand which first-order shocks are most amplified, and how this depends on the nature of strategic interactions.

8. Concluding remarks

We study the problem of a planner who seeks to optimally target individuals in a network of interaction. Our framework allows for a broad class of strategic and non-strategic spillovers across individuals. The analysis builds on the singular value decomposition of the matrix that summarizes strategic interactions—this yields the principal components—and offers a general approach for understanding the amplification and attenuation of interventions.

We briefly mention two further applications.

In some circumstances, the planner seeks a budget-balanced tax/subsidy scheme in order to improve the economic outcome. In a supply chain, for example, a planner could tax some suppliers, thereby increasing their marginal costs, and then use that tax revenue to subsidize other suppliers. The planner will solve a problem similar to the one we have studied here, with the important difference that she will face a different constraint, namely, a budget-balance constraint. In ongoing work, Galeotti et al. (2018) show that the decomposition of the network that we employed in this paper is useful in deriving the optimal taxation scheme and, in turn, in determining the welfare gains that can be achieved in supply chains.

We have focused on interventions that alter the standalone marginal returns of individuals. Another interesting problem is the study of interventions that alter the matrix of interaction. We hope this paper stimulates further work along these lines.

References

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2016): "Networks, Shocks, and Systemic Risk," in *Oxford Handbook of the Economics of Networks*, ed. by Y. Bramoullé, A. Galeotti, and B. Rogers, Oxford University Press.

Akbarpour, M., S. Malladi, and A. Saberi (2017): "Diffusion, Seeding, and the Value of Network Information," Available at SSRN: https://ssrn.com/abstract=3062830.

Allouch, N. (2015): "On the private provision of public goods on networks," *Journal of Economic Theory*, 157, 527–552.

——— (2017): "Aggregation in Networks," University of Kent, School of Economics Discussion Paper 1718.

Alon, N. and N. Kahale (1997): "A Spectral Technique for Coloring Random 3-Colorable Graphs," *SIAM Journal on Computing*, 26, 1733–1748.

Angeletos, G.-M. and A. Pavan (2007): "Efficient Use of Information and Social Value of Information," *Econometrica*, 75, 1103–1142.

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

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013): "The Diffusion of Microfinance," *Science*, 341, 1236498.

Baqaee, D. R. and E. Farhi (2017): "The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem," Working paper, London School of Economics.

Belhaj, M. and F. Deroian (2017): "Contracting on Networks," Mimeo.

Bloch, F. and N. Querou (2013): "Pricing in Social Networks," *Games and Economic Behavior*, 80, 263–281.

Borgatti, S. (2006): "Identifying Sets of Key Players in a Social Network," *Computational and Mathematical Organization Theory*, 12, 21–34.

Bramoulle, 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," *The American Economic Review*, 104, 898–930.

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

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

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

Desai, M. and V. Rao (1994): "A Characterization of the Smallest Eigenvalue of a Graph," *Journal of Graph Theory*, 18, 181–194.

Dessein, W., A. Galeotti, and T. Santos (2016): "Rational Inattatention and Organziational Focus," *American Economic Review*, 106, 1522–36.

Elliott, M. and B. Golub (2018): "A Network Approach to Public Goods," *Journal of Political Economy*, forthcoming.

Fainmesser, I. and A. Galeotti (2017): "Pricing Network Effects," *Review of Economic Studies*, 83, 165–198.

Galeotti, A., B. Golub, S. Goyal, E. Talamàs, and O. Tamuz (2018): "Optimal Taxation in Supply Chains," *mimeo.*

Galeotti, A. and S. Goyal (2009): "Influencing the Influencers: A Theory of Strategic Diffusion," *The Rand Journal of Economics*, 40, 509–532.

——— (2010): "The Law of the Few," *American Economic Review*, 100, 1468–92.

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

Galeotti, A. and B. W. Rogers (2013): "Strategic Immunization and Group Structure," *American Economic Journal: Microeconomics*, 5, 1–32.

Golub, B. and M. O. Jackson (2012a): "Does homophily predict consensus times? Testing a model of network structure via a dynamic process," *Review of Network Economics*, 11.

——— (2012b): "How Homophily Affects the Speed of Learning and Best-Response Dynamics," *The Quarterly Journal of Economics*, 127, 1287–1338.

——— (2012c): "Network Structure and the Speed of Learning: Measuring Homophily Based on its Consequences," *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, 33–48.

Golub, B. and C. Lever (2010): "The Leverage of Weak Ties: How Linking Groups Affects Inequality," Working paper, Harvard University.

Golub, B. and S. Morris (2017): "Expectations, Networks and Conventions," Available at SSRN: https://ssrn.com/abstract=2979086.

Golub, G. H. and C. F. Van Loan (1996): *Matrix Computations*, The Johns Hopkins University Press.

Goyal, S., J. Moraga, and M. van der Leij (2006): "Economics: an emerging small world?" *Journal of Political Economy*, 114, 403–412.

Goyal, S. and J. Moraga-Gonzalez (2001): "R&D Networks," *The Rand Journal of Economics*, 32, 686–707.

Hartfiel, D. and C. D. Meyer (1998): "On the Structure of Stochastic Matrices with a Subdominant Eigenvalue Near 1," *Linear Algebra and its Applications*, 272, 193–203.

Horn, R. A. and C. R. Johnson (2012): *Matrix Analysis*, New York, NY, USA: Cambridge University Press, 2nd ed.

Jackson, M., B. W. Rogers, and Y. Zenou (2017): "The Economic Consequences of Social-Network Structure ," *Journal of Economic Literature*, 55, 49–95.

Kempe, D., J. Kleinberg, and E. Tardos (2003): "Maximizing the Spread of Influence through a Social Network." in *Proceedings 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.*

Lambert, N. S., G. Martini, and M. Ostrovsky (2018): "Quadratic Games," Available at NBER: https://www.nber.org/papers/w24914.

Leduc, M. V., M. O. Jackson, and R. Johari (2017): "Pricing and referrals in diffusion on networks," *Games and Economic Behavior*, 104, 568–594.

Levin, D. A., Y. Peres, and E. L. Wilmer (2009): *Markov Chains and Mixing Times*, Providence, RI: American Mathematical Society.

Marschak, T. and R. Radner (1972): *Economic Theory of Teams*, New Haven, USA: Cowles Foundation.

Meyer, C. D. (2000): *Matrix Analysis and Applied Linear Algebra*, Philadelphia, PA, USA: Society for Industrial and Applied Mathematics.

Morris, S. and H. S. Shin (2002): "Social Value of Public Information," *American Economic Review*, 92, 1521–1534.

Rogers, E. (1983): *Diffusion of Innovations. Third Edition*, Free Press. New York.

Valente, T. (2012): "Network Interventions," *Science*, 337, 49–53.

Zenou, Y. (2016): "Key Players," in *Oxford Handbook of the Economics of Networks*, ed. by Y. Bramoullé, A. Galeotti, and B. Rogers, Oxford University Press.

Appendix A. Proofs

Proof of Theorem 1. We wish to solve

$$\begin{aligned} \max_{\boldsymbol{b}} \quad & w\boldsymbol{a}^{\mathsf{T}}\boldsymbol{a} \\ \text{s.t.:} \quad & [\boldsymbol{I} - \beta\boldsymbol{G}]\boldsymbol{a}^* = \boldsymbol{b}, \\ & \sum_i (b_i - \hat{b}_i)^2 \leq C \end{aligned}$$

The first step is to transform the maximization problem into the basis given by the principal components of $\boldsymbol{G}$. To this end, we first rewrite the cost and the objective in the principal components basis, using the fact that norms do not change under the orthogonal transformation $\boldsymbol{U}^{\mathsf{T}}$. (Unless otherwise noted, the norm symbol $\|\cdot\|$ always refers to the Euclidean norm.) Letting $\boldsymbol{y} = \boldsymbol{b} - \hat{\boldsymbol{b}}$,

$$K(\boldsymbol{b}, \hat{\boldsymbol{b}}) = \sum_i y_i^2 = \|\boldsymbol{y}\|_2^2 = \sum_\ell \underline{y}_\ell^2$$

and

$$w\boldsymbol{a}^{\mathsf{T}}\boldsymbol{a} = w\|\boldsymbol{a}\|^2 = w\|\underline{\boldsymbol{a}}\|^2 = w\underline{\boldsymbol{a}}^{\mathsf{T}}\underline{\boldsymbol{a}}.$$

By recalling that, in equilibrium, $\underline{\boldsymbol{a}}^* = [\boldsymbol{I} - \beta\boldsymbol{\Lambda}]^{-1}\underline{\boldsymbol{b}}$, and using the definition $\alpha_\ell = \frac{1}{(1-\beta\lambda_\ell(\boldsymbol{G}))^2}$, the intervention problem (IT) can be rewritten as:

$$\begin{aligned} \max_{\underline{\boldsymbol{b}}} \quad & w\sum_\ell \alpha_\ell \underline{b}_\ell^2 \\ \text{s.t.} \quad & \sum_\ell \underline{y}_\ell^2 \leq C. \end{aligned}$$

We now transform the problem so that the control variable is $\boldsymbol{x}$ where $x_\ell = y_\ell / \underline{\hat{b}}_\ell$. We obtain

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & w \sum_\ell \alpha_\ell (1+x_\ell)^2 \underline{\hat{b}}_\ell \\ \text{subject to} \quad & \sum_\ell \underline{\hat{b}}_\ell^2 x_\ell^2 \leq C \end{aligned}$$

Note that, for all ℓ, α_ℓ are well-defined (by Assumption 1) and strictly positive (by genericity of $\boldsymbol{G}$). This has two implications.[27]

First, at the optimal solution $\boldsymbol{x}^*$ the resource constraint problem must bind. To see this, note that Assumption 3 says that either $w > 0$, or $w < 0$ and $\sum_\ell \underline{\hat{b}}_\ell^2 > C$. Suppose that at the optimal solution the constraint does not bind. Then, without violating the constraint, we can slightly increase or decrease any x_ℓ. If $w > 0$ (resp. $w < 0$) the increase or the decrease is guaranteed to increase (resp. decrease) the corresponding $(x_\ell + 1)^2$ (since the α_ℓ are all strictly positive).

Second, we show that the optimal solution $\boldsymbol{x}^*$ satisfies $x_\ell^* \geq 0$ for every ℓ if $w > 0$, and $x_\ell^* \in [-1, 0]$ for every ℓ if $w < 0$. Suppose $w > 0$ and, for some ℓ, $x_\ell^* < 0$. Then $[-x_\ell^* + 1]^2 > [x_\ell^* + 1]^2$. Since $w > 0$ and every α_ℓ is positive, we can raise the aggregate utility without changing the cost by flipping the sign of x_ℓ^*. Analogously, suppose $w < 0$. It is clear that if $x_\ell^* < -1$, then by setting $x_\ell = -1$ the objective improves and the constraint is relaxed; hence, at the optimum, $x_\ell^* \geq -1$. Suppose next that $x_\ell > 0$ for some ℓ. Then $[-x_\ell^* + 1]^2 < [x_\ell^* + 1]^2$. Since $w < 0$ and every α_ℓ is positive, we can improve the value of the objective function without changing the cost by flipping the sign of x_ℓ^*.

We now complete the proof. Observe that the Lagrangian corresponding to the maximization problem is

$$\mathcal{L} = w \sum_\ell \alpha_\ell (1+x_\ell)^2 \underline{\hat{b}}_\ell + \mu \left[C - \sum_\ell \underline{\hat{b}}_\ell^2 x_\ell^2 \right].$$

Taking our observation above that the constraint is binding at $\boldsymbol{x} = \boldsymbol{x}^*$, together with the standard results on the Karush–Kuhn–Tucker conditions, the first-order conditions must hold exactly at the optimum with a positive μ:

$$0 = \frac{\partial \mathcal{L}}{\partial x_\ell} = 2\underline{\hat{b}}_\ell^2 \left[w\alpha_\ell (1 + x_\ell^*) - \mu x_\ell^* \right] = 0. \tag{12}$$

We take a generic $\hat{\boldsymbol{b}}$ such that $\underline{\hat{b}}_\ell \neq 0$ for all ℓ. If for some ℓ we had $\mu = w\alpha_\ell$ then the right-hand side of the second equality in (12) would be $2\underline{\hat{b}}_\ell^2 w \alpha_\ell$, which, by the generic assumption we just made and the positivity of α_ℓ, would contradict (12). Thus, the following holds with a

[27] Note that if Assumption 3 does not hold (that is, $w < 0$ and $\sum_\ell \underline{\hat{b}}_\ell^2 \leq C$) then the optimal solution is $x_\ell^* = -1$ for all ℓ. This is what we ruled out with Assumption 3, before Theorem 1.

nonzero denominator:

$$x_\ell^* = \frac{w\alpha_\ell}{\mu - w\alpha_\ell},$$

and the Lagrange multiplier μ is therefore pinned down by

$$\sum_\ell w^2 \hat{\underline{b}}_\ell^2 \left(\frac{\alpha_\ell}{\mu - w\alpha_\ell}\right)^2 = C.$$

Note finally that

$$\rho(\boldsymbol{y}^*, \boldsymbol{u}^\ell(\boldsymbol{G})) = \frac{\boldsymbol{y}^* \cdot \boldsymbol{u}^\ell(\boldsymbol{G})}{\|\boldsymbol{y}^*\|\|\boldsymbol{u}^\ell(\boldsymbol{G})\|} = \frac{\underline{y}_\ell^*}{\sqrt{C}} = \frac{\hat{\underline{b}}_\ell x_\ell^*}{\sqrt{C}} = \frac{\|\hat{\boldsymbol{b}}\|}{\sqrt{C}} \rho(\hat{\boldsymbol{b}}, \boldsymbol{u}^\ell(\boldsymbol{G}))x_\ell^* \propto_\ell \rho(\hat{\boldsymbol{b}}, \boldsymbol{u}^\ell(\boldsymbol{G}))x_\ell^*.$$

□

Proof of Proposition 1. Part 1. From expression 6 of Theorem 1, it follows that if $C \to 0$ then $\mu \to \infty$. The result follows by noticing that

$$\frac{r_\ell^*}{r_{\ell'}^*} = \frac{\alpha_\ell}{\alpha_{\ell'}} \frac{\mu - w\alpha'_\ell}{\mu - w\alpha_\ell}.$$

Part 2. Suppose that $\beta > 0$. Using the derivation of the last part of the proof of Theorem 1, we write:

$$\rho(\boldsymbol{y}^*, \boldsymbol{u}^\ell(\boldsymbol{G})) = \frac{\|\hat{\boldsymbol{b}}\|}{\sqrt{C}} \rho(\hat{\boldsymbol{b}}, \boldsymbol{u}^\ell(\boldsymbol{G}))x_\ell^*,$$

with $x_\ell^* = \frac{w\alpha_\ell}{\mu - w\alpha_\ell}$. From expression 6 of Theorem 1, it follows that if $C \to \infty$ then $\mu \to w\alpha_1$. This implies that $x_\ell^* \to \frac{\alpha_\ell}{\alpha_1 - \alpha_\ell}$ for all $\ell \neq 1$. As a result, if $C \to \infty$ then $\rho(\boldsymbol{y}^*, \boldsymbol{u}^\ell(\boldsymbol{G})) \to 0$ for all $\ell \neq 1$. Furthermore, we can rewrite expression 6 of Theorem 1 as

$$\sum_\ell \left(\|\hat{\boldsymbol{b}}\|\rho(\hat{\boldsymbol{b}}, \boldsymbol{u}^\ell(\boldsymbol{G}))\frac{x_\ell^*}{\sqrt{C}}\right)^2 = 1,$$

and therefore

$$\lim_{C\to\infty} \sum_\ell \left(\|\hat{\boldsymbol{b}}\|\rho(\hat{\boldsymbol{b}}, \boldsymbol{u}^\ell(\boldsymbol{G}))\frac{x_\ell^*}{\sqrt{C}}\right)^2 = \lim_{C\to\infty} \left(\|\hat{\boldsymbol{b}}\|\rho(\hat{\boldsymbol{b}}, \boldsymbol{u}^1(\boldsymbol{G}))\frac{x_1^*}{\sqrt{C}}\right)^2 = 1,$$

where the first equality follows because $x_\ell^* \to \frac{\alpha_\ell}{\alpha_1 - \alpha_\ell}$ for all $\ell \neq 1$. The proof for the case of $\beta < 0$ follows the same steps, with the only exception that if $C \to \infty$ then $\mu \to w\alpha_n$.

□

Proof of Proposition 2. We first prove the result on welfare and then turn to the result on cosine similarity.

Welfare. Consider the case of strategic complementarities, $\beta > 0$. Define by $\tilde{\boldsymbol{x}}$ the simple intervention, and note that $\tilde{x}_1 = \sqrt{C}/\hat{\underline{b}}_1$ and that $\tilde{x}_\ell = 0$ for all $\ell > 1$. The aggregate utility obtained under the simple intervention is:

$$W^s = \sum_\ell \hat{\underline{b}}_\ell^2 \alpha_\ell (1 + \tilde{x}_\ell)^2 = \hat{\underline{b}}_1^2 \alpha_1 \tilde{x}_1(\tilde{x}_1 + 2) + \sum_\ell \alpha_\ell \hat{\underline{b}}_\ell^2.$$

The aggregate utility at the optimal intervention is

$$W^* = \sum_{\ell} \hat{\underline{b}}_\ell^2 \alpha_\ell (1 + x_\ell^*)^2 = \hat{\underline{b}}_1^2 \alpha_1 x_1^*(x_1^* + 2) + \sum_{\ell \neq 1} \hat{\underline{b}}_\ell^2 \alpha_\ell x_\ell^*(x_\ell^* + 2) + \sum_{\ell} \alpha_\ell \hat{\underline{b}}_\ell^2$$

Hence

$$\begin{aligned}
\frac{W^*}{W^s} &= \frac{\hat{\underline{b}}_1^2 \alpha_1 x_1^*(x_1^* + 2) + \sum_\ell \alpha_\ell \hat{\underline{b}}_\ell^2}{\hat{\underline{b}}_1^2 \alpha_1 \tilde{x}_1(\tilde{x}_1 + 2) + \sum_\ell \alpha_\ell \hat{\underline{b}}_\ell^2} + \frac{\sum_{\ell \neq 1} \hat{\underline{b}}_\ell^2 \alpha_\ell x_\ell^*(x_\ell^* + 2)}{\hat{\underline{b}}_1^2 \alpha_1 \tilde{x}_1(\tilde{x}_1 + 2) + \sum_\ell \alpha_\ell \hat{\underline{b}}_\ell^2} \\
&\leq 1 + \frac{\sum_{\ell \neq 1} \hat{\underline{b}}_\ell^2 \alpha_\ell x_\ell^*(x_\ell^* + 2)}{\hat{\underline{b}}_1^2 \alpha_1 \tilde{x}_1(\tilde{x}_1 + 2) + \sum_\ell \alpha_\ell \hat{\underline{b}}_\ell^2} && \text{as } \tilde{x}_1 \geq x_1^* \\
&\leq 1 + \frac{\sum_{\ell \neq 1} \hat{\underline{b}}_\ell^2 \alpha_\ell x_\ell^*(x_\ell^* + 2)}{\hat{\underline{b}}_1^2 \alpha_1 \tilde{x}_1^2} && \text{summands in denominator are positive} \\
&= 1 + \frac{\sum_{\ell \neq 1} \hat{\underline{b}}_\ell^2 \alpha_\ell x_\ell^*(x_\ell^* + 2)}{\alpha_1 C} && \underline{b}_1^2 \tilde{x}_1^2 = C; \text{ see below} \\
&\leq 1 + \frac{2\alpha_1 - \alpha_2}{\alpha_1} \frac{\|\hat{\boldsymbol{b}}\|^2}{C} \left(\frac{\alpha_2}{\alpha_1 - \alpha_2} \right)^2 && \text{see calculation below} \\
&\leq 1 + \frac{2\|\hat{\boldsymbol{b}}\|^2}{C} \left(\frac{\alpha_2}{\alpha_1 - \alpha_2} \right)^2.
\end{aligned}$$

The fact $\underline{b}_1^2 \tilde{x}_1^2 = C$, used above follows because the simple policy allocates the entire budget to changing $\underline{b}_1$. The inequality after that statement follows because

$$\begin{aligned}
\sum_{\ell \neq 1} \hat{\underline{b}}_\ell^2 \alpha_\ell x_\ell^*(x_\ell^* + 2) &\leq \alpha_2 \sum_{\ell \neq 1} \hat{\underline{b}}_\ell^2 x_\ell^*(x_\ell^* + 2) && \text{ordering of the } \alpha_\ell \\
&\leq \alpha_2 x_2^*(x_2^* + 2) \sum_{\ell \neq 1} \hat{\underline{b}}_\ell^2 && \text{Corollary 1} \\
&\leq \alpha_2 \frac{w\alpha_2}{\mu - w\alpha_2} \left(\frac{w\alpha_2}{\mu - w\alpha_2} + 2 \right) \sum_{\ell \neq 1} \hat{\underline{b}}_\ell^2 && \text{Theorem 1} \\
&\leq \alpha_2 \frac{w\alpha_2}{w\alpha_1 - w\alpha_2} \left(\frac{w\alpha_2}{w\alpha_1 - w\alpha_2} + 2 \right) \|\hat{\underline{\boldsymbol{b}}}\|^2 \\
&= \left(\frac{\alpha_2}{\alpha_1 - \alpha_2} \right)^2 (2\alpha_1 - \alpha_2) \|\hat{\boldsymbol{b}}\|^2
\end{aligned}$$

Hence, the inequality

$$C > \frac{2\|\hat{\boldsymbol{b}}\|^2}{\epsilon} \left(\frac{\alpha_2}{\alpha_1 - \alpha_2} \right)^2$$

is sufficient to establish that $\frac{W^*}{W^s} < 1 + \epsilon$. The proof for the case of strategic substitutes follows the same steps; the only difference is that we use α_n instead of α_1 and α_{n-1} instead of α_2.

Cosine similarity. We now turn to the cosine similarity result. We focus on the case of strategic complements. The proof for the case of strategic substitutes is analogous. We start by writing a useful explicit expression for $\rho(\Delta b^*, \sqrt{C}u^1)$:

$$\rho(\Delta \boldsymbol{b}^*, \sqrt{C}\boldsymbol{u}^1) = \frac{(\boldsymbol{b}^* - \hat{\boldsymbol{b}}) \cdot (\sqrt{C}\boldsymbol{u}^1)}{\|\boldsymbol{b}^* - \hat{\boldsymbol{b}}\| \|\sqrt{C}\boldsymbol{u}^1\|} = \frac{(\boldsymbol{b}^* - \hat{\boldsymbol{b}}) \cdot (\boldsymbol{u}^1)}{\sqrt{C}}, \tag{13}$$

where the last equality follows because, at the optimum, $\|\boldsymbol{b}^* - \hat{\boldsymbol{b}}\|^2 = C$. At the optimal intervention, by Theorem 1,

$$\underline{b}^*_\ell - \underline{\hat{b}}_\ell = \frac{w\alpha_\ell}{\mu - w\alpha_\ell}\underline{\hat{b}}_\ell;$$

now, using the definition $\underline{\boldsymbol{b}} = \boldsymbol{U}^T\boldsymbol{b}$, we have that

$$b^*_i - \hat{b}_i = w\sum_\ell u^i_\ell \frac{\alpha_\ell}{\mu - w\alpha_\ell}\underline{\hat{b}}_\ell$$

and therefore

$$(\boldsymbol{b}^* - \hat{\boldsymbol{b}}) \cdot \boldsymbol{u}^1 \quad = \quad \sum_i \sum_\ell u^1_i u^\ell_i \frac{w\alpha_\ell}{\mu - w\alpha_\ell}\underline{\hat{b}}_\ell = \sum_\ell \frac{w\alpha_\ell}{\mu - w\alpha_\ell}\underline{\hat{b}}_\ell(\boldsymbol{u}^1 \cdot \boldsymbol{u}^\ell) = \frac{w\alpha_1}{\mu - w\alpha_1}\underline{\hat{b}}_1$$

Hence, using this in equation 13, we can deduce that

$$\rho(\Delta \boldsymbol{b}^*, \boldsymbol{u}^1) = \frac{1}{\sqrt{C}}\frac{w\alpha_1}{\mu - w\alpha_1}\underline{\hat{b}}_1 \geq \sqrt{1-\epsilon} \quad \text{iff} \quad \left(\frac{w\alpha_1}{\mu - w\alpha_1}\right)^2 \underline{\hat{b}}_1^2 - C(1-\epsilon) \geq 0. \tag{14}$$

The following lemma shows that the inequality after the "if and only if" follows from our hypothesis that

$$C > \frac{2\|\hat{\boldsymbol{b}}\|^2}{\epsilon}\left(\frac{\alpha_2}{\alpha_1 - \alpha_2}\right)^2,$$

and thus establishing it completes the proof.

Lemma 1. Assume

$$C > \frac{2\|\hat{\boldsymbol{b}}\|^2}{\epsilon}\left(\frac{\alpha_2}{\alpha_1 - \alpha_2}\right)^2.$$

Then

$$\left(\frac{w\alpha_1}{\mu - w\alpha_1}\right)^2 \underline{\hat{b}}_1^2 \geq C(1-\epsilon) \tag{15}$$

Proof of Lemma 1. Note that

$$C > \frac{2\|\hat{\boldsymbol{b}}\|^2}{\epsilon}\left(\frac{\alpha_2}{\alpha_1 - \alpha_2}\right)^2 \quad \implies \quad \epsilon C > \|\hat{\boldsymbol{b}}\|^2\left(\frac{\alpha_2}{\alpha_1 - \alpha_2}\right)^2,$$

and therefore

$$C(1-\epsilon) < C - \|\hat{\boldsymbol{b}}\|^2\left(\frac{\alpha_2}{\alpha_1 - \alpha_2}\right)^2. \tag{16}$$

But then we have the following chain of statements, explained immediately after the display:

$$\begin{aligned}\left(\frac{w\alpha_1}{\mu - w\alpha_1}\right)^2 \underline{\hat{b}}_1^2 - C(1-\epsilon) &\geq \left(\frac{w\alpha_1}{\mu - w\alpha_1}\right)^2 \underline{\hat{b}}_1^2 - C + \|\hat{\boldsymbol{b}}\|^2 \left(\frac{\alpha_2}{\alpha_1 - \alpha_2}\right)^2 \\ &= \left(\frac{w\alpha_1}{\mu - w\alpha_1}\right)^2 \underline{\hat{b}}_1^2 - \sum_\ell \left(\frac{w\alpha_\ell}{\mu - w\alpha_\ell}\right)^2 \hat{b}_\ell^2 + \|\hat{\boldsymbol{b}}\|^2 \left(\frac{\alpha_2}{\alpha_1 - \alpha_2}\right)^2 \\ &= \|\hat{\boldsymbol{b}}\|^2 \left(\frac{\alpha_2}{\alpha_1 - \alpha_2}\right)^2 - \sum_{\ell \neq 1} \left(\frac{w\alpha_\ell}{\mu - w\alpha_\ell}\right)^2 \hat{b}_\ell^2 \\ &= \left(\frac{\alpha_2}{\alpha_1 - \alpha_2}\right)^2 \sum_\ell \underline{\hat{b}}_\ell^2 - \sum_{\ell \neq 1} \left(\frac{w\alpha_\ell}{\mu - w\alpha_\ell}\right)^2 \hat{b}_\ell^2 > 0.\end{aligned}$$

The first inequality follows from substituting the upper bound on $C(1-\epsilon)$, statement (16) above, which we derived from our initial condition on C. The equality after that follows by substituting the condition on the binding budget constraint at the optimum, which we derived in Theorem 1. The next equality follows by isolating the term for the first component in the sum and by noticing that that cancels with the first term. The next equality follows by noticing that $\|\hat{\boldsymbol{b}}\|^2 = \|\underline{\hat{\boldsymbol{b}}}\|^2$. The final inequality follows because, from the facts that $\mu > w\alpha_1$ and that $\alpha_1 > \alpha_2 > \cdots > \alpha_n$, we can deduce that for each $\ell > 1$

$$\frac{w\alpha_\ell}{\mu - w\alpha_\ell} < \frac{w\alpha_\ell}{w\alpha_1 - w\alpha_\ell} = \frac{\alpha_\ell}{\alpha_1 - \alpha_\ell} < \frac{\alpha_2}{\alpha_1 - \alpha_2}$$

□

This concludes the proof of Proposition 2. □

Proof of Proposition 3. First, we state and prove a lemma.

Lemma 2. Under the conditions of Assumption 5, on any compact set the function $C^{-1}\kappa(C^{1/2}\boldsymbol{z})$ converges uniformly to $k\|\boldsymbol{z}\|^2$, as $C \downarrow 0$, where $k > 0$ is some constant. We call the limit G.

Proof. Consider the Taylor expansion of κ around $\mathbf{0}$ (κ is defined by part (1) of the assumption). We will now study its properties under parts (2) to (5) of Assumption 5. (5) ensures that the Taylor expansion exists. Local separability (4) says that there are no terms of the form $y_i y_j$. Non-negativity (3) (κ is nonnegative and $\kappa(\mathbf{0}) = 0$) implies that all first-order terms are zero. Also, (5) says that terms of the form y_i^2 must have positive coefficients, and symmetry (2) says that their coefficients must all be the same. □

Write $\boldsymbol{y} := \boldsymbol{b} - \hat{\boldsymbol{b}}$. Let $\Delta(\boldsymbol{y})$ denote the change in welfare from the status quo. Fix all parameters of the problem, and recall the main optimization problem:

$$\max_{\boldsymbol{b}} \ \Delta(\boldsymbol{y}) \tag{IT(C)}$$

$$\text{s.t. } \kappa(\boldsymbol{y}) \leq C$$

We maintain, but do not explicitly write, that welfare is evaluated at $\boldsymbol{a}^*(\boldsymbol{y})$, where $\boldsymbol{a}^* = [\boldsymbol{I} - \beta \boldsymbol{G}]^{-1}(\hat{\boldsymbol{b}} + \boldsymbol{y})$.

Let $\boldsymbol{y}(C)$ be the solution of problem IT(C), which is unique for small enough C. Then we claim that, as $C \downarrow 0$, we have

$$\frac{r_\ell^*}{r_{\ell'}^*} \to \frac{\alpha_\ell}{\alpha_{\ell'}},$$

where the similarity ratios are defined at the optimum $\boldsymbol{y}(C)$.

We will prove the result by studying an equivalent problem using Berge's Theorem of the Maximum. Let $\check{\boldsymbol{y}} = C^{-1/2}\boldsymbol{y}$. We will now define a rescaled version of the problem, $\check{\mathrm{IT}}(C)$.

$$\max_{\boldsymbol{b}} \; C^{-1}\Delta(C^{1/2}\check{\boldsymbol{y}}) \qquad (\check{\mathrm{IT}}(C))$$
$$\text{s.t. } C^{-1}\kappa(C^{1/2}\check{\boldsymbol{y}}) \leq 1.$$

This is clearly equivalent to the original problem. Let $\check{\boldsymbol{y}}^*(C)$ be the (possibly set-valued) solution for C.

The problem $\check{\mathrm{IT}}(C)$ is not yet defined at $C = 0$, but we now define it there. Let the objective at $C = 0$ be the limit of $C^{-1}\Delta(C^{1/2}\check{\boldsymbol{y}})$ as $C \downarrow 0$, which we call F. Let the constraint be $G(\check{\boldsymbol{y}}) \leq 1$, where G is from Lemma 2.

Let us restrict $\check{\mathrm{IT}}(C)$ to a compact set $\mathcal{K}$ such that the constraint set $\{\boldsymbol{y} : C^{-1}\kappa(C^{1/2}\check{\boldsymbol{y}}) \leq 1\}$ is contained in $\mathcal{K}$ for all small enough C. Now we claim that the conditions of Berge's Theorem of the Maximum are satisfied: The constraint correspondence is continuous at $C = 0$ because $C^{-1}\kappa(C^{1/2}\check{\boldsymbol{y}})$ converges uniformly to G, while the objective function is jointly continuous in its two arguments.

The Theorem of the Maximum therefore implies that the maximized value is continuous at $C = 0$. Because the convergence of the objective is actually uniform on $\mathcal{K}$ by the Lemma, this is possible if and only if $\check{\boldsymbol{y}}$ approaches the solution of the problem

$$\max_{\boldsymbol{b}} \; F(\check{\boldsymbol{y}})$$
$$\text{s.t. } \|\check{\boldsymbol{y}}\|^2 \leq 1.$$

By the same argument, the same point is the limit of the solutions to

$$\max_{\boldsymbol{b}} \; C^{-1}\Delta(C^{1/2}\check{\boldsymbol{y}})$$
$$\text{s.t. } \|\check{\boldsymbol{y}}\|^2 \leq 1.$$

By Proposition 1, in that limit this satisfies

$$\frac{r_\ell^*}{r_{\ell'}^*} \to \frac{\alpha_\ell}{\alpha_{\ell'}}.$$

□

Proof of Proposition 4. Using expression (10), we can write $\mathbb{E}\left[W(\boldsymbol{b};\boldsymbol{G})\right]$ determined by intervention $\mathcal{B}_{\boldsymbol{y}}$ as follows:

$$\mathbb{E}\left[W(\boldsymbol{b};\boldsymbol{G})\right] = w\sum_{\ell}\alpha_\ell\left(\left\{\mathbb{E}[\hat{\underline{b}}_\ell] + \underline{y}_\ell\right\}^2 + \mathrm{Var}[\underline{b}_\ell]\right).$$

Choosing $\boldsymbol{y}$ to maximize this is identical to the problem analyzed in the deterministic setting in the proof of Theorem 1. Thus, defining $x_\ell = \underline{y}_\ell/\bar{b}_\ell$, with $\bar{b}_\ell = \mathbb{E}[\hat{\underline{B}}_\ell]$, it satisfies the same conditions at the optimum as those derived in Theorem 1. □

Proof of Proposition 5. Given Assumption 7, without loss of generality we can normalize $\bar{\boldsymbol{b}} = \boldsymbol{0}$. Using expression (10) and normalization, we obtain that if the optimal solution is $\mathcal{B}^*$ the expected welfare obtained is

$$\mathbb{E}\left[W(\boldsymbol{b}^*;\boldsymbol{G})\right] = w\sum_{\ell}\alpha_\ell \mathrm{Var}(\underline{b}^*_\ell).$$

Note that the random variable $\underline{\mathcal{B}}^* = \boldsymbol{U}^\mathsf{T}\mathcal{B}^*$, and so the variance–covariance matrix of the random variable $\underline{\mathcal{B}}^*$ is $\boldsymbol{\Sigma}_{\underline{\mathcal{B}}^*} = \boldsymbol{U}^\mathsf{T}\boldsymbol{\Sigma}_{\mathcal{B}^*}\boldsymbol{U}$, where recall that $\boldsymbol{\Sigma}_{\mathcal{B}^*}$ is the variance–covariance matrix of the random variable $\mathcal{B}^*$.

We consider the case of $w > 0$ and $\beta > 0$; the proof of the other cases is analogous and therefore omitted. The expected welfare is a weighted sum of the variances of the principal components, $\mathrm{Var}(\underline{b}^*_\ell) = \mathrm{Var}(\boldsymbol{u}^\ell(\boldsymbol{G})\cdot\boldsymbol{b}^*)$, and the weight α_ℓ on the variance of principal component ℓ of $\boldsymbol{G}$ is an increasing function of its eigenvalue λ_ℓ, because $\beta > 0$.

Suppose the Proposition is violated, that is there exists a ℓ, ℓ' such that $\ell < \ell'$ and $\mathrm{Var}(\underline{b}^*_\ell) < \mathrm{Var}(\underline{b}^*_{\ell'})$. We construct an alternative intervention that has the same cost and does strictly better. Take the permutation matrix (and therefore an orthogonal matrix) $\boldsymbol{P}$ such that $P_{kk} = 1$ for all $k \notin \{\ell, \ell'\}$ and $P_{\ell\ell'} = P_{\ell'\ell} = 1$. Define $\mathcal{B}^{**} = \boldsymbol{O}\mathcal{B}^*$ with $\boldsymbol{O} = \boldsymbol{U}\boldsymbol{P}\boldsymbol{U}^\mathsf{T}$. Clearly, $\boldsymbol{O}$ is orthogonal, as $\boldsymbol{U}$ and $\boldsymbol{P}$ are both orthogonal. Hence, by Assumption 7, $K(\mathcal{D}_{\mathcal{B}^*}) = K(\mathcal{D}_{\mathcal{B}^{**}})$. Furthermore, the matrix

$$\boldsymbol{\Sigma}_{\underline{\mathcal{B}}^{**}} = \boldsymbol{P}\boldsymbol{\Sigma}_{\underline{\mathcal{B}}^*}\boldsymbol{P}^\mathsf{T}$$

and so $\mathrm{Var}(\underline{b}^{**}_k) = \mathrm{Var}(\underline{b}^*_k)$ for all $k \notin \{\ell, \ell'\}$ and $\mathrm{Var}(\underline{b}^{**}_\ell) = \mathrm{Var}(\underline{b}^*_{\ell'}) > \mathrm{Var}(\underline{b}^{**}_{\ell'}) = \mathrm{Var}(\underline{b}^*_\ell)$. Since $\alpha_\ell > \alpha_{\ell'}$ intervention $\mathcal{B}^{**}$ does strictly better than $\mathcal{B}^*$, a contradiction to our initial hypothesis that $\mathcal{B}^*$ was optimal. □

Appendix B. Online appendix

B.1. **Non-strategic externalities.** Section 4 characterizes optimal interventions for network games that satisfy Property A. We explain here how to extend the analysis beyond this assumption. We maintain Assumption 1 and Assumption 2. Recall that player i's utility for action profile $\boldsymbol{a}$ is

$$U_i(\boldsymbol{a}, \boldsymbol{G}) = \hat{U}_i(\boldsymbol{a}, \boldsymbol{G}) + P_i(\boldsymbol{a}_{-i}, \boldsymbol{G}, \boldsymbol{b})$$

where $\hat{U}_i(\boldsymbol{a}, \boldsymbol{G}) = a_i(b_i + \sum_j g_{ij} a_j) - \frac{1}{2} a_i^2$ and $P_i(\boldsymbol{a}_{-i}, \boldsymbol{G}, \boldsymbol{b})$ is a non-strategic externality term that takes the following form:

$$P_i(\boldsymbol{a}_{-i}, \boldsymbol{G}) = m_1 \sum_j g_{ij} a_j + m_2 \sum_j g_{ij} a_j^2 + m_3 \sum_{j \neq i} a_j + m_4 \left(\sum_{j \neq i} a_j \right)^2 + m_5 \sum_{j \neq i} a_j^2.$$

Here we have taken local and global externality terms given by second-order polynomials in actions. (We could also accommodate externalities that depend directly on the b_i in the same sort of way, as will become clear in the proof, but we omit this for brevity.)

The implication of Assumption 4 for our analysis is summarized next.

Lemma 3. Assumption 4 implies that:

1. for any $a \in \mathbb{R}^n$, $\sum_i \sum_j g_{ij} a_j = \sum_i a_i$ and $\sum_i \sum_j g_{ij} a_j^2 = \sum_i a_i^2$
2. $\lambda_1(\boldsymbol{G}) = 1$ and $u_i^1(\boldsymbol{G}) = \sqrt{n}$ for all i
3. $\sum_i a_i^* = \frac{1}{1-\beta} \sum b_i = \frac{\sqrt{n}}{1-\beta} \underline{b}_1 = \sqrt{n \alpha_1} \underline{b}_1$, where $\boldsymbol{a}^*$ is equilibrium action profile.[28]

The proof of Lemma 3 is immediate. Using part 1 of Lemma 3, and that individuals play an equilibrium (actions satisfy expression (3)), we obtain:

$$W(\boldsymbol{b}, \boldsymbol{G}) = w_1 \left(\boldsymbol{a}^*\right)^{\mathsf{T}} \boldsymbol{a}^* + \frac{w_2}{n} \left(\sum_i a_i^* \right)^2 + \frac{w_3}{\sqrt{n}} \sum_i a_i^*,$$

with:

$$\begin{aligned} w_1 &= 1 + m_2 + m_5 + (n-1) m_4 \\ w_2 &= n m_5 (n-2) \\ w_3 &= \sqrt{n}[m_1 + (n-1) m_3]. \end{aligned}$$

Using the decomposition $\boldsymbol{G} = \boldsymbol{U} \boldsymbol{\Lambda} \boldsymbol{U}^{\mathsf{T}}$, together with part 2 and part 3 of Lemma 3, we obtain:

$$W(\underline{\boldsymbol{b}}, \boldsymbol{G}) = w_1 \underline{\boldsymbol{a}}^{*\mathsf{T}} \underline{\boldsymbol{a}}^* + w_2 \alpha_1 \underline{b}_1^2 + w_3 \sqrt{\alpha_1} \underline{b}_1.$$

The intervention problem reads

$$\max_{\underline{\boldsymbol{b}}} \quad w_1 \underline{\boldsymbol{a}}^{*\mathsf{T}} \underline{\boldsymbol{a}}^* + w_2 \alpha_1 \underline{b}_1^2 + w_3 \sqrt{\alpha_1} \underline{b}_1$$

[28] The last equality follows because $\alpha_1 = 1/(1-\beta\lambda_1)^2$, and assumption 4 implies that $\lambda_1 = 1$.

$$\text{subject to} \quad \underline{a}_\ell^* = \sqrt{\alpha_\ell}\underline{b}_\ell$$
$$\sum_\ell (\underline{b}_\ell - \hat{\underline{b}}_\ell)^2 \leq C.$$

Using the expression for equilibrium actions, we obtain:

$$\max_{\underline{\boldsymbol{b}}} \quad w_1 \sum_{\ell=1} \alpha_\ell \underline{b}_\ell^2 + w_2 \alpha_1 \underline{b}_1^2 + w_3 \sqrt{\alpha_1}\underline{b}_1$$
$$\text{subject to} \quad \sum_\ell (\underline{b}_\ell - \hat{\underline{b}}_\ell)^2 \leq C.$$

Recalling the definition $x_\ell = \frac{\underline{b}_\ell - \hat{\underline{b}}_\ell}{\hat{\underline{b}}_\ell}$ for every ℓ, we finally rewrite the problem as:

$$\max_{\boldsymbol{x}} \quad w_1 \sum_{\ell=1} \alpha_\ell \hat{\underline{b}}_\ell^2 (1+x_\ell)^2 + w_2 \alpha_1 \hat{\underline{b}}_1^2 (1+x_1)^2 + w_3 \sqrt{\alpha_1}\hat{\underline{b}}_1 (1+x_1)$$
$$\text{subject to} \quad \sum_\ell \hat{\underline{b}}_\ell^2 x_\ell^2 \leq C.$$

Theorem 2 characterizes optimal interventions for two cases: (i) $w_1 \geq 0$ and (ii) $w_1 < 0$ and $\sum_{\ell=2} \hat{\underline{b}}_\ell^2 > C$. The extension of the analysis for the remaining case $w_1 < 0$ and $\sum_{\ell=2} \hat{\underline{b}}_\ell^2 < C$ is explained in Remark 1, which is presented after the proof of Theorem 2. Taken together, Theorem 2, and Remark 1 following it, constitute our extension of Theorem 1 to games that do not satisfy Property A.

Theorem 2. Suppose Assumptions 1, 2 and 4 hold. Suppose that either: (i) $w_1 \geq 0$ or that (ii) $w_1 < 0$ and $\sum_{\ell=2} \hat{\underline{b}}_\ell^2 > C$. The optimal intervention is characterized as follows:

1. $$x_1^* = \frac{\alpha_1}{\mu - (w_1 + w_2)\alpha_1}\left[w_1 + w_2 + \frac{w_3}{2\sqrt{\alpha_1}\hat{\underline{b}}_1}\right],$$
   and, for all $\ell \geq 2$,
   $$x_\ell^* = \frac{w_1 \alpha_\ell}{\mu - w_1 \alpha_\ell}.$$
   The shadow price of the planner's budget, $\mu > (w_1 + w_2)\alpha_1$, is uniquely determined as the solution of:
   $$\sum_{\ell=2} \hat{\underline{b}}_\ell^2 \left(\frac{w_1 \alpha_\ell}{\mu - w_1 \alpha_\ell}\right)^2 + \hat{\underline{b}}_1^2 \left(\frac{\alpha_1}{\mu - (w_1 + w_2)\alpha_1}\right)^2 \left[w_1 + w_2 + \frac{w_3}{2\sqrt{\alpha_1}\hat{\underline{b}}_1}\right]^2 = C$$
2. a. For all $\ell \neq 1$, $x_\ell^* > 0$ if and only if $w_1 > 0$;
   b. $x_1^* > 0$ if and only if $w_1 + w_2 + \frac{w_3}{2\sqrt{\alpha_1}\hat{\underline{b}}_1} > 0$. 2b. If the game has strategic complements, $\beta > 0$, then $|x_2^*| > |x_3^*| > \cdots > |x_n^*|$. If the game has strategic substitutes, $\beta < 0$, then $|x_2^*| < |x_3^*| < \cdots < |x_n^*|$.

3. Suppose $w_1 \neq 0$. In the limit as $C \to 0$, $\mu \to \infty$ and:

$$\begin{aligned} \frac{x_\ell^*}{x_{\ell'}^*} &\to \frac{\alpha_\ell}{\alpha_{\ell'}} \text{ for all } \ell, \ell' \neq 1 \\ \frac{x_1^*}{x_\ell^*} &\to \frac{\alpha_1}{\alpha_\ell}\left[w_1 + w_2 + \frac{w_3}{2\sqrt{\alpha_1}\hat{\underline{b}}_1}\right] \text{ for all } \ell \neq 1 \end{aligned}$$

4. Suppose the game has strategic complements, $\beta > 0$. In the limit as $C \to \infty$, $\mu \to \max\{w_1\alpha_2, (w_1 + w_2)\alpha_1\}$, and
   a. If $w_1\alpha_2 > (w_1 + w_2)\alpha_1$ then

$$\begin{aligned} x_1^* &\to \frac{\alpha_1}{w_1\alpha_2 - (w_1 + w_2)\alpha_1}\left[w_1 + w_2 + \frac{w_3}{2\hat{\underline{b}}_1\sqrt{\alpha_1}}\right], \\ |x_2^*| &\to \infty, \\ |x_\ell^*| &\to \frac{\alpha_\ell}{\alpha_2 - \alpha_\ell} \text{ for all } \ell > 2. \end{aligned}$$

   b. If $w_1\alpha_2 < (w_1 + w_2)\alpha_1$ then

$$\begin{aligned} |x_1^*| &\to \infty \\ x_\ell^* &\to \frac{w_1\alpha_\ell}{(w_1 + w_2)\alpha_1 - w_1\alpha_\ell} \text{ for all } \ell \geq 2. \end{aligned}$$

5. Suppose the game has strategic substitutes, $\beta < 0$. In the limit as $C \to \infty$, $\mu \to \max\{w_1\alpha_n, (w_1 + w_2)\alpha_1\}$. Hence:
   a. If $w_1\alpha_n > (w_1 + w_2)\alpha_1$ then:

$$\begin{aligned} x_1^* &\to \frac{\alpha_1}{w_1\alpha_n - (w_1 + w_2)\alpha_1}\left[w_1 + w_2 + \frac{w_3}{2\hat{\underline{b}}_1\sqrt{\alpha_1}}\right], \\ |x_\ell^*| &\to \frac{\alpha_\ell}{\alpha_n - \alpha_\ell} \text{ for all } \ell \in \{2, \ldots, n-1\}, \\ |x_n^*| &\to \infty. \end{aligned}$$

   b. If $w_1\alpha_n < (w_1 + w_2)\alpha_1$ then

$$\begin{aligned} |x_1^*| &\to \infty \\ x_\ell^* &\to \frac{w_1\alpha_\ell}{(w_1 + w_2)\alpha_1 - w_1\alpha_\ell} \text{ for all } \ell \geq 2. \end{aligned}$$

Before the proof, we briefly explain the sense in which this extends Theorem 1 and associated results in the basic model. The formula for x_ℓ^* in part 1 is a direct generalization of equation (5), with the shadow price characterized by an equation analogous to (6). The monotonicity relations on x_ℓ^* in part 2 correspond to Corollary 1. The small-C analysis of part 3 corresponds to Proposition 1. The large-C analysis in parts 4 and 5 corresponds to the limits studied in Section 4.1.

Proof of Theorem 2. **Part 1.** For a given $\boldsymbol{x} \in \mathbb{R}^n$, define

$$\begin{aligned} K(x_1) &= (w_1 + w_2)\alpha_1 \hat{\underline{b}}_1^2 (1 + x_1)^2 + w_3 \sqrt{\alpha_1}\hat{\underline{b}}_1 (1 + x_1) \\ C(x_1) &= C - \hat{\underline{b}}_1^2 x_1^2. \end{aligned}$$

The maximization problem can be rewritten as:

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & w_1 \sum_{\ell=2} \alpha_\ell \hat{\underline{b}}_\ell^2 (1 + x_\ell)^2 + K(x_1) \\ \text{subject to} \quad & \sum_{\ell=2} \hat{\underline{b}}_\ell^2 x_\ell^2 \leq C(x_1) \end{aligned}$$

We solve this problem in two steps.

First Step. We fix x_1 so that $C(x_1) \geq 0$; that is, $x_1 \in [-C/\hat{\underline{b}}_1, C/\hat{\underline{b}}_1]$. We then solve

$$\begin{aligned} \max_{\boldsymbol{x}_{-1}} \quad & w_1 \sum_{\ell=2} \alpha_\ell \hat{\underline{b}}_\ell^2 (1 + x_\ell)^2 \\ \text{subject to} \quad & \sum_{\ell=2} \hat{\underline{b}}_\ell^2 x_\ell^2 \leq C(x_1) \end{aligned}$$

In the case in which $w_1 = 0$ we skip this first step. If $w_1 \neq 0$, then we argue in a way exactly analogous to the proof of Theorem 1 that for all $\ell \neq 1$,

$$x_\ell^* = \frac{w_1 \alpha_\ell}{\mu - w_1 \alpha_\ell}$$

where, for all $\ell \neq 1$, $\mu \geq w_1 \alpha_\ell$ and it solves

$$\sum_{\ell=2} \hat{\underline{b}}_\ell^2 \left(\frac{w_1 \alpha_\ell}{\mu - w_1 \alpha_\ell} \right)^2 = C(x_1).$$

Note that, for all $\ell \geq 2$, $x_\ell^* > 0$ if $w_1 > 0$ and $x_\ell^* < 0$ if $w_1 < 0$.

Note also that if $w_1 < 0$ the constraint binds: the bliss point ($x_\ell^* = -1$ for all $\ell \neq 1$) cannot be achieved because $C < \sum_{\ell=2}^n \hat{\underline{b}}_\ell^2$.

Second Step. Substituting into the objective function the expression for x_ℓ^*, for all $\ell \geq 2$, we obtain:

$$\begin{aligned} \max_{x_1} \quad & W = w_1 \sum_{\ell=2} \alpha_\ell \hat{\underline{b}}_\ell^2 \left(\frac{\mu}{\mu - w_1 \alpha_\ell} \right)^2 + K(x_1) \\ \text{subject to} \quad & \sum_{\ell=2} \hat{\underline{b}}_\ell^2 \left(\frac{w_1 \alpha_\ell}{\mu - w_1 \alpha_\ell} \right)^2 = C(x_1) \\ & x_1 \in \left[-\frac{C}{\hat{\underline{b}}_1}, \frac{C}{\hat{\underline{b}}_1} \right] \end{aligned}$$

The following lemma is instrumental to the solution of this problem. It characterizes μ, which is implicitly a function of x_1.

Lemma 4. From the budget constraint in the above problem it follows that

1. $\lim_{x_1 \to -\sqrt{C}/\hat{\underline{b}}_1} \mu = \lim_{x_1 \to \sqrt{C}/\hat{\underline{b}}_1} \mu = \infty$
2. $$\frac{d\mu}{dx_1} = \frac{\hat{\underline{b}}_1^2 x_1}{\sum_{\ell=2} \frac{w_1^2 \hat{\underline{b}}_1^2 \alpha_\ell^2}{(\mu - w_1\alpha_\ell)^3}}$$
3. $\frac{d\mu}{dx_1} > 0$ if $x_1 > 0$ and $\frac{d\mu}{dx_1} < 0$ if $x_1 < 0$;
4. $\lim_{x_1 \to -\sqrt{C}/\hat{\underline{b}}_1} \frac{d\mu}{dx_1} = -\infty$ and $\lim_{x_1 \to \sqrt{C}/\hat{\underline{b}}_1} \frac{d\mu}{dx_1} = \infty$.

Proof of Lemma 4. The proof of part 1 of Lemma 4 follows directly by inspection of the budget constraint. Expression 2 in part 2 of Lemma 4 is derived by implicit differentiation of the budget constraint. Part 3 and part 4 of Lemma 4 follow by inspection of the expression in part 2, and the fact that $\mu > w_1\alpha_\ell$. This concludes the proof of Lemma 4. □

Lemma 4 implies that μ as a function of $x_1 \in \left[-C/\hat{\underline{b}}_1, C\hat{\underline{b}}_1\right]$ is U-shaped; the slope is $-\infty$ at $x_1 = -C/\hat{\underline{b}}_1$ and $+\infty$ at $x_1 = C/\hat{\underline{b}}_1$; and it reaches a minimum at $x_1 = 0$.

For $w_1 \neq 0$, taking the derivative of the objective function W in expression (17) with respect to x_1, we obtain:

$$\frac{dW}{dx_1} = -2\mu \sum_{\ell=2} \frac{w_1^2 \hat{\underline{b}}_1^2 \alpha_\ell^2}{(\mu - w_1\alpha_\ell)^3} \frac{d\mu}{dx_1} + 2(w_1 + w_2)\alpha_1 \hat{\underline{b}}_1^2 (1 + x_1) + w_3\sqrt{\alpha_1}\hat{\underline{b}}_1.$$

Plugging in expression for $\frac{d\mu}{dx_1}$ in part 2 of Lemma 4 we obtain that:

$$\frac{dW}{dx_1} = -2\mu\hat{\underline{b}}_1^2 x_1 + 2(w_1 + w_2)\alpha_1 \hat{\underline{b}}_1^2 (1 + x_1) + w_3\sqrt{\alpha_1}\hat{\underline{b}}_1.$$

Part 1 of Lemma 4 implies that $\frac{dW}{dx_1} \to \infty$ when $x_1 \to -\sqrt{C}/\hat{\underline{b}}_1$, whereas $\frac{dW}{dx_1} \to -\infty$ when $x_1 \to \sqrt{C}/\hat{\underline{b}}_1$. Hence, the optimal x_1 must be interior, which implies that $\frac{dW}{dx_1} = 0$ or, equivalently:

$$x_1^* = \frac{\alpha_1}{\mu - (w_1 + w_2)\alpha_1} \left[w_1 + w_2 + \frac{w_3}{2\sqrt{\alpha_1}\hat{\underline{b}}_1} \right].$$

Substituting x_1^*, in the budget constraint

$$\sum_{\ell=2} \hat{\underline{b}}_\ell^2 \left(\frac{w_1\alpha_\ell}{\mu - w_1\alpha_\ell} \right)^2 = C(x_1^*),$$

we obtain that the Lagrange multiplier μ must solve:

$$\sum_{\ell=2} \hat{\underline{b}}_\ell^2 \left(\frac{w_1\alpha_\ell}{\mu - w_1\alpha_\ell} \right)^2 + \hat{\underline{b}}_1^2 \left(\frac{\alpha_1}{\mu - (w_1 + w_2)\alpha_1} \right)^2 \left[w_1 + w_2 + \frac{w_3}{2\sqrt{\alpha_1}\hat{\underline{b}}_1} \right]^2 = C.$$

The conclusion for $w_1 = 0$ are obtained by taking the limits as $w_1 \to 0$ of the expression x_1^* and the expression determining μ. This concludes the proof of part 1 of Theorem 2.

Part 2. We have already proved that, for all $\ell \geq 2$, $x_\ell^* > 0$ if and only if $w_1 > 0$. We now claim that $x_1^* > 0$ if and only if $w_1 + w_2 + \frac{w_3}{2\hat{\underline{b}}_1\sqrt{\alpha_1}} > 0$. Suppose, toward a contradiction, that $x_1^* < 0$. Suppose, toward a contradiction, that $x_1^* < 0$. By inspection of the maximization problem

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & w_1 \sum_{\ell=2} \alpha_\ell \hat{\underline{b}}_\ell^2 (1+x_\ell)^2 + K(x_1) \\ \text{subject to} \quad & \sum_{\ell=2} \hat{\underline{b}}_\ell^2 x_\ell^2 \leq C(x_1) \end{aligned}$$

note that if $w_1 + w_2 + \frac{w_3}{2\hat{\underline{b}}_1\sqrt{\alpha_1}} > 0$ and $x_1^* < 0$, then, by flipping the sign of x_1^*, $K(x_1)$ increases and the constraint is unaltered; this is a contradiction to our initial assumption that x_1^* was optimal.

We have just established that $x_1^* > 0$. Now, by (B.1) above, $x_1^* > 0$ if and only if $w_1 + w_2 + \frac{w_3}{2\hat{\underline{b}}_1\sqrt{\alpha_1}} > 0$. And since

$$x_1^* = \frac{\alpha_1}{\mu - (w_1 + w_2)\alpha_1} \left[w_1 + w_2 + \frac{w_3}{2\sqrt{\alpha_1}\hat{\underline{b}}_1} \right]$$

it follows that $\mu > \alpha_1(w_1 + w_2)$. Finally, if the game has strategic complements then $\alpha_2 > \cdots > \alpha_n$ and so $|x_2^*| > |x_3^*| > \cdots > |x_n^*|$, and if the game has strategic substitutes then $\alpha_2 < \cdots < \alpha_n$ and so $|x_2^*| < |x_3^*| < \cdots < |x_n^*|$.

Part 3. This follows by using the characterization in part 1 and by noticing that if $C \to 0$ then $\mu \to \infty$.

Part 4 and Part 5. Both parts follow by using the characterization together with the following fact, which we will now establish.

$$\lim_{C \to \infty} \mu = \max\{w_1 \max\{\alpha_2, \alpha_n\}, (w_1 + w_2)\alpha_1\}.$$

To show this, recall from above that we have the following equation for the Lagrange multiplier:

$$\sum_{\ell=2} \hat{\underline{b}}_\ell^2 \left(\frac{w_1 \alpha_\ell}{\mu - w_1 \alpha_\ell} \right)^2 + \hat{\underline{b}}_1^2 \left(\frac{\alpha_1}{\mu - (w_1 + w_2)\alpha_1} \right)^2 \left[w_1 + w_2 + \frac{w_3}{2\sqrt{\alpha_1}\hat{\underline{b}}_1} \right]^2 = C$$

If C tends to ∞ it must be that either the first denominator $(\mu - w_1\alpha_\ell)$ or the second denominator $(\mu - (w_1 + w_2)\alpha_1)$ tends to zero. Concerning the first one, this is true if either $w_1\alpha_2$ or $w_1\alpha_n$ (depending on which one is positive) approaches μ. The second denominator tends to 0 if $(w_1 + w_2)\alpha_1$ tends to μ. Both denominators are positive by definition of the Lagrange multiplier, so it will be the greater of $w_1 \max\{\alpha_2, \alpha_n\}$ and $(w_1 + w_2)\alpha_1$ which tends to μ. This concludes the proof of Theorem 2. □

A special case of Theorem 2 is one where the planner wants to maximize the sum of equilibrium actions. This occurs when $w_1 = w_2 = 0$. In this case we obtain

Corollary 2. Suppose Assumption 1, 2 and 4 hold. Suppose that $w_1 = w_2 = 0$ and $w_3 > 0$, i.e., the planner wants to maximize the sum of equilibrium actions. Then the optimal intervention is $\boldsymbol{b}^* = \hat{\boldsymbol{b}} + \boldsymbol{u}^1\sqrt{C}$.

Remark 1. Suppose $w_1 < 0$ and $\sum_{\ell=2} \underline{b}_\ell^2 < C$, in contrast to what was assumed in the theorem. If x_1 is sufficiently small, the solution in Step 1 in the proof of Theorem 2 entails $x_\ell = -1$ for all $\ell \geq 2$. That is, fixing x_1, the bliss point can be achieved with the remaining budget after x_1 is paid for, $C(x_1)$. This implies that when we move to Step 2 and optimize over x_1, we need to take into account that, for small values of x_1. Step 1 yields a corner solution. Hence, the analysis of how the network multiplier changes when x_1 changes will need to be adapted to take this fact into account.

Example 4, continued. Social interaction and peer effects

We conclude this Appendix by applying Theorem 2 to Example 4 from Section 5.1. In this example $w_1 = 1$, $w_2 = 0$ and $w_3 = -\gamma\sqrt{n}(n-1)$.

Corollary 3. The optimal intervention in Example 4 is characterized by

$$x_1^* = \frac{\alpha_1}{\mu - \alpha_1}\left[1 - \gamma\frac{\sqrt{n}(n-1)}{2\sqrt{\alpha_1}\hat{\underline{b}}_1}\right]$$

and, for all $\ell \geq 2$:

$$x_\ell^* = \frac{\alpha_\ell}{\mu - \alpha_\ell}$$

where the Lagrange multiplier μ solves

$$\sum_{\ell=2} \hat{\underline{b}}_\ell^2 \left(\frac{\alpha_\ell}{\mu - \alpha_\ell}\right)^2 + \hat{\underline{b}}_1^2 \left(\frac{\alpha_1}{\mu - \alpha_1}\right)^2 \left[1 - \gamma\frac{\sqrt{n}(n-1)}{2\sqrt{\alpha_1}\hat{\underline{b}}_1}\right]^2 = C.$$

Corollary 4. Consider the optimal intervention in Example 4. It has the following properties.

1. $x_2^* > \cdots > x_n^* > 0$; $x_1^* > 0$ if and only $\gamma < \frac{2\sqrt{\alpha_1}\hat{\underline{b}}_1}{\sqrt{n}(n-1)}$
2. If $C \to 0$

$$\frac{x_\ell^*}{x_{\ell'}^*} \to \frac{\alpha_\ell}{\alpha_{\ell'}}, \text{ for all } \ell, \ell' \neq 1$$

$$\frac{x_1^*}{x_\ell^*} \to \frac{\alpha_1}{\alpha_\ell}\left[1 - \gamma\frac{\sqrt{n}(n-1)}{2\sqrt{\alpha_1}\hat{\underline{b}}_1}\right], \text{ for all } \ell \neq 1$$

3. If $C \to \infty$ then $|x_1^*| \to \infty$ and $x_\ell^* \to \frac{\alpha_\ell}{(\alpha_1 - \alpha_\ell)}$ for all $\ell \geq 2$.

B.2. **Beauty contest with local interactions.** Each individual has a preferred action, b_i, but also cares about coordinating with her neighbors. The utility of individual i is

$$U_i(\boldsymbol{a}, \boldsymbol{G}) = -(a_i - b_i)^2 - \gamma \sum_j g_{ij}(a_j - a_i)^2.$$

(In Morris and Shin (2002) and Angeletos and Pavan (2007), every individual's target action is the same, b, but there is uncertainty in b. Furthermore, in their environment each individual cares about coordinating with the action of all other individuals.) We maintain the assumption that $\sum_j g_{ij} = 1$ for all i.

Denote by $\beta = \frac{\gamma}{1+\gamma}$. Expanding the utility above we can write

$$\frac{U_i(\boldsymbol{a}, \boldsymbol{G})}{2(1+\gamma)} = a_i \left(\frac{b_i}{1+\gamma} + \beta \sum_j g_{ij} a_j \right) - \frac{1}{2} a_i^2 - \frac{1}{2(1+\gamma)} [b_i^2 + \gamma \sum_j g_{ij} a_j^2]$$

The Nash equilibrium conditions read

$$[\boldsymbol{I} - \beta \boldsymbol{G}] \boldsymbol{a}^* = \frac{1}{1+\gamma} \boldsymbol{b}$$

and player i's equilibrium utility is

$$U_i(\boldsymbol{a}^*, \boldsymbol{G}) = (1+\gamma)[a_i^*]^2 - b_i^2 - \gamma \sum_j g_{ij} a_j^2$$

We obtain that

$$W(\boldsymbol{b}, \boldsymbol{G}) = (\boldsymbol{a}^*)^{\mathsf{T}} \boldsymbol{a}^* - \boldsymbol{b}^{\mathsf{T}} \boldsymbol{b}$$

The intervention problem of the planner is

$$\max_{\boldsymbol{b}} W(\boldsymbol{b}, \boldsymbol{G}) \tag{IT-BC}$$
$$\text{s.t. } [\boldsymbol{I} - \beta \boldsymbol{G}] \boldsymbol{a}^* = \frac{1}{1+\gamma} \boldsymbol{b},$$
$$K(\boldsymbol{b}; \hat{\boldsymbol{b}}) = \sum_{i \in \mathcal{N}} \left(b_i - \hat{b}_i \right)^2 \leq C,$$

We can solve problem (IT-BC) using our method. Specifically, we rewrite:

$$W(\underline{\boldsymbol{b}}, \boldsymbol{G}) = [\underline{\boldsymbol{a}}^{*\mathsf{T}} \underline{\boldsymbol{a}}^* - \underline{\boldsymbol{b}}^{\mathsf{T}} \underline{\boldsymbol{b}}]$$

with equilibrium actions projected in the principal components being equal to:

$$\underline{a}_\ell^* = \frac{1}{(1+\gamma)(1-\beta\lambda_\ell)} \underline{b}_\ell = \frac{1}{(1+\gamma-\gamma\lambda_\ell)} \underline{b}_\ell$$

where the second equality follows by using the definition of β. Hence, define $\alpha_\ell = \frac{1}{(1+\gamma-\gamma\lambda_\ell)^2}$, we obtain

$$W(\underline{\boldsymbol{b}}, \boldsymbol{G}) = \sum_\ell (\alpha_\ell - 1) \underline{b}_\ell^2$$

Since $\lambda_1 = 1$ it follows that $\alpha_1 - 1 = 0$. Moreover, $\alpha_\ell < 1$ for all ℓ and $\alpha_2 > \alpha_3 > \ldots > \alpha_n$ and so

$$W(\underline{\boldsymbol{b}}, \boldsymbol{G}) = -\sum_{\ell=2}(1-\alpha_\ell)\underline{b}_\ell^2$$

or

$$W(\underline{\boldsymbol{x}}, \boldsymbol{G}) = -\sum_{\ell=2}(1-\alpha_\ell)(1+x_\ell)^2$$

where $x_\ell = (\underline{b}_\ell - \hat{\underline{b}}_\ell)/\hat{\underline{b}}_\ell$. The respective budget constraint is

$$\sum_\ell \hat{\underline{b}}_\ell^2 x_\ell^2 \leq C$$

Hence, at the optimal intervention the first component is unchanged, $x_1^* = 0$. Intuitively, the planner would like to align individuals' ideal points so that each individual can match it and in the same time coordinate with others. The first component is a scalar of the sum of ideal points and its value does not affect expected welfare. When $C > \sum_{\ell \neq 1} \hat{\underline{b}}_\ell^2$ the planner can set $x_\ell = -1$ for all $\ell \neq 1$ and in this way the first best is achieved. When $C < \sum_{\ell \neq 1} \hat{\underline{b}}_\ell^2$ the planner cannot achieve the first best. In this case the optimal intervention has the property that for all $\ell > 1$, $x_\ell^* < 0$ and $|x_2^*| < |x_3^*| < \cdots < |x_n^*|$.

B.3. **More general costs of intervention.** We conclude by imposing an additional restriction on the structure of the costs of intervention and we show that this new restriction together with Assumption 5 fully characterizes the cost functions that we used in our main analysis.

Assumption 8. There is a function $f : \mathbb{R}_+ \to \mathbb{R}_+$ so that $\kappa(s\boldsymbol{y}) = f(s)\kappa(\boldsymbol{y})$.

Proposition 6. Consider a cost function that satisfies Assumptions 5 and 8. There is a function $f : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\kappa(\boldsymbol{y}) = f(\|\boldsymbol{y}\|).$$

Proposition 6 says that the costs of intervention $\boldsymbol{y}$ are the same as the cost of an intervention obtained as an orthogonal transformation of $\boldsymbol{y}$; that is $\kappa(\boldsymbol{y}) = \kappa(\boldsymbol{O}\boldsymbol{y})$ with $\boldsymbol{O}$ be an orthogonal matrix. This allows to rewrite the intervention problem using the orthogonal decomposition of welfare and costs that we employ in Section 4, and all the results developed there extends to this more general environment.

B.4. **Linear costs of intervention.** In this section we solve optimal intervention in Example 1 and for the case of a linear costs of intervention. We also retain Assumption 1 and Assumption 2. The analysis can be easily extended to general network games. We consider the following intervention problem:

$$\max_{\boldsymbol{b}} \; (\boldsymbol{a}^*)^\mathsf{T} \boldsymbol{a}^* \qquad \text{(IT-Linear Cost)}$$

$$\text{s.t. } \boldsymbol{a}^* = [\boldsymbol{I} - \beta \boldsymbol{G}]^{-1}\boldsymbol{b},$$
$$K(\boldsymbol{b}; \hat{\boldsymbol{b}}) = \sum_{i \in \mathcal{N}} |b_i - \hat{b}_i| \leq C,$$

Proposition 7. The solution to problem IT-Linear Cost has the property that there exists i^* such that $b_i^* \neq \hat{b}_{i^*}$ and $b_i^* = \hat{b}_i$ for al $i \neq i^*$.

Proof of Proposition 7. Define $W(\boldsymbol{b}) = \boldsymbol{a}(\boldsymbol{b})^\mathsf{T}\boldsymbol{a}(\boldsymbol{b})$. Let F be the set of *feasible* $\boldsymbol{b}$, those satisfying the budget constraint $K(\boldsymbol{b}; \hat{\boldsymbol{b}}) \leq C$. Suppose the conclusion does not hold and let $\boldsymbol{b}^*$ be the optimum, with $W^* = W(\boldsymbol{b}^*)$. Then, because by hypothesis the optimum is not at an extreme point, F contains a line segment L such that $\boldsymbol{b}^*$ is in the interior of L.[29]

Now restrict attention to a plane P containing this L and the origin. Note that L is contained in a convex set

$$E = \{\boldsymbol{b} : W(\boldsymbol{b}) \leq W^*\}.$$

The point $\boldsymbol{b^*}$ is contained in the interior of L; thus $\boldsymbol{b^*}$ is in the interior of E. On the other hand, $\boldsymbol{b^*}$ must be on the (elliptical) boundary of E because U is strictly increasing in each component (by irreducibility of the network) and continuous. This is a contradiction. □

We now characterize the optimal target for the case of strategic complements, i.e., $\beta > 0$. Remark 2 explains how to extend the analysis for the case of strategic substitutes.

In the case of strategic complements, it is clear that the planner uses all the budget C to increase the standalone marginal benefit of i^*, i.e., $b_i^* = \hat{b}_i + C$; reducing someone's effort can never help. Thus, the planner changes the status quo $\hat{\boldsymbol{b}}$ into $\boldsymbol{b} = \hat{\boldsymbol{b}} + C\boldsymbol{1}_{i^*}$ where $\boldsymbol{1}_{i^*}$ is a vector of 0 except for entry i^* that takes value 1. Let $\boldsymbol{a}(\boldsymbol{1}_i)$ be the Nash equilibrium when all individuals have $b_j = 0$ and $b_i = 1$, i.e., $\boldsymbol{a}(\boldsymbol{1}_i) = [\boldsymbol{I} - \beta\boldsymbol{G}]^{-1}\boldsymbol{1}_i$. It is easy to verify that the solution to problem IT-Linear Cost is:

$$i^* = \underset{i}{\operatorname{argmax}} \left\{ \boldsymbol{a}(\hat{\boldsymbol{b}} + C\boldsymbol{1}_i)^\mathsf{T}\boldsymbol{a}(\hat{\boldsymbol{b}} + C\boldsymbol{1}_i) - \boldsymbol{a}(\hat{\boldsymbol{b}})^\mathsf{T}\boldsymbol{a}(\hat{\boldsymbol{b}}) \right\}.$$

This is equivalent to

$$i^* = \underset{i}{\operatorname{argmax}} \left\{ C\|\boldsymbol{a}(\boldsymbol{1}_i)\| \left[2\|\boldsymbol{a}(\hat{\boldsymbol{b}})\|\rho(\boldsymbol{a}(\boldsymbol{1}_i), \boldsymbol{a}(\hat{\boldsymbol{b}})) + C\|\boldsymbol{a}(\boldsymbol{1}_i)\| \right] \right\}. \tag{17}$$

where recall that $\rho(\boldsymbol{a}(\boldsymbol{1}_i), \boldsymbol{a}(\hat{\boldsymbol{b}}))$ is the cosine similarity between vectors $\boldsymbol{a}(\boldsymbol{1}_i)$ and $\boldsymbol{a}(\hat{\boldsymbol{b}})$. There are two characteristics of a player that determines whether the player is a good target.

The first characteristic is $\|\boldsymbol{a}(\boldsymbol{1}_i)\|$. This is the square root of the aggregate equilibrium utility in the game with $\boldsymbol{b} = \boldsymbol{1}_i$, i.e., the squared root of $\boldsymbol{a}(\boldsymbol{1}_i)^\mathsf{T}\boldsymbol{a}(\boldsymbol{1}_i)$. So, a player with a high $\|\boldsymbol{a}(\boldsymbol{1}_i)\|$ is a player who induces a large welfare in the game in which he is the only player with positive standalone marginal benefit. We call this the *welfare centrality* of an individual. It is

[29]Formally, for some $z > 0$ there is a linear map $\varphi : [-z, z] \to F$ such that $\varphi(0) = \boldsymbol{b}^*$.

convenient to express the welfare centrality of individual i in terms of principal components of $\boldsymbol{G}$. Note that

$$\|\boldsymbol{a}(\mathbf{1}_i)\| = \|\underline{\boldsymbol{a}}(\underline{\mathbf{1}}_i)\| = \sqrt{\sum_\ell \alpha_\ell (u_i^\ell)^2}.$$

Recall that under strategic complement $\alpha_1 > \alpha_2 > .. > \alpha_n$ and so an individual with a high welfare centrality is one that is highly represented in the main principal components of the network.

The second factor is $\rho(\boldsymbol{a}(\mathbf{1}_i), \boldsymbol{a}(\hat{\boldsymbol{b}}))$. This measures the vector similarity between (i) the equilibrium action profile in the game with $\boldsymbol{b} = \mathbf{1}_i$; and (ii) the status quo equilibrium action profile. A player with a large $\rho(\boldsymbol{a}(\mathbf{1}_i), \boldsymbol{a}(\hat{\boldsymbol{b}}))$ is a player that, in the game in which he is the only player with positive standalone marginal benefit, leads a distribution of effort similar to the distribution of effort in the status quo.

Small C. Suppose $C \approx 0$. Then the optimal target is selected based on the first term of expression (17); that is:

$$i^* = \underset{i}{\operatorname{argmax}} \|\boldsymbol{a}(\mathbf{1}_i)\| \rho(\boldsymbol{a}(\mathbf{1}_i), \boldsymbol{a}(\hat{\boldsymbol{b}}))$$

For small budgets, the optimal intervention focuses on the player who has a large welfare centrality and that, in the same time, leads to a distribution of effort not too different from the status quo equilibrium effort.

Large C. Suppose $C \approx 0$. Then the optimal target is selected based on the first term of expression (17); that is:

$$i^* = \underset{i}{\operatorname{argmax}} \|\boldsymbol{a}(\mathbf{1}_i)\| \rho(\boldsymbol{a}(\mathbf{1}_i), \boldsymbol{a}(\hat{\boldsymbol{b}}))$$

For small budgets, the optimal intervention focuses on the player who has a large welfare centrality and that, in the same time, leads to a distribution of effort not too different from the status quo equilibrium effort.

Large C. For C sufficiently large, the last term of expression (17) dominates and therefore the player that is targeted is the player with the higest welfare centrality.

Remark 2 (Extension to the case of strategic substitutes)**.** In the case of strategic substitutes, we know for the targeted player i^*, $b^*_{i^*} = \hat{b}_i \pm C$, but we cannot say, a priori, which (positive or negative), and indeed it is easy to provide examples that both can happen. Under this qualification, the analysis developed for the case of strategic complements extends

B.5. **Monetary intervention.** We provide an outline of the analysis of optimal intervention in this game. It is immediate to see that the best reply of each individual in population i is a cutoff strategy: there exists a cutoff $a_i \in \mathcal{I}$ so that $q(\tau_i) = 1$ for all $\tau_i \leq a_i$ and $q(\tau_i) = 0$ otherwise. The equilibrium condition for these cutoffs is that, for all $i \in \mathcal{N}$,

$$\widetilde{\beta} \sum_j g_{ij} P[\tau_j \leq a_j^*] + b_i - a_i^* = 0 \quad \Longleftrightarrow \quad a_i = b_i + \frac{\widetilde{\beta}}{\overline{\tau}} \sum_j g_{ij} a_j^*.$$

Denoting by $\beta = \widetilde{\beta}/\overline{\tau}$, the equilibrium threshold profile $\boldsymbol{a}^*$ solves

$$[\boldsymbol{I} - \beta \boldsymbol{G}]\boldsymbol{a}^* = \boldsymbol{b}.$$

The equilibrium expected payoff to group i is:

$$\begin{aligned} U_i(\boldsymbol{a}^*, \boldsymbol{b}) &= \int_0^{a_i^*} \left(\beta \sum_j g_{ij} a_j^* + b_i - \tau_i \right) d\tau_i \\ &= \int_0^{a_i^*} (a_i^* - \tau_i)\, d\tau_i = \frac{1}{2} a_i^{*2}, \end{aligned}$$

where the second equality follows by using the best response of each population. So aggregate equilibrium utility is

$$W(\boldsymbol{b}, \boldsymbol{G}) = \frac{1}{2} (\boldsymbol{a}^*)^{\mathsf{T}} \boldsymbol{a}^*.$$

Suppose the planner, before the players choose their action, commits to the a subsidy scheme. The subsidy scheme depends on realized actions, which are taken after the scheme is announced. More precisely, the planner selects a vector $\boldsymbol{y} \in \mathbb{R}^n$ and offers the following scheme:

Subsidizing action 1. If $y_i > 0$ then the planner gives a subsidy of $s_i^1(\tau_i) = \tau_i - [a_i(\boldsymbol{y}) - y_i]$ to all population i's types $\tau_i \in [a_i(\boldsymbol{y}) - y_i, a_i(\boldsymbol{y})]$ who take action 1.

Subsidizing action 0. If $y_i < 0$ then the planner gives a subsidy of $s_i^0(\tau_i) = [a_i(\boldsymbol{y}) + |y_i|] - \tau_i$ to all $\tau_i \in [a_i(\boldsymbol{y}), a_i(\boldsymbol{y}) + |y_i|]$ who do not adopt the new technology (take action 0).

We make three observations. First, under intervention $\boldsymbol{y}$ the profile of thresholds $\boldsymbol{a}(\boldsymbol{y})$ is a Nash equilibrium. Furthermore, the planner does not waste resources in the sense that she uses the minimum amount of resources to implement $\boldsymbol{a}(\boldsymbol{y})$. To see this note that, by construction, the planner provides monetary payments to take action 1 or to take action 0 only to types who need such transfers to satisfy their incentive compatibility constraint. The monetary payments make these incentive compatible constrains just bind. Finally, let $\mathbf{1}_{y_i>0}$ be an indicator function that takes value 1 if $y_i > 0$ and 0 otherwise, then note that the cost of intervention $\boldsymbol{y}$ is

$$\begin{aligned} K(\boldsymbol{y}) &= \frac{1}{2} \sum_i \mathbf{1}_{y_i>0} \int_{a_i(\boldsymbol{y})-y_i}^{a_i(\boldsymbol{y})} s_i^1(\tau_i) d\tau_i + \sum_i (1 - \mathbf{1}_{y_i>0}) \int_{a_i(\boldsymbol{y})}^{a_i(\boldsymbol{y})+|y_i|} s_i^0(\tau_i) d\tau_i \\ &= \frac{1}{2} \sum_i y_i^2 \end{aligned}$$

We can now define the intervention problem of the planner as follows. Starting from the status quo $\hat{\boldsymbol{b}}$, the planner chooses intervention $\boldsymbol{y}$ to maximize aggregate equilibrium utility under the constraint that individuals play according to equilibrium and that the cost of the intervention cannot exceed C. Formally,

$$\begin{aligned} &\max_{\boldsymbol{y}\in\mathbb{R}^n} \frac{1}{2}\boldsymbol{a}^{\mathsf{T}}\boldsymbol{a} \\ &\text{s.t. } [\boldsymbol{I}-\beta\boldsymbol{G}]\boldsymbol{a} = \hat{\boldsymbol{b}}+\boldsymbol{y}, \\ &K(\boldsymbol{y}) = \frac{1}{2}\sum_i y_i^2 \leq C, \end{aligned} \tag{IT-P}$$

Intervention problem (IT-P) is equivalent to the intervention problem (IT) defined in Section 2.