Social Connectedness and Local Contagion*†

C. Matthew Leister‡ Yves Zenou§ Junjie Zhou¶

March 24, 2021

Abstract

We study a coordination game among agents in a network. The agents choose whether to take action (e.g., adopting a new technology) in an uncertain environment that yields increasing value in the actions of neighbors. We develop an algorithm that fully partitions the network into communities (coordination sets) within which agents have the same propensity to adopt. Our main finding is that a novel measure of network connectedness, which we term "social connectedness", determines the propensity to adopt for each agent. Social connectedness captures both the number of links each agent has within her community (interconnectedness) as well as the number of links she has with members of other communities who have a higher propensity to adopt (embeddedness). There is a single coordination set if and only if the network is balanced —that is, the average degree of each subnetwork is no larger than the average degree of the network. Finally, we demonstrate that contagion is localized within coordination sets, such that a shock to an agent uniformly affects this agent and all members of her coordination set but has no impact on the other agents in the network.

Keywords: global games, networks, social connectedness, welfare.
JEL codes: D85, C72, Z13.

*This paper has been previously circulated under the following title "Coordination on Networks."

†We are grateful for comments from the editor, Adam Szeidl, four anonymous referees, Francis Bloch, Arthur Campbell, Ozan Candogan, Bogachan Celen, Yi-Chun Chen, Ying-Ju Chen, Chris Edmond, Matt Elliott, David Frankel, Ben Golub, Sanjeev Goyal, Matt Jackson, Jun-Sung Kim, Ryota Iijima, Barton Lipman, Mihai Manea, Andrew McLennan, Stephen Morris, Alex Nichifor, Alessandro Pavan, Joel Sobel, Satoru Takahashi, Takashi Ui, Rakesh Vohra, Muhamet Yildiz, and seminar participants at HKU, NUS, Monash, University of Queensland, UNSW, City University of Hong Kong, Wuhan University, Fifth Annual Conference on Network Science and Economics (Bloomington), Workshop on Game Theory (NUS), Symposium on Social and Economic Networks (Melbourne), Workshop on Information and Social Economics (WISE18), and the Coalition Theory Network Conference (Marseille).

‡Monash University, Australia. Email: matthew.leister@monash.edu
§Monash University, Australia. Email: yves.zenou@monash.edu
¶National University of Singapore. Email: zhoujj03001@gmail.com

1 Introduction

In numerous real-world situations, agents make binary choices in an unknown state of the world while being positively influenced by their neighbors. For example, in developing countries, agents choose whether or not to adopt a new technology, such as a new crop (Bandiera and Rasul 2006), weather insurance (Cai et al. 2015), or a new rice technology (Islam et al. 2018), but the value/benefit of the technology is only partially known and increases with peer adoption. Similarly, individuals decide whether to join a gym (Babcock et al. 2020) or to participate in crime when the proficiency of the criminal, and thus the likelihood of not getting caught, is unknown and increasing in terms of the criminality of accomplices (Chalfin and McCrary 2017). In each of these examples, uncertainty in a common state influences the value of adoption: the underlying value of the technology, the strength or presence of the police force, and each binary decision is influenced by peers' decisions.

This paper is the first to study coordination issues with binary choices in these uncertain environments in a network setting.[1] The model employs the tools of global games embedded in a network game. Players' positions in the network define their preferences regarding the action choices of others. Using the language of technology adoption, the total value an agent receives from adopting the technology increases in the technology's underlying value (the state) and in its adoption by neighbors. Agents receive noisy signals that inform them of the state. In equilibrium, agents further use their private information to infer the observations and actions of neighboring agents and to anticipate the ultimate value that they are likely to enjoy from adopting the technology. The classic equilibrium selection of global games is obtained. In our setting with binary actions, the equilibrium selected in the noiseless limit comes in the form of cutoff strategies. Agents adopt the technology when their private signal exceeds their equilibrium cutoff, which is determined by the agent's position in the network.

We explore the role of the network's architecture in determining who coordinates their adoption choices with whom. Currently, there are no results in the literature describing exactly who, as a function of the agents' position in the network, takes a common action together in the noiseless limit. The main methodological contribution of this paper is to provide an exact prediction by putting forward the notion of *coordination sets*, which provides a unique endogenous partitioning of the agents in a network. Agents within a coordination set are path-connected and take a common cutoff strategy in order to adopt the technology together.

Further, we develop an algorithm called the Sequential Average Network Density (SAND) to fully characterize the equilibrium partitions and thresholds using the primi-

[1] In numerous situations, agents may have more than two choices. For example, when someone has to decide whether or not to get insured, she may also decide which insurance to choose from a non-singleton menu. In this paper, we focus our attention on binary actions because our aim is not to address general coordination problems in discrete games with an arbitrary number of actions but rather to investigate the impact of network structure on binary choices in coordination games.

tives of the model: intrinsic valuations for adoption, network topology, and the strength of network effects. In order to understand the manner in which the SAND algorithm works, assume that all agents have the same intrinsic valuations for adoption; thus, their only heterogeneity stems from their position in the network. First, the SAND algorithm identifies the community (or coordination set) with the highest *interconnectedness* —that is, the highest average number of links— as these members obtain the highest level of network effects; thus, they will have the highest propensity to adopt in the network. Then, for the remaining agents in the network, SAND determines the community that maximizes both interconnectedness and the number of links to the first community; this is referred to as *embeddedness*. Indeed, in order to have a high propensity to adopt, agents need to be part of a well-connected community as well as to have a high average number of edges with those who have a higher propensity to adopt. Communities that have both high interconnectedness and embeddedness are referred to as highly *socially connected*.[2] The SAND then continues until every agent in the network is part of a community (or coordination set). In a nutshell, the SAND algorithm creates an order of a set of players based on their index of social connectedness and shows how to divide the agents in terms of their marginal contributions to this index. This order defines a partition of coordination sets within which cutoff strategies are uniform. When we introduce different intrinsic valuations for adoption, the SAND algorithm proceeds in the same manner but maximizes the sum of network effects and these valuations.

Contrary to the standard network games with complementarities in actions and perfect information (e.g., Ballester et al. 2006; Bramoullé et al. 2014) in which Katz-Bonacich centrality totally determines agents' actions, we show that, in certain networks, agents with a lower Katz-Bonacich centrality can have a higher propensity to adopt if they are well embedded with agents in coordination sets that have an even higher propensity to adopt. The novel conceptual contribution of our model is that it enables us to create a precise connection between network structure and the propensity to adopt; specifically, it reveals how community structure and an agent's particular links interact in determining her "social connectedness." This provides a more nuanced picture of adoption patterns under vanishing uncertainty as compared to the conventional Katz-Bonacich centrality (Ballester et al. 2006). Moreover, the distinction from the Katz-Bonacich centrality provides a general insight regarding why underlying uncertainty complicates agents' adoption decisions and the consequent adoption pattern. Indeed, it turns out that both links *within* a coordination set (interconnectedness) and *between* coordination sets (embeddedness) are of significance when agents face sufficient incentive to adopt. For the former, agents must form expectations regarding their neighbors adopting within the same coordination set, while for the latter, they put a probability of either 0 (lower-propensity coordination sets) or 1 (higher-propensity coordination sets) on these neighbors adopting. The SAND algorithmic solution untangles this added complication due to uncertainty by identifying

[2] Bailey et al. (2018) were the first to empirically introduce the notion of *social connectedness*; however, it is not microfounded and corresponds to our notion of *embeddedness*.

the community boundaries that are of significance for the adoption decision. In particular, the notion of "social connectedness" simplifies the network coordination analysis and enables us to condense the problem down to dividing the network into communities.

We study the role of network structure on adoption by considering the case of homogeneous values, where the network alone introduces (*ex-ante*) heterogeneity across agents. We provide an exact characterization for which a single coordination set exists in the network. This condition requires that the network structure must be *balanced*, that is, the average degree of each subnetwork (comprising any nonempty subset of agents in the network) is not greater than the average degree of the entire network.[3] This characterization implies that, remarkably, agents with very different positions in the network may belong to the same coordination set. For example, in a star network, regardless of the number of peripheral agents, all agents coordinate together, thereby implying that they have the same propensity to adopt. The same is true for regular networks, tree networks, regular bipartite networks, and those that have a maximum of four agents.

Thereafter, we explore the policy implications of our model. We first investigate how equilibrium coordination is adjusted as intrinsic valuations change. We show that small shocks to intrinsic valuations do not change the coordination set partition, but they do change the adoption thresholds for the agents within the coordination sets that are affected by the shocks. Strikingly, perturbations are shown to influence equilibrium adoption only across members within the perturbed agents' coordination set. Thus, the contagion of such perturbations extends within coordination sets but does not spread across other coordination sets: contagion is *local*.

We then study the marginal gains of subsidizing intrinsic valuations to a policy designer aiming at maximizing either (i) ex-ante adoption or (ii) welfare under a budget constraint. We show that, for a sufficiently small budget, the marginal impact of these targeted policies is independent of the particular target selected from the target's coordination set. In other words, optimal policy design becomes a problem of targeting a given coordination set rather than a particular agent. Strikingly, under certain conditions on primitives, the adoption- and welfare-maximizing planner targets opposite extremes; the former targets the coordination set where agents take the highest cutoff, while the latter targets the coordination set where agents take the lowest cutoff. Indeed, the welfare-maximizing planner incorporates the effect of the intervention on both equilibrium adoption and externalities, while the adoption-maximizing planner values only its effect on the former.

We also study policies that aim at changing the structure of the network. First, we investigate the key-player policy —that is, finding players who, once removed, reduce

[3] Observe that our balance condition characterizing networks with a single coordination set resembles the necessary and sufficient conditions for a single social status class to emerge in the setting of Immorlica et al. (2017) and for the law of one price to hold in Manea (2011). However, these models are rather different from ours. In Immorlica et al. (2017), the actions are continuous and the information is complete; thus, there are multiple equilibria, which lead to different implications (see our discussion in Section J). In addition, the shortage ratio of Manea (2011) shares the averaging functional of our SAND algorithm.

total crime or adoption the most. In the standard key-player policy (Ballester et al. 2006), what matters the most is the complementarity in actions between agents; thus, the key players are the agents who generate the highest level of spillovers to themselves and to others. In our model, we also have complementarity in actions. However, in addition to this, there is a coordination problem since agents do not know with certainty whether their neighbors will adopt (or commit a crime); thus, they do not know the level of complementarity in the actions to which they are exposed. As stated above, the social connectedness of each agent determines who has the highest propensity to adopt; it also indicates who the key players are. Second, we determine how adding links in a network affects adoption. We show that, under certain conditions, adding links between agents from different coordination sets does not affect the cutoff of agents from the coordination set with the lower cutoff. However, it reduces the cutoff of agents from the coordination set with the higher cutoff.

Finally, we show that our results significantly differ from those of the studies on complete information (Milgrom and Roberts 1990; Vives 1990), in which multiple equilibria prevail and where the customary equilibrium selection, which focuses on either the biggest or the smallest possible sets of players taking action, is rather arbitrary. Compared to these studies, we provide a clean, elegant method for selecting equilibria. We show that our unique equilibrium can correspond to neither the maximum nor minimum equilibrium and that the comparative statics properties as well as the policy implications can be very different in these two types of information settings.

Our paper contributes to the global games literature. Carlsson and van Damme (1993) first exhibited a selection device for global games of two players and two actions.[4] Frankel, Morris, and Pauzner (2003) extended the result to arbitrary games of strategic complements. In a two-sided environment, Morris and Shin (1998) provided closed forms to their common limit-equilibrium cutoff to studying the interaction of a government defending a currency from a continuum of currency speculators. In the global games literature, most papers consider binary action global games with *symmetric* and supermodular payoffs. However, there are remarkably few papers that look at the extensions with *asymmetric* payoffs across players. There are certain papers with heterogeneous payoffs but parametric assumptions that ensured a single coordination set (Corsetti et al. 2004; Guimaraes and Morris 2007; Sákovics and Steiner 2012 are examples). In particular, Sákovics and Steiner (2012) studied the policy impact of a global game with a continuum of agents who value an agent-weighted average action. However, a number of papers have examined cases in which there are multiple coordination sets but with different parameterized classes of payoffs than ours —including Dai and Yang (2019), Abadi and Brunnermeier (2018), Drozd and Serrano-Padial (2018), and Serrano-Padial (2018).[5]

[4] They show that the risk-dominant equilibrium is selected in these games.

[5] In particular, Dai and Yang (2019) studied a similar model to that of Sákovics and Steiner (2012), in which the continuum of agents carries private information regarding idiosyncratic costs of adoption, incorporating multiple cutoffs in the noiseless limit. In this setting, the authors focus on the role of organizations in mitigating miscoordination.

Compared to this literature, the network structure introduces heterogeneity through the agents' network positions. In addition, the application to a network coordination setting grounds the abstract results from general global games (e.g., Frankel et al. 2003) in an interpretable environment that enables us to more clearly interpret the implications of the global games structure.[6]

Our equilibrium characterization in terms of coordination sets is also related to other network models, which also partition agents into endogenous community structures, including risk-sharing (Ambrus, Mobius, and Szeidl 2014), information resale and intermediation (Manea 2021), and interaction between market and community (Gagnon and Goyal 2017). However, the driving forces and policy implications are very different.

More broadly, this paper adds to the growing literature on network games.[7] In numerous situations in which networks are of significance, agents make both binary decisions (*extensive margin*) and quantity decisions (*intensive margin*). Consider, for example, crime. First, an individual has to decide whether to become a criminal; this is a binary decision (extensive margin). Then, if he becomes a criminal, he must decide how many crimes to commit (intensive margin). The literature on network games has mostly focused on the intensive margin by assuming that actions are continuous. With strategic complements in actions and complete information, a unique equilibrium is usually obtained under certain condition on the eigenvalues of the network (see, e.g., Ballester et al. 2006 and Bramoullé et al. 2014);[8] agents with more connections (in terms of degree/Katz-Bonacich/eigenvector centrality) exert higher effort.[9] In the present paper, we focus on the extensive margin by assuming binary actions. We also assume strategic complements under incomplete information. We show that, in the noiseless limit of our game, there is a unique equilibrium that we are able to characterize in terms of coordination sets. Further, we show that it is not only a larger number of connections (interconnectedness) but also connections to coordination sets with lower cutoffs (embeddedness) that are of significance.

Finally, we contribute to the literature on both network and global games by making predictions in terms of adoption curves that substantially differ from leading alternatives

[6] Our results also have bearing on those of the studies on network contagion and community structure (Chwe 2000; Morris 2000; Oyama and Takahashi 2015; Jackson and Storms 2017); however, our focus is mainly on partitioning the network into coordination sets, while these studies examined coordination games on a network to determine when a convention spreads contagiously from a finite subset of players to the entire population in certain networks.

[7] See Jackson (2008) chapter 9, Jackson and Zenou (2015), and Bramoullé and Kranton (2016) for surveys.

[8] Indeed, in network games with strategic complementarities and complete information (e.g., Ballester et al. 2006), the uniqueness of equilibrium is guaranteed by an (largest) eigenvalue condition, which requires that the strength of complementarities (positive cross-effects) is smaller than own-concavity. Ollar and Penta (2017, 2019) generalized this result for general incomplete information structure, both with and without a common prior, and for general payoff functions that satisfy weak concavity and smoothness requirements.

[9] There is also more recent literature on network games with incomplete information (see, in particular, Calvó-Armengol and de Martí 2007, Galeotti et al. 2010, Calvó-Armengol et al. 2015, de Martí and Zenou 2015, Golub and Morris 2017a,b, Myatt and Wallace 2019, Leister 2020).

in these two literatures. Indeed, our model generates adoption that is neither smooth nor fully coordinated but comes in *batches*. In other words, the share of agents who adopt as a function of the price is neither a continuous function (as, e.g., in Ballester et al. (2006)) nor a jump at a single threshold (as, e.g., in Sákovics and Steiner (2012)), but instead a step function. This prediction is unique to our model.

2 Model setup

A finite set of agents N, connected via a network $\mathcal{G} = (N, E)$, simultaneously choose whether or not to adopt a technology.[10] E defines the set of edges between unordered pairs ij taken from N. We assume a connected and undirected graph: $i \in N_j$ if and only if $j \in N_i$, where $N_i \equiv \{j : (i, j) \in E\}$ is the set of i's neighbors, and $d_i \equiv |N_i|$ her degree. Let $a_i \in \Omega_i \equiv \{0, 1\}$ denote agent i's binary choice: $a_i = 1$ implies that agent i adopts, while $a_i = 0$ implies that she does not adopt.

Payoffs from adopting the technology depend on the action profile $\mathbf{a} = (a_1, \dots, a_{|N|}) \in \prod_{i \in N} \Omega_i$ and the underlying fundamental state $\theta \in \Theta$, where Θ is a bounded interval in $\mathbb{R}$. Specifically, we assume each i obtains the following payoff:[11]

$$u_i(\mathbf{a}, \theta) = \begin{cases} v_i + \theta + \phi \sum_{j \in N_i} a_j & \text{if } a_i = 1, \\ 0 & \text{if } a_i = 0, \end{cases} \tag{1}$$

where $v_i \in \mathbb{R}$, and $\phi > 0$. Here, v_i provides the intrinsic (state-independent) value to i from adopting, θ the state dependent value, with the adoption of each of i's neighbors having a positive influence on the value of technology. The externality benefit in (1) is equal to $\phi \sum_{j \in N_i} a_j$, which is referred to as the *local-aggregate model* in the network games literature (Topa and Zenou 2015). Alternatively, we can consider the *local-average model*, in which the network externality is equal to the *proportion* of adopting neighbors rather than to their absolute number —that is, $\phi(\sum_{j \in N_i} a_j)/d_i$. We discuss this alternative model in Section H.6 in the Online Appendix. In particular, we show that, in the local-average model, if $v_i = v$ for all i, for *any* network structure, all agents adopt or do not adopt together.[12] Further, a more general model with *weighted* networks is analyzed in Section H.3 in the Online Appendix.

Dominance regions and multiplicity of equilibria. In the stage game where θ is commonly

[10] For the sake of the exposition, we use the example of technology adoption but, of course, all results subsume arbitrary binary action sets.

[11] In Section H.4 in the Online Appendix, we consider a slightly different utility function in which agents face miscoordination costs in their adoption choices, while in Section H.5, we introduce non-strategic externalities in the utility function. In both cases, we show that our main results still hold.

[12] Ushchev and Zenou (2020) obtained a similar network-independent result in a continuous action, local-average network model. Among other things, they showed that, if all agents are *ex-ante* identical ($v_i = v$ for all i), the network position does not matter, and all agents exert the same effort and adopt the same social norm.

known among agents, each agent i has a dominant strategy to adopt (not to adopt) when θ is sufficiently high (low). For each i, we assume v_i and ϕ are such that there exist $\underline{\theta}_i$ and $\overline{\theta}_i$ in the interior of Θ, such that $v_i + \theta + \phi d_i < 0$ when $\theta < \underline{\theta}_i$ and $v_i + \theta > 0$ when $\theta > \overline{\theta}_i$. Thus, there exist dominant regions $[\min \Theta, \underline{\theta}]$ and $[\overline{\theta}, \max \Theta]$, with $\underline{\theta} \equiv \min_{i \in N}\{\underline{\theta}_i\}$ and $\overline{\theta} \equiv \max_{i \in N}\{\overline{\theta}_i\}$ such that not adopting and adopting the technology (respectively) are dominant strategies for all players. Obviously, when θ lies within the dominance regions, a unique Nash equilibrium is obtained; when θ lies outside the dominance regions, multiple pure strategy Nash equilibria may occur.

In this paper, we adopt the global game approach for equilibrium selection (see, e.g., Carlsson and van Damme 1993 and Frankel et al. 2003). In Section G.2 in the Online Appendix, we provide a detailed discussion and justification of these selection criteria and present an alternative yet equivalent selection from the perspective of potential games.

Information structure and the perturbed game. Following the global games literature, we perturb the stage game of complete information into a Bayesian game of incomplete information. In the perturbed game, the common state, θ, is observed with noise by all agents.[13] Each i receives a private signal $s_i = \theta + \nu\epsilon_i$, $\nu > 0$, where ϵ_i is distributed via density function g and cumulative function G with support $[-1, 1]$. All signals are independently drawn across agents conditional on θ. The agents share a common prior belief regarding θ, which is denoted by the CDF $H(\cdot)$ with continuously differentiable density $h(\cdot) > 0$.[14] For each $\nu > 0$, we write Γ^ν for the corresponding global game.[15] We are interested in the perturbed game Γ^ν for ν close to zero.[16]

3 Equilibrium analysis

3.1 Limiting equilibrium and coordination sets

We formally introduce the global games machinery to determine the limiting equilibrium. We first analyze the Bayesian Nash equilibrium of Γ^ν for each $\nu > 0$. Then, we determine the limit of the equilibrium as ν approaches zero. Here, we apply standard results in the literature (see, e.g., Carlsson and van Damme 1993 and Frankel et al. 2003) to obtain

[13] Here, we introduce small noises to θ but not to $\{v_i\}_{i \in N}$. In Section I in the Online Appendix, we consider an extension in which both θ and v_i are perturbed. All of our results continue to hold in this extension as perturbations shrink to zero.

[14] The prior distribution H is inconsequential for the limiting equilibrium characterized in Theorem 1 (e.g., Sákovics and Steiner (2012) assume agents hold a dispersed prior of θ). However, H affects the analysis of policy interventions in Section 5.1.

[15] The assumption of a common noise structure is without loss of generality, as the limit-equilibrium selection is robust to arbitrary, idiosyncratic F_i (see Section H.2 in the Online Appendix). Moreover, all results in the limit hold under unbounded supports —for example, Gaussian state and noise (see Section H.1 in the Online Appendix).

[16] Because, as in all global games models, we use incomplete information as a means of equilibrium refinement, the critique of Weinstein and Yildiz (2007) applies here: introducing even small amounts of higher-order uncertainty can enable us to select other equilibria. This is a common critique of the global games literature that we acknowledge.

the existence and essential uniqueness of the limiting equilibrium, thereby deferring the explicit characterization of the limiting equilibrium to Section 3.2.

A strategy for i is a measurable function π_i that assigns a probability mixture on $\Omega_i = \{0, 1\}$ to each signal s_i of player i in Γ^ν. We let $S_i^\nu = [\min \Theta - \nu, \max \Theta + \nu]$ be the set of all possible signals for i and Π_i^ν denote the set of strategies for i, and we write $\pi_i(s_i)$ for the probability that π_i is assigned to action 1 at s_i. We write $\boldsymbol{\pi} \equiv (\pi_1, \dots, \pi_{|N|})$.

Let $U_i^\nu(\boldsymbol{\pi}_{-i}|s_i)$ denote the expected payoff for action $a_i = 1$ when agent i observes s_i and each $j \neq i$ uses strategy π_j:[17]

$$U_i^\nu(\boldsymbol{\pi}_{-i}|s_i) = v_i + \mathbb{E}\left[\theta|s_i\right] + \phi \sum_{j \in N_i} \mathbb{E}[\pi_j(s_j)|s_i].$$

In other words, i takes expectations of the state and of her neighbors' actions, conditioning on her private signal s_i. Since the payoff for action a_i is zero, an optimal strategy for i is to set $\pi_i(s_i) = 1$ when $U_i^\nu(\boldsymbol{\pi}_{-i}|s_i) > 0$ and to set $\pi_i(s_i) = 0$ when $U_i^\nu(\boldsymbol{\pi}_{-i}|s_i) < 0$. We use the indicator function

$$\pi_i(s_i) = \mathbb{I}(s_i \geq c_i) = \begin{cases} 1 & \text{if } s_i \geq c_i \\ 0 & \text{if } s_i < c_i \end{cases}$$

to denote agent i's cutoff strategy at $c_i \in S_i^\nu$. Then, a Bayesian Nash Equilibrium $\boldsymbol{\pi}^{*\nu}$ of Γ^ν in cutoff strategies with $\pi_i^{*\nu} = \mathbb{I}(s_i \geq c_i^{*\nu})$ for each i must satisfy[18]

$$U_i(\boldsymbol{\pi}_{-i}^{*\nu}|s_i = c_i^{*\nu}) = 0, \quad \forall i \in N, \tag{2}$$

with each i indifferent between adopting and not adopting when observing signal s_i equal to her equilibrium cutoff $c_i^{*\nu}$. Indeed, a standard fixed-point argument reveals that a pure Bayesian Nash Equilibrium of Γ^ν in cutoff strategies always exists for $\nu > 0$. The properties of the limit of $\boldsymbol{\pi}^{*\nu}$ as ν tends to zero are formally presented in Theorem 1 in Frankel et al. (2003) (which is stated in Proposition 0 in Appendix B). The vector of limiting state cutoffs $\boldsymbol{\theta}^* = (\theta_1^*, \cdots, \theta_{|N|}^*) \equiv (\lim_{\nu \to 0} c_i^{*\nu})_{i \in N}$ fully determines the limiting equilibrium $\boldsymbol{\pi}^* = (\pi_1^*, \cdots, \pi_{|N|}^*)$, with each i choosing to adopt when θ rises above θ_i^* —that is, $\pi_i^* = \mathbb{I}(\theta \geq \theta_i^*)$. The uniqueness of $\boldsymbol{\pi}^*$ in Proposition 0 in Appendix B implies that the limiting cutoffs $\boldsymbol{\theta}^*$ are unique.

The ranking of the limiting cutoffs determines who has a higher propensity to adopt by taking a lower cutoff and who has a lower propensity to adopt with a higher cutoff. The next definition reveals a nice structure that captures the behavior of agents in the limiting equilibrium.

[17]Observe that, in our model, there is *no* learning over time. Each individual makes a *once-and-for-all* decision that involves coordination problems with her friends' decision. We do not model the learning and dynamic aspect of this decision. Hak (2019) develops a dynamic learning model by showing how establishments slowly adopt slot machines over a few years.

[18]In other words, $U_i(\boldsymbol{\pi}_{-i}^{*\nu}|s_i = c_i^{*\nu}) = v_i + \mathbb{E}\left[\theta|c_i^{*\nu}\right] + \phi \sum_{j \in N_i} \mathbb{E}[\mathbb{I}(s_j \geq c_j^{*\nu})|c_i^{*\nu}] = 0$ for every $i \in N$.

Definition 1 (Coordination sets)**.** *The limiting cutoffs* $\boldsymbol{\theta}^*$ *map to a unique partition* $\mathcal{C}^* = \{C_1^*, \ldots C_M^*\}$ *of* N *that satisfy:*

(i) $\forall m \in \{1, \cdots M\}$, $\forall i, j \in C_m^*$, *we have* $\theta_i^* = \theta_j^*$*;*

(ii) for all $(i,j) \in E$ *satisfying* $\theta_i^* = \theta_j^*$, *there is some* $m \in \{1, \ldots, M\}$, *such that* $i, j \in C_m^*$.

Each element C_m^*, $m = 1, 2, \cdots, M$, *of the partition* $\mathcal{C}^*$ *is called a coordination set.*

This definition (i) implies that agents in the same coordination set have the same cutoffs. Item (ii) requires that any pair of directly connected agents sharing the same cutoff must be in the same coordination set. Item (i) guarantees that agents with distinct cutoffs cannot coexist in the same coordination set, while item (ii) eliminates the fact that one could have singletons for reasons other than everybody adopting a different threshold. Obviously, the limiting cutoffs $\boldsymbol{\theta}^*$ naturally map to a unique partition $\mathcal{C}^*$ thereby satisfying Definition 1 (i) and (ii).[19]

Remark 1. *The global games approach that we employ functions as an equilibrium selection device. Indeed,* $\boldsymbol{\pi}^*(\theta)$ *is a Nash equilibrium of the stage game (with complete information) at* θ*.*[20]

The result of Frankel et al. (2003) proves the existence and essential uniqueness of the equilibrium, but it cannot be used to derive a characterization of equilibrium cutoffs and coordination sets in our network setting. This is the subject of the next section.

3.2 An algorithmic approach

We provide an algorithm to explicitly characterize the equilibrium cutoffs. To this end, we define $F(\cdot)$, a mapping from 2^N to $\mathbb{R}$, in the following manner:

$$F(S) \equiv v(S) + \phi e(S), \qquad \text{for any } S \subset N, \tag{3}$$

where $v(S) \equiv \sum_{i \in S} v_i$, and $e(S)$ is the number of edges between members of S. Formally, $e(S) \equiv \frac{1}{2} \sum_{i \in S} d_i(S)$, where $d_i(S) \equiv |N_i \cap S|$ denotes the *within-degree* of i, or the number of edges between i and the members of agent set S. In other words, $F(S)$ captures the total internal intrinsic value plus network effects for agents in S. The interpretation of $F(\cdot)$ will be evident after we explain the close link between the equilibrium conditions for agents at cutoffs and the SAND algorithm introduced below.[21]

[19]Since $\boldsymbol{\theta}^*$ is unique according to Theorem 1 in Frankel et al. (2003) (which is stated in Proposition 0 in Appendix B), the partition $\mathcal{C}^*$ is well defined by Definition 1.

[20]This well-known result follows from the upper semicontinuity of the equilibrium correspondence, which relies on the continuity of prior density and noise density. In our setting, the selected equilibrium corresponds to the maximizer of a potential function of the stage game and has several appealing features (see further discussion in Remark 4 and Section G.3 in the Online Appendix).

[21]A major property of $F(\cdot)$ is supermodularity. In other words, for any $S, T \in 2^N$, $F(S \cap T) + F(S \cup T) \geq F(S) + F(T)$. See Lemma 4 in Appendix C. Observe that A_1^* is well defined because we prove that, by

Algorithm 1 (Sequential Average Network Density (SAND))**.**
Step 1.

$$A_1^* = \underset{S \supsetneq \emptyset}{\operatorname{argmax}} \frac{F(S)}{|S|}.$$

(If there are multiple maximizers, we set A_1^* to the largest maximizer.

Step k.

$$A_k^* = \underset{S \supsetneq A_{k-1}^*}{\operatorname{argmax}} \frac{F(S) - F(A_{k-1}^*)}{|S| - |A_{k-1}^*|}. \tag{4}$$

(If there are multiple maximizers, we set A_k^* to the largest maximizer.)

Continue until $A_k^* = N$.

The SAND algorithm partitions the network into communities or coordination sets (Definition 1), where, within each community, all agents have the same propensity to adopt. Indeed, when the noise of the signal vanishes, the total expected value of adoption within a community S is equal to $F(S)$; thus, $F(S)/|S|$ is the average of this value. Consequently, the SAND will first search for A_1^*, the subset S of agents that maximizes $F(S)/|S|$. After knowing A_1^*, the algorithm continues and, in step 2, searches for $S = A_2^*$, a proper superset of A_1^*, with the next largest $\left(F(S) - F(A_1^*)\right)/|S - A_1^*|$ conditioning on the inclusion of A_1^*. Indeed, contrary to step 1, in step 2, the algorithm must take into account the fact that certain members of $A_2^* \backslash A_1^*$ have links to members of A_1^* because the former benefit in terms of network spillovers of being neighbors to agents who have a higher propensity to adopt. Therefore, the expected value of adoption depends both on links within the coordination set $A_2^* \backslash A_1^*$ and on links between coordination sets —that is, between $A_2^* \backslash A_1^*$ and A_1^*. The SAND then continues until step k, when all agents in the network have been selected —that is, when $A_k^* = N$.

Let $t_{[k]}^*$ denote the maximum value obtained in step k:

$$t_{[k]}^* = \frac{F(A_k^*) - F(A_{k-1}^*)}{|A_k^*| - |A_{k-1}^*|}. \tag{5}$$

Observation 1. *The SAND algorithm terminates in finite $K \geq 1$ steps. Furthermore, it yields a sequence of strictly nested sets,*

$$\emptyset \subsetneq A_1^* \subsetneq A_2^* \subsetneq \cdots \subsetneq A_K^* = N, \tag{6}$$

and a sequence of strictly decreasing numbers,

$$t_{[1]}^* > t_{[2]}^* > \cdots > t_{[K]}^*. \tag{7}$$

the supermodularity of $F(\cdot)$, the union of any two maximizers is again a maximizer and, therefore, the maximal maximizer exists; a similar argument holds for each step k (see Observation 3 in Appendix C). Further, for generic $\mathbf{v}$, multiple maximizers cannot occur and, thus, a unique maximizer is obtained in each step (see Observation 5 in Appendix C).

Observation 1 is intuitive. First, $1 \leq K \leq |N|$, as at each step the cardinality of A_k^* increases by at least one. Second, the strict monotonicity of the maximum value, $t_{[k]}^*$, in k follows from a simple optimization argument. To illustrate this point, consider $k = 1$. Since A_1^* is the largest maximizer in step 1 of the SAND algorithm, A_2^* cannot be a maximizer in step 1, thereby implying that $F(A_1^*)/|A_1^*| > F(A_2^*)/|A_2^*|$, and $t_{[1]}^* > t_{[2]}^*$. Extending this logic to any step k yields the monotonicity of sequence $t_{[k]}^*$ in (7). The sequence of sets in (6) is closely related to coordination sets in Definition 1, while the sequence of numbers in (7) is tightly linked to cutoffs, as illustrated in Theorem 1.[22]

Definition 2. *Agent i is said to be found in step k if $i \in A_k^* \backslash A_{k-1}^*$ (defining $A_0^* = \emptyset$).*

Theorem 1 (Main Equilibrium Characterization)**.** *Suppose agent i is found in step k in the SAND algorithm. Then, her equilibrium cutoff θ_i^* satisfies*

$$\theta_i^* = -t_{[k]}^*. \tag{8}$$

Theorem 1 provides a simple algorithmic characterization of the limiting equilibrium using the model primitives $(\mathbf{v}, \phi, \mathcal{G})$ and indicates that an agent's cutoff is exactly equal to the negative value of the maximum obtained in the step of the SAND algorithm where she is "found."

We discuss several features of the SAND algorithm and illustrate how it works in two examples. The economic and empirical implications of Theorem 1 are investigated in Section 3.4.

Remark 2. *The computational complexity of the SAND algorithm is polynomial, which is an appealing feature that makes the computation of equilibrium cutoffs and coordination sets very fast, even for very large networks.*[23]

Remark 3. *Theorem 1 also confirms that the limiting equilibrium $\boldsymbol{\pi}^*$ is unique and is independent of the noise distribution, as the algorithm does not use any details about the noise distribution. Although these uniqueness and noise independence results are established in a more general setting in Frankel et al. (2003), our approach using the SAND algorithm provides more direct and informative proofs of such results in our setting.*

Remark 4. *The SAND algorithm is closely related to an alternative equilibrium characterization using the potential function in Section G.3. By Frankel et al. (2003), the global game machinery selects the maximizer of the potential $P(\mathbf{a}, \theta)$ for generic θ (see equation*

[22] Observe that SAND is similar to the algorithms in Manea (2011) and Immorlica et al. (2017) in that it identifies the players with the most extreme equilibrium actions/payoffs, and then applies the same procedure to the remaining network, which can either be analyzed independently or taking the equilibrium actions/payoffs of "extreme" players as given. The measures driving equilibrium outcomes in our model —interconnectedness and embeddedness— have analogues in the algorithms of these papers: shortage ratio in the case of Manea (2011) and cohesion in Immorlica et al. (2017).

[23] Note that, in each step of SAND, the program (4) is a combinatorial optimization problem that is solvable in polynomial time by supermodularity of $F(\cdot)$ (see, e.g., Fujishige 2005). The total number of steps, K, is at most $|N|$. Thus, SAND can be computed in polynomial time.

([G.7](#)). *Thus, we can effectively infer the cutoffs of agents by identifying the potential maximizers at every state. In Theorem G2 in the Online Appendix, we formally show how the SAND algorithm is used to determine the state-wise potential maximizer under this alternative characterization.*

Remark 5. *Theorem G1 in the Online Appendix G.1 provides an alternative characterization of the equilibrium thresholds $\boldsymbol{\theta}^*$ by solving a simple constrained convex minimization program.*[24] *In contrast to the discrete nature of the SAND algorithm, this minimization problem is continuous. This program is less demanding to compute than the SAND algorithm because it is a simple convex program with linear constraints. The solution can be viewed as the dual to the SAND algorithm and provides an alternative interpretation to our limit equilibrium.*

To illustrate the SAND algorithm, we provide two examples using the networks in Figure 1. For simplicity, in both examples, we set $\phi = 1$ and $v_i = 0$ for every i, so $F(S) = e(S)$.

(a) Star network. (b) Quad-core-periphery network.

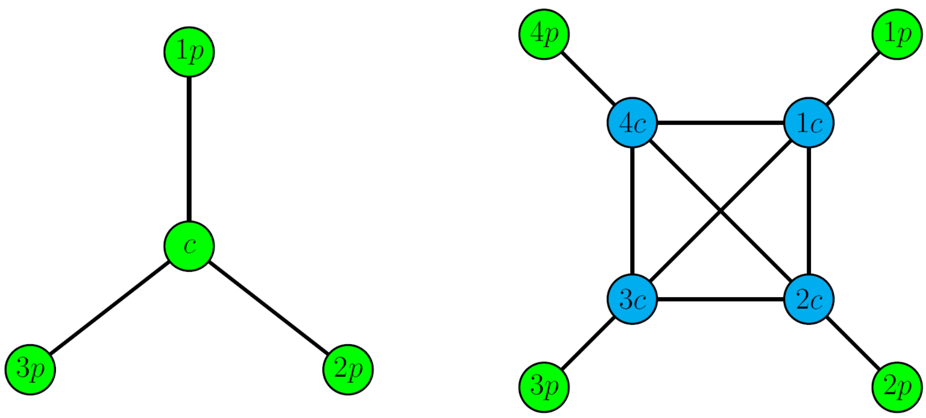

Figure 1: Coordination and network structure

Example 1. *For the star network in Figure 1(a), the SAND algorithm terminates in step 1:*[25]

S	$\{c\}$	$\{c, 1p\}$	$\{c, 1p, 2p\}$	**N**
$e(S)/\lvert S\rvert$	0	1/2	2/3	**3/4**

Example 2. *For the quad-core-periphery network in Figure 1(b), the SAND algorithm terminates in two steps. In step 1, $A_1^* = \{1c, 2c, 3c, 4c\}$, as shown in the table below:*

[24] In a different context, Nguyen (2015) also developed a new technique based on convex programming to characterize the unique stationary equilibrium payoff in the setting of Manea (2011).

[25] There is a unique coordination set as $A_1^* = N$. Note that $t_{[1]}^* = e(N)/|N| = 3/4$.

S	$\{1c, 2c, 3c\}$	$\{\mathbf{1c}, \mathbf{2c}, \mathbf{3c}, \mathbf{4c}\}$	$\{1c, 2c, 3c, 4c, 1p\}$	N
$e(S)/\lvert S\rvert$	$3/3$	$\mathbf{6/4}$	$7/5$	$10/8$

In step 2, the algorithm terminates with $A_2^* = N$.[26]

3.3 The main idea of the proof of Theorem 1

The intuition underlying Theorem 1 lies in connecting the agents' indifference conditions at the equilibrium cutoffs with the combinatorial maximization programs in SAND.

First, each agent i must be indifferent between adopting or not at her cutoff. Taking the limit of equilibrium condition (2) yields the following equation for i:

$$\theta_i^* + v_i + \phi \sum_{j \in N_i} w_{ij}^* = 0, \quad \forall i \in N. \tag{9}$$

In other words, the intrinsic value of adopting v_i plus the aggregate of w_{ij}^* over i's neighbors must equal $-\theta_i^*$, the minus of i's cutoff, where

$$w_{ij}^* = \lim_{\nu \to 0} \mathbb{E}[\mathbb{I}(s_j \geq c_j^{*\nu}) | s_i = c_i^{*\nu}] \tag{10}$$

yields the limit of the conditional probability of agent j adopting (i.e., j's signal s_j is above $c_j^{*\nu}$) when agent i's signal s_i is at her cutoff $c_i^{*\nu}$. Clearly, $0 \leq w_{ij}^* \leq 1$.

Lemma 1. *For each $(i, j) \in E$,*

(i) the following identity holds:

$$w_{ij}^* + w_{ji}^* = 1, \tag{11}$$

(ii) If, in addition, $\theta_i^* < \theta_j^*$, then

$$w_{ij}^* = 0, \text{ and } w_{ji}^* = 1. \tag{12}$$

An identity similar to (11) is presented in Carlsson and van Damme (1993) and plays a key role in the proof of their main theorem of selecting risk dominance equilibrium in the two-player setting.[27] We provide the intuition of (11) along similar lines as in Carlsson and van Damme, as Lemma 1, similar to their setting, involves a pair of players. Consider two thresholds x_1 and x_2, and let $\epsilon = \nu(\epsilon_2 - \epsilon_1)$ be the difference of the two players' idiosyncratic signal errors. Let G be the CDF of ϵ. Upon receiving signal $s_1 = x_1$, agent 1 attaches probability $1 - G(x_2 - x_1)$ to player 2 adopting.[28] Upon observing signal

[26]Note that $t_{[1]}^* = e(A_1^*)/|A_1^*| = 6/4 = 1.5$ and $t_{[2]}^* = \frac{e(N)-e(A_1^*)}{|N|-|A_1^*|} = (10-6)/(8-4) = 1$.

[27]See equation (4.1) on page 997 and Lemma 4.1 on page 999 in Carlsson and van Damme (1993). Technically, this Lemma is fundamentally driven by the common belief assumption among players. See Morris, Shin, and Yildiz (2016) and Morris and Yildiz (2019) for more discussion.

[28]That is, $\mathbb{E}[s_2 \geq x_2 | s_1 = x_1] = \Pr(x_1 - \nu\epsilon_1 + \nu\epsilon_2 \geq x_2) = 1 - G(x_2 - x_1)$.

$s_2 = x_2$, agent 2 attaches probability $G(x_2 - x_1)$ to player 1 adopting.[29] Thus, these two probabilities add up to 1. To understand the intuition of (12), assume that $\theta_i^* < \theta_j^*$. Then, when player i observes the signal at i's threshold, she infers that the state is almost θ_i^* for $\nu > 0$ small and that j is not going to adopt almost surely (recall j's threshold is strictly higher than i's). Consequently, in the limit as $\nu \to 0$, $w_{ij}^* = 0$.

The pairwise information in Lemma 1 regarding these weights w_{ij}^* is central to our analysis. We differ from Carlsson and van Damme (1993) with two players in that in our network setting, we would like to aggregate this pairwise information over a subset of agents to build a link between the SAND algorithm and equilibrium cutoffs.

Second, we use Lemma 1 to show that, for any S,

$$F(S)/|S| \leq -\theta_{[1]},$$

that is, the negative of the lowest cutoff $\theta_{[1]} \equiv \min_{j \in N} \theta_j^*$ provides an upper bound of the average F over S. This inequality, shown in Lemma 5 in Appendix C, builds an interesting link between the lowest cutoff and the maximization program in step 1 of SAND, and it follows simply by summing up the agents' equilibrium conditions (equation (9)) over S:

$$0 = \underbrace{\left(\sum_{i \in S} \theta_i^* \right)}_{\geq |S| \theta_{[1]}} + \underbrace{\left(v(S) + \phi \sum_{i \in S} \sum_{j \in N_i} w_{ij}^* \right)}_{\geq v(S) + \phi e(S) = F(S)}. \tag{13}$$

The first inequality in (13) is obvious, since $\theta_{[1]}$ is the lowest cutoff. The second inequality follows, as each link among the members of S contributes ϕ in the double summation above by Lemma 1, and these extra terms in the summation are non-negative.

In fact, the upper bound $-\theta_{[1]}$ is achievable for $S = \mathcal{A}_1 = \{i \in N | \theta_i = \theta_{[1]}\}$ (the set of agents with the lowest cutoff), as both inequalities in (13) are equalities for $\mathcal{A}_1$:

$$\frac{F(\mathcal{A}_1)}{|\mathcal{A}_1|} = -\theta_{[1]} \geq \frac{F(S)}{|S|}, \quad \forall S.$$

In other words, $\mathcal{A}_1$, the set of agents who have the highest propensity to adopt (with the lowest cutoff), is a maximizer to the program in step 1 of SAND. It turns out that $\mathcal{A}_1$ is the largest maximizer, that is, $\mathcal{A}_1 = A_1^*$.[30] Further, the minus of the maximum value in step 1 corresponds to the lowest cutoff in the equilibrium ($\theta_{[1]} = -t_{[1]}^*$).

After establishing that $\mathcal{A}_1 = A_1^*$ (i.e., the output in step 1 of the SAND algorithm is exactly the set of nodes with the lowest cutoff), the algorithm continues and in step 2 selects the subset of nodes with the second-lowest cutoff. The process then continues until it stops when all the players are included.

[29] That is, $\mathbb{E}[s_1 \geq x_1 | s_2 = x_2] = G(x_2 - x_1)$.

[30] To show the maximality, we note that the second inequality in (13) is strict whenever S contains any node with a strictly higher cutoff than the lowest cutoff; thus, such S cannot be a maximizer.

3.4 Social connectedness in networks

The algorithmic approach in Theorem 1 provides a new method of determining who has the highest propensity to adopt and why. In the standard network games with continuous actions (see, e.g., Jackson 2008, Jackson and Zenou 2015, Bramoullé and Kranton 2016), the most central agents (in terms of degree, Katz-Bonacich, or eigenvector centrality) or, equivalently, the most *interconnected* agents are the ones who exert the highest efforts. This is not necessarily true in our model. In fact, in our model, the agents with the highest *social connectedness* —that is, those with high *interconnectedness* and high *embeddedness*— are the ones who have the highest propensity of adopting. Let us first explain the notions of social connectedness in our framework and then define it formally.

Consider the network in Figure 2 (for simplicity, we set $\phi = 1$ and $v_i = 0$ for every i; thus, $F(S) = e(S)$). The SAND stops in three steps with

$$A_1^* = S_1,\ A_2^* = S_1 \cup S_2, A_3^* = S_1 \cup S_2 \cup S_3 = N,$$

where $S_1 = \{1, 2, 3, 4, 5, 6\}$, $S_2 = \{i, j, k\}$, and $S_3 = \{i', j', k'\}$. Indeed, in step 1, the SAND algorithm finds the subset that has the highest average degree $e(S)/|S|$, which is clearly $A_1^* = S_1$ with $t_{[1]}^* = e(S_1)/|S_1| = 5/2$. Next, in step 2, the SAND algorithm determines the subset S, which has the highest value of $\frac{e(S)-e(A_1^*)}{|S|-|A_1^*|}$ (see (4)). It turns out that $A_2^* = S_1 \cup S_2$ and $t_{[2]}^* = 5/3$. In step 3, the SAND algorithm terminates with $A_3^* = N$ and $t_{[3]}^* = \frac{4}{3}$. Thus, we end up with three coordination sets: $C_1^* = S_1$, $C_2^* = S_2$, and $C_3^* = S_3$.

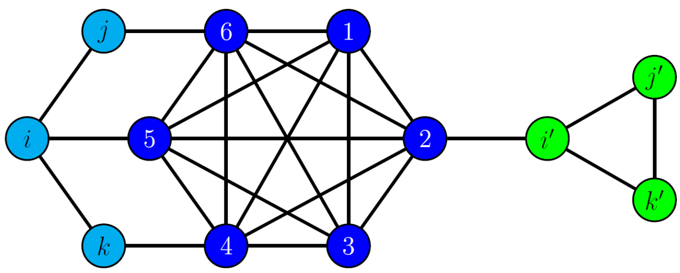

Figure 2: Interconnectedness versus embeddeness

What is notable is that agents in S_2, instead of S_3, are found in step 2 since

$$t_{[3]}^* = \frac{e(S_1 \cup S_3) - e(A_1^*)}{|S_1 \cup S_3| - |A_1^*|} = \frac{4}{3} < \frac{e(S_1 \cup S_2) - e(A_1^*)}{|S_1 \cup S_2| - |A_1^*|} = \frac{5}{3} = t_{[2]}^*.$$

This implies, in particular, that agent $j \in S_2$ has a higher propensity to adopt than agent $i' \in S_3$, although j has fewer links than i'. Note that this observation holds for any value of ϕ in our model (see Remark 6 and footnote 35). On the contrary, using Katz-Bonacich centrality (Ballester et al. 2006), i' is more influential than j when the peer effect is not overly strong.[31]

To understand this puzzling observation, we formally define the concept of social connectedness, which encompasses two (new) notions —interconnectedness and embeddedness— through the lens of SAND and show that they are both of significance in determining who has the higher propensity to adopt in the network. Define $L(S', S'')$ for $S' \cap S'' = \emptyset$ to be the number of links from agents in $S' \subset N$ to agents in $S'' \subset N$.[32] Then, for $S \supsetneq A^*_{k-1}$, we have

$$e(S) - e(A^*_{k-1}) = e(S \backslash A^*_{k-1}) + L(S \backslash A^*_{k-1}, A^*_{k-1}),$$

such that step k in SAND in (4) can be written in the following manner by setting $T = S \backslash A^*_{k-1}$:[33]

$$\max_{\emptyset \neq T \subseteq N \backslash A^*_{k-1}} \overbrace{\frac{e(T)}{|T|}}^{\text{interconnectedness}} + \overbrace{\frac{L(T, A^*_{k-1})}{|T|}}^{\text{embeddedness}}. \tag{14}$$

In other words, $e(T)/|T|$, which measures the average number of links *within* $T = S \backslash A^*_{k-1}$, corresponds to our concept of "interconnectedness," while $L(T, A^*_{k-1})/|T|$, which measures the average number of links *between* T and A^*_{k-1} (note that the agents in A^*_{k-1} have lower cutoffs and adopt earlier than those in T), corresponds to our concept of "embeddedness." Together, the sum $e(T) + L(T, A^*_{k-1})$ divided by the size of the set T, captures our notion of "social connectedness" (see (14)).

Using this new concept, we can now easily explain why agents in S_2 have a higher propensity of adopting than agents in S_3. Indeed, S_2 is less interconnected than S_3, as $\frac{e(S_2)}{|S_2|} = \frac{2}{3} < \frac{e(S_3)}{|S_3|} = 1$, but it is more embedded than S_3, as $\frac{L(S_2, A^*_1)}{|S_2|} = 1 > \frac{L(S_3, A^*_1)}{|S_3|} = \frac{1}{3}$. Because $\frac{2+3}{3} > \frac{3+1}{3}$, S_2 clearly has higher *social connectedness* than S_3.[34]

We believe that this result —that is, more socially connected groups of people are more likely to adopt— is demonstrably new and has intuitions that resonate and extend beyond our specific model. Let us explain this intuition and why it is different from other standard network games with complementarities, like that of Ballester et al. (2006) (BCZ hereafter). In BCZ, what matters the most is the complementarity in actions between agents, so the agents who exert the highest efforts are the ones who generate and receive the largest spillovers from their neighbors. This is captured by how *interconnected* each agent is or, equivalently, her Katz-Bonacich centrality. In our model, we also have

[31] Ballester et al. (2006) provide a microfoundation of Katz-Bonacich centrality using a network game with continuous actions. Since i' has a higher degree than j, i' has a higher Katz-Bonacich centrality than j when the peer effect $\phi > 0$ is not too strong (see Ballester et al. 2006).

[32] Formally, $L(S', S'') = \sum_{i \in S'} d_i(S'')$, which also equals $\sum_{j \in S''} d_j(S')$.

[33] See Observation 2 in Appendix C for the reformulation in the general case with $v_i > 0$ and ϕ that can take any value.

[34] To further see this point, let us consider another subset with three agents, $\tilde{S} = \{i, k, i'\}$ with $e(\tilde{S}) = 1$ and $L(\tilde{S}, A^*_1) = 3$. Note that $\tilde{S}$ is equally embedded as S_1 but is less interconnected than S_1.

complementarity in actions but, on top of it, we have a coordination problem since agents do not know with certainty whether certain neighbors adopt and, thus, the level of spillovers to which they are exposed. This is why both links *within* a coordination set (interconnectedness) and *between* coordination sets (embeddedness) are of significance. In the former, agents need to form expectations regarding their neighbors adopting within the same coordination set (see Lemma 1 (i)), while in the latter, they place a probability of either 0 (higher-cutoff agents) or 1 (lower-cutoff agents) on their neighbors adopting within other coordination sets (see Lemma 1 (ii)).

4 Network topology and coordination sets

Throughout this section, we assume that Assumption 1 holds.

Assumption 1 (Homogeneous intrinsic valuations)**.** *$v_i = v$ for each $i \in N$.*

Remark 6. *Under Assumption 1, $\mathcal{C}^*$ is independent of v and of ϕ.*

Under homogeneous intrinsic values, scaling the size of common valuation v or the size of network effects ϕ has no effect on the coordination sets. Moreover, the cutoff $\boldsymbol{\theta}^*$ is linearly augmented by v and ϕ.[35] Under such homogeneity, we analyze how the network topology in and of itself shapes the coordination sets in subsequent subsections.

4.1 Networks with a single coordination set

In this section, we identify which classes of networks have a single coordination set. The following proposition provides the necessary and sufficient conditions for this property.

Proposition 1 (Single coordination set)**.** *Under Assumption 1, a network $\mathcal{G}$ yields a single coordination set if and only if it is* ***balanced****, in the sense that for every nonempty $S \subset N$,*

$$\frac{e(S)}{|S|} \leq \frac{e(N)}{|N|}. \tag{15}$$

From the lens of the SAND algorithm, the balanced condition in Proposition 1 is a necessary and sufficient condition for SAND to terminate in one step —that is, $K = 1$. The condition in (15) states that a network $\mathcal{G}$ is *balanced* if the average degree of each subnetwork $\mathcal{G}_S$ is no greater than the average degree of the original network $\mathcal{G}$. The following alternative condition for balancedness is useful for certain applications.

Remark 7. *$\mathcal{G}$ is balanced if and only if there exists $\mathbf{w} = \{w_{ij}, (i,j) \in E\}$, such that for all $i, j \in N$: (i) $w_{ij} \geq 0$, (ii) $w_{ij} + w_{ji} = 1$, and (iii) $\sum_{k \in N_i} w_{ik} = \frac{e(N)}{|N|}$.*

[35] Moreover, under Assumption 1, $\boldsymbol{\theta}^* = -v\mathbf{1} + \phi\hat{\boldsymbol{\theta}}^*$, where $\hat{\boldsymbol{\theta}}^*$ denotes the cutoffs at $\mathbf{v} = \mathbf{0}$ and $\phi = 1$.

Indeed, suppose $\mathcal{G}$ is balanced. Then, the SAND algorithm terminates in one step and, thus, we set w_{ij} equal to the w^*_{ij} defined in (9). Conditions (i), (ii), and (iii) are then clearly satisfied by Lemma 1 and equilibrium condition (9). Conversely, if there exists $\mathbf{w}$ which satisfies these three conditions, then we can directly verify that

$$e(S) = 1/2 \sum_{i,j\in S:(i,j)\in E} (w_{ij} + w_{ji}) \leq \sum_{i\in S}\sum_{j\in N_i} w_{ij} = \sum_{i\in S} \frac{e(N)}{|N|} = |S|\frac{e(N)}{|N|}, \text{ for all } S \subset N.$$

The first equality follows by (ii), and the inequality follows by (i). Thus, $\mathcal{G}$ is balanced.[36]

Applying Proposition 1 or Remark 7, we find that seemingly disparate classes of network structures do satisfy this balancedness condition. Hence, for economies in which intrinsic valuations are *the same across individuals*, each of these networks would induce a single coordination set. Network $\mathcal{G}$ is called *regular* if $d_i = d$ for all i. A *tree* is any connected network without cycles. We say network $\mathcal{G}$ is a *regular bipartite network* with disjoint within-set symmetric agent sets B_1 and B_2, with $B_1 \cup B_2 = N$, and of sizes $n_s \equiv |B_s|$ and degrees $d_s \equiv d_i$ for each $i \in B_s$ for sides $s = 1, 2$.

Proposition 2 (Classes of balanced networks)**.** *Under Assumption 1, there exists a single coordination set if $\mathcal{G}$ has at least one of the following properties:*[37]

(1) it is a regular network; or

(2) it is a tree network; or

(3) it is a regular bipartite network; or

(4) it has a unique cycle; or

(5) it has a maximum of four agents.

Proposition 2 provides a (non-exhaustive) list of networks that have one coordination set. Members of all trees, regardless of their size and complexity, adopt together using a common limit cutoff. This generalizes the star network of Example 1. Parts (2) and (4) establish the existence of at least two distinct cycles in $\mathcal{G}$ as a necessary (but generally insufficient) condition for multiple limit cutoffs in equilibrium.[38] More generally, for any network of each of these families, there exists no proper subset of agents who have greater average within-degree (i.e., interconnectedness) than the entire network.

In summary, to determine the (possible) presence of a single coordination set in a network, the only aspect that needs to be checked is that the balanced condition given

[36]The $\mathbf{w}$ satisfying these three conditions might not necessarily correspond to $\mathbf{w}^*$ defined in (9) using the equilibrium cutoffs. Moreover, Gale's Demand Theorem (Gale 1957) provides direct proof of equivalence between condition (15) and the existence of $\mathbf{w}$ satisfying (i), (ii), and (iii) (see Section G.4 in the Online Appendix).

[37]Note that we assume that $\mathcal{G}$ is connected throughout the paper.

[38]The statement in Proposition 2(4) cannot be extended to networks with multiple non-overlapping cycles. Certain networks with multiple non-overlapping cycles are balanced, while some are not.

in (15) or Remark 7 is verified. If this condition is verified for each subset of agents, then we know that there is a single, network-wide coordination set with a common threshold given by Theorem 1. If this condition is not satisfied for at least a subset of agents, then one can use Algorithm 1 (SAND) to construct the coordination sets and cutoffs and the notion of social connectedness to determine who adopts at each step (see, e.g., Example 2).

4.2 Who has the highest propensity to adopt?

When the balanced condition is met, all the agents adopt together. When this condition fails, we determine who has the highest propensity to adopt in the network. We provide a characterization of the first coordination set in the following proposition.

Proposition 3. *Under Assumption 1, the set of agents in N who have the highest propensity of adopting, that is, A_1^* is the unique nonempty set $A \subset N$ that simultaneously satisfies the following conditions:*

(i) For any nonempty subset $\underline{A}$ of A, the average density of $\underline{A}$ is no greater than that of A, that is,

$$\frac{e(\underline{A})}{|\underline{A}|} \leq \frac{e(A)}{|A|}, \quad \forall \emptyset \subsetneq \underline{A} \subseteq A.$$

(ii) For any nonempty subset T of $N \backslash A$, the average number of links across T and A is smaller than the difference in average densities between A and T, that is,

$$\frac{L(T,A)}{|T|} < \frac{e(A)}{|A|} - \frac{e(T)}{|T|}, \quad \forall \emptyset \subsetneq T \subseteq N \backslash A.$$

Condition (i) implies that the subnetwork $\mathcal{G}_A$ itself must be balanced; otherwise, agents in A cannot adopt together. Condition (ii) states that, for any set of agents who have a lower propensity of adopting than A, it must be the case that either the density among themselves (i.e., $\frac{e(T)}{|T|}$) is low or the average number of links connecting to A (i.e.,$\frac{L(T,A)}{|T|}$) is small. Interestingly, the former corresponds to our notion of *interconnectedness*, while the latter corresponds to that of *embeddedness*. In fact, condition (ii) is equivalent to saying that, for any strict superset $\bar{A}$ of A, the average density of $\bar{A}$ is strictly lower than that of A, i.e., $\frac{e(\bar{A})}{|\bar{A}|} < \frac{e(A)}{|A|}$.

Clearly, if A is indeed A_1^* (the agents found in step 1 of the SAND algorithm), then it has the largest density among all S. In particular, conditions (i) and (ii) must hold. Interestingly, these two conditions are sufficiently strong to uniquely pin down A_1^*. Practically, Proposition 3 provides a way to determine who has the highest propensity of adopting by checking these two conditions instead of solving the full limiting equilibrium and cutoffs.

We revisit Examples 1 and 2 to illustrate Proposition 3. In Example 1, the star network is balanced; hence, item (ii) is automatically satisfied for $A = N$. Indeed, all agents adopt together. In Example 2, let $A = \{ic, i = 1, \cdots, 4\}$ denote the core nodes. The subnetwork $\mathcal{G}_A$ is regular with degree 3; therefore, it is balanced (see Proposition 2 case (1)). Condition (ii) in Proposition 3 also holds, as each spoke node has only one link to the core, which has an average density of 3/2. Consequently, the members in the core have the highest propensity of adopting in the unique limiting equilibrium for this example.

As we have seen, the SAND algorithm is central to our analysis. If we are concerned about knowing who has the second-highest propensity to adopt, A_2^*, we can formulate similar conditions as those in Proposition 3 using the intuition guided by SAND. Even though we focus on homogeneous intrinsic values to highlight the impact of network topology, the analysis throughout this section can easily be extended by incorporating heterogeneous intrinsic values. To provide an example, the balanced condition with heterogeneous intrinsic values becomes $\frac{F(S)}{|S|} \leq \frac{F(N)}{|N|}$ for any S.

5 Policy interventions

In this section, we consider the policy implications of our model by first studying the changes in the intrinsic valuations and then studying the changes in network structure.

5.1 Changes to intrinsic valuations

We first investigate how the equilibrium adjusts as intrinsic valuations change.

Proposition 4 (Local contagion)**.**

(i) For a generic $\mathbf{v}$*, there exists a nonempty open neighborhood* $\mathcal{N}(\mathbf{v})$ *around* $\mathbf{v}$*, such that, for any* $\mathbf{v}', \mathbf{v}'' \in \mathcal{N}(\mathbf{v})$*, the coordination sets under* $\mathbf{v}'$ *are the same as those under* $\mathbf{v}''$*.*

(ii) The mapping $\boldsymbol{\theta}^*(\mathbf{v})$ *is piecewise linear, Lipschitz continuous, and monotone. For generic* $\mathbf{v}$*, for each* $i, j \in C_m^*$*, and* $k \notin C_m^*$*:*

$$\frac{\partial \theta_j^*}{\partial v_i} = \frac{-1}{|C_m^*|}, \quad \text{and} \qquad \frac{\partial \theta_k^*}{\partial v_i} = 0. \tag{16}$$

Proposition 4(i) states that, in each step of the SAND algorithm, the solution to the maximization problem is locally invariant in $\mathbf{v}$. Indeed, $\mathbf{v}$ affects the values of $F(\cdot)$ continuously, while the argument of $F(\cdot)$ is discrete. For generic $\mathbf{v}$, in each step of SAND, the solution A_k^* as well as the partition $\mathcal{C}^*$ remain fixed.

Proposition 4(ii) indicates that, increasing the intrinsic value of agent i, v_i, reduces the common cutoff value θ_j^* for any agent j in the coordination set C_m^* containing i such that all these individuals are now more likely to adopt. However, for each k not in C_m^*, the effect of the shock to i does not affect k's cutoff. Thus, this shock to i remains *local* to i's coordination set. Strikingly, the effect of increasing v_i lowers cutoffs at a rate inversely proportional to the size of the coordination set, $|C_m^*|$. In other words, as the number of agents coordinating together increases, the coordination set responds more slowly to an increase in v_i. Intuitively, the agents in C_m^* continue to coordinate together for small shocks (Proposition 4(i)). As we change v_i by one (infinitesimal) unit, the aggregate intrinsic valuations within C_m^* adjust by exactly one unit, while the network effects —captured by the partition— remain the same. Thus, the average incentive to adopt across the coordination set C_m^* increases by exactly one unit.[39] We also see with (16) that the cutoffs depend on network $\mathcal{G}$ only through the network's impact on the partition $\mathcal{C}^*$.

Our local contagion result in Proposition 4(ii) is related to that of Ambrus et al. (2014), who developed a model in which the network captures the connections between individuals and serves as social collateral to enforce informal insurance payments. They showed that the network can be partitioned into endogenously organized connected groups called "risk-sharing islands." This partition has the property that agents fully share shocks *within* but only imperfectly across islands; thus, insurance and spillovers are *local*.[40]

In Proposition 4, we established that, following perturbations to intrinsic value v_i, there is a homogeneous response for members of the same coordination set but unresponsiveness to these perturbations for agents outside the perturbed agent's coordination set. Given these results, we ask a few important questions from a planner's perpective. What marginal benefits are realized with adoption subsidies? And which agents' adoptions must be subsidized?

To address these questions, we first put forward the following planner's problems. Consider a planner who shares prior $H(\cdot)$ over θ and holds no private information.[41] The planner has a fixed budget B to be allocated to the users. Recall that, for each given $\mathbf{v}$, Theorem 1 explicitly gives the limiting equilibrium $\boldsymbol{\pi}^* = (\pi_1^*, \cdots, \pi_{|N|}^*)$, where $\pi_i^*(\theta) = \mathbb{I}(\theta \geq \theta_i^*), i \in N$. Here, we did not explicitly mention the dependence of θ_i^* on $\mathbf{v}$.

[39] This observation that the size of the coordination set is not relevant for the overall impact of a marginal subsidy is related to a similar observation in Sákovics and Steiner (2012), who showed that the population size of the target groups is not relevant for subsidy targeting.

[40] The local effect of "contagion" is consistent with the results of several empirical studies. For example, Angelucci et al. (2018) showed that Progresa, a conditional cash transfer program in rural Mexico, had a positive impact on the consumption of the *untreated* (i.e., spillover effects), but this was mainly due to increase in the consumption of households who were *relatives* of the treated. Similarly, using an educational program targeted to very young kids (early childhood) in disadvantaged neighborhoods in Chicago, List et al. (2020) showed that there were very strong spillover effects from treated to untreated kids, but these spillover effects were *local*: they mainly operated within a 5 km radius. Rosenthal and Strange (2008) also indicated that human-capital spillover effects (i.e., proximity to college-educated workers) attenuate with geographical distance, and most of the spillover effects occur within five miles.

[41] One can interpret this to capture either the agents' expertise relative to the planner or the fact that the policy faced by the planner takes place before the realization of private information.

Using this equilibrium, we define $TA(\mathbf{v})$ as the expected aggregate adoption:

$$TA(\mathbf{v}) \equiv \sum_{j\in N} \mathbb{E}[\pi_i^*(\theta)] = \sum_{j\in N}(1 - H(\theta_j^*)), \tag{17}$$

and $TW(\mathbf{v})$ as the expected aggregate welfare:

$$TW(\mathbf{v}) \equiv \mathbb{E}[W(\boldsymbol{\pi}^*(\theta), \theta)], \tag{18}$$

where $W(\mathbf{a}, \theta) = \sum_{i\in N} u_i(\mathbf{a}, \theta)$.

We assume that the planner can subsidize but cannot tax agents. Given a budget $B > 0$, the planner can allocate funds subject to the following feasible budget set:

$$K(\mathbf{v}, B) \equiv \left\{ \tilde{\mathbf{v}} \in \mathbb{R}^{|N|} : \tilde{v}_j \geq v_j, \forall j \in N, \text{ and } \sum_{j\in N} \tilde{v}_j - v_j \leq B \right\}.$$

We say that a policy $\tilde{\mathbf{v}} \in K(\mathbf{v}, B)$ is "expended" if $\sum_{j\in N} \tilde{v}_j - v_j = B$.

An adoption-maximization planner, henceforth, an "A-planner," solves $\max_{\tilde{\mathbf{v}}\in K(\mathbf{v},B)} TA(\tilde{\mathbf{v}})$. A welfare-maximization planner, henceforth, a "W-planner," solves $\max_{\tilde{\mathbf{v}}\in K(\mathbf{v},B)} TW(\tilde{\mathbf{v}})$. Recall the sequence of nested sets $(A_1^*, A_2^*, \ldots, N)$ derived in SAND; we define $A_0^* \equiv \emptyset$. By Proposition 4(i), for generic $\mathbf{v}$, there exists an open neighborhood $\mathcal{N}(\mathbf{v})$ of $\mathbf{v}$ such that the coordination sets are locally constant in $\mathcal{N}(\mathbf{v})$. When B is sufficiently small, the solutions to these problems are simple and are characterized as shown below.

Proposition 5 (Optimal policies)**.** *For sufficiently small B, the set of solutions to the A-planner's problem is given by the set of expended $\tilde{\mathbf{v}}$ satisfying $\tilde{v}_i > v_i$, if and only if i maximizes $H'(\theta_i^*)$, and the set of solutions to the W-planner's problem is given by the set of expended $\tilde{\mathbf{v}}$ satisfying $\tilde{v}_i > v_i$ if and only if $i \in C_m^* \subseteq A_k^*$ maximizes:*

$$1 - H(\theta_i^*) + \phi\left(\frac{L(C_m^*, A_{k-1}^*) + e(C_m^*)}{|C_m^*|}\right) H'(\theta_i^*). \tag{19}$$

Let us interpret (19). First, the aggregate marginal welfare is now decreasing in θ_i^* through the direct effect on target i's ex-ante welfare, quantified by $(1 - H(\theta_i^*))$, for $i \in C_m^*$. Second, the W-planner values the additional externalities among members of the targeted coordination set, as these agents jointly increase their total adoption. Third, adoption subsidies can generate positive welfare gains to coordination sets that may not contain the target agent i. More precisely, provided that an agent j is either in C_m^* (along with i) or takes cutoff $\theta_j^* < \theta_i^*$ and is a neighbor to a member of C_m^*, agent j obtains additional value in all additional state realizations in which her neighbors in C_m^* begin to adopt. Collectively, these components determine (19).

Remark 8. *The targeting problems of the A-planner and the W-planner each reduce to targeting key coordination sets rather than key players when B is sufficiently small.*

The choice of targeting coordination sets that maximize $H'(\theta_i^*)$ can be interpreted in the following manner. A subsidy to member i's adoption increases adoption among other members of C_m^*, while having no influence on members of other coordination sets. The effect on the adoption of each member of C_m^* is inversely proportional to $|C_m^*|$ by Proposition 4. Therefore, the aggregate marginal effect of these adoption-based policies is left as a function of the targeted coordination set C_m^* through the steepness of H at θ_i^* (which captures the probability of occurrence of the state).

The next remark follows immediately as a corollary to Proposition 5. For this, assume that SAND terminates in $K > 1$ steps.[42]

Remark 9. *Assume that B is sufficiently small. For $H'(.)$ decreasing, the A-planner and the W-planner both target members of A_1^*. For $H'(.)$ increasing, (i) the A-planner targets members of $A_K^* \backslash A_{K-1}^*$, and (ii) there exists $\bar{\phi} > 0$, such that, if $\phi < \bar{\phi}$, then the W-planner targets members of A_1^*.*

The A-planner maximizes adoption and, therefore, $H'(\theta_i^*)$ (Proposition 5), the probability density function of the occurence of θ. Thus, when $H'(.)$ is decreasing —that is, *low* values of θ are more likely to occur— the A-planner will always subsidize individuals belonging to the coordination set with the lowest θ and, thus, targets members of A_1^*. When $H'(.)$ is increasing, the reverse is true and, thus, the A-planner targets members of $A_K^* \backslash A_{K-1}^*$.

The W-planner maximizes welfare, which is equivalent to maximizing (19) (Proposition 5). In other words, contrary to the A-planner, the W-planner takes into account both the *direct* effect of the subsidy of individual i on the probability $1 - H(\theta_i^*)$ that she adopts as well as the *indirect* effect on the adoption of i's neighbors, which depends on ϕ, the intensity of the spillover effect, $\left(\frac{L(C_m^*, A_{k-1}^*) + e(C_m^*)}{|C_m^*|}\right)$, the social connectedness of C_m^* that contains i, which is the sum of its "interconnectedness" $e(C_m^*)/|C_m^*|$ and its "embeddedness" $L(C_m^*, A_{k-1}^*)/|C_m^*|$, and $H'(\theta_i^*)$, the probability of occurence of θ_i^*. If $H'(\cdot)$ is decreasing, then the W-planner will always target members of A_1^* since they are more likely to adopt because the probability that the state θ has a low value is very high. In this case, only i and members of i's coordination set A_1^* will benefit from this subsidy policy since there are no spillover effects to higher coordination sets (local contagion). If $H'(\cdot)$ is increasing, then there is trade off. On the one hand, because of the direct effect of the subsidy, the W-planner wants to target individuals with the lowest θ —that is, members of A_1^*— because they have the highest probability to adopt. On the other hand, because of the indirect effect of i's neighbors, the W-planner wants to target members of the coordination set that has high social connectedness, —that is, numerous links within the same coordination set C_m^* (i.e., high "interconnectedness") and numerous links with coordination sets with lower cutoffs (i.e., high "embeddedness"). Remark 9 reveals that, when ϕ is sufficiently small, the indirect effect is negligible and only the direct effect matters. Thus, the W-planner will always target members of A_1^*, which is the extreme opposite of what the A-planner will target in this case.

[42] By Proposition 4(i), K, like the coordination sets, does not vary with $\mathbf{v}$ locally.

Clearly, when ϕ is sufficiently large (i.e., $\phi > \bar{\phi}$), Remark 9 will no longer hold for the W-planner when $H'(\cdot)$ is increasing. The W-planner's key coordination set will lie in some A_k^*, where the value of k will depend on the social connectedness of the coordination set. Let us illustrate this with the network in Figure 2 and, for the sake of simplicity, set $v_i = 0$ for every i, such that $F(S) = \phi e(S)$. It is evident that the members of $S_1 = \{1, 2, 3, 4, 5, 6\}$ have the *highest* propensity of adopting ($\theta_1^* = -5\phi/2$), those in $S_2 = \{i, j, k\}$ have the second highest propensity of adopting ($\theta_2^* = -5\phi/3$), and those in $S_3 = \{i', j', k'\}$ have the *lowest* propensity of adopting ($\theta_3^* = -4\phi/3$).

Assume that $H'(.)$ is increasing (states with higher θ are more likely to occur) and $\phi > \bar{\phi}$. We have seen in Section 3.4 that

$$\frac{L(S_2, A_1^*) + e(S_2)}{|S_2|} = \frac{3+2}{3} = \frac{5}{3} > \frac{4}{3} = \frac{1+3}{3} = \frac{L(S_3, A_1^*) + e(S_3)}{|S_3|}.$$

In other words, the coordination set S_2 has higher social connectedness than S_1. Consequently, the W-planner may want to target the coordination set S_2 because, compared to members of S_3, they have a higher social connectedness and a higher propensity to adopt. Further, compared to members of S_1, they have a lower probability of adopting but generate more externalities. Indeed, when the W-planner subsidizes an individual i member of S_2, it reduces the cutoff θ_i^* of all members of S_2 but also generates positive externalities on members of S_1. On the contrary, when it subsidizes a member of S_1, there are no positive externalities on members of other coordinations sets (local contagion).

Observe that, if the planner had a bigger budget, B, she would then be able to offer larger subsidies. In this case, the planner would be more likely to subsidize nodes rather than coordination sets because she will influence the manner in which the coordination sets are determined. In particular, the planner will subsidize highly interconnected and embedded agents to maximize either total welfare or total adoption. The problem is that, in this case, we would not have obtained the simple analytical results as in Proposition 5 and Remark 8.

5.2 Changes to the network structure

In this subsection, we consider two policies that affect the network structure.

5.2.1 Key players

In the previous subsection, we analyze a policy that aims at subsidizing agents in the network. The network literature has also focused on targets in networks, particularly key players. In the context of crime, Ballester et al. (2006) defined the key player(s) as the player who, once removed, reduces total crime the most (see Zenou 2016 for an overview of the key-player policies). Let us now consider this policy in our model.

Consider a crime model (instead of a technology adoption model) in which $a_i = 1$ implies that agent i commits a crime, and $a_i = 0$ implies that agent i does not commit a crime. The objective of the planner is to find the key player(s) —that is, the player(s) whose removal from the network leads to the highest total crime reduction.

Contrary to Ballester et al. (2006), we cannot calculate a precise formula that determines the key players in each network (the so-called intercentrality measure). However, we can still determine the key players using the SAND algorithm. The process is to start with a network of n players and first remove player 1. Then, run the SAND algorithm, and obtain the different coordination sets. Then, for a given distribution of $H(\theta)$, we calculate the total expected crime rate. Next, remove player 2 (while adding player 1 back), and proceed in the same manner to determine the total expected crime. Continue until we reach player n. The key player is the one that leads to the highest reduction in total expected crime. We then ascribe a key-player ranking of all players in the network in terms of reduction of total expected crime.

In Appendix E.1, we illustrate this methodology with an example and compare the rankings of the key players in our model with those in Ballester et al. (2006)'s model. We show that the key-player rankings are in fact relatively similar but may differ because our concept of *social connectedness* (i.e., interconnectedness and embeddedness) is different but related to the intercentrality measure of key players in Ballester et al. (2006).

5.2.2 Adding links

Next, we investigate how adding a link in the network affects adoption. Consider the network $\mathcal{G}_{+ij}$, which is defined as the network created by adding the additional link ij in $\mathcal{G}$, and $\mathcal{C}^*_{+ij}$, which is the limit partition under $\mathcal{G}_{+ij}$. While adding links may affect the limit partition, SAND can be employed to verify when the coordination set is unchanged ($\mathcal{C}^*_{+ij} = \mathcal{C}^*$) or changed ($\mathcal{C}^*_{+ij} \neq \mathcal{C}^*$). Let θ^*_k and $\theta^*_{k,+ij}$ correspond to the cutoffs of agent k under networks $\mathcal{G}$ and $\mathcal{G}_{+ij}$, respectively.

Proposition 6 (linkage)**.** *Take i, j with $i \in C^*_m$, $ij \notin E$.*

- (i) *Assume that $\mathcal{C}^*_{+ij} = \mathcal{C}^*$. If*
  - $(i1)$ *$j \notin C^*_m$ with $\theta^*_i > \theta^*_j$, then*
  $$\theta^*_i - \theta^*_{i,+ij} = \phi\frac{1}{|C^*_m|}, \qquad \textit{and} \qquad \theta^*_{j,+ij} = \theta^*_j;$$
  - $(i2)$ *$j \in C^*_m$, then*
  $$\theta^*_i - \theta^*_{i,+ij} = \phi\frac{1}{|C^*_m|}.$$
- (ii) *Assume that $\mathcal{C}^*_{+ij} \neq \mathcal{C}^*$. If*
  - $(ii1)$ *$j \notin C^*_m$ with $\theta^*_i > \theta^*_j$, then $\theta^*_i > \theta^*_{i,+ij} \geq \theta^*_{j,+ij}$, where $\theta^*_{j,+ij} = \theta^*_j$ if $\theta^*_{i,+ij} \neq \theta^*_{j,+ij}$;*
  - $(ii2)$ *$j \in C^*_m$, then i and j are in the same coordination set in $\mathcal{C}^*_{+ij}$.*

First, consider part (i) of Proposition 6 when $\mathcal{C}^*_{+ij} = \mathcal{C}^*$. We establish a disparity in the effects of adding links on equilibrium cutoffs. While the addition of a link between agents belonging to two different coordination sets (part $(i1)$) unambiguously encourages adoption among agents who take higher cutoffs, the equilibrium adoption of the agent taking a lower cutoff is not influenced by the additional link. In other words, the inclusion of links between agents i and j in distinct coordination sets will likely expand the adoption outcomes within i's coordination set, taking a higher cutoff, but have zero influence on adoption within j's coordination set, taking a lower cutoff. Now, consider the case in which the additional link is between agents belonging to the same coordination set (part $(i2)$). While the inclusion of link ij for agents i and j in the same coordination set directly influences incentives to adopt within the coordination set, the decrease in the equilibrium cutoff is comparable to that resulting from a single link to an agent who takes a lower cutoff.

Now consider part (ii) of Proposition 6, when $\mathcal{C}^*_{+ij} \neq \mathcal{C}^*$. The additional link between i and j always decreases the cutoff of i, taking a higher cutoff, toward the cutoff of j. If the link does not bring the two agents into the same coordination set, then j's cutoff is unchanged. Otherwise, the link brings the two agents into the same coordination set. In other words, the addition of a link never causes a linking agent who initially takes a higher cutoff to jump the other agent's cutoff. Intuitively, linkage only increases the incentives of the two agents to coordinate their adoption choices.

We illustrate these results with an example in Appendix E.2 and show the importance of social connectedness when adding links.

5.3 Complete versus incomplete information

We now explain our contribution with respect to the complete-information case. First, in most real-world situations, incomplete information is ubiquitous and agents have limited information on the state variable θ (e.g., the benefits of a new technology, a new crop, or the value of migrating). Second, in supermodular games with complete information (Topkis 1998; Vives 1990; Milgrom and Roberts 1990) —where agents can perfectly observe the state of the world defined as a common component to the technology's value— the set of pure strategy Nash equilibria is a complete, non-empty sub-lattice and contains the greatest (maximal equilibrium)[43] and lowest (minimal equilibrium)[44] elements. In these types of games in which multiple equilibria prevail, the minimal or the maximal equilibrium is usually selected.[45]

In Online Appendix J, using a simple example, we show that the unique equilibrium we select using the global game does not always correspond to the minimal or maximal

[43] The equilibrium in which the largest possible set of people (in set inclusion) take the action, given the underlying state of the world.

[44] The equilibrium in which the smallest possible set of people (in set inclusion) take the action, given the underlying state of the world.

[45] For exmple, the selection of extreme equilibrium is employed in Elliott, Golub, and Jackson (2014) on financial networks and in Immorlica, Kranton, Manea, and Stoddard (2017) on social status.

equilibrium. We also show that the comparative statics properties of these equilibria can be rather different between the complete and incomplete information cases.

6 New predictions on adoption

In this section, we compare the predictions of our model (referred to as LZZ) to leading alternatives in the literature of network games (e.g., Ballester et al. (2006) (BCZ)), and of global games (e.g., Sákovics and Steiner (2012) (SS)). We introduce prices in these three models and show that our model generates adoption that is neither smooth nor fully coordinated but comes in *batches*.[46] In other words, the share of agents who adopt as a function of the price is neither a continuous function (as in BCZ; see Figure 3) nor a jump at a single threshold (as in SS; see Figure 4) but instead a step function (Figure 6). We then derive the comparative statics results of these models by showing how the adoption curves change for increased spillover effects, ϕ, which is a common parameter in all three models. We provide the intuition of our results and show how our model can guide future empirical research.

6.1 Price, aggregate demand, and comparative statics in Ballester et al. (2006) (BCZ)

In Appendix F.1, we introduce price p in the model of BCZ and derive the aggregate demand $D^{BCZ}(p)$ (given by (33)). Typically, $D^{BCZ}(p)$ is first linear in p with slope $b(\mathbf{G}, \phi)$, the aggregate Katz-Bonacich centrality in the network, and then truncated at the cutoff point r (see Figure 3), above which the aggregate demand is zero. In particular, $D^{BCZ}(p)$ is *continuous* in price p. When the spillover effect ϕ increases, $b(\mathbf{G}, \phi)$ increases and, thus, the demand becomes more responsive to price change as the slope becomes steeper (see Figure 3).

6.2 Price, aggregate demand, and comparative statics in Sákovics and Steiner (2012) (SS)

In Appendix F.2, we introduce price p in the model of SS and derive the aggregate demand $D^{SS}(p)$ in (35) for a given realization of the state θ and for $\nu \to 0$. For the adoption curve, there is a *jump at a single threshold* $p = p^*$ (given by (34)), for which all agents adopt if p is below p^*, and do not adopt otherwise. This feature follows from the common limiting threshold property in SS. Further, when ϕ —the benefit of adopting— increases, the threshold p^* increases and, thus, the aggregate demand $D^{SS}(p)$ shifts to the right (Figure 4).

[46] For the technical details of these three models with price, see Appendix F.

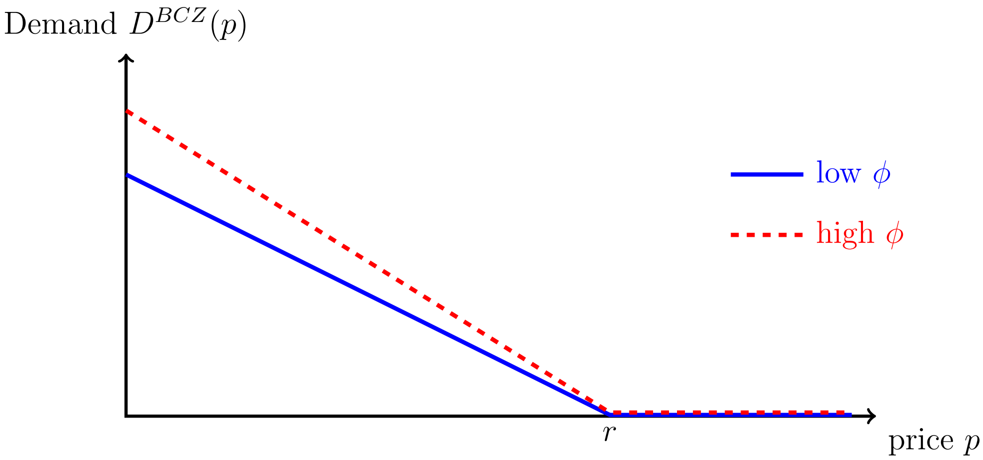

Figure 3: Aggregate demand in Ballester et al. (2006)

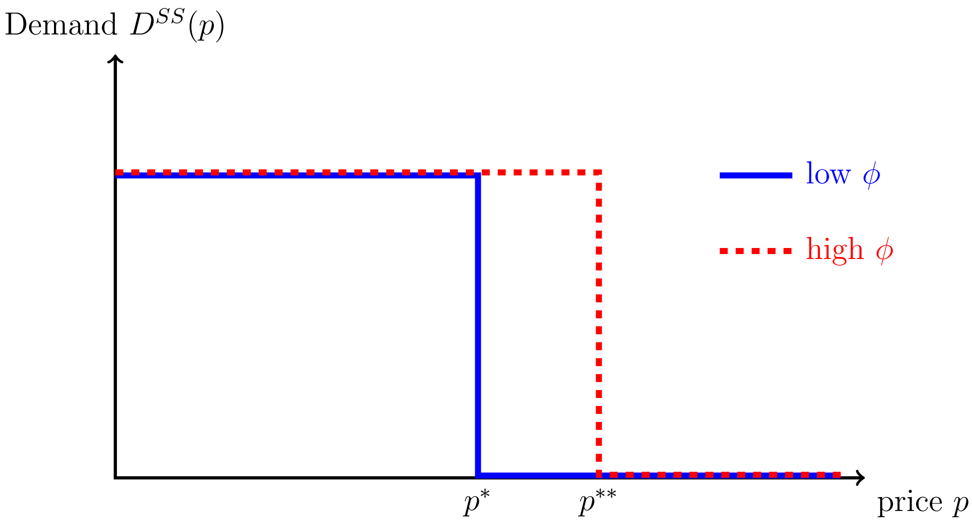

Figure 4: Aggregate demand in Sákovics and Steiner (2012)

6.3 Price, aggregate demand, and comparative statics in our model (LZZ)

In Appendix F.3, we introduce price p in our model and derive the aggregate demand $D^{LZZ}(p)$ in (36) for a given realization of the state θ and for $\nu \to 0$. Unlike SS, in our model, there are multiple threshold prices $p^*_{[i]}$, where the subscript $[i]$ refers to the coordination set A^*_i. This implies that agents in A^*_i all adopt if the adoption price p is below $p^*_{[i]}$ and none of them adopt if it is above $p^*_{[i]}$. Consequently, the aggregate demand $D^{LZZ}(p)$ is a step function with *multiple jumps* at different thresholds.

For the sake of concreteness, we use one real-world network in rural India, which is studied in Banerjee et al. (2013) and displayed in Figure 5,[47] to plot the aggregate demand $D^{LZZ}(p)$ in Figure 6. It is evident that small price changes occasionally have no effect on adoption and other times have large abrupt effects, depending on whether or not the price change occurs in the vicinity of a threshold $p^*_{[i]}$. This prediction in terms of multiple batches of different sizes cannot occur in either BCZ or SS. Indeed, in BCZ, a change in price always has a small (continuous) impact on aggregate demand (Figure 3). In SS, a small change of p has either no effect or a huge effect (Figure 4).[48] When ϕ —the intensity of the spillover effects— increases, the demand curve of LZZ shifts to the right (Figure 6) but the general pattern of multiple jumps remains.

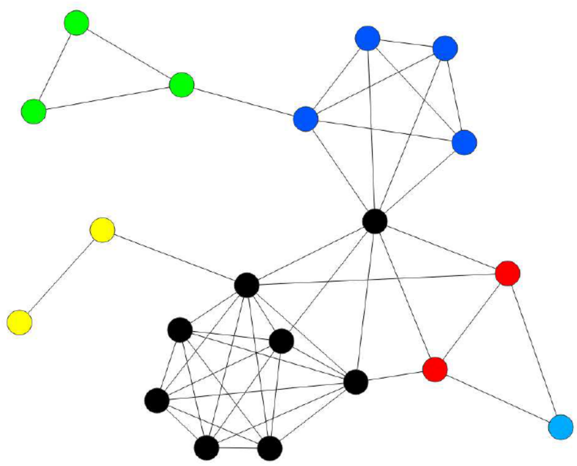

Figure 5: A network from Banerjee et al. (2013)

[47]In Figure 5, different colors indicate different coordination sets. See Appendix F.3 for details.

[48]It must be noted that the general pattern of the aggregate demand with mutiple jumps displayed in Figure 6 will be similar for any network with multiple coordination sets.

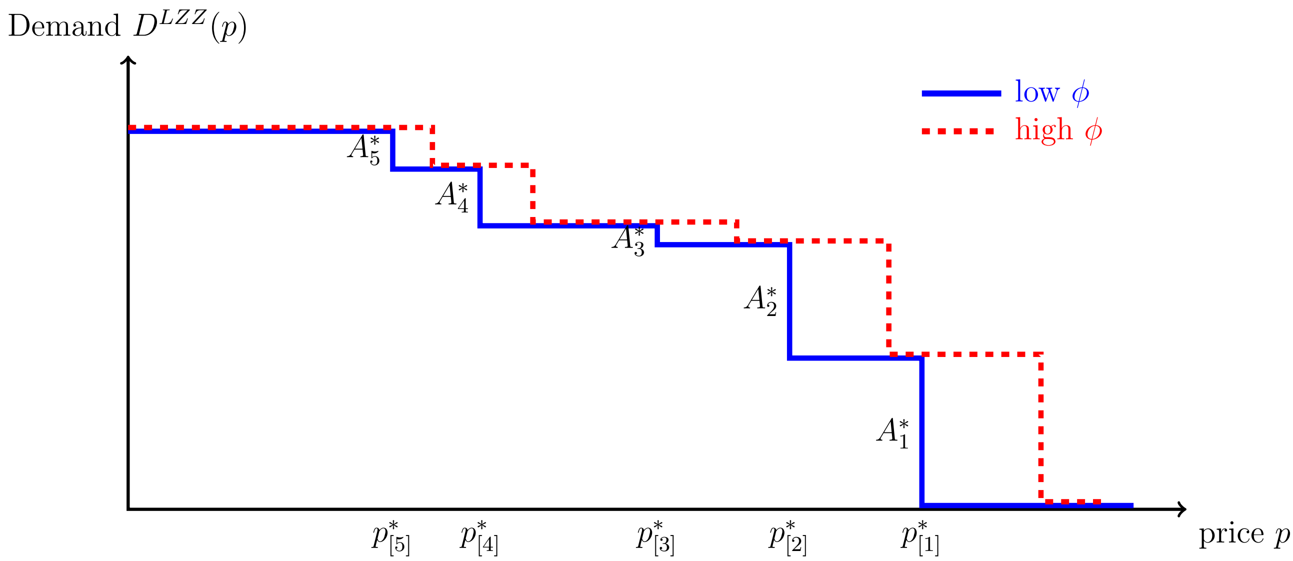

Figure 6: Aggregate demand in our model

6.4 Predictions of our model: Intuition and empirical predictions

The predictions of our model in terms of the adoption curve are, thus, very different from that of BCZ or SS, since adoption is neither *smooth* (BCZ) nor *fully coordinated* (SS) but comes in *batches*.

Intuition: In our model, in order to have a high propensity to adopt, or, equivalently, a low price threshold, agents must have a high interconnectedness and embeddedness —that is, a high social connectedness— because they both generate network effects. Interconnectedness is determined by links to agents within the same coordination set, while embeddedness is determined by links to agents with lower price threshold. Consequently, all agents belonging to a coordination set whose threshold is below (above) the price p will (not) adopt, thereby generating adoption curves with multiple batches. This prediction is unique to our model and did not occur in either BCZ or SS. In particular, BCZ focus on the intensive margin (continuous decisions), while we focus on the extensive margin (discrete decisions); moreover, in SS, players consider a common weighted-average action and, thus, face no heterogeneity in terms of network position.

Empirical predictions: We are not aware of any direct empirical test of our predictions.[49] However, we believe that our model can guide future empirical research on adoption, particularly on the impact of price on adoption rate. Indeed, to test our model, one could design a field experiment (or a lab experiment) where the network is measured as done in Banerjee et al. (2013). At time t, a new technology is introduced in different

[49]Some papers have tested the impact of prices on adoption but without networks. For example, see Ashraf et al. (2010) for a field experiment in Zambia of a home water purification solution and Dupas (2014) for a field experiment in Kenya of a long-lasting insecticide-treated bed net.

villages at a given price p_t. We can design this price such that $p^*_{[2]} < p_t < p^*_{[1]}$, where the subscript $[i]$ refers to the coordination set A^*_i. One could record how many people adopt at time t. We should expect that only agents belonging to coordination set A^*_1 adopt. Then, at time $t+1$, one could go back to the villages and announce that the new technology is now available at a lower price $p_{t+1} < p_t$, such that $p^*_{[3]} < p_{t+1} < p^*_{[2]}$, and expect that agents from coordination sets A^*_1 and A^*_2 adopt. Then, at time $t+2$, we can announce that the new technology is available at a lower price $p_{t+2} < p_{t+1}$ and examine how many new individuals adopt, and so forth. A similar experiment could be implemented in a laboratory with different rounds at different prices.[50] We expect that price changes lead to adoption that comes in batches, as predicted by our model.

7 Concluding remarks

This paper studies a coordination model in networks within a global game environment. The main contributions of this paper are to provide a clean and elegant means of selecting equilibria when information is incomplete by providing an algorithm that computes the limiting cutoffs and to characterize the properties of the cutoffs as a function of the network structure. This characterization enables a partition of the agents into coordination sets —that is, sets of path-connected agents with the same cutoffs. In particular, we show that the individuals who have a higher propensity to adopt are those who have high social connectedness; these are the agents with a high degree —that is, high interconnectedness— and who are well connected to other individuals belonging to coordination sets with lower cutoffs —that is, high embeddedness. We also show that there is a single coordination set if and only if the network is "balanced" —that is, the average degree of each subnetwork is smaller than the average degree of the full network. Importantly, the set of coordination sets is revealed to be instrumental to the comparative statics and welfare properties of the model. In particular, we show that contagion is localized within coordination sets. We also demonstrate that, with a small budget, the planner is indifferent to whom she wants to subsidize within a coordination set and, therefore, targets coordination sets rather than individuals. In addition, we derive the key-player policy and investigate the effect of adding links to the agents' adoption rate.

A possible direction for future research is to study the effects of signaling (Angeletos et al. 2006) or signal jamming (Edmond 2013) on equilibrium properties, such as limit uniqueness and coordination partitioning. Dahleh et al. (2016) studied information exchange through a social network in a symmetric global game; however, the implications of information transmission under a general network game remain an open question. Equilibrium characterizations under more extensive departures from idiosyncratic noise, like the introduction of a public signal, also remain a subject for future research.[51]

[50]The laboratory can be leveraged to control for potential unobserved confounds (e.g., information diffusion) by applying price treatments across distinct subject pools for a given network structure.

[51]See Weinstein and Yildiz (2007) and Morris, Shin, and Yildiz (2016) for contributions.

References

Abadi, J. and M. Brunnermeier (2018). Blockchain economics. Unpublished manuscript, Princeton University.

Ambrus, A., M. Mobius, and A. Szeidl (2014). Consumption risk-sharing in social networks. *American Economic Review 104*(1), 149–82.

Angeletos, G.-M., C. Hellwig, and A. Pavan (2006). Signaling in a global game: Coordination and policy traps. *Journal of Political Economy 114*(3), 452–484.

Angelucci, M., G. de Giorgi, and I. Rasul (2018). Consumption and investment in resource pooling family networks. *Economic Journal 128*, 2613–2651.

Ashraf, N., J. Berry, and J. Shapiro (2010). Can higher prices stimulate product use? evidence from a field experiment in zambia. *American Economic Review 100*(5), 2383–2413.

Babcock, P., K. Bedard, S. Fischer, and J. Hartman (2020). Coordination and contagion: Individual connections and peer mechanisms in a randomized field experiment. *Journal of Public Economics 185*, 104069.

Bailey, M., R. Cao, T. Kuchler, A. Wong, and J. Stroebel (2018). Social connectedness: Measurement, determinants, and effects. *Journal of Economic Perspectives 32*(3), 259–280.

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006). Who's who in networks. wanted: The key player. *Econometrica 74*(5), 1403–1427.

Bandiera, O. and I. Rasul (2006). Social networks and technology adoption in Northern Mozambique. *Economic Journal 116*, 869–902.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013). The diffusion of microfinance. *Science 341*(6144), 1236498.

Basteck, C., D. Tijmen, and F. Heinemann (2013). Characterising equilibrium selection in global games with strategic complementarities. *Journal of Economic Theory 148*(6), 2620–2637.

Bramoullé, Y. and R. Kranton (2016). Games played on networks. In: Y. Bramoullé, Rogers, B.W. and Galeotti, A. (Eds.)*, The Oxford Handbook of the Economics of Networks,* Oxford: Oxford University Press.

Bramoullé, Y., R. Kranton, and M. D'Amours (2014). Strategic interaction and networks. *American Economic Review 104*(3), 898–930.

Cai, J., A. De Janvry, and E. Sadoulet (2015). Social networks and the decision to insure. *American Economic Journal: Applied Economics 7*(2), 81–108.

Calvó-Armengol, A. and J. de Martí (2007). Communication networks: Knowledge and decisions. *American Economic Review, Papers and Proceedings 97*(2), 86–91.

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

Carlsson, H. and E. van Damme (1993). Global games and equilibrium selection. *Econometrica 61*(5), 989–1018.

Chalfin, A. and J. McCrary (2017). Criminal deterrence: A review of the literature. *Journal of Economic Literature 55*(1), 5–48.

Chwe, M. (2000). Communication and coordination in social networks. *Review of Economic Studies 67*(1), 1–16.

Corsetti, G., A. Dasgupta, S. Morris, and H. S. Shin (2004). Does one soros make a difference? A theory of currency crises with large and small traders. *Review of Economic Studies 71*(1), 87–113.

Dahleh, M. A., A. Tahbaz-Salehi, J. N. Tsitsiklis, and S. Zoumpoulis (2016). Coordination with local information. *Operations Research 64*(3), 622–637.

Dai, L. and M. Yang (2019). Organizations and coordination in a diverse population. *Available at SSRN 3051192*.

de Martí, J. and Y. Zenou (2015). Network games with incomplete information. *Journal of Mathematical Economics 61*, 221–240.

Drozd, L. and R. Serrano-Padial (2018). Financial contracting with enforcement externalities. *Journal of Economic Theory 178*, 153–189.

Dupas, P. (2014). Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment. *Econometrica 82*(1), 197–228.

Edmond, C. (2013). Information manipulation, coordination, and regime change. *Review of Economic Studies 80*, 1422–1458.

Elliott, M., B. Golub, and M. O. Jackson (2014). Financial networks and contagion. *American Economic Review 104*(10), 3115–3153.

Frankel, D. M., S. Morris, and A. Pauzner (2003). Equilibrium selection in global games with strategic complementarities. *Journal of Economic Theory 108*(1), 1–44.

Fujishige, S. (2005). *Submodular Functions and Optimization.* Amsterdam: Elsevier Science.

Gagnon, J. and S. Goyal (2017). Networks, markets, and inequality. *American Economic Review 107*(1), 1–30.

Gale, D. (1957). A theorem of flows in networks. *Pacific Journal of Mathematics 7*(2), 1073–1082.

Galeotti, A., S. Goyal, M. O. Jackson, F. Vega-Redondo, and L. Yariv (2010). Network games. *Review of Economic Studies 77*(1), 218–244.

Golub, B. and S. Morris (2017a). Expectations, networks, and conventions. *Available at SSRN: https://ssrn.com/abstract=2979086*.

Golub, B. and S. Morris (2017b). Higher-order expectations. *Available at SSRN: https://ssrn.com/abstract=2979089*.

Guimaraes, B. and S. Morris (2007). Risk and wealth in a model of self-fulfilling currency attacks. *Journal of Monetary Economics 54*(8), 2205–2230.

Hak, N. (2019). Estimation of learning, adoption and diffusion over a network. Unpublished manuscript, Harvard University.

Immorlica, N., R. Kranton, M. Manea, and G. Stoddard (2017). Social status in networks. *American Economic Journal: Microeconomics 9*(1), 1–30.

Islam, A., P. Ushchev, Y. Zenou, and X. Zhang (2018). The value of information in technology adoption: Theory and evidence from bangladesh. CEPR Discussion Paper No. 13419.

Jackson, M. and E. Storms (2017). Behavioral communities and the atomic structure of networks. *mimeo*.

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

Jackson, M. O. and Y. Zenou (2015). Games on networks. In: P. Young and S. Zamir (Eds.), *Handbook of Game Theory, Vol. 4,* Amsterdam: Elsevier Publisher, 91–157.

Kajii, A. and S. Morris (1997). The robustness of equilibria to incomplete information. *Econometrica 65*(6), 1283–1309.

Leister, C. M. (2020). Information acquisition and welfare in network games. *Games and Economic Behavior 122*, 453–475.

Leister, C. M., Y. Zenou, and J. Zhou (2019). Coordination on networks. *SSRN 3082671*.

List, J., F. Momeni, and Y. Zenou (2020). The social side of early human capital formation: Using a field experiment to estimate the causal impact of neighborhoods. NBER Working Paper No. 28283.

Manea, M. (2011). Bargaining in stationary networks. *American Economic Review 101*(5), 2042–2080.

Manea, M. (2021). Bottleneck links, essential intermediaries, and competing paths of diffusion in networks. *Theoretical Economics,* forthcoming.

Milgrom, P. and J. Roberts (1990). Rationalizability and learning in games with strategic complementarities. *Econometrica 58*, 1255–1278.

Milgrom, P. and C. Shannon (1994). Monotone comparative statics. *Econometrica*, 157–180.

Monderer, D. and L. Shapley (1996). Potential games. *Games and Economic Behavior 14*(1), 124–143.

Morris, S. (2000). Contagion. *Review of Economic Studies 67*, 57–78.

Morris, S. and H. S. Shin (1998). Unique equilibrium in a model of self-fulfilling currency attacks. *American Economic Review 88*(3), 587–597.

Morris, S., H. S. Shin, and M. Yildiz (2016). Common belief foundations of global games. *Journal of Economic Theory 163*, 826–848.

Morris, S. and T. Ui (2005). Generalized potentials and robust sets of equilibria. *Journal of Economic Theory 124*(1), 45–78.

Morris, S. and M. Yildiz (2019). Crises: Equilibrium shifts and large shocks. *American Economic Review 109*(8), 2823–54.

Myatt, D. P. and C. Wallace (2019). Information acquisition and use by networked players. *Journal of Economic Theory 182*, 360–401.

Nagurney, A. (1992). *Network Economics: A Variational Inequality Approach.* Amsterdam: Kluwer Academic Publishers.

Nguyen, T. (2015). Coalitional bargaining in networks. *Operations Research 63*(3), 501–511.

Ollar, M. and A. Penta (2017). Full implementation and belief restrictions. *American Economic Review 107*(8), 2243–2277.

Ollar, M. and A. Penta (2019). Implementation via transfers with identical but unknown distributions. Barcelona GSE Working Paper No. 1126.

Oury, M. (2013). Noise-independent selection in multidimensional global games. *Journal of Economic Theory 148*(6), 2638–2665.

Oyama, D. and S. Takahashi (2015). Contagion and uninvadability in local interaction games: The bilingual game and general supermodular games. *Journal of Economic Theory 157*, 100–127.

Oyama, D. and S. Takahashi (2020). Generalized belief operator and robustness in binary-action supermodular games. *Econometrica 88*(2), 693–726.

Rosenthal, S. S. and W. C. Strange (2008). The attenuation of human capital spillovers. *Journal of Urban Economics 64*(2), 373–389.

Sákovics, J. and J. Steiner (2012). Who matters in coordination problems? *American Economic Review 102*(7), 3439–3461.

Serrano-Padial, R. (2018). Coordination in global games with heterogeneous agents. Unpublished manuscript, Drexel University University.

Topa, G. and Y. Zenou (2015). Neighborhood versus network effects. In: G. Duranton, V. Henderson W. Strange (Eds.), *Handbook of Regional and Urban Economics, Vol. 5A,* Amsterdam: Elsevier Publisher, 561–624.

Topkis, D. M. (1998). *Supermodularity and Complementarity.* Princeton: Princeton University Press.

Ui, T. (2001). Robust equilibria of potential games. *Econometrica 69*(5), 1373–1380.

Ushchev, P. and Y. Zenou (2020). Social norms in networks. *Journal of Economic Theory 185*, 104969.

Vives, X. (1990). Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics 19*, 305–321.

Weinstein, J. and M. Yildiz (2007). A structure theorem for rationalizability with application to robust predictions of refinements. *Econometrica 75*, 365–400.

Zenou, Y. (2016). Key players. *In:* Y. Bramoullé and B.W. Rogers and A. Galeotti (Eds.), *The Oxford Handbook of the Economics of Networks, Vol. 4,* Oxford: Oxford University Press, 244–274.

APPENDIX

A Auxiliary results

Here, we present some auxillary results that are used in the proofs.

A.1 Preliminary results on supermodularity

Definition 3. *X is a finite set. $f : 2^X \to R$ is supermodular if for any $S, T \in 2^X$,*

$$f(S \cap T) + f(S \cup T) \geq f(S) + f(T).$$

An equivalent definition of superdmodularity is that for any $S \subseteq T \subseteq N \backslash \{i\}$

$$f(S \cup \{i\}) - f(S) \leq f(T \cup \{i\}) - f(T).$$

In other words, the marginal contribution of i to set S: $f(S \cup \{i\}) - f(S)$, is monotone in S.

Lemma 2. *Suppose $f(S) : 2^X \to \mathbf{R}$ is supermodular. Then A is a maximizer of f over 2^X, if and only if*

$$(\textit{internal optimality}) \quad f(A) \geq f(\underline{A}), \quad \forall \underline{A} \subseteq A.$$

and

$$(\textit{external optimality}) \quad f(A) \geq f(\bar{A}), \quad \forall \bar{A} \supseteq A.$$

When the inequalities for internal and external optimality are strict for $\underline{A} \neq A \neq \bar{A}$, then A is the unique maximizer.

Proof of Lemma 2: The *only if* direction is obvious. For the *if* direction, suppose f satisfies both internal and external optimality. Take any set B; we have $A \cap B \subseteq A, A \cup B \supseteq A$. Therefore, $f(A \cap B) \leq f(A)$ and $f(A \cup B) \leq f(A)$. By supermodularity of f:

$$f(B) \leq f(A \cup B) + f(A \cap B) - f(A) \leq f(A) + f(A) - f(A) = f(A).$$

Since the set B is arbitrary, A must be a maximizer of f.

Note that $f(B) = f(A)$ only if $f(A \cap B) = f(A)$ and $f(A \cup B) = f(A)$. Therefore, when the inequalities for internal and external optimality are strict for $\underline{A} \neq A \neq \bar{A}$, then A is in fact the unique maximizer. □

Lemma 3. *Suppose f is supermodular with $f(\emptyset) = 0$. The following statements are true for the set of maximizers of the program $\max_{A \neq \emptyset} \frac{f(A)}{|A|}$.*

(i) Suppose that A_1 and A_2 are both maximizers of the program $\max_{A\neq\emptyset} \frac{f(A)}{|A|}$. Then, $A_1 \cup A_2$ is also a maximizer of $\max_{A\neq\emptyset} \frac{f(A)}{|A|}$. If, in addition, $A_1 \cap A_2$ is not empty, $A_1 \cap A_2$ is also a maximizer of $\max_{A\neq\emptyset} \frac{f(A)}{|A|}$.

(ii) The largest maximizer exists.

Proof of Lemma 3: For (i), suppose A_1 and A_2 are both maximizers of the program $\max_{A\neq\emptyset} \frac{f(A)}{|A|}$. Denote $\beta = \frac{f(A_1)}{|A_1|} = \frac{f(A_2)}{|A_2|}$; then by optimality of β, $f(A_1 \cap A_2)) \leq \beta|A_1 \cap A_2|$. As a result,

$$\begin{aligned} f(A_1 \cup A_2) &\geq f(A_1) + f(A_2) - f(A_1 \cap A_2)) && (20)\\ &\geq \beta|A_1| + \beta|A_2| - \beta|A_1 \cap A_2| = \beta|A_1 \cup A_2|, && (21)\end{aligned}$$

that is, $\frac{f(A_1\cup A_2)}{|A_1\cup A_2|} \geq \beta$. By optimality of β, so, $\frac{f(A_1\cup A_2)}{|A_1\cup A_2|} = \beta$, thus $A_1 \cup A_2$ must be a maximizer too.

When $A_1 \cap A_2$ is not empty, since we know $\frac{f(A_1\cup A_2)}{|A_1\cup A_2|} = \beta$,

$$\begin{aligned} f(A_1 \cap A_2) &\geq f(A_1) + f(A_2) - f(A_1 \cup A_2)) && (22)\\ &= \beta|A_1| + \beta|A_2| - \beta|A_1 \cup A_2| = \beta|A_1 \cap A_2|, && (23)\end{aligned}$$

so $\frac{f(A_1\cap A_2)}{|A_1\cap A_2|} \geq \beta$. By optimality of β, $A_1 \cap A_2$, which is not empty by assumption, must be a maximizer too.

The result in (ii) directly follows from (i). □

B Uniqueness of limit equilibrium

Proposition 0 (Frankel et al. 2003). *There exists an essentially unique strategy profile $\boldsymbol{\pi}^*$, which is in cutoff strategies, such that any $\boldsymbol{\pi}^{*\nu}$ surviving iterative elimination of strictly dominated strategies in Γ^ν satisfies $\lim_{\nu\to 0} \boldsymbol{\pi}^{*\nu} = \boldsymbol{\pi}^*$.*[52]

Any Bayesian Nash equilibrium of Γ^ν obviously survives iterative elimination of strictly dominated strategies so Proposition 0 applies to $\boldsymbol{\pi}^{*\nu} = (\pi_i^{*\nu} = \mathbb{I}(s_i \geq c_i^{*\nu}))_{i\in N}$ defined by (2). Given an equilibrium in cutoff strategies in Γ^ν always exists for $\nu > 0$, $\boldsymbol{\pi}^*$ is also in cutoff strategies by Proposition 0.

Though there may exist multiple equilibria in cutoff strategies in Γ^ν for a fixed $\nu > 0$, Proposition 0 shows that any selection of such equilibria in Γ^ν must converge to $\boldsymbol{\pi}^*$ as ν goes to zero.

[52]Observe that the elimination of strictly dominated strategies refers to the interim strategic form of the game.

C Proof of Theorem 1 and discussions

C.1 Auxiliary results about SAND

We present several preliminary results before proving Theorem 1. We start with some observations of SAND.

Lemma 4. *$F(S) \equiv \sum_{i\in S} v_i + \phi e(S)$ is supermodular in S.*

Proof of Lemma 4: For any S not including i,

$$F(S \cup \{i\}) - F(S) = v_i + \phi(e(S \cup \{i\}) - e(S)) = v_i + \phi d_i(S)$$

Since $d_i(S) = |N_i \cap S|$ is obviously increasing in S, F is supermodular. $\square$

Observation 2. *For each $T \subseteq N \backslash A^*_{k-1}$, define*

$$\hat{v}(T) \equiv \sum_{i\in T} \hat{v}_i, \quad \hat{F}(T) \equiv \hat{v}(T) + \phi e(T), \tag{24}$$

where $\hat{v}_i = v_i + \phi d_i(A^*_{k-1}), i \in N\backslash A^*_{k-1}$. Then we can equivalently reformulate the program (4) in step k of SAND as follows:

$$\max_{\emptyset \subsetneq T \subseteq N\backslash A^*_{k-1}} \frac{\hat{F}(T)}{|T|} = \frac{v(T) + \phi(e(T) + L(T, A^*_{k-1}))}{|T|}, \tag{25}$$

Furthermore, $A^*_k \backslash A^*_{k-1}$ is the largest maximizer of (25) and $\frac{\hat{F}(A^*_k\backslash A^*_{k-1})}{|A^*_k\backslash A^*_{k-1}|} = t^*_{[k]}$.

Proof of Observation 2: The reformulation is obtained by changing variables. For each $T \subseteq N\backslash A^*_{k-1}$,

$$F(T \cup A^*_{k-1}) - F(A^*_{k-1}) = v(T) + \phi(e(T) + \sum_{i\in T} d_i(A^*_{k-1})) = v(T) + \phi(e(T) + L(T, A^*_{k-1})) = \hat{F}(T),$$

and $|T \cup A^*_{k-1}| - |A^*_{k-1}| = |T|$. The results just follow by observing $\frac{F(T\cup A^*_{k-1}) - F(A^*_{k-1})}{|T\cup A^*_{k-1}| - |A^*_{k-1}|} = \frac{\hat{F}(T)}{|T|}$. $\square$

Observation 3. *The largest maximizers to program (4) exist in step k of SAND.*

Proof of Observation 3: We first show the case for $k = 1$. Since F is supermodular, the union of two maximizers to the program $max_{S \supsetneq \emptyset} \frac{F(S)}{|S|}$ is also a maximizer by Lemma 3; hence, the largest maximizer must exist. The proof for any k follows similarly, using the reformulation of the program (4) in step k of SAND in Observation 2 (note $\hat{F}$ is also supermodular). $\square$

Observation 4. *For program* (4) *in step k of SAND, any connected component of subgraph $\mathcal{G}_{A_k^*\backslash A_{k-1}^*}$ is a maximzier.*[53] *In particular, suppose the subgraph $\mathcal{G}_{A_k^*\backslash A_{k-1}^*}$ is not connected. Then there are multiple maximizers to program* (4) *in step k.*

Proof of Observation 4: We use the reformulation of the program (4) in step k of SAND in Observation 2 using $\hat{F}$. Suppose the subgraph $\mathcal{G}_{A_k^*\backslash A_{k-1}^*}$ is not connected; then we can find B, B' such that $B \cup B' = A_k^*\backslash A_{k-1}^*$, $B' \cap B = \emptyset$, and $L(B, B') = 0$, i.e., the agents in B are not connected to agents in B'. Then $e(B \cup B') = e(B) + e(B')$, and $\hat{F}(B \cap B') = \hat{F}(B) + \hat{F}(B')$. Therefore,

$$t_{[k]}^* = \frac{\hat{F}(B \cup B')}{|B \cup B'|} = \frac{\hat{F}(B) + \hat{F}(B')}{|B| + |B'|} \leq \max(\frac{\hat{F}(B)}{|B|}, \frac{\hat{F}(B')}{|B'|}).$$

By the optimality of $A_k^*\backslash A_{k-1}^*|$, $\frac{\hat{F}(B)}{|B|} \leq t_{[k]}^*$ and $\frac{\hat{F}(B')}{|B'|} \leq t_{[k]}^*$. These inequalities together show

$$\frac{\hat{F}(B)}{|B|} = \frac{\hat{F}(B')}{|B'|} = t_{[k]}^*,$$

implying B and B' are also maximizers of (25). Continuing this process yields that each connected component of subgraph $\mathcal{G}_{A_k^*\backslash A_{k-1}^*}$ is in fact a maximizer of (25). The rest just follows. $\square$

Observation 5. *A unique maximizer exists in each step of SAND for any $\mathbf{v} \in \mathbf{R}^{|N|}$, with the exception of $\mathbf{v}$ in the union of a finite number of hyperplanes in $\mathbf{R}^{|N|}$. Thus for generic $\mathbf{v}$, there is a unique maximizer in each step of SAND.*

Proof of Observation 5: Multiple maximizers exist at $\mathbf{v}$ in a step of SAND only when we can find a triple $A, B, B \in 2^N$ satisfying (i) $A \subsetneq B, A \subsetneq B', B \neq B'$ and (ii) $\frac{F(B)-F(A)}{|B|-|A|} = \frac{F(B')-F(A)}{|B'|-|A|}$. As the network is finite, there exists a finite number of triple $A, B, B \in 2^N$ satisfying condition (1). For each such triple, condition (ii) imposes a linear restriction defining a hyperplane in the space of $\mathbf{v}$. Together, we show that the set of $\mathbf{v}$ generating multiple maximizers is included in the union of a finite number of hyperplanes; thus, it has zero measure. The result just follows. $\square$

Observation 6. *For generic $\mathbf{v}$, the subnetwork $\mathcal{G}_{A_k^*\backslash A_{k-1}^*}$ is connected; hence, $A_k^*\backslash A_{k-1}^*$ is a coordination set by itself, and the partition $\mathcal{C}^* = \{A_1^*, A_2^*\backslash A_1^*, \cdots, A_K^*\backslash A_{K-1}^*\}$.*

[53] Recall that any connected component of subgraph $\mathcal{G}_{A_k^*\backslash A_{k-1}^*}$ is a separate coordination set by Lemma 1 and Theorem 1. So each coordination set C_i^* contained in $A_k^*\backslash A_{k-1}^*$ must be be a maximzier to program (4) in step k of SAND, i.e.,

$$\frac{F(C_i^* \cup A_{k-1}^*) - F(A_{k-1}^*)}{|C_i^*|} = t_{[k]}^* = \frac{F(A_k^*) - F(A_{k-1}^*)}{|A_k^*| - |A_{k-1}^*|}.$$

Proof of Observation 6: It follows directly from Observations 5 and 4. $\square$

The next Lemma will be used repeatedly in the proof.

Lemma 5. *For any $S \subseteq N$:*

$$-\sum_{i \in S} \theta_i^* \geq F(S). \tag{26}$$

Moreover, suppose for any $i \in S$ and $j \in N \backslash S$, $\theta_i^* < \theta_j^*$, the above inequality holds with equality.

Proof of Lemma 5: Taking the summation of (9) over S, we obtain

$$\begin{aligned}
-\sum_{i \in S} \theta_i^* &= \sum_{i \in S} v_i + \phi \sum_{i \in S} \sum_{j \in N_i} w_{ij}^* \\
&= \sum_{i \in S} v_i + \phi \sum_{i \in S} \sum_{j \in N_i \cap S} w_{ij}^* + \phi \sum_{i \in S} \sum_{j \in N_i \cap (N \backslash S)} w_{ij}^* \\
&\geq \sum_{i \in S} v_i + \phi e(S) = F(S).
\end{aligned}$$

The final inequality is obtained by noticing that $\sum_{i \in S} \sum_{j \in N_i \cap S} w_{ij}^* = e(S)$ by Lemma 1 (i) and $\sum_{i \in S} \sum_{j \in N_i \cap (N \backslash S)} w_{ij}^*$ is non-negative and equal to zero when, for any $i \in S$ and $j \in N \backslash S$, $\theta_i^* < \theta_j^*$ by Lemma 1 (ii). $\square$

C.2 Proof of Theorem 1

Now we proceed to prove Theorem 1.

Proof of Theorem 1: Given the equilibrium cutoffs $(\theta_i^*)_{i \in N}$, we define $\theta_{[k]}$, $\mathcal{A}_k, \mathcal{B}_k$, $k = 1, 2, \cdots$ as follows:

$$\begin{array}{ll}
\theta_{[1]} = \min_{j \in N} \theta_j^* & \mathcal{B}_1 = \{s \in N | \theta_s = \theta_{[1]}\}, \quad \mathcal{A}_1 = \mathcal{B}_1, \\
\theta_{[2]} = \min_{j \in N \backslash \mathcal{A}_1} \theta_j^* & \mathcal{B}_2 = \{s \in N | \theta_s = \theta_{[2]}\}, \quad \mathcal{A}_2 = \mathcal{A}_1 \cup \mathcal{B}_2, \\
\cdots & \cdots
\end{array} \tag{27}$$

Intuitively, $\theta_{[1]}$ is the lowest cutoff, while $\theta_{[2]}$ is the second-lowest cutoff, and so on: $\theta_{[1]} < \theta_{[2]} < \cdots$. And $\mathcal{B}_k$ is the agents with cutoff $\theta_{[k]}$, while $\mathcal{A}_k$ is the agents with cutoffs in $\{\theta_{[1]}, \cdots, \theta_{[k]}\}$.

To prove Theorem 1, it suffices to show, for $k = 1, 2, \cdots$,,

$$(\dagger) \qquad\qquad \theta_{[k]} = -t_{[k]}^*, \quad \mathcal{A}_k = A_k^*, \quad \mathcal{B}_k = A_k^* \backslash A_{k-1}^*.$$

The case with $k = 1$ is already proven in subsection 3.3. Now we use induction to show ($\dagger$). Assume it holds for $k-1$. To show the case for k, we first need a similar version of Lemma 5: for any $T \subseteq N \backslash A^*_{k-1}$,

$$-\sum_{i \in T} \theta^*_i \geq \underbrace{v(T) + \phi(e(T) + \sum_{i \in T} d_i(A^*_{k-1}))}_{:=F(T \cup A^*_{k-1}) - F(A^*_{k-1})} \tag{28}$$

Moreover, suppose for any $i \in T$ and $j \in N \backslash (T \cup A^*_{k-1})$, $\theta^*_i < \theta^*_j$, the above inequality holds with equality. The proof is exactly parallel to the proof of Lemma 5 with the exception that each link from $i \in T$ to agent $j \in A^*_{k-1}$ contributes ϕ by Lemma 1 (ii) (note $\theta^*_i > \theta^*_j$), which explains the extra term $d_i(A^*_{k-1})$ in (28).

Following the same logic in step 1 in subsection 3.3 , we show (28) implies

$$\frac{F(T \cup A^*_{k-1}) - F(A^*_{k-1})}{|T|} \leq -\theta_{[k]} = \frac{F(\mathcal{B}_k \cup A^*_{k-1}) - F(A^*_{k-1})}{|\mathcal{B}_k|}$$

Therefore, $\mathcal{B}_k \cup A^*_{k-1}$ is a maximizer in step k of SAND (the fact that it is the largest maximizer follows the same logic as $k = 1$). Therefore, (†) is proven for k. □

Remark 10. *Although the SAND algorithm does not explicitly generate the equilibrium* $\mathbf{w}^* = \{w^*_{ij}, (i,j) \in E\}$ *as outputs, the construction of the algorithm guarantees that such* $\mathbf{w}^*$ *always exists and is consistent with* $\boldsymbol{\theta}^*$ *(Lemma 1).*[54]

D Proofs of other results in the main text

Proof of Observation 1: It follows from the discussion in the main text. □

Proof of Lemma 1: It largely follows from the discussion in the main text.[55] □

Proof of Remark 6: Under homogeneous intrinsic values, $F(S)/|S| = v + \phi e(S)/|S|$. Therefore, in step 1 of SAND:

$$A^*_1 = \underset{S \supsetneq \emptyset}{\text{argmax}} \frac{e(S)}{|S|}.$$

Thus, A^*_1 is independent of (v, ϕ), and $t^*_{[1]} = v + \phi e(A^*_1)/|A^*_1|$. Moreover, assume that in step $k > 1$, A^*_1 through A^*_{k-1} are each independent of (v, ϕ). Therefore:

$$A^*_k = \underset{S \supsetneq A^*_{k-1}}{\text{argmax}} \frac{e(S) - e(A^*_{k-1})}{|S| - |A^*_{k-1}|}.$$

[54]See also Theorem G1.

[55]An identity similar to (11) is shown in Carlsson and van Damme (1993) (see equation (4.1) on page 997 and Lemma 4.1 on page 999). We give a detailed proof of Lemma 1 in the working paper version of this paper (Leister et al. 2019).

Thus, A_k^* is also independent of (v, ϕ), and $t_{[k]}^* = v + \phi\frac{e(A_k^*)-e(A_{k-1}^*)}{|A_k^* \setminus A_{k-1}^*|}$. The result follows by induction on k. $\square$

Proof of Proposition 1: Under homogeneous intrinsic values, $F(S)/|S| = v + \phi e(S)/|S|$. The rest just follows from observing that the network is balanced if and only if the SAND algorithm terminates in step 1. In Section G.4 in the Online Appendix, we provide an alternative proof of Proposition 1 using Gale's Demand Theorem (Gale 1957). $\square$

Proof of Remark 7: The remark follows from the discussion following Remark 7. $\square$

Proof of Proposition 2: For any regular network with degree d, for any non-empty subset S, $2\frac{e(S)}{|S|} = \frac{\sum_{i \in S} d_i(S)}{|S|} \leq \frac{\sum_{i \in S} d}{|S|} = d = 2\frac{e(N)}{|N|}$, so a regular graph is always balanced.

For trees, there are no cycles, so $e(N) = N - 1$, while for each subset S the resulting subnetwork G_S is still cycle-free. Therefore, the number of edges within S is at most $|S| - 1$, so $e(S) \leq |S| - 1$; thus:

$$\frac{e(S)}{|S|} \leq \frac{|S|-1}{|S|} \leq \frac{e(N)}{|N|} = \frac{|N|-1}{|N|}.$$

For regular bipartite networks with two disjoint groups B_1, B_2 with sizes n_1, n_2 and degrees d_1, d_2, clearly $d_1 n_1 = d_2 n_2 = e(N)$. We set $w_{ij} = \frac{n_1}{n_1+n_2}$ if $i \in B_1, j \in B_2$, and $w_{ij} = \frac{n_2}{n_1+n_2}$ if $i \in B_2, j \in B_1$. We use Remark 7 to check that $\mathcal{G}$ is balanced. For each $i \in B_1$:

$$\sum_{k \in N_i} w_{ik} = \frac{n_1}{n_1+n_2} d_1 = \frac{e(N)}{n_1+n_2} = \frac{e(N)}{|N|}.$$

And for each $i \in B_2$:

$$\sum_{k \in N_i} w_{ik} = \frac{n_2}{n_1+n_2} d_2 = \frac{e(N)}{n_1+n_2} = \frac{e(N)}{|N|}.$$

Therefore, $\mathcal{G}$ is balanced.

If $\mathcal{G}$ is a network with a unique cycle, then $e(N) = N$. For each subset S, the resulting subnetwork $\mathcal{G}_S$ contains at most one cycle. This means the number of edges within S is at most $|S|$, so that $e(S) \leq |S|$; thus:

$$\frac{e(S)}{|S|} \leq \frac{|S|}{|S|} = 1 = \frac{e(N)}{|N|}.$$

When $\mathcal{G}$ contains at most four nodes, all networks with three or fewer nodes contain at most one cycle. The only network structures over four nodes that contain more than one cycle are the circle with a link connecting one non-adjacent pair i and j (two networks) and the complete network. For the former, we can show these networks to have one coordination set with weights: $w_{ij} = w_{ji} = 1/2$, $w_{ki} = w_{kj} = 5/8$, and $w_{ij} = w_{ik} = 3/8$

for each $k \neq i, j$. The complete network with 4 nodes and 6 edges is regular; hence, it is balanced. Note that when $N = 5$, there exists a network such that two coordination sets emerge. For example, a core with 4 nodes plus one periphery node has one link to one of the core nodes. $\square$

Proof of Proposition 3: We can rewrite condition (ii) as $\frac{F(T \cup A) - F(A)}{|T|} < \frac{F(A)}{|A|}$, for any $\emptyset \neq T \subseteq N \backslash A$, which is equivalent to saying that $\frac{e(\bar{A})}{|\bar{A}|} < \frac{e(A)}{|A|}, \forall \bar{A} \supsetneq A$. Under homogeneous intrinsic values, A_1^* is uniquely characterized as the largest maximizer of the following:

$$A_1^* = \underset{S \supsetneq \emptyset}{\operatorname{argmax}} \frac{e(S)}{|S|}.$$

Obviously, any solution to the above program, A_1^*, must satisfy both conditions stated in Proposition 3. To prove the other direction, we assume set A satisfies (i) and (ii). Denote $\beta = \frac{e(A)}{|A|}$. Take any set B; we have $A \cap B \subseteq A, A \cup B \supseteq A$, so $e(A \cap B) \leq \beta |A \cap B|$ by condition (i)[56] and $e(A \cup B) \leq \beta |A \cup B|$ by condition (ii). By supermodularity of $e(\cdot)$ (see Observation 4),

$$e(B) \leq e(A \cap B) + e(A \cup B) - f(A) < \beta(|A \cap B| + |A \cup B| - |A|) = \beta |B|,$$

implying that $e(B)/|B| \leq \beta = \frac{e(A)}{|A|}$. Since it holds for any B, it shows that A is indeed a maximizer. To show that it is in fact the largest maximizer, we need on more step. Note that for any B such that $A \cup B$ is a proper superset A, we have a strict inequality $e(A \cup B) < \beta |A \cup B|$ by condition (ii); hence $e(B)/|B| < \beta = \frac{e(A)}{|A|}$. This effectively proves that any B having non-empty intersection with $N \backslash A$ cannot be a maximizer; thus A is indeed the largest maximizer. $\square$

Proof of Proposition 4:

To prove part *(i)*, we first note that, by Observation 5 in Appendix C, for generic $\mathbf{v}$, multiple maximizers cannot occur in each step (see also footnote 21). In other words, in fixing a generic $\mathbf{v}$, suppose SAND gives the sequence of nested sets, $A_1^*, A_2^*, \cdots, A_K^*$ at $\mathbf{v}$ in K steps; then A_k^* is the unique maximizer in step $k = 1, \cdots, K$ of SAND. We claim that for $\mathbf{v}'$ sufficiently close to $\mathbf{v}$, SAND also terminates K steps with (the same) A_k^* being the unique maximizer in step $k = 1, \cdots, K$ of SAND at $\mathbf{v}'$ as well. To prove this claim, we note that, in step 1 of SAND, at $\mathbf{v}$

$$\frac{\sum_{i \in S} v_i + \phi e(S)}{|S|} < \frac{\sum_{i \in A_1^*} v_i + \phi e(A_1^*)}{|A_1^*|}$$

for any nonempty S that is not equal to A_1^* (the inequality is strict by the uniqueness of A_1^*). Since the above inequality is an open condition on v_i for each i, it must hold for any $\mathbf{v}'$ that is sufficiently close to $\mathbf{v}$, implying A_1^* is the unique maximizer in step 1 for

[56]Note this also holds even if $A \cap B$ is empty, as $e(\emptyset) = 0$.

SAND at $\mathbf{v}'$. By simple induction, we show that by any $k = 1, 2, \cdots, K$,A_k^* is the unique maximizer in SAND at $\mathbf{v}'$ as long as $\mathbf{v}'$ is sufficiently close to $\mathbf{v}$.

Since the coordination sets are uniquely pinned down by the sequences $A_k^*, k = 1 \cdots, K$, part *(i)* is proven.

We now show each claim of *(ii)* below.

Lipschitz continuity. By Theorem G1 in the Online Appendix G.1, $-\boldsymbol{\theta}^*(\mathbf{v})$ is the projection of $\mathbf{0}$ onto the space $\Phi(\mathcal{W})$:

$$-\boldsymbol{\theta}^*(\mathbf{v}) = \mathbf{Proj}_{\Phi(\mathbf{W})}[\mathbf{0}],$$

where $\Phi(\cdot)$ and $\mathbf{W}$ are defined in Online Appendix G.1. Since Φ depends on $\mathbf{v}$ in a linear way, we let $\mathbf{K} = \Phi(\mathcal{W})$ when $\mathbf{v} = \mathbf{0}$. Then for any $\mathbf{v}$:

$$\Phi(\mathcal{W}) = \mathbf{v} + \mathbf{K}.$$

We can rewrite the projection problem as follows:

$$-\boldsymbol{\theta}^*(\mathbf{v}) = \underset{\mathbf{z} \in \mathbf{v}+\mathbf{K}}{\operatorname{argmin}} ||\mathbf{z}||^2 = \mathbf{v} + \underset{\mathbf{y} \in \mathbf{K}}{\operatorname{argmin}} ||(-\mathbf{v}) - \mathbf{y}||^2 = \mathbf{v} + \mathbf{Proj}_{\mathbf{K}}[-\mathbf{v}].$$

The projection mapping is nonexpansive (see chapter 1 of Nagurney 1992), i.e: $||\mathbf{Proj}_{\mathbf{K}}[\mathbf{x}] - \mathbf{Proj}_{\mathbf{K}}[\mathbf{y}]|| \leq ||\mathbf{x} - \mathbf{y}||, \forall \mathbf{x}, \mathbf{y} \in \mathbf{R}_n$. So for any $\mathbf{v}$ and $\mathbf{v}'$, we have $||\boldsymbol{\theta}^*(\mathbf{v}) - \boldsymbol{\theta}^*(\mathbf{v}')|| \leq 2||\mathbf{v} - \mathbf{v}'||$. Hence, $\boldsymbol{\theta}^*(\mathbf{v})$ is Lipschitz continuous in $\mathbf{v}$.

Comparative Statics. By Lipschitz continuity, $\mathbf{q}^*(\mathbf{v})$ is differentiable for almost all $\mathbf{v}$. By part *(i)*, for generic $\mathbf{v}$, $\mathcal{C}^*$ is locally constant in $\mathbf{v}$, and we have the following relationship between $\mathbf{v}$ and $t_{[k]}^*$:

$$t_{[k]}^* = \frac{F(A_k^*) - F(A_{k-1}^*)}{|A_k^*| - |A_{k-1}^*|} = \frac{\sum_{i \in A_k^* \backslash A_{k-1}^*} v_i + \phi(e(A_k^*) - e(A_{k-1}^*))}{|A_k^*| - |A_{k-1}^*|}. \tag{29}$$

(Recall that from part *(i)* that the sequences $A_k^*, k = 1, \cdots, K$ in the above equation do not vary with $\mathbf{v}$ when $\mathbf{v}$ changes marginally.) The derivative results follow directly by differentiating equation (29) and applying Theorem 1.[57]

Piecewise linearity. This follows from the results of comparative statics.

Monotonicity. $\partial \boldsymbol{\theta}^* / \partial \mathbf{v}$ is nonpositive, so $\boldsymbol{\theta}^*$ is monotonically decreasing in $\mathbf{v}$. $\square$

Proof of Proposition 5 and Remark 8:

[57]One caveat is that for generic $\mathbf{v}$, $A_k^* \backslash A_{k-1}^*$ is a coordination set by itself, and the partition $\mathcal{C}^* = \{A_1^*, A_2^* \backslash A_1^*, \cdots, A_K^* \backslash A_{K-1}^*\}$ (see Observation 6 in Appendix C).

Giving $\mathbf{v}$, we choose $B > 0$, which is small enough that the budget set $K(\mathbf{v}, B)$ is contained in the neighborhood $\mathcal{N}(\mathbf{v})$ around $\mathbf{v}$. Such a B always exists as $\mathcal{N}(\mathbf{v})$ is open and nonempty by Proposition 4 (i). Furthermore, by Proposition 4 (ii), $\boldsymbol{\theta}^*$ is linear in $\mathbf{v}$ for $\mathbf{v}$ in $\mathcal{N}(\mathbf{v})$; hence, it is in $K(\mathbf{v}, B)$ as well. From now on, we fix such a B.

First, by Proposition 4 (ii), for any j, we have

$$\frac{\partial \theta_j^*}{\partial v_i} = -\frac{1}{|C_m^*|}\mathbb{I}(j \in C_m^*), \tag{30}$$

where C_m^* is the coordination set containing i, and $\mathbb{I}(j \in C_m^*)$ is the indicator function, which equals 1 only when j lies in C_m^*, and zero, otherwise.

We use (30) to evaluate the effects of v_i on the total adoption $TA(\mathbf{v})$ and total welfare $TW(\mathbf{v})$. Differentiating $TA(\mathbf{v}) = \sum_{j \in N}(1 - H(\theta_j^*))$ and using (30) yields

$$\frac{\partial TA(\mathbf{v})}{\partial v_i} = H'(\theta_i^*)$$

as a marginal increase in v_i affects the cutoffs of agents contained in set C_m^* by Proposition 4 (ii) with slopes given by (30).

For TW, we first rewrite it as

$$TW(\mathbf{v}) = \sum_{j \in N}\left\{\int_{\theta_j^*}^{\infty}(v_j + \theta)dH(\theta) + \phi\sum_{k \in N_j}\int_{\max(\theta_j^*, \theta_k^*)}^{\infty} dH(\theta)\right\}, \tag{31}$$

where each term in $\{\}$ corresponds to the equilibrium payoff of agent j. We differentiate $TW(\mathbf{v})$ and use (30) and Theorem 1 to obtain

$$\frac{\partial TW(\mathbf{v})}{\partial v_i} = 1 - H(\theta_i^*) + \phi\left(\frac{L(C_m^*, A_{k-1}^*) + e(C_m^*)}{|C_m^*|}\right)H'(\theta_i^*)$$

after some simple algebra.[58]

Note that the computed formulae for $\frac{\partial TA(\mathbf{v})}{\partial v_i}$ and $\frac{\partial TW(\mathbf{v})}{\partial v_i}$ are valid as long as $\mathbf{v}$ is in the budget $K(\mathbf{v}, B)$ for the selected B. Now we consider the A-planner's problem. Note that the marginal effect of increasing v_i on $TA(\mathbf{v})$ is given by $\frac{\partial TA(\mathbf{v})}{\partial v_i} = H'(\theta_i^*)$. Since the budget set is a simplex, the A-planner must never allocate positive budget to j that does not maximize $H'(\theta_i^*)$; hence, the statement for the A-planner in the Proposition is proven. The solution for a W-planner is similar.

Remark 8 is easy to show. Note that for any i, j in the same coordination set, their equilibrium cutoffs must be the same, i.e., $\theta_i^* = \theta_j^*$. Therefore, we have $\frac{\partial TA(\mathbf{v})}{\partial v_i} = H'(\theta_i^*) = H'(\theta_j^*) = \frac{\partial TA(\mathbf{v})}{\partial v_j}$, and by the same logic, $\frac{\partial TW(\mathbf{v})}{\partial v_i} = \frac{\partial TW(\mathbf{v})}{\partial v_i}$, which implies Remark 8. $\square$

Proof of Proposition 6:

[58]The middle steps are tedious; hence, they are omitted.

First, consider part (i) of Proposition 6. Given $\mathcal{C}^*_{+ij} = \mathcal{C}^*$, C^*_m is unchanged upon inclusion of link (i,j). Moreover, if $\theta^*_i > \theta^*_j$, then this ordering must be maintained upon inclusion of (i,j); otherwise, it contradicts $\mathcal{C}^*_{+ij} = \mathcal{C}^*$. For $i \in C^*_m \subseteq A^*_k$ from SAND, we can define $A^*_0 \equiv \emptyset$ and write:

$$
\begin{aligned}
-\theta^*_i &= \frac{v(C^*_m) + \phi(e(C^*_m) + L(C^*_m, A^*_{k-1}))}{|C^*_m|}, \\
-\theta^*_{i,+ij} &= \frac{v(C^*_m) + \phi(e(C^*_m) + L(C^*_m, A^*_{k-1}) + 1)}{|C^*_m|}, \\
\theta^*_{j,+ij} &= \theta^*_j \ \text{ if } \theta^*_i > \theta^*_j,
\end{aligned}
$$

where the second equality holds whether $j \notin C^*_m$ with $\theta^*_i > \theta^*_j$ (for (i1)) or $j \in C^*_m$ (for (i2)). Taking differences gives the result. The proof of part (ii) of Proposition 6 is straightforward and just omitted. □

E Examples: Removing a node (key player) versus adding a link

E.1 Key players

Consider the network displayed in Figure 7. Assume that all agents have the same $v_i = v = 1$ and $\phi = 0.4$ (the same parameter values will be used to study the key player in Ballester et al. 2006) and that $H(\cdot)$ follows a uniform distribution on the support $\Theta = [-2, 0]$. Since the network is a tree, there is only one coordination set by Proposition 2(2) with $t^*_{[1]} = \frac{F(N)}{|N|} = \frac{vN+\phi e(N)}{N} = 1 + 0.4\frac{4}{5} = 1.32$, so that the common cutoff for all i is given by $\theta^*_i = -1.32$. Assume that $H(\cdot)$ follows a uniform distribution on the support $\Theta = [-2, 0]$, so the probability of adopting for each agent i is equal to $1 - H(\theta^*_i) = -\theta^*_i/2 = 0.66$. The total expected adoption is then equal to $0.66 \times 5 = 3.3$.

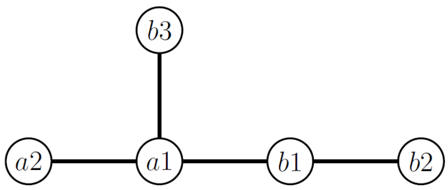

Figure 7: A tree network

Let us determine the key players. Let us start with players $a2$, $b3$, and $b2$ (peripheral agents), so that when one of these players is removed, there is still a connected tree left and thus one coordination set. The removal of players $a2$, $b3$, or $b2$ leads to a common cutoff equal to $-\left(1 + 0.4\frac{3}{4}\right) = -1.30$. The probability of committing crime is now equal to 0.65, i.e., all agents now have a 65% chance of committing crime, which implies that the total expected crime is $0.65 \times 4 = 2.64$. Thus, after the removal of player $a2$, $b3$, or $b1$, the total expected crime has been reduced by 20%.

Now consider the removal the removal of $b1$. There will be one network g_a consisting of agents $a1$, $a2$, and $b3$, and one isolated agent $b2$. The common cutoff for g_a is $\theta_a^* = -1.27$, so that the probability of committing a crime is 0.635. The isolated agent will have a probability of 0.5 of committing a crime. Thus, following the removal of player $b1$, the total expected crime is given by $3 \times 0.635 + 1 \times 0.5 = 2.4$. Consequently, after the removal of player $b1$, the total expected crime has been reduced by 27.27%. Performing a similar analysis for the removal of player $a1$ leads to a reduction of the total expected crime by 33.33%. In summary, the ranking of key players in terms of reduction of total expected crime is: (1) $a1$; (2) $b1$; (3) $a2$, $a3$, $b2$.

Let us now calculate the key-player ranking in Ballester et al. (2006) using the key-player centralities or intercentralities. It is easily verified that the ranking of key players is very similar to the one above, that is: (1) $a1$; (2) $b1$; (3) $a2$, $a3$, (4) $b2$.[59] Indeed, in Ballester et al. (2006), the key players are the ones that generate the most spillover effects, so high degree and high Katz-Bonacich centrality agents are good predictors of key players. In our paper, degree, which is related to "interconnectedness," matters, but connections to "higher" coordination sets, as captured by "embeddedness," also matter. This is precisely because when we remove a player, we change the network density, which affects both aspects, and therefore the cutoffs of all agents in the network. Thus, it is possible that the ranking of key players in Ballester et al. (2006) and in our model differ, because the concepts of intercentrality and of social connectedness (i.e., interconnectedness and embeddedness) are different, albeit related.

E.2 Adding a link

Let us illustrate the results in Section 5.2.2 with an example using the network in Figure 8:

[59] Assume a standard utility function for player i, i.e., $u_i = v_i x_i - \frac{1}{2} x_i^2 + \phi x_i \sum_{j \in N_i} x_j$. Assume that $v_i = 1$ for all i and $\phi = 0.4$, which guarantees that there is a unique interior equilibrium. The intercentrality for each agent i is defined as $c_i = b_i^2 / m_{ii}$, where b_i is the Katz-Bonacich centrality of player i (see Ballester et al. 2006). We easily obtain: $(c_{a_1},\ c_{a_2},\ c_{b3},\ c_{b_1},\ c_{b_2}) = (12.43,\ 6.85,\ 6.85,\ 10.00,\ 5.46)$.

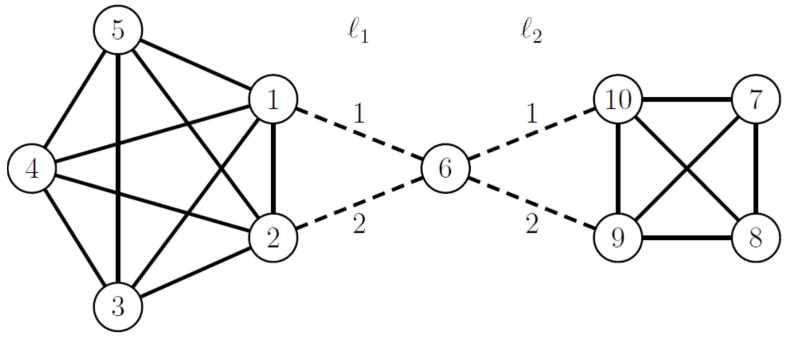

Figure 8: Coordination and bridging

Assume homogeneous values $v_i = v = 0$ for each $i \in N$ and $\phi = 1$. Denote $S_1 = \{1, 2, 3, 4, 5\}$, $S_2 = \{7, 8, 9, 10\}$ and $S_3 = \{6\}$, with $S_1 \cup S_2 \cup S_3 = N$. Agents in S_1 as well as agents in S_2 form cliques with agent 6 bridging the two cliques with varying connectivity to each clique. We denote ℓ_1 as the number of links that 6 has with agents in S_1, and ℓ_2 as the number of links that 6 has with agents in S_2. Table 1 summarizes the equilibrium coordination sets and provides the $t^*_{[k]}$ in step k defined in equation (5) for various values of (ℓ_1, ℓ_2). Observe that $\theta^*_i = -t^*_{[k]}$ if agent i is born in step k. Denote $\mathbf{t}^* = \left(t^*_{[1]};\ t^*_{[2]};\ \cdots ; t^*_{[K]}\right)$.

When agent 6 has either no link or one link, she always adopts last. This is because she has both low interconnectedness and low embeddedness. When agent 6 has two links, then it depends on to whom she is linked. If her two links are not with agents in S_1, agent 6 stays in the second coordination set with S_2. If these two links are with agents in S_1, then she is in the coordination with S_1. Eventually, when agent 6 forms two links with each of the two cliques, all of the agents coordinate together on a common cutoff. While the total number of links that agent 6 carries with each clique is below the number of links that members of each respective clique have with other members, agent 6 functions as a coordination bridge, synchronizing adoption strategies through the economy. When the number of links to either clique drops below two, agent 6 either coordinates with one of the two cliques or coordinates with neither when holding only one link. This illustrates the importance of interconnectedness and embeddedness of each node defined in Section 3.4 with regard to who adopts earliest and why. In particular, while agent 6 is less interconnected than clique members, she is well embedded in the network once she has at least two links.

(ℓ_1, ℓ_2)	$\mathcal{C}^*$	$\mathbf{t}^*$
$(0, 0)$		$(2; 1.5; 0)$
$(0, 1)$	$\{S_1, S_2, S_3\}$	$(2; 1.5; 1)$
$(1, 0)$		$(2; 1.5; 1)$
$(1, 1)$		$(2; 1.6)$
$(0, 2)$	$\{S_1, S_2 \cup S_3\}$	$(2; 1.6)$
$(1, 2)$		$(2; 1.8)$
$(2, 0)$	$\{S_1 \cup S_3, S_2\}$	$(2; 1.5)$
$(2, 1)$		$(2; 1.75)$
$(2, 2)$	$\{N\}$	(2)

Table 1: Coordination sets $\mathcal{C}^*$ and cutoffs $t^*_{[k]}$ for different link additions in the network in Figure 8.

Let us now illustrate the results in Proposition 6. First, consider part (i) of this proposition when $\mathcal{C}^*_{+ij} = \mathcal{C}^*$, that is, the addition of a link does not change the network partition in terms of coordination sets. Start with agent 6 having no link, i.e., $(\ell_1, \ell_2) = (0, 0)$, so that there are three coordination sets: $C^*_1 = S_1$, $C^*_2 = S_2$ and $C^*_3 = S_3 = \{6\}$. Then, when we increase ℓ_1 by one or ℓ_2 by one, we see that agents in the lower coordination sets C^*_1 and C^*_2 are not affected and still have a cutoff of 2 and 1.5, respectively. However, agent 6 now increases her cutoff from 0 to 1. Indeed, as part (i) of Proposition 6 shows, the difference in cutoffs for agent 6 is equal to $\theta^*_6 - \theta^*_{6,+1} = \phi/|C^*_3| = 1$. Likewise, if agent 6 holds two links with clique S_1 and no link to S_2, i.e., $(\ell_1, \ell_2) = (2, 0)$, there are two coordination sets: $C^*_1 = S_1 \cup S_3$ and $C^*_2 = S_2$. Now, if we add a link between agent 6 and S_2, the cutoff of individuals in C^*_1 is not affected and is still equal to 2. However, for $i \in S_2$, we see a decrease in θ^*_i of $0.25 = \phi/|S_2|$, specifically from -1.5 to -1.75, as predicted by part (i) of Proposition 6.

Now, consider part (ii) of Proposition 6 when $\mathcal{C}^*_{+ij} \neq \mathcal{C}^*$, that is, the addition of a link changes the network partition in terms of coordination sets. For example, when agent 6 holds 0 links to S_1 and 1 link to S_2, i.e., $(\ell_1, \ell_2) = (0, 1)$, there are three coordination sets: $C^*_1 = S_1$, $C^*_2 = S_2$, and $C^*_3 = S_3$. Then, when we add a link between agent 6 and S_1, so that $(\ell_1, \ell_2) = (1, 1)$, the partition changes to contain only two coordination sets: $C^*_1 = S_1$ and $C^*_2 = S_2 \cup S_3$. As predicted by part (ii) of Proposition 6, agents in $C^*_1 = S_1$ have the same cutoff 2 and are unaffected by this additional link. Now, however, agents in $C^*_2 = S_2 \cup S_3$ have a cutoff of 1.6; this means that any agent k in S_2 sees a decrease in θ^*_k of 0.1 while agent 6 (S_3) sees a decrease in θ^*_6 of 0.6.

F New predictions on adoption: Technical details

F.1 Price, aggregate demand, and comparative statics in Ballester et al. (2006) (BCZ)

Let us introduce price in the model of BCZ. Each agent i chooses effort a_i to maximize the following payoff:

$$u_i(\mathbf{a}) = v_i a_i - \frac{1}{2}a_i^2 + \phi \sum_{j \in \mathcal{N}} g_{ij} a_i a_j, \tag{32}$$

where $v_i = r - p$ for each i, where p is the (common) price and $r > 0$ is the marginal private returns of exerting a_i. Observe that, contrary to our model (LZZ), effort a_i is continuous, i.e., $a_i \in \mathbf{R}_+$. BCZ show that, when $\phi\lambda_{\max}(\mathbf{G}) < 1$, there exists a unique Nash equilibrium given by:

$$\mathbf{a}^*(p) = (a_1^*(p), \cdots, a_n^*(p))' = \begin{cases} [\mathbf{I} - \phi\mathbf{G}]^{-1}(r-p)\mathbf{1} & \text{if } p \leq r; \\ \mathbf{0} & \text{if } p > r, \end{cases}$$

This equilibrium determines the individual demand for each player i for a given price p. Thus, the *aggregate demand* is equal to:

$$D^{BCZ}(p) = \sum_{i \in \mathcal{N}} a_i^*(p) = \begin{cases} b(\mathbf{G}, \phi)(r-p) & \text{if } p \leq r; \\ 0 & \text{if } p > r, \end{cases} \tag{33}$$

where $b(\mathbf{G}, \phi) = \mathbf{1}'[\mathbf{I} - \phi\mathbf{G}]^{-1}\mathbf{1}$ is the (unweighted) aggregate Katz-Bonacich centrality of $\mathbf{G}$ with parameter ϕ.

F.2 Price, aggregate demand, and comparative statics in Sákovics and Steiner (2012) (SS)

Let us now introduce price in Sákovics and Steiner (2012). The utility function of agent i is given by:

$$u_i(a, \theta, p) = \begin{cases} \phi_i + v_i - p & \text{if } a \geq 1 - \theta \\ v_i - p & \text{if } a < 1 - \theta \end{cases}$$

where $a = \int_0^m w_i a_i di$, $v_i < 0$, $\phi_i + v_i - p > 0$ for all p, and $w_i > 0$. As in our model, effort a_i can only take two values: either adopt ($a_i = 1$) or not adopt ($a_i = 0$), and θ is the state. Each group g of players has measure m_g, $\sum_g m_g = m$, $\sum_g w_g m_g = 1$ (a normalization), and each $v_i = v_j$, $\phi_i = \phi_j$ and $w_i = w_j$ for $i, j \in g$. Players are endowed with signals $s_i = \theta + \nu\epsilon_i$, $\nu \in (0, 1]$ and ϵ_i follows cdf $F(\cdot)$ with support $[-1, 1]$. Each player i's group membership is private information with $\Pr(g_i = g) = m_g/m$.

Sákovics and Steiner (2012) show that, for each $\nu \in (0, 1]$, there is a unique Bayes-Nash equilibrium and each player follows the following threshold strategy:

$$a_i(s_i, g) = \begin{cases} 1 & \text{if } s_i \geq s_g^* \\ 0 & \text{if } s_i < s_g^* \end{cases}$$

Moreover, as $\nu \to 0$, all thresholds s_g^* converge to a common limit θ^*, where:

$$\theta^* = \sum_g m_g \frac{w_g}{\phi_g}(p - v_g).$$

For a given realization of θ, as $\nu \to 0$, the above maps to a p^* such that $p \leq p^*$ implies $a_i = 1$ and $p > p^*$ implies $a_i = 0$ for all i, with

$$p^* = \frac{\theta + \sum_g \frac{m_g w_g v_g}{\phi_g}}{\sum_g \frac{m_g w_g}{\phi_g}}$$

Under the assumption that, for all g, $\phi_g = \phi$, p^* becomes:

$$p^* = \theta\phi + \sum_g m_g w_g v_g. \tag{34}$$

When all players adopt, aggregate demand $D^{SS}(p)$ is equal to m, otherwise demand is 0. That is, for a given realization of the state θ, as $\nu \to 0$,

$$D^{SS}(p) = \sum_{i \in \mathcal{N}} a_i^*(p) = \begin{cases} m & \text{if } p \leq p^*; \\ 0 & \text{if } p > p^*. \end{cases} \tag{35}$$

F.3 Price, aggregate demand, and comparative statics in our model (LZZ)

Let us introduce price p in our model. We consider the following payoff function (see (1)):

$$u_i(\mathbf{a}, \theta) = (v_i + \theta + \phi \sum_{j \in \mathcal{N}} g_{ij} a_j) a_i,$$

where a_i is either 0 or 1 and θ is the state. Assume $v_i = r - p$ for each i, where p is the price and $r > 0$ is a constant. We can use the SAND Algorithm in Theorem 1 to characterize the limiting cutoffs for players for a given p and a given realization of θ. The aggregate demand function in this framework is given by:

$$D^{LZZ}(p) = \sum_{i \in \mathcal{N}} a_i^*(p) = \sum_{i \in \mathcal{N}} \mathbf{1}_{\{\theta + r - p + \phi q_i^* \geq 0\}}, \tag{36}$$

where $\mathbf{q}^* = (q_1^*, \cdots, q_n^*)$ is computed using the SAND algorithm in Theorem 1 with $v_i = 0, \phi = 1$.[60] In other words, all agents i for which $p \leq p_{[i]}^* := \theta + r + \phi q_i^*$ will adopt while those for which $p > p_{[i]}^* := \theta + r + \phi q_i^*$ will not adopt.

As stated in the main text, we use the real-world network in rural India studied in Banerjee et al. (2013)[61] to plot the aggregate demand $D^{LZZ}(p)$ in Figure 6.

This network of 20 agents is displayed in Figure 5 where different colors indicate different coordination sets. Using the SAND algorithm, we can divide the network into different coordination sets. Let A_1^* be the set of black nodes who share a common q_i^* value of $q_{[1]}^* = 3$, A_2^* the set of dark blue and red nodes with the same cutoff $q_{[2]}^* = 2.5$, A_3^* the set of the single node of light blue color with $q_{[3]}^* = 2$, A_4^* the set of the green nodes with $q_{[4]}^* = 4/3$, and, finally, A_5^* the set of the yellow nodes with $q_{[5]}^* = 1$. Thus, there are 5 coordination sets.[62]

[60] Observe that $q_i^* = -\theta_i^*$ in Theorem 1.

[61] This is an "help decision" network. The exact question is: "If you had to make a difficult personal decision, whom would you ask for advice?"

[62] There is clearly a one-to-one relationship between q_i^* and $p_{[i]}^*$ since $p_{[i]}^* = \theta + r + \phi q_i^*$.

(NOT FOR PUBLICATION) ONLINE APPENDIX

G Alternative equilibrium characterizations and alternative proofs

In this section, we first provide an alternative characterization of the equilibrium cutoffs by solving a simple constrained convex minimization program (Section G.1). Then, we provide another characterization of the limiting equilibrium by showing both that the stage game has a potential function and that the limiting equilibrium coincides with the strategy profile, which maximizes its potential (Section G.2). In Section G.3, we discuss alternative equilibrium characterizations. Finally, we provide an alternative proof of Proposition 1 (balance condition for a single coordination set) using Gale's Demand Theorem (Section G.4). Though we employ the SAND algorithm as the main tool for both equilibrium characterization and comparative statics analysis in the main text, the results and the techniques behind their proofs reported here are of independent interest.

G.1 Projection mapping and alternative equilibrium characterization

In this section, we provide an alternative route to calculating $\mathcal{C}^*$ and $\boldsymbol{\theta}^*$. The solution can be viewed as the dual to SAND, and provides an alternative interpretation to our limit equilibrium. Define the set of *consistent weighting functions* for $\mathcal{G}$:

$$\mathcal{W} \equiv \{\mathbf{w} = (w_{ij}, (i,j) \in E) | w_{ij} \geq 0, w_{ji} \geq 0, w_{ij} + w_{ji} = 1; \forall (i,j) \in E\}. \tag{G.1}$$

Clearly, $\mathcal{W}$ is compact, convex. For each $i \in N$, given the intrinsic value v_i, scale factor ϕ, and edges E, we can define the affine mapping $\Phi_i : \mathcal{W} \to \mathbb{R}$:

$$\Phi_i(\mathbf{w}) =: v_i + \phi \sum_{j \in N_i} w_{ij}, \quad \forall i \in N. \tag{G.2}$$

Let $\Phi(\mathcal{W}) \subset \mathbb{R}^{|N|}$ denote the image of $\mathcal{W}$ under the mapping $\Phi = (\Phi_i)_{i \in N}$. Given the linearity of $\Phi(\cdot)$ and the structure of $\mathcal{W}$, it is easy to show that $\Phi(\mathcal{W})$ is a compact, convex polyhedron. Denote $\langle \cdot, \cdot \rangle$ the inner product in $\mathbb{R}^n$ and $||\mathbf{x} - \mathbf{y}|| \equiv \sqrt{\langle \mathbf{x}, \mathbf{y} \rangle}$ the Euclidean norm. The following theorem gives the limit cutoffs $\boldsymbol{\theta}^*$.

Theorem G1. *For any $\mathbf{v}$, ϕ, and network $\mathcal{G}$, the equilibrium limit cutoffs $\boldsymbol{\theta}^*$ are given by:*

$$\theta_i^* + q_i^* = 0, \quad \forall i \in N, \tag{G.3}$$

where $\mathbf{q}^* = (q_1^*, \cdots, q_n^*)$ *is the unique solution to:*

$$\mathbf{q}^* = \underset{\mathbf{z}\in\Phi(\mathcal{W})}{\operatorname{argmin}} ||\mathbf{z}||. \tag{G.4}$$

The program (G.4) in Theorem G1 solves for a set of cutoffs by way of finding the vector $\mathbf{q}^*$ of minimal Euclidean length. The minimization program in (G.4) is related to the projection mapping.

Definition G1. *Let K be a closed convex set in $\mathbb{R}^n$. For each $\mathbf{x} \in \mathbb{R}^n$, the* orthogonal projection *(or,* projection*)*[A1] *of $\mathbf{x}$ on the set K is the unique point $\mathbf{y} \in K$ such that:*

$$||\mathbf{x} - \mathbf{y}|| \leq ||\mathbf{x} - \mathbf{z}||, \quad \forall \mathbf{z} \in K.$$

We denote $\mathbf{Proj}_K[\mathbf{x}] \equiv \mathbf{y} = \operatorname{argmin}_{\mathbf{z}\in K} ||\mathbf{x} - \mathbf{z}||$.

Using Definition G1, we observe that the solution vector $\mathbf{q}^*$ to (G.4) is the projection of the origin onto the compact, convex space $\Phi(\mathcal{W})$, which is the image of $\mathcal{W}$ under the mapping Φ. Equivalently, we can reformulate the equilibrium cutoffs as

$$-\boldsymbol{\theta}^* = \mathbf{q}^* = \mathbf{Proj}_{\Phi(\mathcal{W})}[\mathbf{0}].$$

Here, we discuss the projection mapping Theorem G1 in more detail. Like the SAND algorithm, Theorem G1 yields the equilibrium cutoff value for each agent i. The solution of Theorem G1 is expressed in terms of $\mathbf{q}^*$, which is directly solved for with problem (G.4). Strikingly, $\mathbf{q}^*$ solves a simple quadratic program with linear constraints. $\mathbf{q}^*$ maps back to weights $\mathbf{w}^*$, via $\Phi(\cdot)$. That is, $\mathbf{q}^* = \Phi(\mathbf{w}^*)$. In the proof of Theorem G1, we show that consistency in beliefs (i.e., the conditions imposed on w_{ij}^* defined in (10) by Lemma 1), together with the system of indifference conditions at cutoffs (as shown in (9)), imply that the cutoffs satisfy the necessary conditions for a simple quadratic minimization program with linear constraints specified in (G.4). Since the program (G.4) is convex and constraints are linear, those necessary conditions are also sufficient to show that the cutoffs indeed solve program (G.4).

In term of computation complexity, for large networks, the quadratic problem (G.4), which converges in approximately cubic time, can perform much better than SAND. However, the approach using SAND is more convenient; for example, the comparative statics analysis using SAND gives the coordination sets directly, which play a key role in the analysis of shock propagation and policy interventions.

Remark G11. *Theorem G1 and SAND in Theorem 1 can now be viewed as dual problems. The former calculates $\mathbf{q}^*$ yielding the partition $\mathcal{C}^*$ as a bi-product, while the latter constructs $\mathcal{C}^*$ yielding $\mathbf{q}^*$ as a bi-product. Interestingly, Theorem 1 gives the weights w_{ij}^**

[A1] See Chapter 1 of Nagurney (1992) for characterization and properties of this projection operator.

explicitly in the projection step. Alternatively, in the Algorithm 1, these weights are set either to either 0 or 1 for agents in different coordination sets by construction, while they are implied by Gale's Demand Theorem for agents residing within the same coordination set.

G.1.1 Proofs

Proof of Theorem G1: Recall Equation (9) for equilibrium cutoff of agent i,

$$\theta_i^* + v_i + \phi \sum_{j \in N_i} w_{ij}^* = 0, \quad \forall i \in N, \tag{G.5}$$

where $w_{ij}^* = \lim_{\nu \to 0} \mathbb{E}[\mathbb{I}(s_j \geq c_j^{*\nu})|s_i = c_i^{*\nu}]$ is defined in (10). Define $\mathbf{w}^* = (w_{ij}^*, (i,j) \in E)$, then $\mathbf{w}^* \in \mathcal{W}$ since $w_{ij}^* + w_{ji}^* = 1$ by Lemma 1 (i).

Define $\mathbf{q} = (q_1^*, \cdots, q_N^*)$ where $q_i^* = -\theta_i^*$. Then, by (G.5), $q_i^* = v_i + \phi \sum_{j \in N_i} w_{ij}^*$, $\forall i$; hence, $\mathbf{q}^*$ is in $\Phi(\mathcal{W})$. Clearly, $\theta_i^* < \theta_j^*$ if and only if $q_i^* > q_j^*$.

Next, we verify that this $\mathbf{q}^*$ satisfies the two conditions stated in Lemma G6; therefore, $\mathbf{q}^*$ must be the projection of $\mathbf{0}$ onto $\Phi(\mathcal{W})$ by Lemma G6, which proves the theorem. Now we check conditions (C1) and (C2) of Lemma G6. Clearly, condition (C1) holds from the definition of this $\mathbf{q}^*$. Condition (C2) is easy to check. Moreover, for any connected agent i and j, suppose $\theta_i^* < \theta_j^*$, then $q_i^* > q_j^*$, and $w_{ij}^* = 0$ and $w_{ji}^* = 1$ by Lemma 1 (ii). □

Lemma G6. *Let $\mathbf{q}^*$ denote the projection of $\mathbf{0}$ onto the $\Phi(\mathcal{W})$, i.e.,*

$$\mathbf{q}^* = \mathbf{Proj}_{\Phi(\mathcal{W})}[\mathbf{0}].$$

Then $\mathbf{q}^*$ is uniquely characterized by the following two conditions:

(C1 There exists $\mathbf{w}^* \in \mathcal{W}$ such that $q_i^* = v_i + \phi \sum_{j \in N_i} w_{ij}^*$, $\forall i$, so that $\mathbf{q}^* \in \Phi(\mathcal{W})$.

(C2) In addition, for any edge $(i,j) \in E$ and for any $z_{ij} \in [0,1]$,

$$(q_i^* - q_j^*)(z_{ij} - w_{ij}^*) \geq 0.$$

Moreover, we can replace (C2) by the equivalent form:

(C2') $(i,j) \in E, (q_i^* - q_j^*) > 0 \implies w_{ij}^* = 0, w_{ji}^* = 1$.

Proof of Lemma G6: We first show necessity. Clearly, *(C1)* is necessary from the definition of $\Phi(\mathcal{W})$. For *(C2)*, for any $\mathbf{w}' \in \mathcal{W}$, by optimality of $\mathbf{q}^*$, the following must be true:

$$\eta(t) \equiv ||\Phi((1-t)\mathbf{w}^* + t\mathbf{w}')||^2 \geq ||\Phi(\mathbf{w}^*)||^2 = ||\mathbf{q}^*||^2 = \eta(0)$$

for any $t \in [0,1]$.

Since $\Phi(\cdot)$ is an affine mapping, $\frac{\partial}{\partial t}\Phi((1-t)\mathbf{w}^* + t\mathbf{w}') = \Phi(\mathbf{w}') - \Phi(\mathbf{w}^*)$. Taking the derivative of $\eta(t)$ at $t = 0$, we obtain:

$$0 \leq \eta'(0) = 2\langle \mathbf{q}^*, \Phi(\mathbf{w}') - \Phi(\mathbf{w}^*)\rangle. \tag{G.6}$$

Now for any $z'_{ij} \in [0,1]$, we construct a special $\mathbf{w}'$ by only modifying the weights w^*_{ij} and $w^*_{ji} = 1 - w^*_{ij}$ on the edge between i and j in $\mathbf{w}^*$ to $w'_{ij} = z_{ij}$ and $w'_{ji} = 1 - z_{ij}$. Clearly, $\mathbf{w}'$ is still in $\mathcal{W}$. Inequality (G.6) becomes:

$$\phi(q^*_i(z_{ij} - w^*_{ij}) + q^*_j(z_{ji} - w^*_{ji})) \geq 0.$$

However, $z_{ji} - w^*_{ji} = (1 - z_{ij}) - (1 - w^*_{ij}) = -(z_{ij} - w^*_{ij})$. So the above inequality is equivalent to:

$$(q^*_i - q^*_j)(z_{ij} - w^*_{ij}) \geq 0.$$

Let us show sufficiency. For any $\mathbf{w}' \in \mathcal{W}$, a simple calculation shows that:

$$\langle \mathbf{q}^*, \Phi(\mathbf{w}') - \Phi(\mathbf{w}^*)\rangle = \phi \sum (q^*_i - q^*_j)(w'_{ij} - w^*_{ij}) \geq 0,$$

as each term in the summation is nonnegative. Therefore, $\eta'(0) \geq 0$; moreover, $\eta(\cdot)$ is clearly convex in $t \in [0,1]$.[A2] Therefore,

$$\eta(1) - \eta(0) \geq (1-0)\eta'(0) \geq 0,$$

that is:

$$||\Phi(\mathbf{w}')||^2 \geq ||\Phi(\mathbf{w}^*)||^2 = ||\mathbf{q}^*||^2$$

since $\mathbf{w}'$ is arbitrary, and indeed $\mathbf{q}^*$ is the projection of $\mathbf{0}$ onto $\Phi(\mathcal{W})$.

Now we need to verify that for any edge ij with $(i,j) \in E$, *(C2)* is equivalent to *(C2')*:

$$\left\{(q^*_i - q^*_j)(z_{ij} - w^*_{ij}) \geq 0, \forall z_{ij} \in [0,1]\right\} \Leftrightarrow \left\{(i,j) \in E, (q^*_i - q^*_j) > 0 \Rightarrow w^*_{ij} = 0, w^*_{ji} = 1\right\}.$$

If so, then $q^*_i > q^*_j \implies w^*_{ij} = 0$ and $w^*_{ji} = 1$; $q^*_i < q^*_j \implies w^*_{ij} = 1$ and $w^*_{ji} = 0$.

From *(C2)* to *(C2')*: Suppose $q^*_i > q^*_j$, and let $z_{ij} = 0$. We have $(q^*_i - q^*_j)(0 - w^*_{ij}) \geq 0$, and by $w^*_{ij} \geq 0$ it must be the case that $w^*_{ij} = 0$. Similarly, assuming $q^*_i < q^*_j$ and picking $z_{ij} = 1$ shows that $w^*_{ij} = 1$.

From *(C2')* to *(C2)*: If $q^*_i > q^*_j$ and $w^*_{ij} = 0$, then for any $\forall z_{ij} \in [0,1]$, $(q^*_i - q^*_j)(z_{ij} - w^*_{ij}) = (q^*_i - q^*_j)(z_{ij}) \geq 0$. Similarly, if $q^*_i < q^*_j$ and $w^*_{ij} = 1$, then for any $\forall z_{ij} \in [0,1]$, $(q^*_i - q^*_j)(z_{ij} - w^*_{ij}) = -(q^*_i - q^*_j)(1 - z_{ij}) \geq 0$. $\square$

[A2] As Φ is affine and $||\mathbf{x}||^2$ is a convex function of $\mathbf{x}$.

G.2 Potential function and alternative equilibrium characterization

In this section, we discuss an alternative approach of equilibrium characterization from the perspective of potential function. Here, we derive the results formally.

Define

$$P(\mathbf{a}, \theta) = \sum_{i \in N} a_i(\theta + v_i) + \frac{1}{2}\phi \sum_{(i,j) \in E} a_i a_j. \tag{G.7}$$

It is easy to check the following

$$u_i(1, \mathbf{a}_{-i}, \theta) - u_i(0, \mathbf{a}_{-i}, \theta) = P(1, \mathbf{a}_{-i}, \theta) - P(0, \mathbf{a}_{-i}, \theta) \tag{G.8}$$

for any $\mathbf{a}_{-1} \in \{0, 1\}^{|N|-1}$. In other words, the difference of i's payoff between adopting and not adopting is exactly equal to the difference of P, holding other players' strategy $\mathbf{a}_{-i}$ and state θ fixed. Therefore, $P(\cdot, \theta)$ is an exact potential function of the stage game G at θ (see Monderer and Shapley 1996).

With slight abuse of notation, we also use $P(\mathbf{a}, \theta)$ by identifying $S \subset N$ as the set of players who choose to adopt under $\mathbf{a} \in \{0, 1\}^N$, i.e.,

$$S \leftrightarrow \{i \in N, a_i = 1\}.$$

Using this notation, the potential takes the following form:

$$P(S, \theta) \equiv v(S) + \phi e(S) + |S|\theta = F(S) + |S|\theta.$$

Define the potential maximizer

$$\mathcal{M}(\theta) \equiv \arg\max_{S \subset N} P(S, \theta). \tag{G.9}$$

It is possible to have multiple maximizers, in which case $\mathcal{M}$ contains all the maximizers.

Lemma G7. *Define $\theta^*_{[k]} = -t^*_{[k]}$. The SAND algorithm finds the potential maximizer $\mathcal{M}$ systematically as follows:*

$$\mathcal{M}(\theta) = \begin{cases} \emptyset & \text{if } \theta < \theta^*_{[1]} \\ A^*_1 & \text{if } \theta^*_{[1]} < \theta < \theta^*_{[2]} \\ \cdots & \cdots \\ A^*_K = N & \text{if } \theta > \theta^*_{[K]}. \end{cases} \tag{G.10}$$

Moreover, at the cutoffs $\theta^*_{[i]}$, there might be multiple maximizers. Indeed, $\{A^*_{i-1}, A^*_i\} \subseteq \Gamma(\theta^*_{[i]})$ for $i = 1, 2, \cdots, K$.

Lemma G8. *The following statements are true.*

(i) Except at a finite number of points of θ, the maximizer of (G.9) is unique, so $\mathcal{M}(\theta)$ is a singleton.

(ii) For any $\theta' < \theta''$, $\mathcal{M}(\theta') \subseteq \mathcal{M}(\theta'')$.

$P(S, \theta)$ is supermodular in S for each θ, and it satisfies increasing difference in (S, θ).[A3] By the Monotone Comparative Statics Theorem (Milgrom and Shannon 1994), Lemma G8 follows. $\mathcal{M}$ is *"increasing"* in θ in the sense of strong set order. The potential maximizer $\mathcal{M}(\theta)$ is monotone in θ. So, for each agent i, if $i \in \mathcal{M}(\theta)$ at $\theta = \theta'$, it remains in the set $\mathcal{M}(\theta)$ for any state $\theta'' > \theta'$ by Lemma G8 (ii). Therefore, the set $\{\theta, i \in \mathcal{M}(\theta)\}$ is an interval by Lemma G8. So we record the smallest state to make i included in the $\mathcal{M}$ as follows:

$$\gamma_i = \inf\{\theta : i \in \mathcal{M}(\theta)\}.$$

We call γ_i the switching point of agent i: when θ is below γ_i, agent i is not included in $\mathcal{M}(\theta)$; however, when θ is above γ_i, i is included in $\mathcal{M}(\theta)$. Conversely, we can almost recover the mapping $\Gamma(\cdot)$ from these $\{\gamma_i\}$s by observing that

$$\{k : \theta < \gamma_k\} \subseteq \mathcal{M}(\theta) \subseteq \{k : \theta \leq \gamma_k\}. \tag{G.11}$$

When θ is not equal to these cutoff values γ_k, the left and right sides are the same, which determines $\mathcal{M}(\theta)$.

The connection between switching points for the potential maximizer and the limiting equilibrium threshold characterized in Theorem 1 is present in the following.

Theorem G2.

(1) For every i, the switching point exactly equals the limiting threshold, i.e.,

$$\gamma_i = \theta_i^*.$$

(2) The limiting equilibrium selected by the global game $\boldsymbol{\pi}^*(\theta) = \{\mathbb{I}(\theta \geq \theta_i), i \in N\}$ (see Proposition 0) is the maximizer of the potential P for generic θ.[A4]

Theoretically, we can view Theorem G2 (1) as an equivalent characterization of the limiting threshold since $\mathcal{M}$ fully determines the thresholds of every player, and hence the

[A3] It is easy to show that:
(1) $P(\cdot, \theta)$ is supermodular, at each θ:

$$P(S \cup T, \theta) + P(S \cap T, \theta) \geq P(S, \theta) + P(T, \theta), \forall S, T.$$

(2) P satisfies increasing difference in (S, θ):

$$P(T, \theta') - P(S, \theta') < P(T, \theta'') - P(S, \theta''), \quad \forall \theta'' > \theta', \ \forall T \supsetneq S.$$

[A4] With the exception of $\theta \in \{-t^*_{[1]}, \cdots, -t^*_{[K]}\}$.

partition and coordination sets. Formally, for each agent i,

$$\theta_i^* = \gamma_i = \inf\{\theta : i \in \mathcal{M}(\theta)\}.$$

Furthermore, Theorem G2 (2) reveals the special feature of the equilibrium selected by the global game approach.

The connection between the two theorems (Theorem 1 and Theorem G2) is captured by the following observation in the global game literature (see Frankel et al. 2003): the limiting equilibrium for a potential game equals the potential maximizers for generic θ, as confirmed by Theorem G2 (2). In other words, the SAND can be treated as an algorithm that gives a systematic solution to $\mathcal{M}(\theta)$. Due to the combinatorial feature of the potential maximization problem, finding the maximizer for all θ might not be a simple task. The SAND algorithm achieves it as a by-product.

Remark G12. *The equilibrium we selected using the global game approach is robust in the sense of Kajii and Morris (1997). Furthermore, the selection is noise-independent. To see that, recall that our coordination game has a potential $P(\mathbf{a}, \theta)$ (see (G.7)), which is supermodular in $(a_i, \mathbf{a}_{-i})$ for fixed θ, and strictly supermodular in (a_i, θ) for fixed $\mathbf{a}_{-i}$. As a result, according to Frankel et al. (2003), Oyama and Takahashi (2020), and Basteck et al. (2013), the game Γ^0 has an exact potential; therefore, the maximizer of the potential is selected by the global game, and this selection is independent of noise distribution F.*[A5][A6]

While the potential approach in Theorem G2 requires solving the maximization of P for each θ, which makes it challenging for comparative statics due to the discreteness of $\mathbf{a}$, SAND has the advantage in that more precise information about the equilibrium cutoff points θ_i^* is obtained using Theorem 1 and the projection method in Theorem G1. Moreover, the information about the coordination set, i.e., who coordinates with whom, is also directly decoded using the cutoff values, which enables us to conduct comparative statics with respect to network structure and valuations in a much simpler manner.

Proofs of Lemma G7, Lemma G8, and Theorem G2: It is sufficient to prove that the $\mathcal{M}$ takes the specific form given in (G.10) in Lemma G7, as Lemma G8 and Theorem G2 directly follow from it. Recall $P(S, \theta) \equiv F(S) + \theta|S|$, and $\theta_{[k]}^*$ and $t_{[k]}^*$ satisfy the following condition by SAND:

$$-\theta_{[k]}^* = \frac{F(A_k^*) - F(A_{k-1}^*)}{|A_k^*| - |A_{k_1}^*|} \equiv t_{[k]}^* = \max_{S \supsetneq A_{k-1}^*} \frac{F(S) - F(A_{k-1}^*)}{|S| - |A_{k-1}^*|}.$$

[A5] Moreover, Ui (2001) shows that the selected equilibrium is *robust* in the sense of Kajii and Morris (1997). See Morris and Ui (2005), Oyama and Takahashi (2020) for further discussions.

[A6] In the simple case of a two-player, binary action coordination game (dyad case in our paper), as shown in Carlsson and van Damme (1993), the risk-dominant equilibrium is selected by the global game and it is independent of noise distribution. Frankel et al. (2003) generalized this result to n-player supermodular games that yield a potential, which applies to our setting under arbitrary network structures.

Step 1: Fix $\theta < \theta^*_{[1]}$, for any nonempty S,

$$P(S,\theta) = |S|(\frac{F(S)}{|S|} + \theta) \leq |S|(t^*_{[1]} + \theta) < |S| \underbrace{(t^*_{[1]} + \theta^*_{[1]})}_{=0} = 0 = P(\emptyset, \theta)$$

as $\frac{F(S)}{|S|} \leq \frac{F(A^*_1)}{|A^*_1|} = t^*_{[1]}$ by step 1 of the SAND algorithm. So, the unique maximizer is $\emptyset$ in such a state, i.e.,

$$\mathcal{M}(\theta) = \emptyset, \quad \forall \theta < \theta^*_{[1]}.$$

By the continuity of P in θ, $\emptyset$ must also be a maximizer when $\theta = \theta^*_{[1]}$, i.e., $\mathcal{M}(\theta^*_{[1]}) \supseteq \{\emptyset\}$ (although it may not be the unique maximizer). Moreover, by the definition of $\theta^*_{[1]}, t^*_{[1]}$, we have

$$0 = F(A^*_1) + \theta^*_{[1]}|A^*_1|, \text{ as } \theta^*_{[1]} = -\frac{F(A^*_1)}{|A^*_1|} = -t^*_{[1]}.$$

Therefore,

$$P(\emptyset, \theta^*_{[1]}) = 0 = P(A^*_1, \theta^*_{[1]}).$$

In other words, $\mathcal{M}(\theta^*_{[1]}) \supseteq \{\emptyset, A^*_1\}$.

Step 2: Next, consider $\theta^*_{[1]} < \theta < \theta^*_{[2]}$. Then, $\theta \in (-t^*_{[1]}, -t^*_{[2]})$.

(2a) For any proper subset $\underline{A}_1 \subset A^*_1$, we have

$$P(\underline{A}_1, \theta^*_{[1]}) \leq P(A^*_1, \theta^*_{[1]})$$

by the optimality of A^*_1 at $\theta^*_{[1]}$. Then, by increasing difference of F in (S,θ) (see footnote A3 in the Online Appendix), we have $P(\underline{A}_1, \theta) < P(A^*_1, \theta)$ as $\theta > \theta^*_{[1]}$.

(2b) Consider any proper superset $\bar{A}_1 \supset A^*_1$. We have

$$P(\bar{A}_1, \theta) - P(A^*_1, \theta) = (|\bar{A}_1| - |A^*_1|)(\theta + \frac{F(\bar{A}_1) - F(A^*_1)}{(|\bar{A}_1| - |A^*_1|)})$$

$$\leq (|\bar{A}_1| - |A^*_1|) \underbrace{(\theta + t^*_{[2]})}_{<0} < 0$$

as $\frac{F(\bar{A}_1) - F(A^*_1)}{(|\bar{A}_1| - |A^*_1|)} \leq \frac{F(A^*_2) - F(A^*_1)}{(|A^*_2| - |A^*_1|)} = t^*_{[2]}$ by step 2 of SAND, and $\theta^*_{[1]} < \theta < \theta^*_{[2]}$.

(2c) Combing (2a) and (2b), we apply Lemma 2 to show that A^*_1 is the unique maximizer of P when $\theta^*_{[1]} < \theta < \theta^*_{[2]}$ (recall $P(S,\theta)$ is supermodular in S for fixed θ). Therefore,

$$\mathcal{M}(\theta) = \{A^*_1\}, \qquad \forall \theta^*_{[1]} < \theta < \theta^*_{[2]}.$$

Again, by continuity of P in θ, A_1^* remains a maximizer of P at $\theta_{[2]}^*$:

$$\mathcal{M}(\theta_{[2]}^*) \supseteq \{A_1^*\}.$$

By the definition of $z_{[2]}^*$ and $\theta_{[2]}^*$, we can verify that

$$P(A_1^*, \theta_{[2]}^*) = P(A_2^*, \theta_{[2]}^*) \text{ as } \theta_{[2]}^* = -\frac{F(A_2^*) - F(A_1^*)}{|A_2^*| - |A_1^*|} = -t_{[2]}^*.$$

As a result,

$$\mathcal{M}(\theta_{[2]}^*) \supseteq \{A_1^*, A_2^*\}.$$

The result of the proof follows by the same logic. □

G.3 Discussion of alternative equilibrium characterizations

Here, we briefly discuss two alternative characterizations using different perspectives.

Alternative equilibrium characterization using Projection Theorem G1 in the Online Appendix G.1 provides an alternative characterization of the equilibrium thresholds $\boldsymbol{\theta}^*$ by solving a simple constrained convex minimization program. In contrast to the discrete nature of SAND, this minimization problem is continuous. The constraints to the program are a function of the network structure and the consistency requirements of Lemma 1. In particular, this program is less demanding to compute than SAND because it is a simple convex program with linear constraints.

Alternative approach using potential games The global game approach we used in Section 3.1 gives a selection of equilibrium $\boldsymbol{\pi}^*(\theta)$ for the stage game for each state θ. To be precise, the adoption profile in which each i adopts if and only if $\theta \geq \theta_i^*$ is a Nash equilibrium of the stage game given θ. In fact, as shown in equation (G.7) in the Online Appendix G.3, our coordination game has an exact potential function. Thus, sharper predictions about this equilibrium selection using a global game approach can be established. Formally, in Theorem G2 in the Online Appendix G.3, we prove that the selected equilibrium exactly corresponds to the unique maximizer of the potential, for any state. Furthermore, the cutoff for each agent can be decoded from the maximizers for different states, which provides a new method to determine equilibrium cutoffs.

Observe that the equilibrium we selected using the global game is *robust* in the sense of Kajii and Morris (1997). Conversely, any equilibrium in the complete information, which differs from our equilibrium (the potential maximizer), is not robust (Morris and Ui 2005). For instance, in Section J, we study the extreme equilibria, i.e., the largest and the smallest ones. We observe that at least one extreme equilibrium is not robust for all states, except when there is a unique Nash equilibrium.

G.4 Alternative proof of Proposition 1 using Gale's Demand Theorem

By Theorem G1, the existence of a single coordination set is equivalent to:

$$\frac{T}{n}\mathbf{1} \in \Phi(\mathcal{W}),$$

where $T = \sum_i v_i + \phi e(N) = F(N)$. This can be reformulated as a feasibility condition to the following linear programming problem:

$$\begin{aligned} &v_i + \phi \sum_{j \in N_i} w_{ij} = \frac{T}{n}, \ \forall i \in N, \\ &w_{ij} \geq 0, \ w_{ij} + w_{ji} = 1, \ \forall (i,j) \in E, \end{aligned}$$

given $v_i = v$, $\forall i$, and $T = \sum v_i + \phi e(N) = nv + \phi e(N)$. So the above system is equivalent to:

$$\begin{aligned} &\sum_{j \in N_i} w_{ij} = \frac{e(N)}{|N|}, \ \forall i \in N, \\ &w_{ij} \geq 0, \ w_{ij} + w_{ji} = 1, \ \forall (i,j) \in E. \end{aligned} \tag{G.12}$$

To show the necessity, suppose there exists a solution $\mathbf{w}^*$ to system (G.12). Then:

$$|S|\frac{e(N)}{|N|} = \sum_{i \in S}(\sum_{j \in N_i} w^*_{ij}) \geq \sum_{i,j \in S:(i,j) \in E} w^*_{ij} = e(S) \cdot (1) = e(S)$$

where the first inequality is trivial, and the second inequality follows from the fact that for each edge with two end nodes i, j both in S, $w^*_{ij} + w^*_{ji} = 1$, there are exactly $e(S)$ such links in the summation.

To show sufficiency, we first reformulate the above condition as a feasibility condition to a network flow problem; we then apply Gale's Demand Theorem (see Gale 1957). From the original network $G = (N, E)$, we construct a specific bipartite network $\tilde{G} = (V, A)$, where the set of nodes is the union $V = V_1 \cup V_2$ where $V_1 = E$ and $V_2 = N$. The arcs (flow) in $\tilde{G}$ are only from V_1 to V_2. In particular, $f \in E = V_1$ is connected to $i \in N = V_2$ in the bipartite graph $\tilde{G} = (V, A)$, if and only if i is one of the endpoints of this edge f in the original network G. Clearly, $|V_1| = e(N)$, and $|V_2| = |N|$.

Each vertex $i \in V_2$ is a demand vertex, demanding $d_i = \frac{e(N)}{|N|}$ units of a homogeneous good. Each vertex in $j \in V_1$ is a supply vertex, supplying $s_j = 1$ unit of the same good. Supply can be shipped to demand nodes only along the arcs A in the constructed bipartite network $\tilde{G}$. Gale's Demand Theorem states that there is a feasible way to match demand

and supply if and only if for all $S \subset V_2$:

$$\sum_{i \in S} d_i \leq \sum_{j \in N(S)} s_j,$$

where $N(S)$ is the set of neighbors of vertices in S in $\tilde{G}$. Substituting the values of s_j, d_i yields the following equivalent condition:

$$|S|\frac{e(N)}{|N|} \leq |N(S)|, \quad \forall\, \emptyset \subset S \subset V_2.$$

Clearly, the above condition holds when S is either empty or the whole set N. For any other case of S, from the construction of $\tilde{G}$, the set $N(S)$ is only the edges in E such that at least one endpoint belongs to S. Therefore:

$$|N(S)| = e(N) - e(S^c)$$

where $S^c = N \backslash S$ is the complement set of S. Recall that:

$$|N| = |S| + |S^c|, e(N) = |N(S)| + e(S^c),$$

It is easy to see that:

$$|S|\frac{e(N)}{|N|} \leq |N(S)| \Longleftrightarrow \frac{e(N)}{|N|} \leq \frac{|N(S)|}{|S|} \Longleftrightarrow \frac{e(N)}{|N|} \leq \frac{e(N) - e(S^c)}{|N| - |S^c|} \Longleftrightarrow \frac{e(S^c)}{|S^c|} \leq \frac{e(N)}{|N|}.$$

So the feasibility condition is equivalent to the following:

$$\frac{e(S^c)}{|S^c|} \leq \frac{e(N)}{|N|}, \quad \forall\, \emptyset \neq S^c \subset N.$$

Since S is an arbitrary subset of N, and S^c is also arbitrary, the sufficiency direction is proven. This establishes Proposition 1. $\square$

H Additional results

We now consider six extensions and variations of our model.

H.1 Unbounded noise

The above model takes agents' noise supports to be contained within the bounded interval $[-\nu, \nu]$.[A7] The positive and normative implications of the model are maintained in the

[A7] This assumption conveniently yields equilibrium properties near the noiseless limit that are commensurate with the properties of $\boldsymbol{\pi}^*$. In particular, local contagion (Proposition 4) and the reach of noiseless limit under unbounded noise. Consider, for example, the perturbed game where θ is observed with Gaussian noise by all agents: each i observes signal $s_i = \theta + \epsilon_i$, where each $\epsilon_i \sim N(0, \nu)$, $\nu > 0$, and all signals are independently drawn conditional on θ. Theorem G1 continues to describe the limit equilibrium $\vec{\boldsymbol{\pi}}$.[A8] Therefore, all limiting characterizations, including those of sticky coordination, linkage, and local contagion, as well as the model's welfare properties, are intrinsic to the equilibrium selected from the complete information game Γ^0.

H.2 Noise-independent selection

The equilibrium selection in the noiseless limit is not sensitive to the commonality of the noise distribution F. Here, we consider the robustness of equilibrium selection to heterogeneous noise structures. Consider the following extension.[A9]

Information structure. In the perturbed game, each i realizes signal $s_i = \theta + \nu\epsilon_i$, $\nu > 0$, where ϵ_i is distributed via density function f_i and cumulative function F_i with support within $[-1, 1]$. Signals are independently drawn across agents conditional on θ.

As shown in Theorem 1, the limiting cutoff θ_i^* are fully determined by the parameters $\mathbf{v}, \phi$, and $\mathcal{G}$; in particular, the cutoffs are independent of the noise distribution F_i and the prior distribution of θ.

H.3 Weighted links

We can extend our results to allow for edge-weights $e_{ij} > 0$ for each $(i, j) \in E$. In this extension, the noise-independent selection is maintained. All results can be extended upon adjusting for edge-weights.

First, we write i's ex-post payoff as follows:

$$u_i(\mathbf{a}, \theta) = a_i(v_i + \theta + \phi \sum_{j \in N_i} e_{ij} a_j). \tag{H.1}$$

Theorem G1 remains unchanged.[A10] We extend the following definitions to allow for weighted links. First, define i's *weighted degree* $d_i(S) \equiv \sum_{j \in N_i \cap S} e_{ij}$. The definitions of $L(\cdot, \cdot)$ and $e(\cdot)$ then go through:

$$L(S, S') = \sum_{i \in S} d_i(S'),$$

policy interventions (Proposition 5) extend throughout but remain contained within coordination sets, provided ν is sufficiently small.

[A8] An analogous proof to Lemma 1 can be constructed. Beyond this, the theorem's proof is identical.

[A9] Frankel, Morris, and Pauzner (2003) Section 6 addresses such an enrichment.

[A10] The proof of Theorem G1 requires only modest adjustments; we leave this for the reader.

$$e(S) = \frac{1}{2}\sum_{i\in S} d_i(S), \qquad F(S) = v(S) + \phi e(S).$$

Theorem 1 and SAND solving for equilibrium cutoffs and Proposition 1 characterizing network-wide coordination remain unchanged.[A11] Proposition 2 is reserved for unweighted graphs, as the following three-agent counterexample can easily be constructed. If $e_{ij} \gg e_{jk}$ with $e_{ik} = 0$, then agents i and j will coordinate together, and agent k will take a strictly higher cutoff. And when intrinsic values are heterogeneous, the comparative statics results of Proposition 4 also go through.

Importantly, the noise-independent selection (Section H.2) continues to hold when links are weighted and symmetric. The relevant potential function becomes:

$$P(\mathbf{a}, \theta) \equiv \sum_{i\in N}(v_i + \theta)a_i + \frac{1}{2}\phi \sum_{(i,j)\in E} e_{ij}a_i a_j, \text{ where } \mathbf{a} \in \{0,1\}^N, \tag{H.2}$$

and the results of Frankel et al. (2003) and Basteck et al. (2013) carry through.

H.4 Miscoordination costs

We can set $v_i = v - \phi d_i$ to give:

$$u_i(1, \mathbf{a}_{-i}, \theta) = v + \theta - \phi \sum_{j\in N_i}(1 - a_j). \tag{H.3}$$

The agents then face miscoordination costs in their adoption choices. More connected coordination sets take *higher* cutoffs, and agents' links to coordination sets taking lower cutoffs carry zero weight; these miscoordination costs are avoided with a probability of one. Links to others within one's coordination set are penalized according to limit likelihoods placed on the neighbors not adopting. And links to coordination sets taking higher cutoffs are penalized with a weight of one, with these costs being borne with a probability of one.

When a single coordination set is obtained, the common cutoff is $\theta_1^* = (-v + \phi\frac{e(N)}{|N|})$, by Theorem 1. Moreover, one can use the SAND algorithm to reconstruct Proposition 1 as the equivalent condition for a single coordination set. To show this, SAND terminates in one step if and only if:

$$\frac{F(S)}{|S|} = \frac{|S|v - \phi\sum_{i\in S} d_i + \phi e(S)}{|S|} \leq \frac{|N|v - \phi\sum_{i\in N} d_i + \phi e(N)}{|N|} = \frac{F(N)}{|N|}, \quad \forall \emptyset \neq S \subset N,$$

which, given $-\sum_{i\in S} d_i + e(S) = -e(S, S^c) - e(S)$, is equivalent to:

$$\frac{e(S, S^c) + e(S)}{|S|} \geq \frac{e(N)}{|N|}, \quad \forall \emptyset \neq S \subset N.$$

[A11] Again, these require modest adjustments to proofs and updated definitions of $F(\cdot)$, $e(\cdot)$, etc.

Because $E = e(S) \cup e(S^c) \cup e(S, S^c)$, for this inequality to hold it must be that:

$$\frac{e(S^c)}{|S^c|} \leq \frac{e(N)}{|N|}, \quad \forall \emptyset \neq S^c \subset N.$$

As this is true for all nonempty $S^c \subset N$, we are free to drop the complement superscripts. Thus, Proposition 1 continues to hold.

H.5 Non-strategic externalties

We can incorporate a non-strategic externalties function $w_i(\mathbf{a}_{-i}, \theta)$ to augment both $u_i(1, \mathbf{a}_{-i}, \theta)$ and the payoffs to not adopting (now equal to $w_i(\mathbf{a}_{-i}, \theta)$ instead of zero). Under this extension, the equilibrium selected in the limit along with all of the positive results remain.

H.6 Local-average model

Here, we show the stark result that a single coordination set obtains for all network structures under *homogeneous intrinsic values* and when agents value the average of neighbors' adoption choices (i.e., the "local-average" model). To be precise, we assume that agent i holds the payoff function:

$$u_i(\mathbf{a}, \theta) = a_i \left(v + \theta + \phi \frac{1}{|N_i|} \sum_{j \in N_i} a_j \right). \tag{H.4}$$

With this payoff function, $U_i^\nu(\boldsymbol{\pi}_{-i}|s_i)$ for action $a_i = 1$ when i observes s_i and each $j \neq i$ uses strategy π_j takes the form:

$$U_i^\nu(\boldsymbol{\pi}_{-i}|s_i) = v + \mathbb{E}\left[\theta|s_i\right] + \phi \frac{1}{|N_i|} \sum_{j \in N_i} \mathbb{E}[\pi_j(s_j)|s_i].$$

Now, consider the profile of cutoff strategies in which all neighbors of i take cutoff c. For the realization $s_i = c$, agent i calculates $\mathbb{E}[\pi_j(s_j)|s_i = c] = 1/2$, yielding a posterior expected payoff to adopting of:

$$U_i^\nu(\boldsymbol{\pi}_{-i}|s_i = c) = v + \mathbb{E}\left[\theta|s_i = c\right] + \phi \frac{1}{2}.$$

The Bayesian-Nash Equilibrium (BNE) indifference condition is then:

$$v + \mathbb{E}\left[\theta|s_i = c^*\right] + \phi \frac{1}{2} = 0,$$

which is independent of the number of neighbors that i has. We have verified that a symmetric BNE in which each player takes cuttoff c^* solving $\mathbb{E}\left[\theta|s_i = c^*\right] = -(v + \phi/2)$

exists. By Theorem 1 of Frankel et al. (2003), the unique limit equilibrium is symmetric and satisfies $\theta_i^* = \theta^* = -(v + \phi/2)$ for each $i \in N$.

I Private intrinsic values

In this section, we show that our limit equilibrium is robust to private information in the perturbed game $\Gamma(\nu)$ regarding intrinsic values v_i. In the following subsection, we formalize such an extension. In the second subsection, we show this extension defines a limiting case of Oury (2013), and apply their main result.

I.1 Model with i.i.d. intrinsic values

Payoffs. Payoffs from adopting the technology are as in the baseline model:

$$u_i(\mathbf{a}, \theta) = a_i \left(v_i + \theta + \phi \sum_{j \in N_i} a_j \right), \tag{I.1}$$

but now with v_i treated as a random variable with θ.

Information structure. In the perturbed game, θ is observed with noise by all agents. Each i realizes signal $s_i = \theta + \nu\epsilon_i \in S$, $\nu > 0$, where ϵ_i is distributed via density function f^ϵ and cumulative function F^ϵ with support $[-1, 1]$. Each intrinsic value v_i is distributed via density function f^v and cumulative function F^v with bounded interval support $V \subset \mathbb{R}$. Player i privately observes her v_i before choosing action a_i. All signals $s_1, \ldots, s_n$ are independently drawn across agents conditional on θ. Intrinsic values $v_1, \ldots, v_n$ are independent of each other, independent of θ, and independent on all signals.

Dominance regions. For each i, we assume $\min(V)$, $\max(V)$, and ϕ are such that there exist $\underline{\theta}_i$ and $\overline{\theta}_i$ such that $\max(V) + \theta + \phi d_i < 0$ when $\theta < \underline{\theta}_i$ and $\min(V) + \theta > 0$ when $\theta > \overline{\theta}_i$. Thus, there exist dominant regions $[\min\Theta, \underline{\theta}]$ and $[\overline{\theta}, \max\Theta]$, with $\underline{\theta} \equiv \min_i\{\underline{\theta}_i\}$ and $\overline{\theta} \equiv \max_i\{\overline{\theta}_i\}$, such that not adopting and adopting the technology (respectively) are dominant strategies for all players. When the realization of θ is common knowledge among agents, with $\phi > 0$ there can exist a strictly positive measure of θ realizations within $[\underline{\theta}, \overline{\theta}]$ at which multiple pure strategy Nash equilibria occur.

I.2 Mapping to Oury (2013)

We now show that the above extension gives a limiting case of the "N-dimensional global games" setting of Oury (2013). For this, we incorporate a passive $(n+1)$'th player by setting $A_{n+1} = \{1\}$, the singleton set. Conceptually, this additional passive player allows for the state $\mathbf{t}$ to incorporate both active players' private intrinsic values in addition to the common state θ.

To facilitate what follows, we adopt the notation of Oury (2013) wherever possible, and state when otherwise. $B^{\mu}(\mathbf{x})$ gives the $(n+1)$-dimensional ball in $\mathbb{R}^n$ around vector $\mathbf{x}$ under euclidean norm $||\cdot||$. $G_\nu(u,\Phi,\Psi)$ denotes the perturbed game given $\nu > 0$, where each player i receives an $(n+1)$-dimensional private signal $\mathbf{s}_i = \mathbf{t} + \nu\boldsymbol{\eta}_i^t$ of payoff relevant state $\mathbf{t} \in T \subseteq \mathbb{R}^{n+1}$, Φ defines the density function to the errors $\boldsymbol{\eta}_i^t \in \mathbb{R}^{n+1}$, and Ψ defines the prior distribution of $\mathbf{t}$. $(\boldsymbol{\eta}_1^t,\dots,\boldsymbol{\eta}_n^t)$ are assumed to be independently distributed. The density functions Φ_i^t are assumed to have support within $B^{1/2}(0)$.

We now define θ as in our baseline model and in the extension of the preceding subsection.[A12] For each $i = 1,\dots,n$ we define $t_i = v_i$, and set $t_{n+1} = \theta$. For each $i \in N$, we capture i's private information of v_i and θ by assuming that the i'th and $(n+1)$'th components of joint distribution Φ_i^t have supports within $B^{\delta}(0)$ and $B^{\delta'}(0)$, respectively, for arbitrarily small but positive $\delta, \delta' \ll 1/2$, while all other components of Φ_i^t have common support $B^{1/2}(0)$. To further capture player i's highly precise information of v_i, we may further impose that $\delta \ll \delta'$. Indeed, in the limit $\delta \to 0$, this model limits on the above extension where players perfectly observe their own intrinsic values.[A13]

We can now establish that $G_\nu(u,\Phi,\Psi)$ indeed embeds the complete information game defined by g in the sense of Oury (2013). For this, we show that for each $t^\star \in T$, we can define a continuous function $\Gamma : \Theta \mapsto \mathbb{R}^{n+1}$ such that (i) $\Gamma(\theta^\star) = \mathbf{t}^\star$ for some $\theta^\star \in \Theta$, and (ii) Assumptions 1 to 3 specified in Oury (2013) hold for this Γ. For (i), for any realization of $\mathbf{v}^\star$, we define $\Gamma(\theta) = (v_1^\star,\dots,v_n^\star,\theta)$; clearly, $\Gamma(\theta^\star) = \mathbf{t}^\star$. For (ii), we further define $g_i(\cdot) = u_i(\cdot,\mathbf{t}^\star) \equiv u_i(\cdot|v_i^\star,\theta^\star)$ and $\Delta g_i(a_i \to a_i', \mathbf{a}_{-i})$. Then, setting $K = 1$ satisfies Assumption 1. Assumption 2 follows by definition of the pair $\underline{\theta},\overline{\theta}$ and the assumption that these cutoffs are in the interior of Θ. Assumption 3 follows given strategic complementarities hold for all $\mathbf{t}$.

This notion of embedding applies to our benchmark model if we instead require that Ψ gives the one-dimensional prior distribution of θ, and v_i is fixed for each $i \in N$. In our model, Ψ corresponds to the distribution H. This one-dimensional embedding corresponds to the "FMP global games embedding" discussion in Oury (2013) and is referred to below.

Along with the following definition of noise-independent selection, we now restate the main result of Oury (2013) verbatim (modulo the above notation).

Definition. *We say that the action profile $\mathbf{a}^\star$ is noise-independently selected in the global games $G_\nu(u,\Phi,\Psi)$ embedding the complete information game g at state parameter $\mathbf{t}^\star$, if for each noise structure Φ, there exists $\hat{\nu} > 0$ such that for each $\nu < \hat{\nu}$ and each strategy profile $\mathbf{s}$, $s_i : \mathbb{R}^{n+1} \mapsto \{0,1\}$, surviving iterative strict dominance in $G_\nu(u,\Phi,\Psi)$, $s(\mathbf{t}^\star) = \mathbf{a}^\star$.*

In the context of our baseline model, the set of players adopting in action profile $\mathbf{a}^\star$ is equal to the union of all coordination sets adopting in state $\theta^\star$ in the noiseless limit where

[A12]Note that in Oury (2013), $\underline{\theta}$ and $\overline{\theta}$ define the lower and upper bounds of θ, while in our setting, these define the cutoffs to the dominance regions; in our setting, we denote the interval domain of θ by Θ.
[A13]Moreover, for large ν, their posteriors place high weight on their priors of others' intrinsic values.

intrinsic values are set to $\mathbf{v}^\star$.

Theorem I3 (Oury 2013). *The action profile* $\mathbf{a}^\star$ *is noise-independently selected in all FMP global games embedding the complete information game* g *if and only if* $\mathbf{a}^\star$ *is noise-independently selected in all multidimensional global games embedding* g.

With Theorem 1 of Oury (2013), the unique equilibrium selected in noiseless limit $\nu \to 0$ is equivalent to that selected in the extended model where players hold private information on intrinsic values.

J Differences betweeen complete and incomplete information: An Example

Consider the network of five agents depicted in the left panel of Figure J.1. It is easily verified that the maximal and the minimal equilibria are very different; see the middle panel of Figure J.1.[A14] Indeed, in the *minimal equilibrium*, all five agents adopt together when the technology's value (the state of the world θ) rises above the common threshold θ^- and do not adopt below this value. However, θ^- corresponds to the adoption cutoff for an agent in isolation, without network effects. So, the agents entirely fail to coordinate and capitalize on the positive externalities between their adoption choices.

In the *maximal equilibrium*, on the other hand, agents $\{1, 2, 3, 4\}$ succeed in capitalizing on their positive network effects. The clique now adopts together for technology values θ above some cutoff $\theta^+_{-5} < \theta^-$, while agent 5 joins for values above θ^+_5, where $\theta^+_{-5} < \theta^+_5 < \theta^* < \theta^-$. That is, the clique perfectly coordinates their adoption, with agent 5 benefiting as a consequence.

If we now consider our model with incomplete information about the quality of the technology (state of the world) and agents observe private signals of the state, with even the mildest noise, things change considerably. In particular, in the noiseless limit, a unique Nash equilibrium of the complete information game is selected. As depicted in the middle panel of Figure J.1, the unique global games equilibrium selection involves all agents adopting when θ rises above the common cutoff θ^* (note that the network is balanced). Thus, for each value of θ, such that $\theta^+_{-5} < \theta < \theta^-$, this equilibrium is distinct from either the maximal or the minimal equilibrium. Importantly, since there is a unique equilibrium in the incomplete information case, we do not need to arbitrarily choose the maximal or the minimal equilibrium, as is usually the case in this literature.

Let us now compare the comparative statics of the maximal or minimal equilibrium with that of the unique, global games equilibrium.

[A14] In the middle panel of Figure J.1, the sets of adopting agents in maximal and minimal equilibria under complete information are depicted in green and red, respectively, while that of the global games equilibrium is displayed in blue.

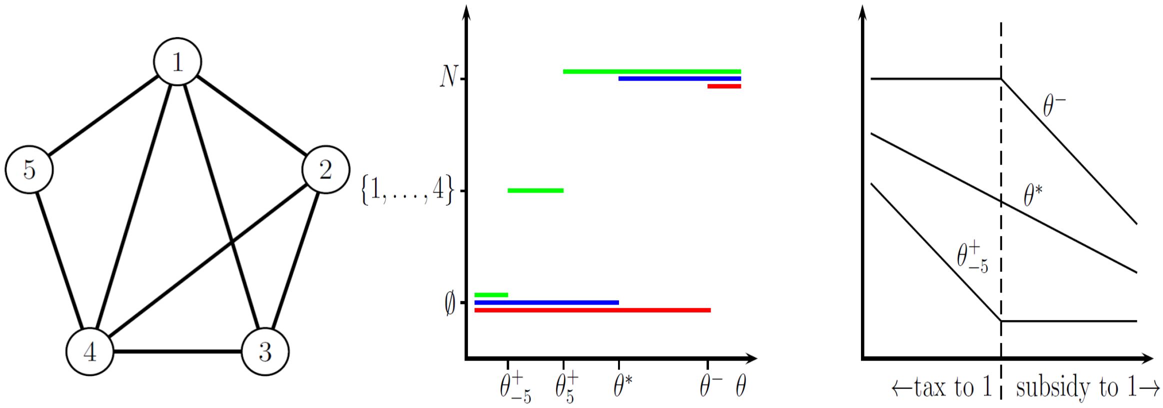

Figure J.1: Complete versus incomplete information

First, the comparative statics of the maximal and minimal equilibria are notably distinct from each other, as depicted in the right panel of Figure J.1.[A15] Indeed, consider the effect of a small, publicly observed adoption *tax* on one agent, say agent 1 (see Figure J.1) in the *minimal* equilibrium. Because θ^- defines the adoption threshold under isolation, 1's decreased value to adoption carries no consequence on the adoption choices of others. In fact, agent 1 *herself* continues to adopt when the technology value rises above θ^-, because the network effects from agent 1's neighbors' adoptions compensate 1 for the tax. As a result, this policy of taxing agent 1 has *no* effect on the adoption rate of all agents. Conversely, if agent 1's adoption is *subsidized*, 1 will begin to adopt at values below θ^-, which effectively subsidizes *others'* adoptions; the intervention affects *all* players as they all adopt at a lower threshold than θ^-. Moreover, a subsidy to a second agent only carries an effect once that subsidy rises above the subsidy to agent 1. In summary, we see very different behaviors in the effects of subsidies and taxes to individual agents in the minimal equilibrium.

Similar distinct behaviors are obtained in the *maximal equilibrium*, though in a reversed manner. A small *subsidy* to agent 1 carries *no* effect on equilibrium actions. On the contrary, a small *tax* on agent 1's adoption causes her (who belongs to the clique) to adopt at values strictly above θ^+_{-5}, which causes the other three agents in the clique to increase their cutoffs as well. The agents in the clique, who coordinate their adoptions with agent 1's adoption, remain unresponsive to taxation until their tax surpasses the tax on agent 1. Hence, we see a *reversal* in the effects of subsidy and taxation on adoption.[A16]

[A15]In the right panel of Figure J.1, resulting from tax/subsidy to agent 1, comparative statics analyses of maximal and minimal equilibria under complete information are given by θ^+_{-5} and θ^-, respectively, while that of global games equilibrium is given by θ^*.

[A16]In Proposition D1 in the Online Appendix of the working paper version (Leister et al. 2019), we

Second, as depicted in the right panel of Figure J.1, the comparative statics of our unique equilibrium using global games are far removed from those of the maximal and minimal equilibria. By Proposition 4, *all* agents who coordinate with agent 1 respond to the intervention. In fact, all that is required for an agent to respond to the subsidy/tax is that she be path connected with the recipient through other agents who also coordinate their adoptions with the recipient.

formally derive the comparative statics results for the maximal and minimal equilibria in the *complete information* case.