Strategic Diffusion: Public Goods vs. Public Bads*

Arthur Campbell† D.J. Thornton‡ Yves Zenou§

September 6, 2024

Abstract

We study the role of influence in a model of the diffusion of social behaviors in a network. Individual behavior creates either positive spillovers (public goods) or negative spillovers (public bads). Our notion of influence captures the causal effect of an agent's adoption decision on the adoption decision of others in the network. We study a phase transition in equilibrium behavior around which viral equilibria—where diffusion occurs among a nontrivial fraction of the population—emerge. Public goods exhibit a continuous phase transition in equilibrium adoption, while public bads exhibit a discontinuous transition– they emerge suddenly. Our findings reconcile disparate evidence that attending a public protest is a strategic complement in some settings and a substitute in others.

Keywords: Diffusion, public goods, public bads, narratives, networks, bystander effect, mob-rule.

JEL classification codes: D43, D85, L13.

*We are grateful to the participants at the 2023 Brown Bag seminar at Monash University, the 2024 Australasian Economic Theory Workshop at the University of Melbourne, and the Barcelona School of Economics Summer Forum Networks Workshop in 2024. The financial support from the Australian Research Council (DP200102547) is gratefully acknowledged.

†Monash University, Australia. E-mail: arthur.campbell@monash.edu

‡University of New South Wales, Australia. E-mail: d.thornton@unsw.edu.au

§Monash University, Australia, and CEPR. E-mail: yves.zenou@monash.edu

1 Introduction

On September 13, 2022, a 22-year-old Iranian Kurdish woman named Mahsa Amini was arrested by the morality police in Tehran for failing to comply with Iran's mandatory headscarf (hijab) laws. She passed away in a hospital three days later under suspicious circumstances. This sparked a revolt in Iran, marking a significant period of civil unrest. There were street protests in which some women tore off their hijabs and demonstratively twirled them in the air, while others threw their hijabs into a bonfire. A common element of protest movements are observable symbols of participation / support for the movement.[1] These symbols raise awareness for the movement and help to facilitate the spread of support through a population.

Participation in protest movements are privately costly but potentially generate wide-spread benefits (public goods) if the movement achieves its goals. The observable nature of participation raises awareness of the movement and thereby potentially *influences* other individual's decision to participate.[2] If those individuals then choose to join the movement their decision may lead others to join and so forth. Thus, an individual may benefit through how their decision influences the diffusion of the action to others, thereby creating positive externalities. The aim of this paper is to develop a model of privately provided public goods and bads where the observable nature of the action serves to spread awareness. We seek to understand how the strategic consideration introduced by observational learning determines the equilibrium level of provision and systematic differences in the emergence of public goods versus public bads.[3]

In our model, agents make a single adoption decision (e.g., whether or not to join a protest movement), which—in the case of a public good—entails a private cost (e.g., being arrested) that exceeds the privately created benefit of the action, but creates a positive externality for everyone else in the population (e.g., increases the probability the movement is successful). In the case of a public bad, the private benefit exceeds the private cost and adoption creates a negative externality. The key innovation in our paper is to explicitly model the externality as something which arises as an equilibrium object that depends both on equilibrium strategies and on the structure of the network on which diffusion takes place. Individuals are connected to each other in a network (modeled as a random graph) where a connection in the graph allows

[1] Such examples are plentiful and include the use of a specific color, gesture, or display of personal items. For example, consider the Orange Revolution in Ukraine, the 3-finger salute in Thailand, the raised clenched fist in the U.S. Black Lives Matter protests, the Umbrella Movement in Hong Kong, Guy Fawkes masks used by Occupy Wall Street protesters, etc.

[2] The importance of observational learning in the diffusion of a protest is exacerbated by governmental and media censorship.

[3] The types of public goods and bads we have in mind are ones that are created through private decisions/actions. Examples of public goods include participation in public protests, mask wearing in public when ill, responsible consumption, engagement in the democratic process and other pro-social behaviors. Examples of public bads include littering, graffiti, rioting/looting, spreading mis-information and other anti-social behaviors.

an individual to observe their neighbors adopting the behavior, at which point they become aware of the action (e.g., joining the protest). Our model can be viewed as a diffusion game (Sadler (2020)) wherein a randomly chosen individual initially adopts the action, then, in each subsequent period, an individual who has observed a neighbor adopt the action in the previous period makes a once-and-for-all decision to take or not take the action.[4] The process continues until no one else adopts.

In our model, the diffusion introduces a new strategic consideration for an individual taking the action through its potential effect on other people's awareness and subsequent decision to take the action (e.g. joining the protest movement). This strategic consideration is what we call an individual's *influence*. An individual's influence is defined as the number of additional people that take the action by the end of the game when the individual takes the action relative to the number of people who would take the action if the individual did not take it. Individual behavior creates either positive spillovers (public goods) or negative spillovers (public bads) in that agents receive positive utility from influencing others to adopt a public good, while they receive disutility from influencing others to adopt a public bad. Thus, influence encourages public good behaviors and discourages public bad behaviors.

Central to our analysis is the interaction between influence and network density (average number of connections) and their effects on the size of diffusion. Network density has two countervailing effects on influence. First, more connections may increase influence because more people observe any one individual's decision to adopt (the "local effect" of density). Second, a network with more connections has more potential paths of via which agents can become informed, reducing each individual's influence (the "global effect" of density). When there are multiple disjoint paths along which diffusion can occur from the initial seed to an individual, none of the people on either one of those paths has influence over that individual's decision. This occurs because if any agent in one of the paths does not adopt, the individual will still become aware through the other path. It is the interaction between these two forces that generates non-trivial effects of network density on equilibrium behavior. To the best of our knowledge, our notion of influence is new to the networks literature.[5] We use random graphs to model the interaction among the players. This gives us a degree of tractability in explicitly characterizing the two countervailing forces of influence that would seem otherwise intractable using alternative network models. In the spirit of Erdös and Renyi's (1959) classic result on the emergence of the giant component in

[4]In this setting, an individual has only one opportunity to take the action upon first becoming aware of it, so there is no possibility of strategically delaying a decision to act.

[5]Our notion of influence captures a kind of *exclusive influence* where an agent's payoff depends only on the subset of others that cannot be influenced by any other agent the original agent does not influence. Put differently, influence captures a kind of *pivotality* over others' adoption decisions. In the simple setting where a graph is a line or a tree, our definition of influence coincides with that of Bénabou et al. (2020); however, in our setting we are also able to capture the global effect of density.

random graphs, we study a phase transition in *equilibrium behavior* as the density of the random graph increases. In particular, there exists a critical threshold level of network density below which the diffusion of actions are contained to a negligible fraction of the population (*non-viral* equilibria) and above which the diffusion of actions may spread to a nontrivial fraction of the population (*viral* equilibria).

In our analysis, we first characterize the expected influence as a function of symmetric strategies and of the network density. In sparse networks, increasing network density increases the expected influence, while in dense networks, increasing the density decreases the expected influence. The properties of influence implies that adoption exhibits strategic compelmentarities in sparse networks and strategic substitutability in dense networks. This finding can reconcile disparate evidence that attending a public protest is a strategic complement in some settings and a substitute in others.[6] We then offer a full characterization of symmetric equilibria and classify them into four different cases: No-adoption, full-adoption, viral mixed, and nonviral-mixed.

One of the significant consequences of our analysis is the stark difference between the phase transition in equilibrium behavior for public goods compared to public bads, despite the fact that the utility of adopting a public bad is simply the negative of the utility of adopting a public good. Public goods exhibit a continuous phase transition in equilibrium adoption, while public bads exhibit a discontinuous transition—they emerge suddenly.

Finally, we study the choice of network density by a social planner. The planner chooses the network density to maximize the ex-ante expected diffusion of public goods and minimize the diffusion of public bads *before* knowing which of these will spread on the network. We show that regardless of the relative likelihood that a public good versus public bad diffuses through the population, the socially optimal density of the network is achieved at precisely the critical point where the phase transition for public bads occurs. Consequently, if a social planner sets a level of network density below but arbitrarily close to this level it may achieve arbitrarily close to the welfare maximizing level of utility.

1.1 Related literature

Our model is related to several different strands of the literature.

Network games on fixed networks and diffusion

There is a large literature on games on networks (Jackson and Zenou, 2015; Bramoullé and Kranton, 2016; Jackson et al., 2017) where the network is fixed and not random. In these

[6]See Cantoni et al. (2019) and Bursztyn et al. (2021).

models, the links between players represent a strategic externality from one player's action on another player's utility. Within this class of models, games with strategic complements and/or strategic substitutes with continuous actions have been studied (e.g.(Ballester et al., 2006) and Bramoullé et al. (2014)). The case of a privately provided public good has also been explicitly considered in Bramoullé and Kranton (2007), Allouch (2015), Allouch (2017) and Elliott and Golub (2019). There are also many papers that study network games with discrete actions (Granovetter, 1978; Schelling, 1978; Blume, 1993; Ellison, 1993; Morris, 2000; Brock and Durlauf, 2001; Leister et al., 2022; Langtry et al., 2024) and study contagion on a fixed network. These are games with strategic complementarities in which the benefits of taking action "1" (e.g., adopting a new technology) increase with other individuals taking the same action. In such a case, an agent will take action 1 if a sufficient fraction of their neighbors has taken action 1 (i.e., a threshold decision). These models show that the possibility of contagion depends on the network structure. In particular, Morris (2000) shows how the spread of behavior in a network can depend on the network's *cohesiveness.*

Our framework differs from these fixed-graph models in two key ways: first, we use random graphs, and second—and more importantly—in our model, network connections allow observation of another person's adoption but do not indicate the presence of an externality from one person's decision on another. In our case, externalities are global and, conditional on the action, independent of the network. This means we do not have a threshold rule based on the fraction of neighbors who adopt; instead, the decision to take an action is primarily influenced by the expected impact on others, whether they are direct neighbors or not.

Network games on random graphs and diffusion

There is also a literature on network games using random graphs to study diffusion.[7] In particular, economists have been using the random graphs techniques involving phase transitions to study diffusion in economic contexts (Watts, 2002; Campbell, 2013; Sadler, 2020; Akbarpour et al., 2023; Langtry, 2023; Campbell et al., 2024). In these models, adoption decisions depend on the structure of the network and its components.[8] There is also a literature using threshold models on random graphs, so that agents decide whether or not to take a binary action depending on the fraction of neighbors who take the same action (Jackson and Yariv, 2005; López-Pintado, 2006, 2008; Jackson and Yariv, 2007, 2011; Jackson and López-Pintado, 2013; Campbell et al., 2024).

The closest model to ours is that of Sadler (2020), who studies the emergence of a giant com-

[7]For an overview, see Newman et al. (2001), Kleinberg (2007), Vega-Redondo (2007), Jackson (2008) and Easley and Kleinberg (2010).

[8]Dasaratha (2023) instead focuses on network formation and examines the interaction between strategic incentives and network structure.

ponent in a game with strategic complementarities between direct neighbors. Our model is different in that strategic considerations arise from *global* rather than local externalities and so adoption is not necessarily complementary. That is, the number of adopting individuals in the entire network (whether or not they are direct neighbors) affects an individual's utility.[9] Moreover, our emphasis on the distinction between the diffusion of public goods and public bads has not been studied in this literature.

Collective action

There is a large literature on collective action. The theoretical literature usually assumes *strategic complementarity* because the cost of taking (for example) joining a protest is anticipated to be lower when a protest is larger (e.g., Kuran, 1989, 1991, 1997, Chwe, 2000, Bueno de Mesquita, 2010, Edmond, 2013, Battaglini, 2017, Passarelli and Tabellini, 2017, Barberà and Jackson, 2020, and Egorov and Sonin, 2021). There is also a literature which assumes that actions are strategic substitutes because agents have an incentive to free-ride on the costly participation of others, and may thus be less willing to turn out when they believe more others will do so (Olson, 1965, Tullock, 1971, Palfrey and Rosenthal, 1984, and Shadmehr and Bernhardt, 2011).

Here, as with many network models, the collective action problem faced by agents is usually modeled as a threshold phenomenon. However, most of this literature does not focus on the role of networks on diffusion (Chwe (2000) is an exception). We have a model of collective action in which the network is explicitly modeled and in which diffusion is related to the emergence of a giant component in a random network. An interesting new feature of our model is that, in the case of public goods (e.g., protests), adoption strategies are *strategic complements* when the network is sparse and *strategic substitutes* when the network is dense. In the context of Hong Kong's anti-authoritarian movement, Cantoni et al. (2019) find evidence of strategic substitutability in protest participation while Bursztyn et al. (2021) find that incentives to attend one protest within a political movement increase subsequent protest attendance but only when a sufficient fraction of an individual's social network is also incentivized to attend the initial protest (i.e., attendance seems to exhibit strategic complementarities).

[9] In Sadler (2020), adoption externalities only arise when a person also adopts. Hence, the number of adopting neighbors when an individual does not adopt is an irrelevant strategic consideration.

2 The Model

2.1 Outline

In this section, we present a model of diffusion on an Erdös-Renyi random graph.[10] Let $\{G(n,p)\}_{n\in\mathbb{N}}$ be a sequence of Erdös-Renyi (or *binomial*) random graphs, where $p = p(n) = \frac{\lambda}{n}$ for some $\lambda \geq 0$. We endow $\{G(n,p)\}_{n\in\mathbb{N}}$ with the structure of a game $\mathcal{G}^{(n)} = (G(n,p), \mathcal{A}, u)$ as follows. Each vertex $i \in \{1,\ldots,n\}$ represents an agent, and each agent faces a binary adoption decision $a_i \in \mathcal{A} = \{0,1\}$. If player i chooses the action $a_i = 1$, we say that player i is an *adopter*.[11]

The game $\mathcal{G}^{(n)}$ is played over $n+1$ time periods t,. Denote by $a_i(t)$ the action player i at time t. At $t = 0$, every player has action $a_i(0) = 0$. At $t = 1$, nature makes three moves:

1. First, nature draws a graph $G = G(n,p)$ by including each possible edge independently with probability p.

2. Second, nature chooses a "seed" uniformly at random and,

3. Third, the seed adopts the public good or public bad $\pi \in \{g,b\}$.[12]

For all $t \geq 2$, neighbors of an adopter in the previous period are *exposed* and make a once and for all decision to adopt or not. Hence, the adoption decision occurs upon the first exposure for an individual but not again on any future occurrences.

2.2 Payoffs

At the end of the game, an agent i's payoff is a function of their own action a_i and the actions a_{-i}'s of the other agents. It is given by:

$$u_\pi(a_i, a_{-i}) = \begin{cases} \overbrace{(v-c)a_i}^{\text{intrinsic cost}} + \overbrace{v\sum_{j\neq i} a_j}^{\text{externality}}, & \text{if } \pi = g \\ (c-v)a_i - v\sum_{j\neq i} a_j, & \text{if } \pi = b. \end{cases} \tag{1}$$

where v captures the common value to all agents from each one of them who adopts and $c > v$

[10] We extend our model to an arbitrary degree distribution in Section 5.1.

[11] This model along with its more general analogue are similar to that of a *single-type diffusion game* (Sadler, 2020). We use Sadler's framework as a convenience to analyze the role of influence in the diffusion of public goods and bads.

[12] One can think of the seed as a non-strategic individual who exogenously wishes to engage in the particular behavior.

is the private cost (benefit) to an individual from adopting a public good (bad) behvavior. An agent's utility can be written as a combination of an intrinsic payoff from the action and an externality from the actions of others. We assume that the intrinsic payoff has the opposite sign to the externality; hence, the intrinsic payoff is negative (positive) in the case of a public good (bad).

2.3 Strategies

We make a number of assumptions about the information set available to a player at the time when they are first exposed. First, agents do not know their degree but believe correctly that it is distributed according to a binomial distribution with parameter $p = \lambda/n$.[13] Second, they do *not* know which of their neighbors has chosen to adopt or not, other than at least one has chosen to adopt. Third, they do not know the time period that they act (i.e. the time that has passed since the seed adopted). These guarantee that each agent acts at a single information set, all agents are symmetrically informed at the moment each acts, and each agent i's strategy may be characterized by a potentially mixed strategy over a single action $\sigma_i = \mathbb{P}(a_i = 1) \in [0, 1]$. The strategy σ_i tells an agent i the probability with which they take the action if they are exposed. We confine our attention to *symmetric equilibria,* that is, where all players play the same strategy $\sigma_i^* = \sigma^*$ for all i.

2.4 Influence

For any agent i (adopter or otherwise), we can consider the random number of other agents who adopt if i were to adopt, but do not have the opportunity to adopt otherwise (i.e., the information set where those agents have the opportunity to take an action is not reached). This is what we refer to as an agent i's influence: the number of agents that agent i *causes* to adopt by choosing $a_i = 1$. Consider the example in Figures 1 and 2, which shows the connected component containing the seed in a potential adopter network. The seed is colored blue, agents colored black have zero influence, and agents colored green have nonzero influence over the agents colored red that are adjacent to them.[14] In this example, if any of the green colored nodes chose to not adopt then the red nodes adjacent to them would become disconnected from the component containing the seed. In total, the reduction in the number of individuals adopting would be equal to the green node and the adjacent red nodes. Furthermore, to illustrate how fewer network links may increase the influence of agents, consider how influence changes after removing the links associated with a single vertex (as we do in Figure 2). In Figure 1, the

[13] We discuss the case where agents know their own degrees in Section 5.3.2.

[14] Some of the red nodes also have nonzero influence, but for illustration we are focusing on the neighbors of the seed.

first green node to the right of the seed has influence over only a single agent. This is because the other nodes adjacent to it are on an alternate path between the green node and the seed. But once we remove the second path, we increase the influence of this agent from 1 to 3. This is an example of what we earlier referred to as the global effect of density: decreasing the connectivity of a graph can *increase* the influence of agents within that graph. This happens precisely because removing links may disconnect loops thereby increasing the influence of agents on those loops. This idea is formalized in Proposition 2.

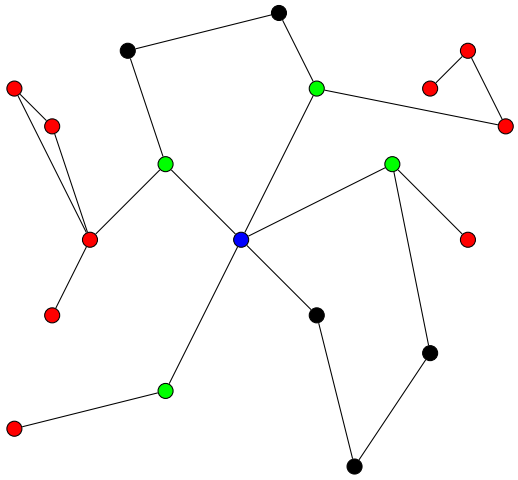

Figure 1: Influence

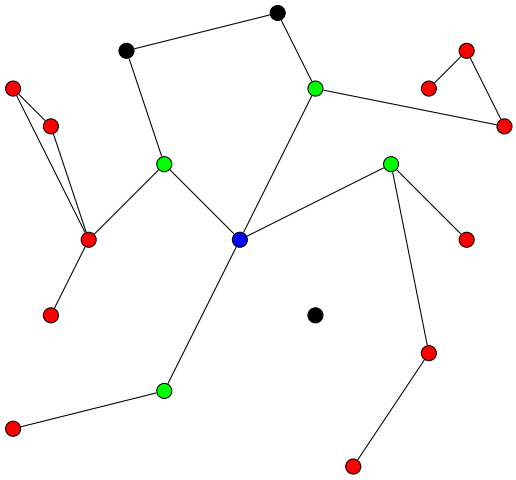

Figure 2: Removing a node

In expectation, the number of agents over which i has influence depends on the strategy σ which i expects others to play, and on the structure of the network. We write $\mathbb{E}_{\sigma}^{(n)}$ for an expectation taken in $\mathcal{G}^{(n)}$ assuming players follow the strategy σ, and define $\mathbb{E}_{\sigma} = \lim_{n\to\infty} \mathbb{E}_{\sigma}^{(n)}$. Similarly, define $\mathbb{E}_{\mathcal{D}_n}^{(n)}$ to be an expectation taken in $\mathcal{G}^{(n)}$ over the degree distribution $\mathcal{D}_n = \text{Bin}(n, p)$, and define $\mathbb{E}_{\mathcal{D}}$ as its limit, where $\mathcal{D} = \text{Po}(\lambda)$—the Poisson distribution. We then define $\mathcal{I}_{\lambda}^{(n)}(\sigma)$ and $\mathcal{I}_{\lambda}(\sigma)$ (read "the expected influence under σ") by

$$\mathcal{I}_{\lambda}^{(n)}(\sigma) \equiv \mathbb{E}_{\mathcal{D}_n}^{(n)}\left[\mathbb{E}_{\sigma}^{(n)}[\sum_{j\neq i} a_j | a_i = 1, h_i] - \mathbb{E}_{\sigma}^{(n)}[\sum_{j\neq i} a_j | a_i = 0, h_i]\right], \quad \text{and} \tag{2}$$

$$\mathcal{I}_{\lambda}(\sigma) \equiv \lim_{n\to\infty} \mathcal{I}_{\lambda}^{(n)}(\sigma) = \mathbb{E}_{\mathcal{D}}\left[\mathbb{E}_{\sigma}[\sum_{j\neq i} a_j | a_i = 1, h_i] - \mathbb{E}_{\sigma}[\sum_{j\neq i} a_j | a_i = 0, h_i]\right]. \tag{3}$$

Suppose all players are playing strategy σ. We normalize the utility of not adopting to 0 and we denote the expected change in utility from adopting by $u_{\pi}^{(n)}(\sigma) \equiv \mathbb{E}_{\sigma}^{(n)}[u_{\pi}(1, a_{-i}) - u_{\pi}(0, a_{-i})]$, and $u_{\pi}(\sigma)$ in the limit as $n \to \infty$.[15] Now, using the notation in (3) and the utility function (1), we can write the expected difference in utility to an agent from adopting $a_i = 1$ versus not

[15]The expectation here is taken over a given fixed graph so that influence is well defined.

adopting $a_i = 0$ at the information set where they have the opportunity to act—in particular, when they hold imperfect information about the graph—as:

$$\mathbb{E}_{\mathcal{D}}^{(n)}[u_\pi^{(n)}(\sigma)] = \begin{cases} v - c + v\mathcal{I}_\lambda^{(n)}(\sigma), & \pi = g \\ c - v - v\mathcal{I}_\lambda^{(n)}(\sigma), & \pi = b. \end{cases} \tag{4}$$

Observe that our definition of (expected) influence $\mathcal{I}_\lambda(\sigma)$ captures the *causal effect* of adoption on others' adoption—that is, the expected influence of agent i is the (random) number of other agents who adopt if i adopts but would not adopt otherwise.

3 Analysis

3.1 Existence of Equilibrium

A symmetric perfect Bayesian equilibrium[16] in the game $\mathcal{G}^{(n)}$ is a single strategy σ^* satisfying:

$$\begin{aligned} \sigma^* = 1 &\implies \mathbb{E}_{\mathcal{D}_n}^{(n)}[u_\pi^{(n)}(\sigma)] \geq 0 \\ \sigma^* \in (0,1) &\implies \mathbb{E}_{\mathcal{D}_n}^{(n)}[u_\pi^{(n)}(\sigma)] = 0 \\ \sigma^* = 0 &\implies \mathbb{E}_{\mathcal{D}_n}^{(n)}[u_\pi^{(n)}(\sigma)] \leq 0. \end{aligned}$$

and beliefs over nodes in the extensive form where an agent is exposed under the strategy σ^* and Nature's choice of a random graph with degree distribution $\mathcal{D}_n$. Beliefs are only payoff relevant for determining the expected influence of each agent and, given the symmetry of our game, this will be identical for each agent. It will be convenient to avoid specifying the beliefs themselves and rather specify the calculation of expected influence as a function of σ^* and $\mathcal{D}_n$ with the understanding this is through a set of beliefs formed via Bayes rule for each agent. The calculation of influence in the finite n case is discussed in Appendix B.

Proposition 1. *There exists a symmetric perfect Bayesian equilibrium in the game* $\mathcal{G}^{(n)}$.

We now focus on symmetric equilibria in the limit as $n \to \infty$.

Definition 1. *Let* σ' *and* σ *be strategies. We say that* σ' *is a* limit best-reply *to* σ, *if*

$$\lim_{n\to\infty} \mathbb{E}_{\mathcal{D}_n}^{(n)}[u_\pi^{(n)}(\sigma)] \geq 0 \text{ whenever } \sigma' = 1, \text{ and} \tag{5}$$

$$\lim_{n\to\infty} \mathbb{E}_{\mathcal{D}_n}^{(n)}[u_\pi^{(n)}(\sigma)] \leq 0 \text{ whenever } \sigma' = 0. \tag{6}$$

[16] All references to equilibria should be understood as meaning "symmetric equilibira".

The strategy σ' is a limit equilibrium *if σ' is a limit best-reply to itself.*

A limit equilibrium is an epsilon equilibrium for any $\epsilon > 0$ and all sufficiently large n. Furthermore, we will show generically that it can be found as a limit of a sequence of symmetric perfect Bayesian equilibria as $n \to \infty$ in the finite player game $\mathcal{G}^{(n)}$ (see Appendix D for more detail).

3.2 Potential Adopter Network

In a symmetric equilibrium, all potential adopters are determined by the realization of their strategy σ^*. Following the realization of the graph $G(n, p)$, one can imagine realizing n independent Bernoulli random variables with parameter σ^*, one for every individual in the network. The realization of these variables establishes the subgraph of G containing all "potential adopters", which we call the *potential adopter network*.[17] The location of the seed determines who actually adopts. The number of adopters is determined by the size of the component containing the seed in the potential adopter network. Hence, central to our analysis will be the component structure of the potential adopter network, which itself is a random graph given by $G(\sigma n, \sigma p)$.

At this point, we need to introduce the notion of a *giant component* in a random graph. A classic result of Erdős and Rényi (1959) is the phase transition in the size of the largest component in $G(n, \lambda/n)$ around $\lambda = 1$. The result states that as $n \to \infty$,

1. *Sub-critical region* $\lambda < 1$, with high probability the largest component of $G(n, \lambda/n)$ has a size which is at most $\underline{c} \log(n)$ for some constant $\underline{c}$.

2. *Super-critical region* $\lambda > 1$, with high probability the largest component of $G(n, \lambda/n)$ contains a fraction $\overline{c}n$ of the vertices for some constant $\overline{c}$. Moreover, the second largest component has a size which is at most $\underline{c} \log(n)$ for some constant $\underline{c}$.

A component which contains a constant fraction of vertices is called a giant component, so the above result states that the binomial random graph contains a giant component that is unique (with high probability) if and only if $\lambda > 1$. Similarly, the potential adopter network contains a giant component if and only if $\lambda\sigma > 1$.

Our analysis is concerned with the limit of large networks i.e., where $n \to \infty$. In this limit, it is well known that the Erdös-Renyi random graph $G(n, \lambda/n)$ converges locally in probability to a Poisson branching process with mean offspring λ (Van der Hofstad, 2023b, see, e.g., Thm 2.18).[18] In the case $\lambda < 1$, the process stops with probability 1; however, in the case $\lambda > 1$, there is a non-zero probability the process never goes extinct. In this case, there are two potential outcomes of

[17] This is commonly known as the *site-percolation network* in random-graph theory.

[18] We describe Poisson branching processes in more detail in Online Appendix A.1.

following a randomly chosen link to one of its ends in the Erdös-Renyi random graph and finding all subsequent connected nodes. The first outcome is that the link leads to a component with a finite expected number of people with probability ρ_σ (*extinction* probability). The second outcome is that the link leads to an infinite path (as $n \to \infty$) that connects to the giant component with probability $1 - \rho_\sigma$ (*survival* probability).[19] An individual with d connections has d independent realizations of these events. If any one of these connections leads to the giant component, then they themselves are part of it. Hence, the probability this person is in the giant component is $1 - \rho_\sigma^d$. Moreover, conditional on this event (at least one connection leads to the giant component) the distribution of extinction events over the $d-1$ remaining connections is given by $\text{Bin}(d-1, \rho_\sigma)$. In our potential adopter network it is well known that the extinction probability ρ_σ is given by the smallest positive solution to:

$$\rho_\sigma = e^{-\lambda\sigma(1-\rho_\sigma)}. \tag{7}$$

Finally, the expected number of people reached by following a link and conditioning on extinction is $\frac{1}{1-\lambda\sigma\rho_\sigma}$.[20]

3.3 Inference and Calculation of Influence

The total number of adopters in an equilibrium corresponds to the number of people in the same component as the seed in the potential adopter network. The expected influence in equilibrium corresponds to the expected number of individuals that are connected to/disconnected from the seed's component when an individual does versus does not adopt. In a diffusion game, exposure confers information about an individual's component, in particular, the agent is in the same component as the seed.[21] This is informative of an agent's expected influence. In the finite network case this inference is complicated and intractable for the purposes of our analysis. However, a consequence of local convergence of the random graph to a branching process in large networks is that it converges to a particularly simple form.

To construct our measure of influence for an arbitrary agent i, we partition the other agents in i's component into those in the forward components found through each of an agent's links.[22] This construction means that agent i is essential for constructing a path between agents in different elements of the forward component partition. In the limit of a large random graph,

[19] See Van der Hofstad 2023b, Thm 2.28 and its application to Erdös-Renyi random graphs Thm 2.34. For an arbitrary degree distribution in the configuration model see also Thm 4.9 and the discussion thereafter.

[20] See Van der Hofstad 2023a, Thm 3.5 combined with Thm 3.15

[21] The probability an individual is in the same component as the seed is in proportion to the number of other people in the same component as the individual.

[22] Of course, some links may result in the same forward component in which case only a single copy is maintained in the partition; hence, there may be fewer forward components created than an agent has links.

these forward components (and an individual's component) are random objects, as discussed earlier, characterized bya Poisson branching process.

First, consider an equilibrium where $\lambda\sigma^* < 1$ (sub-critical region). Each forward component of an individual is with probability one finite and independent of each other. Hence, the probability an individual with d friends is in the same component as the seed is in proportion to their connectivity. Conditional on exposure, the updated probability of having d friends is $\Pr(d|h_i, \lambda\sigma^* < 1) = \frac{dp_d}{\sum dp_d}$ where $\{p_d\}$ are the prior probabilities. Furthermore, when $\{p_d\}$ is Poisson then it is well known that $\mathbb{E}[d|h_i, \lambda\sigma^* < 1] = \lambda\sigma^* + 1$. Upon exposure an individual forms an expectation over the forward components found via following each one of its links excluding the link through which it was exposed. These components are independent realizations of a Poisson branching process with mean offspring $\lambda\sigma^*$ where the extinction probability $\rho_\sigma = 1$. The expected influence of this individual is $\mathbb{E}[d-1|h_i, \lambda\sigma^* < 1] = \lambda\sigma^*$ times the expected forward component size given by $\frac{1}{1-\lambda\sigma^*}$.[23] and hence expected influence is:

$$\mathcal{I}_\lambda(\sigma^*) = \frac{\lambda\sigma^*}{1-\lambda\sigma^*} \tag{8}$$

when $\lambda\sigma^* < 1$.

Second, consider an equilibrium where $\lambda\sigma^* > 1$ (super-critical region). In the super-critical region, the network of potential adopters contains a unique giant component and upon exposure an agent believes, almost surely, that it is in the giant component.[24] In this case an agent updates its beliefs to reflect that its number of links reflects the distribution of friendships for individuals in the giant component and is given by $\Pr(d = k|h_i, \lambda\sigma^* > 1) = \frac{p_k(1-\rho_\sigma^k)}{1-\rho_\sigma}$.[25] Conditional on this event (at least one connection leads to the giant component) the remaining $d-1$ links are either connected to finite forward components (probability ρ_{σ^*}) or connected to a forward component containing everyone in the giant component other than the people contained in i's finite forward components (probability $1-\rho_{\sigma^*}$). We call this forward component the giant forward component. Therefore, i's partition of the giant component consists of the giant forward component and the number of successes of $\text{Bin}(d-1, \rho_\sigma)$ independent realizations of finite forward components. The probability of the seed being in each of these forward components is of course dominated $\Pr \to 1$ by the giant forward component and vanishing $\Pr \to 0$ for the finite components. The calculation of expected influence when the seed is in the giant forward component is simply

[23] Exposure is informative of the size of component found via following the exposure link but, in the sub-critical region, these beliefs do not affect the calculation of expected influence $\mathcal{I}_\lambda(\sigma)$.

[24] The giant component contains a positive fraction of the population as $n \to \infty$ whereas the expected size of all other components $\to 0$. Hence, the probability of being in the same component as the seed is vanishing in the limit $n \to \infty$ for all components other than the giant component.

[25] This is termed "viral" inference in Sadler (2020).

the expected number of people in the remaining finite forward components of the partition, in expectation this is $\frac{\lambda\sigma\rho_\sigma}{1-\lambda\sigma\rho_\sigma} - \rho_\sigma$. Now, despite the vanishing probability of the seed being in one of the finite forward components, these events contribute non-negligibly to expected influence because i is essential for connecting the seed to the giant forward component. The contribution to expected influence of this event is given by the product of the probability of the seed being in one of the finite forward components and the expected influence in this event. If the giant component contains a fraction c of the vertices, then the probability that the seed is in one of the finite forward components is $\frac{\frac{\lambda\sigma\rho_\sigma}{1-\lambda\sigma\rho_\sigma}-\rho_\sigma}{cn}$ (in the limit of a large network), which goes to 0, while the expected size of influence diverges to cn. Hence, the contribution to influence is the product of these and is equal to $\frac{\lambda\sigma\rho_\sigma}{1-\lambda\sigma\rho_\sigma} - \rho_\sigma$.

We summarize our calculation of influence for both cases with the following lemma.

Lemma 1. *The expected influence function $\mathcal{I}_\lambda \colon [0,1] \to \overline{\mathbb{R}}$ is given by*

$$\mathcal{I}_\lambda(\sigma) = \begin{cases} \frac{\lambda\sigma}{1-\lambda\sigma} & \text{if } \lambda\sigma < 1, \\ 2(\frac{\lambda\sigma\rho_\sigma}{1-\lambda\sigma\rho_\sigma} - \rho_\sigma) & \text{if } \lambda\sigma > 1 \end{cases} \tag{9}$$

where ρ_σ is as in (7). In particular, $\mathcal{I}_\lambda$ is continuous.[26]

We further characterize how influence changes with agents' adoption probability σ and network density λ in the following proposition.

Proposition 2. *Let $\sigma \in [0,1]$ be any strategy.*

1. *If σ is nonviral, then $\mathcal{I}_\lambda(\sigma)$ is strictly increasing in σ and strictly increasing in λ.*

2. *If σ is viral, then $\mathcal{I}_\lambda(\sigma)$ is strictly decreasing in σ and strictly decreasing in λ.*

Figure 3 captures the main idea behind Proposition 2, which is a key comparative static result for our analysis.

We see that in the sub-critical region influence diverges to $+\infty$ at the point where the giant component emerges. Beyond this point it is decreasing as the giant component grows. This illustrates one of the consequences of Proposition 2 that adoption is a strategic complement when the graph is sparse, and a strategic substitute when it is dense. Much of the literature on network effects focuses on adoption as a strategic complement, so it is a novel feature of our model that it encompasses both complements and substitutes as a function of the graph and

[26] As $\lambda\sigma \to 1$ the expected influence diverges to $+\infty$ from the left and right and therefore is continuous on the extended real line.

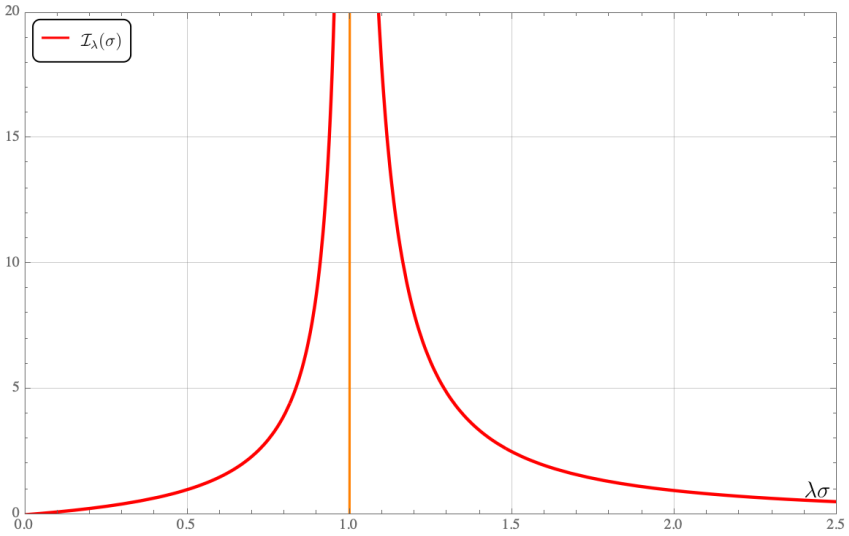

Figure 3: Expected influence is increasing when the network is sparse ($\lambda\sigma < 1$) and decreasing when the network is dense ($\lambda\sigma > 1$).

the strategies.

4 Results

We analyze the behavior of expected influence as a function of the network density and proceed to fully characterize symmetric equilibria for public goods and public bads. Furthermore, we will be interested in the structure of the potential adopter network induced by an equilibrium strategy. We call a strategy *viral* if it induces a giant component in the potential adopter network, and *nonviral* otherwise. In our setting, there are four types of symmetric equilibria:

1. No-adoption ($\sigma^* = 0$)
2. Viral or Nonviral Full-adoption ($\sigma^* = 1$)[27]
3. Nonviral-mix ($\sigma^* \in (0,1),\ \lambda\sigma^* < 1$)
4. Viral-mix ($\sigma^* \in (0,1),\ \lambda\sigma^* > 1$)

All four types of equilibria can arise depending on the structure of the underlying graph (as determined by λ). Clearly there can not be a viral equilibrium when the underlying graph does not contain a giant component ($\lambda < 1$). We will show that there exists a critical value of $\lambda_\pi^{\text{crit}}$ that separates regions where viral equilibria exist ($\lambda > \lambda_\pi^{crit}$) from those where they do not ($\lambda > \lambda_\pi^{crit}$). There is a stark contrast between the transition from nonviral to viral equilibria

[27]When $\sigma = 1$, the potential adopter network is all of $G(n,p)$. Hence the full adoption equilibrium is viral if and only if $\lambda > 1$.

around this threshold for public goods when compared to public bads. To this end, define $\frac{C_{1,\pi}(\lambda)}{n}$ as the fraction of agents in the largest component of the potential adopter network $G_\pi(\bar{\sigma}^*_\pi(\lambda))$, where $\pi \in \{g, b\}$ and $\bar{\sigma}^*_\pi(\lambda)$ is the largest equilibrium. We will be interested in whether the fraction of agents who adopt in the largest equilibrium goes to zero or a limit bounded away from zero as the density approaches the critical density from above i.e. $\lim_{\lambda \to \lambda_\pi^{crit+}} \frac{C_{1,\pi}(\lambda)}{n}$.

4.1 Public Goods

Initially we focus on the public good case $\pi = g$. When other players play according to σ, expected utility from adoption is given by $(v - c) + v\mathcal{I}_\lambda(\sigma)$. Hence, the best response function takes a particular simple form characterized by a threshold level of influence:

$$BR(\sigma) = \begin{cases} 1, & \mathcal{I}_\lambda(\sigma) > \frac{c}{v} - 1 \\ [0,1], & \mathcal{I}_\lambda(\sigma) = \frac{c}{v} - 1 \\ 0, & \mathcal{I}_\lambda(\sigma) < \frac{c}{v} - 1. \end{cases} \tag{10}$$

We characterize the structure of equilibria as a function of λ. We label each case in the proposition below with the "type" of the largest equilibrium—no-adoption, full-adoption, nonviral-mix, or viral-mix.

Proposition 3 (Characterization of equilibria for public goods)**.** *Let $\underline{\lambda} = 1 - \frac{v}{c} \in (0, 1)$, and let $\overline{\lambda}$ be the unique solution in $(1, \infty]$ to $\overline{\lambda}\rho_1 = 1 - \frac{2v}{c + v(1+2\rho_1)}$. Then*

1. *(No-adoption) If $\lambda < \underline{\lambda}$, the unique equilibrium is $\sigma = 0$.*
2. *(Nonviral Full-adoption) If $\underline{\lambda} < \lambda < 1$, there is a no-adoption equilibrium $\sigma = 0$, a nonviral-mixed equilibrium $\lambda\sigma = \underline{\lambda}$, and a nonviral full-adoption equilibrium $\sigma = 1$.*
3. *(Viral Full-adoption) If $1 < \lambda < \overline{\lambda}$, there is a no-adoption equilibrium $\sigma = 0$, a nonviral-mixed equilibrium defined by $\lambda\sigma = \underline{\lambda}$ and a viral full-adoption equilibrium $\sigma = 1$.*
4. *(Viral-mix) If $\lambda > \overline{\lambda}$, there is a no-adoption equilibrium $\sigma = 0$, a nonviral-mixed equilibrium $\lambda\sigma = \underline{\lambda}$ and a viral-mixed equilibrium defined by $\lambda\sigma = \overline{\lambda}$.*

The proposition may be understood through the relationship between expected influence, the threshold $\frac{c}{v} - 1$ in the best response function and the strategic complementarity/substitability of adoption in the sub/super critical regions. When the potential adopter network is sparse (small λ or low levels of adoption σ) then expected influence is small ($< \frac{c}{v} - 1$) and the best response function is zero. Hence, there is always a no-adoption equilibrium for any λ and this

is the unique equilibrium in sufficiently sparse networks (part 1). In the sub-critical region ($\lambda < 1$) adoption is a strategic complement and as the network becomes more dense, expected influence increases and for sufficiently high levels of adoption exceeds the threshold $\frac{c}{v} - 1$ in the best response function (adoption is a best response to sufficiently high adoption). This introduces two more equilibria, a non-viral full adoption equilibrium and a non-viral mixed equilibrium (part 2). As the density is increased further, the underlying network passes into the super-critical region and expected influence remains large (diverging to ∞ around the critical density). Initially, in the super-critical region, the 3 equilibria from part 2 continue to exist. However, the full adoption equilibrium is now viral because the underlying network contains a giant component (part 3). In the super-critical region, expected influence is decreasing in adoption and density so for sufficiently dense networks and high adoption expected influence falls below the threshold $\frac{c}{v} - 1$ in the best response function. Once this occurs full adoption can not be supported as an equilibrium (the best response to everyone else adopting is to not adopt). Rather, the largest equilibrium becomes a viral-mixed strategy equilibrium where the level of mixing maintains the expected influence at the threshold $\frac{c}{v} - 1$ that permits mixing in the best response function.

A consequence of the structure of equilibria for public goods is that one of two forces will always limit a public good's capacity to diffuse through a population. First, full adoption is only a viral equilibrium if the network is not *too* dense ($\lambda < \bar{\lambda}$), hence, although everyone is prepared to adopt, any viral diffusion is limited by the fraction of people in the giant component which is bounded away from 1 in the case $\lambda < \bar{\lambda}$. Second, in more dense networks ($\lambda > \bar{\lambda}$) it is strategic considerations that limit the ability of a public good to diffuse. This occurs because in these dense networks the presence of many alternative paths between individuals and the seed (global effect of density) means there is a strong incentive to free-ride on others' adoption. Hence, full adoption can not be sustained as an equilibrium limiting the fraction of the population that is willing to adopt and the connectedness of the potential adopter network.

We depict the relationship between influence and the best response in Figure 4. For exposition, we fix $\frac{v}{c} = 0.5$ in the figure. The top part of the figure shows expected influence in red and the threshold $1 - \frac{c}{v}$ in blue, around which the best response changes. The bottom part of the figure shows the best response correspondence in red, and the 45-degree line in black—the intersections of the red and black lines are equilibria. We fix $\lambda = 2.5$ so that the figure shows the entire range of possibilities for influence as a function of σ, the vertical orange line is the critical threshold where $\lambda\sigma = 1$, i.e. $\sigma = 0.4$. We also have $\underline{\lambda} = 0.5$, and $\overline{\lambda} \approx 1.97$, and these can be obtained by looking at the values of σ for which the expected influence (red) crosses the threshold $1 - \frac{c}{v}$ (blue). For example, the red and blue lines first intersect at $\sigma = 0.2$, so $\lambda\sigma = \underline{\lambda}$ implies that $\underline{\lambda} = 2.5 \times 0.2 = 0.5$, as claimed. Finally, because influence is small on the tails and large in the

middle, best responses are 0 for small or large σ, and 1 for intermediate values.

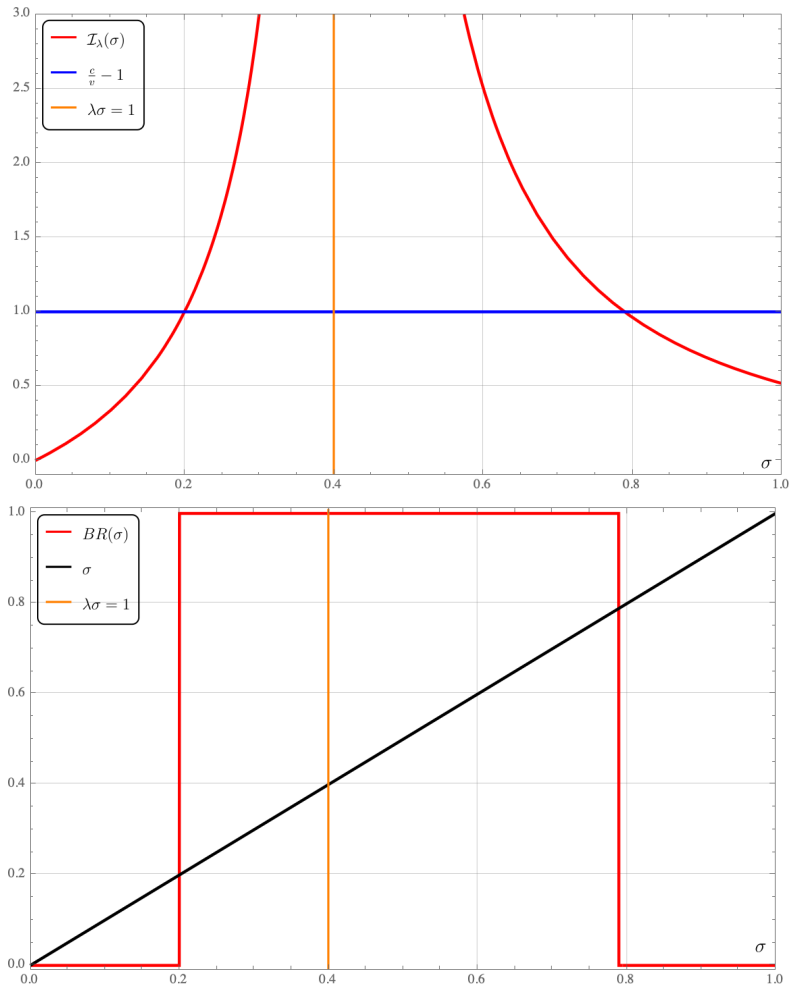

Figure 4: Expected influence and the best response correspondence for public goods. Agents adopt only when expected influence is sufficiently large. Everyone adopts around the critical threshold where $\lambda\sigma = 1$. Maximal influence depends on λ, i.e. if $\lambda < \underline{\lambda}$ then expected influence is always below the blue line $\underline{\lambda}$ and no one ever adopts.

To conclude this section, we translate cases 2 and 3 of Proposition 3 into a statement about the size of diffusion. Observe that $\lambda = 1$ is a critical threshold such that no viral equilibria exist when $\lambda < 1$ and viral equilibria always exist when $\lambda > 1$. A similar threshold exists for public bads, so to unify our notation, we write $\lambda_\pi^{\text{crit}}$, for the critical threshold around which viral equilibria emerge (for $\pi \in \{g, b\}$). With this notation, $\lambda_g^{\text{crit}} = 1$. Finally, recall that we use $C_{1,\pi}(\lambda)$ to denote the fraction of agents in the largest component of the potential adopter network for $\pi \in \{g, b\}$ in the largest equilibrium. Corollary 1 is best understood by looking at Figure 6.

Corollary 1. *When $\pi = g$, there is a continuous phase transition in adoption under the largest*

equilibrium, that is,

$$\lim_{\lambda \to \lambda_g^{crit-}} \frac{C_{1,g}(\lambda)}{n} = 0 = \lim_{\lambda \to \lambda_g^{crit+}} \frac{C_{1,g}(\lambda)}{n}.$$

The upshot of Proposition 3 and Corollary 1 is that around the phase transition in the underlying graph (i.e., $\lambda \in (1-\epsilon, 1+\epsilon)$), there is a parallel phase transition in equilibrium behavior: for $\lambda < 1$, all equilibria are necessarily nonviral, while for $\lambda > 1$ there exists an equilibrium which induces a giant component of potential adopters. Moreover, this transition admits a *continuous* change in the size of the largest component in the potential adopter network with respect to λ. We now contrast this with the case of the diffusion of public bads.

4.2 Public Bads

We now consider the public bad case $\pi = b$. Recall that when other players play according to σ, expected utility from adoption is given by $\mathbb{E}_{\mathcal{D}}[u_b(\sigma)] = -(v-c) - v\mathcal{I}_\lambda(\sigma)$. The best response function again takes a particular simple form characterized by a threshold level of influence:

$$BR(\sigma) = \begin{cases} 1, & \mathcal{I}_\lambda(\sigma) < \frac{c}{v} - 1 \\ [0,1], & \mathcal{I}_\lambda(\sigma) = \frac{c}{v} - 1 \\ 0, & \mathcal{I}_\lambda(\sigma) > \frac{c}{v} - 1. \end{cases} \tag{11}$$

Note that although the utility under public bads is simply the negative of the public good utility and this inverts the best response function, it will not be the case that this results in the public bad equilibria being an inversion of the public good equilibria for each λ. Recall that the thresholds for λ, which we defined for public goods, are equal to

$$\underline{\lambda} = 1 - \frac{v}{c}, \quad \overline{\lambda}\rho_1 = 1 - \frac{2v}{c + v(1+2\rho_1)}$$

We describe the structure of equilibria in the following proposition.

Proposition 4 (Characterization of equilibria for public bads)**.** *Under the diffusion of public bads:*

1. *(Non-Viral Full adoption) When $\lambda < \underline{\lambda}$, the unique equilibrium is $\sigma = 1$.*
2. *(Nonviral-mix) When $\underline{\lambda} < \lambda < \overline{\lambda}$, there is a unique equilibrium determined by the nonviral mixed strategy defined by $\lambda\sigma = \underline{\lambda}$.*
3. *(Viral Full adoption) When $\lambda > \overline{\lambda}$, there is a viral full adoption equilibrium $\sigma = 1$ along*

with a nonviral-mixed equilibrium $\lambda\sigma = \underline{\lambda}$, and a viral-mixed equilibrium $\lambda\sigma = \overline{\lambda}$.

Similarly to earlier Proposition 4 may be understood through the relationship between expected influence, the threshold $\frac{c}{v} - 1$ in the best response function and the strategic complementarity/substitutability of adoption in the sub/super critical regions. When the potential adopter network is sparse (small λ or low levels of adoption σ) then expected influence is small ($< \frac{c}{v} - 1$) and the best response function is to adopt. Hence, full adoption is a unique non-viral equilibrium in sufficiently sparse networks (part 1). In the sub-critical region ($\lambda < 1$) adoption is a strategic substitute and as the network becomes more dense, expected influence increases such that for sufficiently high levels of adoption it exceeds the threshold $\frac{c}{v} - 1$ in the best response function and full adoption is no longer sustainable as an equilibrium (non-adoption is a best response to sufficiently high adoption). Rather, a non-viral mixed equilibrium emerges as the unique equilibrium (part 2). As the density is increased further, the underlying network passes into the super-critical region and expected influence remains large (diverging to ∞ around the critical density). This means the best response is non-adoption to high levels of adoption (σ close to 1) and the unique equilibrium remains the non-viral mixed strategy equilibrium. Hence, unlike the public good case, as the underlying graph moves into the super-critical region no viral equilibrium emerges. In the super-critical region, expected influence is decreasing in adoption and density so for sufficiently dense networks ($\lambda > \bar{\lambda}$) and high adoption expected influence falls below the threshold $\frac{c}{v} - 1$ in the best response function. Once this occurs both a viral mixed and full adoption equilibria emerge (part 4).

A consequence of Proposition 4 is that viral equilibria exist for public bads if and only if the network is sufficiently dense, that is, if and only if λ is large enough. In the case of public bads, the global effect of density encourages adoption and sustains full adoption in sufficiently dense networks. We interpret this as a kind of "mob mentality"—a situation in which a large number of agents are willing to make a bad decision because none of them feel accountable for influencing others to follow suit. However, once a viral full adoption equilibrium emerges it remains in all more dense networks.

As with Figure 4, we depict the relationship between the expected influence and the best response correspondence for public bads. Note that because the utility of public bads is the negative of the utility of public goods, the plot is identical to Figure 4 but with the best response correspondence inverted. We see that viral equilibria do not emerge smoothly alongside the phase transition in the underlying graph because when λ is close to the critical threshold, influence becomes too large for agents to be willing to adopt.

Proposition 4 implies that there is a discontinuous jump in adoption under the largest equilibrium when there is diffusion of a public bad. This is because a viral equilibrium emerges only

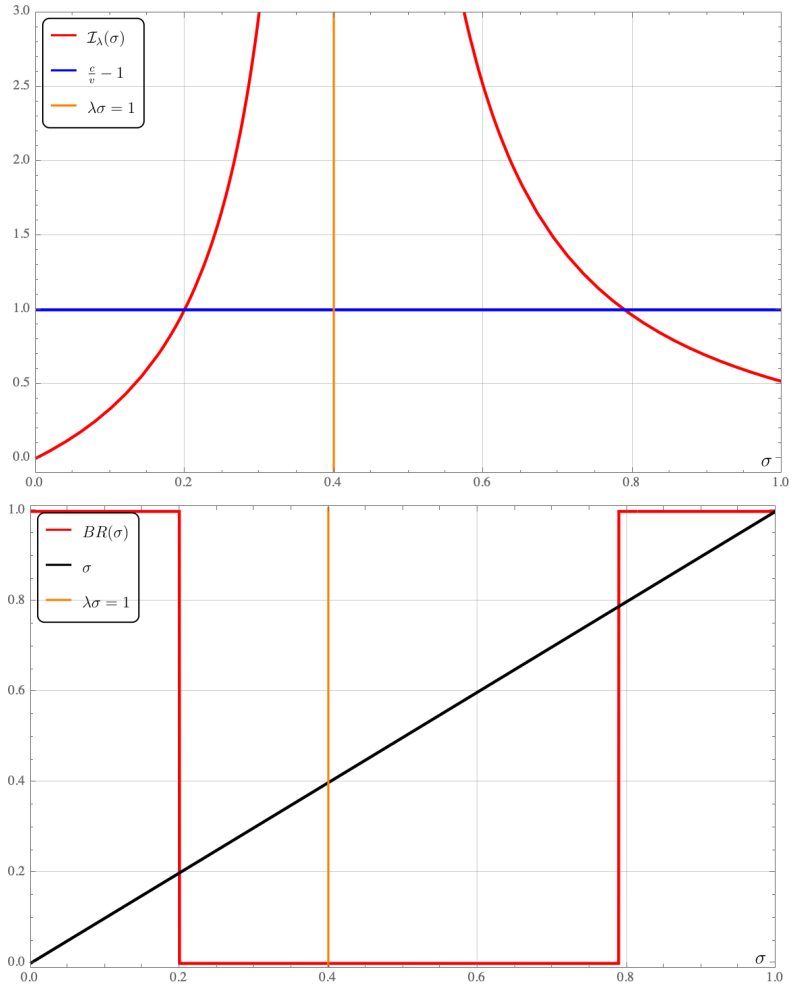

Figure 5: Expected influence and the best response correspondence for public bads. Agents adopt only when expected influence is sufficiently small. No one adopts around the critical threshold where $\lambda\sigma = 1$.

when the giant component is already well-established in the graph—that is, the critical threshold is $\lambda_b^{\text{crit}} = \overline{\lambda} > 1$. In terms of component sizes, we have the following corollary analogous to Corollary 1. Again, the Corollary is best understood by looking at Figure 6.

Corollary 2. *When $\pi = b$, there is a discontinuous phase transition in adoption under the largest equilibrium, that is,*

$$\lim_{\lambda \to \lambda_b^{crit-}} \frac{C_{1,b}(\lambda)}{n} = 0 < \lim_{\lambda \to \lambda_b^{crit+}} \frac{C_{1,b}(\lambda)}{n}.$$

We now provide some comparative statics for our model.

4.3 Comparative Statics

A key object of interest to us is the extent of diffusion in the potential adopter network in an equilibrium, particularly, in the largest equilibrium. If there is a viral equilibrium σ, then the size of a large cascade is given by $\sigma(1-\rho_\sigma)$.

Consider the diffusion of public goods. By Proposition 3, we know that when $1 < \lambda < \overline{\lambda}$, the unique viral equilibrium is the full-adoption equilibrium. It follows that the size of a large cascade in this equilibrium is simply $1-\rho_1$, which is strictly increasing in λ since extinction becomes less likely as the network gets more connected. However, once $\lambda > \overline{\lambda}$, full-adoption is no longer an equilibrium and the unique equilibrium is the mixed strategy equilibrium σ^* determined by

$$\lambda\sigma^*\rho_{\sigma^*} = 1 - \frac{2v}{c + v(1+\rho_{\sigma^*})}. \tag{12}$$

We show in Appendix A (see the proof of Proposition 5) that ρ_{σ^*} is constant for all $\lambda > \overline{\lambda}$. That is, as an equilibrium object, the fraction of potential adopters in the giant component in the potential adopter network is constant whenever the underlying network is sufficiently connected. As a consequence, the equilibrium strategy σ^* must decrease to retain indifference between adoption and non-adoption, and so the overall diffusion $\sigma^*(1-\rho_{\sigma^*})$ is decreasing for $\lambda > \overline{\lambda}$.

Now, consider the diffusion of public bads. Proposition 4 tells us that when $\lambda < \overline{\lambda}$, there are no viral equilibria, while for $\lambda > \overline{\lambda}$, there is a full-adoption equilibrium. Hence by the same reasoning as in with public goods, the size of a large cascade is increasing in λ.

We plot size of a large cascade of public goods and public bads in the largest equilibrium in Figure 6, and formalize the figure in Proposition 5.

Proposition 5. *The dynamics of the size of a large cascade in the largest equilibrium is given in the table below, where* 0 *indicates that there are no large cascades.*

λ	*public goods*	*public bads*
$\lambda < 1$	*0*	*0*
$1 < \lambda < \overline{\lambda}$	*strictly increasing*	*0*
$\lambda > \overline{\lambda}$	*strictly decreasing*	*strictly increasing*

Proposition 5 is our key comparative statics result. It illustrates a *bystander effect* for public goods and a *mob-rule effect* for public bads when $\lambda > \overline{\lambda}$. These effects can be understood through the two effects of connectivity, the direct effect (holding agent's strategies fixed) and the indirect effect through the equilibrium adjustment of agent's strategies. For both public

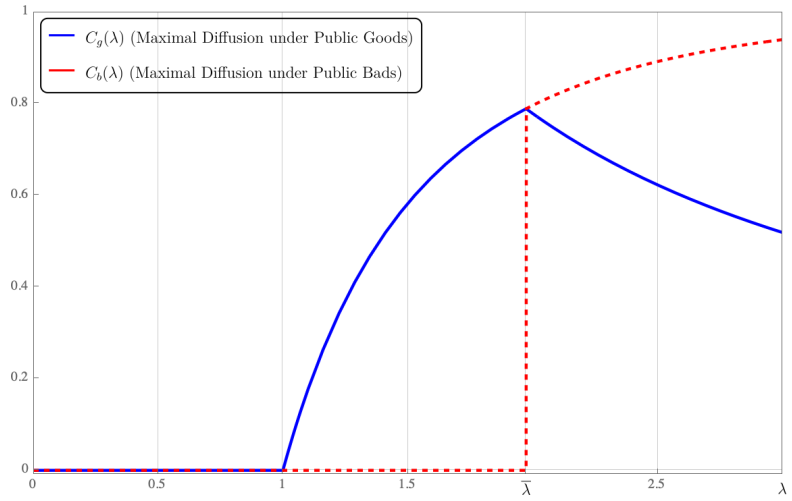

Figure 6: Size of a Large Cascade: Public Goods vs. Public Bads. There is a discontinuous jump in maximal diffusion under public bads at $\overline{\lambda}$.

goods and public bads the direct effect increases the *maximal* size of the diffusion because the giant component becomes larger in the potential adopter network holding agent's adoption decisions fixed. Also for both public goods and public bads connectivity reduces influence when $\lambda > \overline{\lambda}$. For public goods, an agent's equilibrium probability of adoption decreases and this effect dominates the former resulting in a smaller *maximal* diffusion. In contrast for public bads, agent's equilibrium strategy is to always adopt and this continues to be the case for greater connectivity. Hence, the total effect is given by the direct effect leading the *maximal* diffusion to increase. For completeness, we provide comparative statics for diffusion with respect to v, c and their ratio $\frac{v}{c}$ below.

Proposition 6. *When $\lambda > 1$, diffusion of public goods (bads) is weakly increasing (weakly decreasing) in v, weakly decreasing (weakly increasing) in c and weakly increasing (weakly decreasing) in v/c.*

Intuitively the only change that v and c can have on diffusion is through $\overline{\lambda}$, which is a function of both of these parameters. If $\lambda < \overline{\lambda}$, a change in v and/or c can cause $\lambda > \overline{\lambda}$, at which point by Proposition 3 the structure of equilibrium (and hence diffusion) changes. We now proceed to discuss the welfare implications of our comparative static results.

4.3.1 Welfare

Consider a planner who controls the size of the network (λ) and ex-ante knows that a behavior is going to diffuse on the network, but does not know whether that behavior will be a public good or a public bad. Let $\zeta_\lambda \in [0, 1]$ be the ex-ante expected diffusion in the largest equilibrium

as a function of λ.[28] Let $v\colon \{g,b\}\times[0,1]\to\mathbb{R}$ denote the planner's (indirect) utility function for a given choice of λ, so that $v = v(\pi,\zeta_\lambda)$. Call a planner's utility function *in the public interest* if it is monotone increasing in ζ_λ when $\pi = g$, and monotone decreasing in ζ_λ when $\pi = b$. This captures the idea that the diffusion of public goods is socially desirable while the diffusion of public bads is not. Then, Proposition 5 implies that for any utility function that is in the public interest, the socially optimal size of the network is achieved by choosing λ to be just below $\overline{\lambda}$.

Proposition 7. *A planner can achieve arbitrarily close to the maximum welfare under public goods and public bads when $\overline{\lambda} - \epsilon$ for $\epsilon > 0$ sufficiently small.*

Proposition 7 says that a planner wants to make the network large but not too large, lest they risk the emergence of a viral equilibrium in public bads. A converse implication is that if a network is large ($\overline{\lambda} > \lambda$), then the diffusion of a public bad can be eliminated by lowering the connectivity below the critical threshold $\overline{\lambda}$. For example, suppose a society needs a certain number of active voters to maintain a representative democracy, but that the network is sufficiently connected so that non-engagement in the democratic process admits a viral equilibrium. Then, the planner only has to decrease the connectivity of the network by an amount which makes agents "accountable enough" for their action in order to eliminate the behavior entirely. Finally, Proposition 7 implies that the losses on the adoption of public good are small relative to the gains on adoption of public bads near $\overline{\lambda}$, so a planner who is uncertain about the exact threshold is better off "cutting more conservatively". This of course should not be taken too literally, but qualitatively the lesson remains.

In the next section, we will consider a number of extensions to the Poisson model.

5 Extensions

5.1 Arbitrary Degree Distributions

When it comes to real-world networks, there are many properties that the Binomial random graph $G(n,p)$ does not capture, for example "fat tails".[29] As such, it is natural to ask whether our results can be extended beyond the specialised class of Poisson networks. The object of this section is to answer this in the affirmative.

We can easily extend our model to a game on a network with an arbitrary degree distribution constructed using the *configuration model* (Bollobás, 1980; Wormald, 1978). In this setting, so

[28] The analysis is virtually identical if the planner holds a probability distribution over the equilibrium which will be played and puts nonzero probability on the largest equilibrium.

[29] See e.g., Jackson (2008).

long as the degree distribution is sufficiently sparse, the graph is locally tree-like and therefore can be approximated using branching processes (we discuss this in more detail in Online Appendix B).

An approach using generating functions works in this more general setting. Formally, define a game with n agents as a three-tuple $\Gamma^{(n)} = (\mathbf{d}^{(n)}, \mathcal{A}, u)$. The vector $\mathbf{d}^{(n)} = (\mathbf{d}_1^{(n)}, \ldots, \mathbf{d}_n^{(n)})$ is the *degree sequence* for the game and generalizes the binomial random graph of Section 2.[30] The actions and utility functions remain unchanged. We impose standard restrictions on the degree sequence so that the limiting degree distribution, which we denote by $\mathcal{D} \equiv \lim_{n\to\infty} \mathbf{d}^{(n)}$, is well behaved (in particular it must have a finite mean—see Online Appendix C for details).

The timing, payoffs and strategies in the more general model remain unchanged except that expectations will be different because they depend on the degree distribution. Let

$$G_0(z) = \sum_{k=0}^{\infty} p_k z^k$$

be the generating function for $\mathcal{D} = \{p_k\}_{k\geq 0}$. Then, the generating function for $\mathcal{D}'$ is given by

$$G_1(z) = \frac{G_0'(z)}{G_0'(1)},$$

where $G_0'(1) = \mathbb{E}(\mathcal{D})$. If all agents play the strategy σ, then the generating function for the forward adoption degree distribution is

$$G_1(1 - \sigma + \sigma z). \tag{13}$$

It follows that the *forward extinction probability* ρ_σ for the generating function (13) is the smallest solution in $[0, 1]$ to the equation

$$\rho_\sigma = G_1(1 - \sigma + \sigma\rho_\sigma). \tag{14}$$

Note that this implies

$$\frac{d}{dz} G_1(1 - \sigma + \sigma z)\big|_{z=\rho_\sigma} = \sigma G_1'(1 - \sigma + \sigma\rho_\sigma). \tag{15}$$

Equation (15) describes the expected offspring in the subcritical dual branching process with offspring distribution $\mathcal{D}$. In the Poisson model, the critical site percolation threshold was given by $\lambda\sigma^{\text{crit}} = 1$, i.e., $\sigma^{\text{crit}} = \lambda^{-1}$. In general, the critical site percolation threshold for a graph with

[30] This can be viewed as a special case of Sadler's single-type diffusion game.

an arbitrary degree distribution is known to be

$$\sigma^{\text{crit}} \equiv \frac{\mathbb{E}[\mathcal{D}]}{\mathbb{E}[\mathcal{D}(\mathcal{D}-1)]}.$$

We now give an explicit formula for the expected influence function. We have the following analogue of Lemma 1.

Lemma 2. *The expected influence function* $\mathcal{I}_{\mathcal{D}}\colon [0,1] \to \overline{\mathbb{R}}$ *is given by*

$$\mathcal{I}_{\mathcal{D}}(\sigma) = \begin{cases} \frac{\sigma G_1'(1)}{1-\sigma G_1'(1)}, & \sigma < \sigma^{crit} \\ \frac{2\rho_\sigma(\widehat{G}_1'(1)-1)}{1-\sigma G_1'(1-\sigma+\sigma\rho_\sigma)}, & \sigma > \sigma^{crit} \end{cases} \tag{16}$$

where ρ_σ *is as in (14), and*

$$\widehat{G}_1'(1) \equiv \frac{\sum_{k=1}^{\infty} k(1-\rho_\sigma^k)\frac{\sigma^k}{k!}G_0^{(k)}(1-\sigma)}{1-G_0(1-\sigma+\sigma\rho_\sigma)}$$

is the expected number of neighbors who are potential adopters after updating due to viral inference. In particular, $\mathcal{I}_{\mathcal{D}}(\sigma)$ *is continuous.*

It turns out to be difficult in general to determine the behavior of (16) with respect to σ, despite there being a "discrete duality principle" (Molloy and Reed, 1998) analogous to the Poisson case.[31] However, we can show that several cases behave exactly as the Poisson model.

Our results on the characterization of equilibria in Section 2 relied on a continuous parameterization of the underlying density of the graph (λ). The reason for this is that λ is a sufficient statistic for σ^{crit} and a giant component exists in the graph if and only if $\sigma^{\text{crit}} < 1$.

Many distributions of interest can be parameterized in a similar way, for example any degree distribution that scales exponentially in the degree (e.g., a power law), or any family of mixed Poisson distributions. As such, we now restrict our attention to one-parameter families of degree distributions $\{\mathcal{D}_\theta\}_{\theta\in\Theta}$, for which $\theta \in \Theta \subseteq [0,\infty)$ is a sufficient statistic for $\frac{\mathbb{E}[\mathcal{D}_\theta]}{\mathbb{E}[\mathcal{D}_\theta(\mathcal{D}_\theta-1)]}$. We focus on distributions for which $\frac{\mathbb{E}[\mathcal{D}_\theta]}{\mathbb{E}[\mathcal{D}_\theta(\mathcal{D}_\theta-1)]}$ is strictly decreasing in θ, with $\inf_{\theta\in\Theta}\frac{\mathbb{E}[\mathcal{D}_\theta]}{\mathbb{E}[\mathcal{D}_\theta(\mathcal{D}_\theta-1)]} < 1$. This guarantees that there exists a critical threshold θ^c such that a giant component exists in the graph with degree distribution $\mathcal{D}_\theta$ if and only if $\theta > \theta^c$.

In this setting, we can offer a characterization of equilibria in public goods and public bads that

[31] We can characterize exactly what happens in the subcritical regime and around the critical threshold, but as the graph moves further into the supercritical regime we cannot say exactly what happens to influence for an arbitrary degree distribution. We discuss this more in Section 5.3.

is entirely analogous to Propositions 3 and 4, if the following condition is satisfied.[32]

Condition 1. Let $\sigma \in [0,1]$ be any strategy.

(i) If σ is nonviral, then $\mathcal{I}_\theta(\sigma)$ is strictly increasing in σ and strictly increasing in θ.

(ii) If σ is viral, then $\mathcal{I}_\theta(\sigma)$ is strictly decreasing in σ and strictly decreasing in θ.

Condition 1 is identical to Proposition 2 but with λ replaced by the variable θ, which parameterizes the family of distributions. How restrictive is Condition 1? It is not difficult to show that (i) *always* holds, and that (ii) holds around the phase transition (i.e., when $\sigma \approx \sigma^{\text{crit}}$). Whether (ii) always holds away from the phase transition we do not know, however it does hold in all of the examples we provide in the next section.[33]

We write $\underline{\theta}, \overline{\theta}$ for the solutions to

$$\frac{c}{v} - 1 = \mathcal{I}_\theta(1),$$

where $\underline{\theta} < \overline{\theta}$.[34] For any family of distributions satisfying Condition 1, we have the following analogue of Proposition 3.

Proposition 8 (Characterization of equilibria for public goods)**.** *Let $\{\mathcal{D}_\theta\}_{\theta\in\Theta}$ be a family of degree distributions satisfying Condition 1. Then*

1. *(No-adoption) If $\theta < \underline{\theta}$, there is a unique equilibrium $\sigma = 0$.*
2. *(Nonviral Full-adoption) If $\underline{\theta} < \theta < 1$, there is a no-adoption equilibrium $\sigma = 0$, a nonviral-mixed equilibrium $\sigma = \frac{1}{G_1'(1)}\left(1 - \frac{v}{c}\right)$, and a nonviral full-adoption equilibrium $\sigma = 1$.*
3. *(Viral Full-adoption) If $1 < \theta < \overline{\theta}$, there is a no-adoption equilibrium $\sigma = 0$, a nonviral-mixed equilibrium $\sigma = \frac{1}{G_1'(1)}\left(1 - \frac{v}{c}\right)$ and a viral full-adoption equilibrium $\sigma = 1$.*
4. *(Viral-mix) If $\theta > \overline{\theta}$, there is a no-adoption equilibrium $\sigma = 0$, a nonviral-mixed equilibrium $\sigma = \frac{1}{G_1'(1)}\left(1 - \frac{v}{c}\right)$, and a viral-mixed equilibrium defined by the largest solution to $\mathcal{I}_\theta(\sigma) = \frac{c}{v} - 1$.*

The characterization of equilibria for public bads is analogous and therefore omitted here. It is an immediate consequence of our analysis that the comparative statics and welfare analysis provided in Section 4.3 also apply here. In particular our key insight in the Poisson model—that a benevolent social planner wants to make the network "just large enough"—extends to graphs

[32] We can weaken this condition slightly as we discuss in the Online Appendix.

[33] Whether there exists a general class of distributions for which the condition holds is an open question.

[34] It is straightforward to show that $\underline{\theta}$ always exists. Whether $\overline{\theta}$ exists depends on c and v. If $\overline{\theta}$ does not exist then we set it equal to $+\infty$.

with an arbitrary degree distribution that can be parameterized in the way we have described.

5.2 Examples

5.2.1 Zipf Distribution

Consider a configuration model network with degree distribution

$$p_k = \frac{e^{\alpha \underline{k}}}{\Phi(e^{-\alpha}, 1, \underline{k})} \frac{e^{-\alpha k}}{k}, \tag{17}$$

where $\Phi(z, s, k)$ is the Lerch transcendent function and $\underline{k}$ is the smallest degree that occurs with nonzero probability. Equation (17) defines the so-called *Zipf* distribution. The Zipf distribution exhibits a power-law of the form $k^{-\gamma}$ with $\gamma = 1$, and an exponential tail controlled by the parameter $\alpha > 0$.[35] We focus on the case where $\underline{k} = 1$ since this is analytically tractable. In this case it's straightforward to show that

$$G_1(z) = \frac{1 - e^{-\alpha}}{1 - e^{-\alpha} z}.$$

Thus, the extinction probability $\rho_\sigma = G_1(1 - \sigma + \sigma\rho_\sigma)$ is given by

$$\rho_\sigma = \frac{1}{\theta\sigma}, \tag{18}$$

where we define $\theta = (e^\alpha - 1)^{-1}$. It follows that the critical threshold θ^{crit} for the emergence of the giant component is simply $\theta = 1$ (or equivalently, $\alpha = \ln(2)$) so that an equilibrium is viral iff $\theta\sigma > 1$.[36]

We can find an explicit expression for the expected influence in terms of θ. Expected influence is given by (see Online Appendix F.2.1 for details)

$$\mathcal{I}_\theta(\sigma) = \begin{cases} \frac{\theta\sigma}{1-\theta\sigma}, & \theta\sigma < 1 \\ \frac{2(\theta\sigma+1)}{\theta\sigma \ln(\theta\sigma)} - \frac{2}{\theta\sigma - 1}, & \theta\sigma > 1. \end{cases}$$

where we have written $\mathcal{I}_\theta$ to emphasize that θ continuously parameterizes the family of Zipf distributions just as λ does for the Poisson distribution. The parameterization θ also allows us the convenience of being able to plot expected influence under the Zipf degree distribution and under the Poisson degree distribution on the same scale, as we do in Figure 7. We take

[35] The Zipf distribution arises in a number of real-world settings, including the distribution of city sizes. See e.g., Gabaix (1999), Ioannides and Overman (2003), and Arshad et al. (2018).

[36] The reason we define $\theta = (e^\alpha - 1)^{-1}$ is so that can write the results here in a similar way to the Poisson model where virality depended on $\lambda\sigma$.

$\sigma = 1$ in the figure and plot with respect to θ (i.e., $\lambda = \theta$ for the Poisson distribution). Crucially, Condition 1 holds for the Zipf distribution, and therefore Proposition 8 provides a

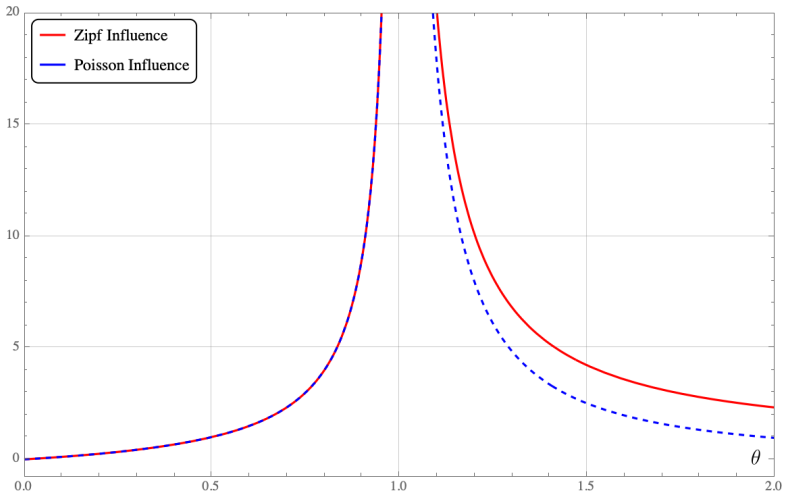

Figure 7: Zipf and Poisson Influence

full characterization of the equilibrium structure as a function of θ. The thresholds $\underline{\theta}, \overline{\theta}$ are determined by the equations

$$\frac{v\underline{\theta}}{1-\underline{\theta}} + v - c = 0 \iff \underline{\theta} = 1 - \frac{v}{c}, \tag{19}$$

$$v\left(\frac{2(\overline{\theta}+1)}{\overline{\theta}\ln(\overline{\theta})} - \frac{2}{\overline{\theta}-1}\right) + v - c = 0, \tag{20}$$

and the second equation can be solved numerically for specific values of v and c. E.g. when $c/v = 2$, we have $\overline{\theta} = 4.92$ (and $\underline{\theta} = 0.5$). In our next example the expected influence can again be solved for analytically, and we use numerical methods to verify Condition 1.

5.2.2 Exponential Distribution

Next, we consider a configuration model network with degree distribution

$$p_k = Ae^{-\alpha k}, \tag{21}$$

where $A = e^{\alpha} - 1$. This is the discrete *exponential distribution* on $k \in \{1, 2 \dots\}$. It is convenient to parameterize the distribution in terms of $\theta = E[\{p_k\}_{k\geq 0}] = (1 - e^{-\alpha})^{-1}$. The generating function for this distribution is given by

$$G_0(z) = \frac{z}{\theta - z(\theta - 1)},$$

and so

$$G_1(z) = \frac{1}{[\theta - z(\theta - 1)]^2}.$$

The extinction probability ρ_σ is therefore the solution of a cubic equation and is solvable by radicals. It is easy to show that $G_1'(1) = 2(\theta - 1)$, and so subcritical influence (when $\rho_\sigma = 1$,) is given by

$$\mathcal{I}_\theta(\sigma) = \frac{4\sigma(\theta - 1)}{1 - 2\sigma(\theta - 1)}.$$

It's clear that this expression is increasing in both θ and σ, and moreover that σ is viral if and only if $\sigma > \frac{1}{2(\theta-1)}$. We can also show that, quite remarkably, that

$$\frac{1}{1 - G_0(1 - \sigma + \sigma\rho_\sigma)} \sum_{k=1}^{\infty} k(1 - \rho_\sigma^k) \frac{\sigma^k}{k!} G_0^{(k)}(1 - \sigma) = 1 + \sigma(\theta - 1).$$

We show in Online Appendix F.2.2 that the supercritical expected influence can be written in closed form as a function of θ and σ,

$$\mathcal{I}_\theta(\sigma) = \frac{2\sigma(\theta - 1)}{(\theta - 1)^2\sigma^2 + 4(\theta - 1)\sigma + \sqrt{4 + \sigma(\theta - 1)}\sqrt{\sigma(\theta - 1)}(\sigma(\theta - 1) - 2)}.$$

Note that $\mu \equiv 2(\theta - 1)\sigma$ is a sufficient statistic for describing $\mathcal{I}_\theta(\sigma)$, and a strategy σ is viral iff $\mu > 1$. We can plot the function to see that it satisfies Condition 1. And so it follows that the

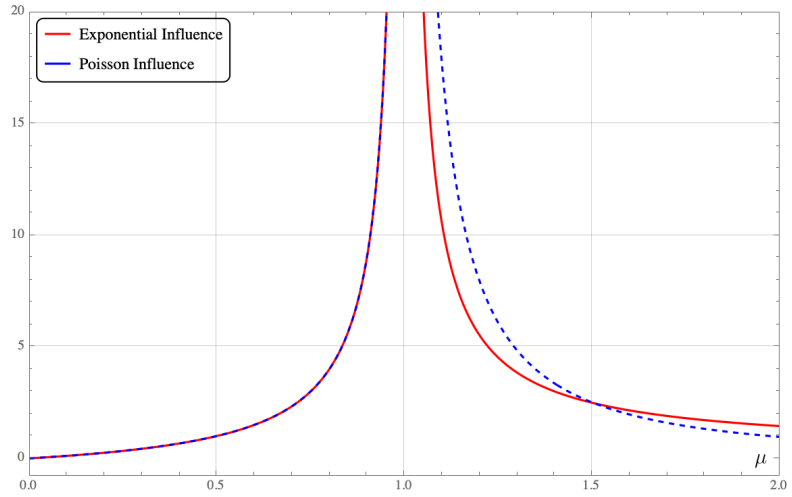

Figure 8: Influence for the Exponential Distribution

qualitative results of our analysis hold for an exponential degree distribution. Moreover, since $\lim_{\mu\to\infty} \mathcal{I}(\mu) \to 0$,[37] both thresholds $\underline{\theta}$ and $\overline{\theta}$ exist.

[37]The numerator has linear order in μ and the denominator quadratic order.

5.2.3 Power Law

Many real-world networks follow a power-law distribution $\mathbb{P}(k) \propto k^{-\gamma}$. When $\gamma \geq 3.5$ there is no giant component in the underlying graph. When $\gamma \leq 3$, the critical percolation threshold is 0, because the second moment is infinite and therefore every strategy $\sigma > 0$ is viral. Moreover, since the expected number of neighbors who are potential adopters depends on derivatives $G^{(k)}$, and these diverge whenever $k > 2$, expected influence is not well defined for this class of distributions. However, $\gamma \in (3, 3.5)$, we can construct a quasi-measure of expected influence by restricting calculations to the first two moments. This also allows us to build some intuition about what happens as the limit as $\gamma \to 3^+$.

Consider a configuration model with degree distribution

$$p_k = Ak^{-\gamma}, \tag{22}$$

where $k \in [1, \infty)$, and

$$A = \zeta(\gamma)^{-1},$$

where ζ is the Riemann zeta function. It is known that

$$G_1(z) = \frac{\mathrm{Li}_{\gamma-1}(z)}{z\zeta(\gamma - 1)},$$

where Li_γ is the polylogarithmic function with parameter γ. The extinction probability cannot be solved for explicitly, but we can solve for it numerically for specific values of gamma. For example, we can verify numerically that $\gamma \approx 3.5$ is the threshold for the emergence of the giant component in the underlying graph, so that a giant component exists if and only if $\gamma < 3.5$. In Figure 9 we take $\gamma = 3.4$ and plot the influence function which we have computed numerically. One can easily see that it exhibits the same qualitative characteristics as the previous two distributions. In Figure 10 we see what happens to influence as $\gamma \to 3^+$, influence explodes closer to $\sigma = 0$, and in the limit all equilibria become viral.

If the evolution of social media has approximately followed a power-law distribution with gradually decreasing parameter γ, then our model provides a new explanation for why social media environments have become increasingly hostile since their inception: network participants have become more and more anonymous.

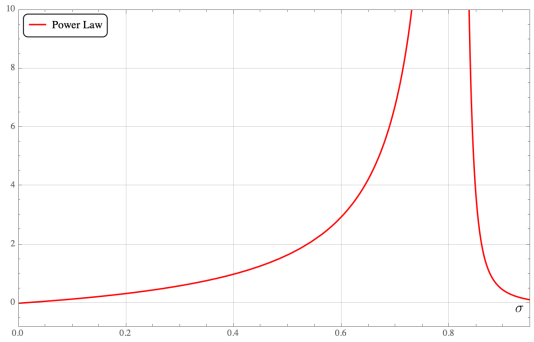

Figure 9: Influence for the Power Law Distribution with $\gamma = 3.4$

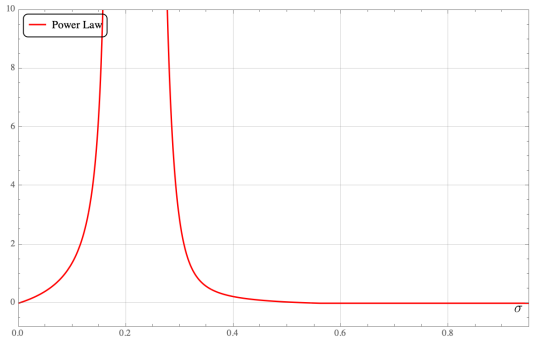

Figure 10: Influence for the Power Law Distribution as $\gamma \downarrow 3$

5.3 Discussion of the Assumptions

5.3.1 No Viral Inference

In our model, agents correctly update their beliefs about their expected degree upon exposure when there is a giant component of potential adopters in the network (this is viral inference). A natural question is whether our results change qualitatively if agents are naive in the sense that they take the degree distribution to the be that of the underlying network regardless of whether they are in the giant component. Since exposure informs an agent that at least one of their links is in the giant component, it stands to reason that viral inference should decrease the expected influence relative to naive beliefs. It is easily shown that expected influence in a naive version of the Poisson model is given by

$$\mathcal{I}_\lambda(\sigma) = \frac{2\lambda\sigma\rho_\sigma}{1-\lambda\sigma\rho_\sigma},$$

so the $-\rho_\sigma$ term in Lemma 1 is the "correction" due to viral inference.

In Figure 11, we plot the expected influence accounting for viral inference, relative to expected influence with naive updating. We plot this only for $\lambda > 1$ since this is the only time viral strategies exist, and we mark $\tilde{\lambda} < \overline{\lambda}$ as the upper thresholds for emergence of viral mixed equilibria under VI and no-VI. We see, as explained in our analysis above, that the expected influence under naive updating lies strictly above the expected influence with viral inference: agents believe they are too influential.

To conclude this section, we emphasize that the introduction of viral inference into the model did not change the qualitative results of our characterization of equilibrium precisely because the expected influence remained qualitatively the same as in the Poisson model. Our analysis of graphs with an arbitrary degree distribution is in many ways analogous (see next section) and we provide further examples illustrating the same point.

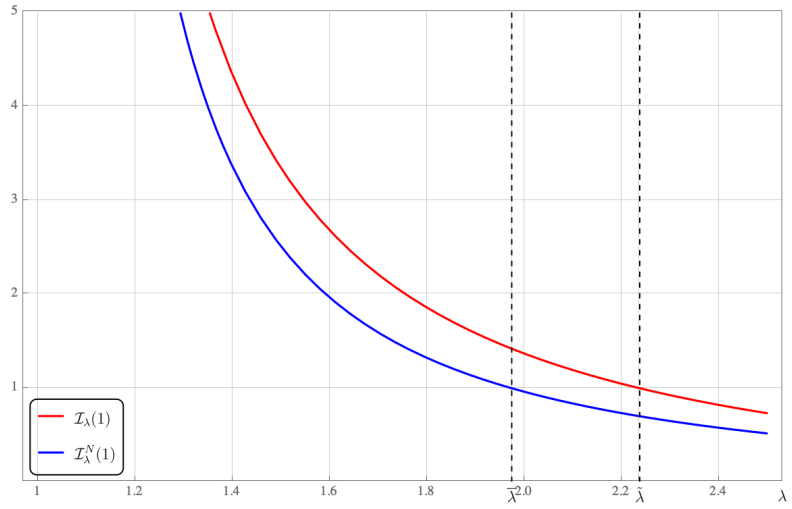

Figure 11: Influence with ($\mathcal{I}_\lambda$) and without ($\mathcal{I}^N_\lambda$) viral inference

5.3.2 Knowledge of Degrees

A critical assumption that simplifies our analysis is that agents do not know their degree. In the settings we are primarily interested in (e.g., participation in a protest), this is a reasonable assumption. In other settings, it may be of interest to know how agents behave when they have knowledge of their degree.

Suppose that agents know their degree. Then, strategies become functions of the degree, i.e., $\sigma\colon \mathbb{N} \to [0,1]$. Since the graph is locally tree-like, higher degrees translate multiplicatively into higher influence. Under degree distribution $\mathcal{D}_\theta$ and strategy *profile* $\boldsymbol{\sigma} \equiv \{\sigma(d)\}_{d\geq 0}$, the expected utility from provision of a public good by an agent of degree d is

$$\mathbb{E}_{\mathcal{D}}[u_\pi(\boldsymbol{\sigma}, d))] = (v + c) + v(d-1)\mathcal{I}_\theta(\boldsymbol{\sigma}).$$

It is straightforward to observe that equilibrium Bayesian strategies will take a threshold form. In the case of a public good, agents adopt (not-adopt) with degrees strictly above (below) a threshold degree k^{thres} and may potentially mix at the threshold.[38] Note that agents with $d = 1$ either never (always) adopt in the case of public goods (bads). Existence of equilibria can be proved from Brouwer's fixed point theorem, and we can write down an explicit form for the expected influence $I_\theta(\boldsymbol{\sigma})$ in terms of generating functions. The exploding nature of influence at the critical threshold $\lambda = 1$ suggests that public goods emerge continuously and public bads emerge discontinuously in this setting as well.

[38] Of course vice-versa in the case of a public bad.

6 Conclusion

We develop and analyze a model of privately provided public good and bad behaviors that spread through populations via observational learning between social contacts. We show that a key strategic consideration is, influence, the *causal effect* of an agent's adoption decision on the adoption decisions of the others in the network. Our analysis of influence identifies a non-trivial effect of greater density on behavior via two competing effects. A direct *positive* effect due to more people being able to observe an individual and a global *negative* effect as greater connectivity leading people to be potentially influenced through multiple sources.

Our results establish that there are some robust properties in the way connectivity is related to influence: at first it increases, then past a critical threshold of connectivity, influence decreases. We also demonstrate that there is a stark difference between the phase transition in equilibrium behavior for public goods compared to public bads. Public goods exhibit a *continuous* phase transition in equilibrium adoption, while public bads exhibit a *discontinuous* transition in which equilibrium adoptions emerge suddenly. We extend our model to graphs with an arbitrary degree distribution and discuss how alternative assumptions such as knowledge of degree may be incorporated into our setting.

References

Akbarpour, M., S. Malladi, and A. Saberi (2023). Just a few seeds more: The inflated value of network data for diffusion. Unpublished manuscript, Stanford University.

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

Allouch, N. (2017). The cost of segregation in (social) networks. *Games and Economic Behavior 106*, 329–342.

Arshad, S., S. Hu, and B. N. Ashraf (2018). Zipf's law and city size distribution: A survey of the literature and future research agenda. *Physica A: Statistical Mechanics and its Applications 492*, 75–92.

Athreya, K. and P. Ney (1972). *Branching Processes*. Springer, Heidelberg.

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006). Who's who in networks. wanted: the key player. *Econometrica 74*(5), 1403–1417.

Barberà, S. and M. O. Jackson (2020). A model of protests, revolution, and information. *Quarterly Journal of Political Science 15*(3), 297–335.

Battaglini, M. (2017). Public protests and policy making. *Quarterly Journal of Economics 132*(1), 485–549.

Bénabou, R., A. Falk, and J. Tirole (2020). Narratives, imperatives, and moral persuasion. Unpublished manuscript, Princeton University.

Blume, L. (1993). The statistical mechanics of strategic interaction. *Games and Economic Behavior 5*, 387–424.

Bollobás, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics 1*(4), 311–316.

Bollobás, B. and O. Riordan (2015). An old approach to the giant component problem. *Journal of Combinatorial Theory, Series B 113*, 236–260.

Bramoullé, Y. and R. Kranton (2007). Public goods in networks. *Journal of Economic Theory 135*(1), 478–494.

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

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

Brock, W. and S. Durlauf (2001). Discrete choice with social interactions. *Review of Economic Studies 68*, 235–260.

Bueno de Mesquita, E. (2010). Regime change and revolutionary entrepreneurs. *American Political Science Review 104*, 446–466.

Bursztyn, L., D. Cantoni, D. Yang, N. Yuchtman, and Y. Zhang (2021). Persistent political engagement: Social interactions and the dynamics of protest movements. *American Economic Review: Insights 3*(2), 233–250.

Callaway, D., M. Newman, S. Strogatz, and D. Watts (2000). Network robustness and fragility: Percolation on random graphs. *Physical Review Letters 85*(25), 5468–5471.

Campbell, A. (2013). Word-of-mouth communication and percolation in social networks. *American Economic Review 103*, 2466–2498.

Campbell, A., C. M. Leister, and Y. Zenou (2024). Competing diffusions in a social network. Unpublished manuscript, Monash University.

Campbell, A., P. Ushchev, and Y. Zenou (2024). The network origins of entry. *Journal of Political Economy,* forthcoming.

Cantoni, D., D. Yang, N. Yuchtman, and Y. Zhang (2019). Protests as strategic games: Experimental evidence from Hong Kong's antiauthoritarian movement. *Quarterly Journal of Economics 134*(2), 1021–1077.

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

Dasaratha, K. (2023, 01). Innovation and strategic network formation. *The Review of Economic Studies 90*(1), 229–260.

Diestel, R. (2000). *Graph Theory, 2nd ed.* Berlin: Springer-Verlag.

Easley, D. and J. Kleinberg (2010). *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge: Cambridge University Press.

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

Egorov, G. and K. Sonin (2021). Elections in non-democracies. *The Economic Journal, 131*(636), 1682–1716.

Elliott, M. and B. Golub (2019). A network approach to public goods. *Journal of Political Economy 127*(2), 730–776.

Ellison, G. (1993). Learning, local interaction, and coordination. *Econometrica 61*, 1047–1071.

Erdős, P. and A. Rényi (1959). On random graphs. *Publicationes Mathematicae 6*, 290–297.

Gabaix, X. (1999). Zipf's law for cities: An explanation. *Quarterly Journal of Economics CXIV*, 739–767.

Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology 83*, 1420–1443.

Ioannides, Y. M. and H. G. Overman (2003). Zipf's law for cities: An empirical examination. *Regional Science and Urban Economics 33*, 127–137.

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

Jackson, M. O. and D. López-Pintado (2013). Diffusion and contagion in networks with heterogeneous agents and homophily. *Network Science 1*, 49–67.

Jackson, M. O., B. W. Rogers, and Y. Zenou (2017). The economic consequences of social-network structure. *Journal of Economic Literature 55*(1), 49–95.

Jackson, M. O. and L. Yariv (2005). Diffusion on social networks. *Economie Publique 16*, 69–82.

Jackson, M. O. and L. Yariv (2007). Diffusion of behavior and equilibrium properties in network games. *American Economic Review, Papers and Proceedings 97*(2), 92–98.

Jackson, M. O. and L. Yariv (2011). Diffusion, strategic interaction, and social structure. In: Benhabib, J., Bisin, A. and M.O. Jackson (Eds.), *Handbook of Social Economics, Volume 1A,* Amsterdam: Elsevier Science, pp. 645–678.

Jackson, M. O. and Y. Zenou (2015). Games on networks. In P. Young and S. Zamir (Eds.), *Handbook of Game Theory, Vol. 4*, pp. 34–61. Amsterdam: Elsevier Publisher.

Janson, S. (2009). On percolation in random graphs with given vertex degrees. *Electronic Journal of Probability 14*(5), 87–118.

Kleinberg, J. (2007). Cascading behavior in networks: Algorithmic and economic issues. In: N. Nisan, T. Roughgarden, E. Tardos and V. Vazirani (Eds.), *Algorithmic Game Theory,* Cambridge: Cambridge University Press.

Kuran, T. (1989). Sparks and prairie fires: A theory of unanticipated political revolution. *Public Choice 61*, 41–74.

Kuran, T. (1991). The East European revolution of 1989: Is it surprising that we were surprised? *American Economic Review Papers and Proceedings 81*(2), 121–125.

Kuran, T. (1997). *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge, MA: Harvard University Press.

Langtry, A. (2023). More connection, less community: Network formation and local public goods provision. Unpublished manuscript, University of Cambridge.

Langtry, A., S. Taylor, and Y. Zhang (2024). Network threshold games. Unpublished manuscript, Cambridge University.

Leister, C. M., Y. Zenou, and J. Zhou (2022). Social connectedness and local contagion. *The Review of Economic Studies 89*(1), 372–410.

López-Pintado, D. (2006). Contagion and coordination in random networks. *International Journal of Game Theory 34*(3), 371–381.

López-Pintado, D. (2008). Diffusion in complex social networks. *Games and Economic Behavior 62*(2), 573–590.

Molloy, M. and B. Reed (1998). The size of the giant component of a random graph with a given degree sequence. *Combinatorics, Probability and Computing 7*(3), 295–305.

Morris, S. (2000). Contagion. *Review of Economic Studies 67*(1), 57–78.

Newman, M., S. Strogatz, and D. Watts (2001). Random graphs with arbitrary degree distributions and their applications. *Physical Review E 64*(2), 026118.

Olson, M. (1965). *The Logic of Collective Action: Public Goods and the Theory of Groups*. New York: Schocken Books.

Palfrey, T. and H. Rosenthal (1984). Participation and the provision of discrete public goods: A strategic analysis. *Journal of Public Economics 24*(105), 171–193.

Passarelli, F. and G. Tabellini (2017). Emotions and political unrest. *Journal of Political Economy 125*, 903–946.

Sadler, E. (2020). Diffusion games. *American Economic Review 110*(1), 225–270.

Schelling, T. (1978). *Micromotives and Macrobehavior*. New York: Norton.

Shadmehr, M. and D. Bernhardt (2011). Collective action with uncertain payoffs: Coordination, public signals and punishment dilemmas. *American Political Science Review 105*, 829–851.

Tullock, G. (1971). The paradox of revolution. *Public Choice 11*, 89–99.

Van der Hofstad, R. (2023a). *Random Graphs and Complex Networks. Volume 1.* Cambridge, MA: MIT press.

Van der Hofstad, R. (2023b). Random graphs and complex networks. volume 2. Unpublished manuscript.

Vega-Redondo, F. (2007). *Complex Social Networks*. Cambridge: Cambridge University Press.

Watts, D. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences of the USA 99*(9), 5766–5771.

Wormald, N. (1978). *Some problems in the enumeration of labelled graphs*. Ph.d. thesis, Newcastle University.

Appendix: Proofs of the results in the main text

6.1 Proof of Proposition 1

We wish to establish the existence of a symmetric perfect Bayesian Equilibrium. We first define a best response correspondence:

$$BR^{\pi}(\sigma) = \begin{cases} 1.\mathbb{1}_{\pi=g}, & \mathcal{I}_{\lambda}^{(n)}(\sigma) > \frac{c}{v} - 1 \\ [0,1], & \mathcal{I}_{\lambda}^{(n)}(\sigma) = \frac{c}{v} - 1 \\ 1.\mathbb{1}_{\pi=b}, & \mathcal{I}_{\lambda}^{(n)}(\sigma) < \frac{c}{v} - 1. \end{cases} \tag{23}$$

Appendix B shows that calculating $\mathcal{I}_{\lambda}^{(n)}(\sigma)$ via Bayes rule results in function that is polynomial and continuous in σ. Hence, $\sigma, BR^{\pi}(\sigma)$ satisfy the conditions of Kakutani's fixed point theorem and there exists a strategy $\sigma^* = BR^{\pi}(\sigma^*)$. Our equilibrium is then this strategy and the associated beliefs that result in $\mathcal{I}_{\lambda}^{(n)}(\sigma^*)$. Off-equilibrium beliefs only arise at an information set in an equilibrium where $\sigma^* = 0$. In this case it is trivial to specify any belief where $\mathcal{I}_{\lambda}^{(n)}(\sigma^*) = 0$ in such an event.

6.2 Proof of Lemma 1

Proved as a special case of Lemma 2.

6.3 Proof of Proposition 2

By Lemma 1, expected influence when $\sigma = 1$ is given by

$$\mathcal{I}_{\lambda}(1) = \begin{cases} \frac{\lambda}{1-\lambda}, & \lambda < 1 \\ 2\left(\frac{\lambda\rho_1}{1-\lambda\rho_1} - \rho_1\right), & \lambda > 1 \end{cases}$$

In what follows we can always recover the same comparative static result for σ by replacing λ with $\lambda\sigma$ (this is essentially invoking the duality result Theorem A1), so it is without loss of generality to focus on $\sigma = 1$. We first consider nonviral strategies ($\lambda < 1$). In this case we have

$$\mathcal{I}_{\lambda}(1) = \frac{\lambda}{1-\lambda},$$

and

$$\frac{d\mathcal{I}_{\lambda}(1)}{d\lambda} > 0 \iff (1-\lambda) + \lambda > 0$$

which is always true. We now consider viral strategies. The derivative of expected influence (after dividing by 2 for ease of exposition) with respect to λ is given by

$$\frac{d}{d\lambda}\left(\frac{\lambda\rho_1}{1-\lambda\rho_1}-\rho_1\right)=\frac{\frac{d(\lambda\rho_1)}{d\lambda}}{(1-\lambda\rho_1)^2}-\frac{d\rho_1}{d\lambda},$$

We begin by finding an explicit formula for $\frac{d(\lambda\rho_1)}{d\lambda}$. By (7) and the implicit function theorem we have

$$\frac{d\rho_1}{d\lambda}=\left((\rho_1-1)+\lambda\frac{d\rho_1}{d\lambda}\right)e^{\lambda(\rho_1-1)}=\frac{\rho_1(\rho_1-1)}{1-\lambda\rho_1}.$$

So,

$$\frac{d(\lambda\rho_1)}{d\lambda}=\rho_1-\frac{\lambda\rho_1(1-\rho_1)}{1-\lambda\rho_1}=\rho_1\left[1-\frac{\lambda(1-\rho_1)}{1-\lambda\rho_1}\right].$$

It follows that expected influence is strictly decreasing (for $\lambda>1$) iff

$$\begin{aligned}\frac{d(\lambda\rho_1)}{d\lambda}&<\frac{d\rho_1}{d\lambda}(1-\lambda\rho_1)^2\\ \iff \rho_1\left[1-\frac{\lambda(1-\rho_1)}{1-\lambda\rho_1}\right]&<-\frac{\rho_1(1-\rho_1)}{1-\lambda\rho_1}(1-\lambda\rho_1)^2\\ \iff \frac{\lambda-1}{1-\rho_1}&>(1-\lambda\rho_1)^2. \qquad (24)\end{aligned}$$

From here the proof takes some work, we sketch the proof strategy here but leave the details to the Online Appendix. First, note that (24) is true whenever $\lambda>2$, since in this case we have the chain of inequalities

$$(1-\rho_1)(1-\lambda\rho_1)^2<1<\lambda-1.$$

It remains to show that the inequality holds when $1<\lambda<2$. To do this, we show that the inequality (24) holds when $\lambda=1$, and then show that both sides of the inequality are strictly increasing for all $\lambda>1$. We conclude the proof by showing that at $\lambda=2$, the RHS of (24) is still smaller than the smallest value of the LHS, so the inequality holds everywhere.

6.4 Proof of Proposition 3

First, $\sigma=0$ is always a best response to itself since

$$v-c+v\mathcal{I}_\lambda(0)=v-c<0,$$

so agents strictly prefer not to adopt if they expect others to do the same. Hence, no adoption ($\sigma=0$) is always an equilibrium. It follows that full adoption ($\sigma=1$) is an equilibrium if and only if a mixed strategy equilibrium exists. A nonviral mixed strategy equilibrium exists if and

only if there is a $\sigma \in (0,1)$ satisfying

$$\begin{aligned} v - c + v\mathcal{I}_\lambda(\sigma) = 0 &\iff v - c + \frac{v\lambda\sigma}{1-\lambda\sigma} \\ &\iff \lambda\sigma = 1 - \frac{v}{c}. \end{aligned}$$

while a viral mixed strategy equilibrium exists if and only if there is a $\sigma \in (0,1)$ with $\rho_\sigma < 1$ satisfying

$$\begin{aligned} v - c + v\mathcal{I}_\lambda(\sigma) = 0 &\iff v - c + \frac{2v\lambda\sigma\rho_\sigma}{1-\lambda\sigma\rho_\sigma} - 2v\rho_\sigma = 0 \\ &\iff \lambda\sigma\rho_\sigma = 1 - \frac{2v}{c + v(1+\rho_\sigma)}. \end{aligned}$$

It follows that in the subcritical regime (when $\lambda < 1$), we have a mixed strategy equilibrium if and only if

$$1 > \lambda > 1 - \frac{v}{c} = \underline{\lambda}.$$

Moreover, when $\lambda < \underline{\lambda}$, we have $\lambda\sigma < 1 - \frac{v}{c}$ for all $\sigma \in [0,1]$ and so $\sigma = 0$ is the only equilibrium. This proves the first two cases of Proposition 3. Next, in the supercritical regime ($\lambda > 1$), we know from Proposition 2 that influence is strictly decreasing for any fixed σ. In particular, there exists a smallest λ, which we call $\overline{\lambda}$ such that

$$\overline{\lambda}\rho_1 = 1 - \frac{2v}{c + v(1+2\rho_1)}.$$

For any $1 < \lambda < \overline{\lambda}$, there is a full adoption equilibrium, and this is the unique viral equilibrium (the nonviral mixed equilibrium and no adoption equilibrium are also still present). However, for $\lambda > \overline{\lambda}$, full adoption is no longer an equilibrium, since $\overline{\lambda}\rho_1 < 1 - \frac{2v}{c+v(1+2\rho_1)}$, and so by Proposition 2 there is a unique viral equilibrium in mixed strategies determined by

$$\lambda\sigma = \overline{\lambda}.$$

This proves the final two cases of Proposition 3.

6.5 Proof of Corollary 1

The first equality holds because the LHS is necessarily 0 for all $\lambda < 1$, and the second equality holds because Proposition 3 implies that a viral full-adoption equilibrium exists when λ is arbitrarily close to (but larger than) 1. In particular, since the fraction of agents in the largest component of $G(n,p)$ can be made arbitrarily small as $\lambda \to 1^+$, and in a full-adoption equilibrium the potential adopter network is all of $G(n,p)$, second equality follows.

6.6 Proof of Corollary 2

The proof follows from Proposition 4 but is essentially identical to the proof of Corollary 1. We omit the details.

6.7 Proof of Proposition 4

The details of the proof are virtually identical to the proof of Proposition 3. We provide a sketch for completeness.

The key difference is that when λ is small, $\sigma = 1$ is the unique nonviral equilibrium, since

$$c - v - v\mathcal{I}_\lambda(1) > 0.$$

However when λ becomes sufficiently large (exceeds $\underline{\lambda}$), full adoption is no longer an equilibrium since agents expect to influence too many others. This leads to a mixed strategy nonviral equilibrium and no adoption equilibrium when $\underline{\lambda} < \lambda < \overline{\lambda}$. Finally, when $\lambda > \overline{\lambda}$, the giant component is large enough such that expected influence is small, and full adoption becomes the unique viral equilibrium.

6.8 Proof of Proposition 5 and Corollary 7

We first note that in the subcritical regime, all equilibria are nonviral and so their size is 0. We proved in the text of Section 4.3 the case where $\lambda < \overline{\lambda}$ for public goods. We also fully characterised the comparative statics for public bads. The only case that remains to prove is when $\lambda > \overline{\lambda}$ for public goods. We now show that for public goods, the size of a large cascade in the largest equilibrium is strictly decreasing when $\lambda > \overline{\lambda}$.

Recall from (12) that

$$\lambda\sigma^*\rho_{\sigma^*} = 1 - \frac{2v}{c + v(1 + 2\rho^*_\sigma)}$$

in equilibrium, where

$$\rho_{\sigma^*} = e^{\lambda\sigma^*(1-\rho_{\sigma^*})}.$$

Hence since ρ_{σ^*} is a function of $\lambda\sigma^*$, and the equilibrium is unique, it follows that we must have $\lambda\sigma^* = \overline{\lambda}$, where

$$\overline{\lambda}\rho_1 = 1 - \frac{2v}{c + v(1 + 2\rho_1)}.$$

In other words, in equilibrium, σ is chosen such that ρ_{σ^*} is held constant. Hence when $\lambda > \overline{\lambda}$, an increase in λ keeps $\lambda\sigma^*$ fixed, and therefore σ^* must be strictly decreasing in λ.

Finally, since ρ^* is constant for all $\lambda > \overline{\lambda}$, and diffusion is determined by $\sigma^*(1-\rho_{\sigma^*})$, we have that diffusion is strictly decreasing in λ in the largest equilibrium, as claimed. Proposition 7 is immediate. The diffusion of public goods is maximized at $\lambda = \overline{\lambda}$, while the diffusion of public bads is 0 if and only if $\lambda < \overline{\lambda}$.

6.9 Proof of Proposition 6

Suppose $1 < \lambda < \overline{\lambda}$ and consider the diffusion of a public good. By Proposition 3, the largest equilibrium is the full adoption equilibrium. Recall that a viral mixed strategy equilibrium exists if and only if $\lambda > \overline{\lambda}$, where

$$\overline{\lambda}\rho_1 = 1 - \frac{v}{c}.$$

The RHS of the above equation is decreasing in v, increasing in c and decreasing in v/c. In particular, since $\lambda\rho_1 > 1 - \frac{v}{c}$ when $\lambda < \overline{\lambda}$, it follows that if the RHS increases, $\overline{\lambda}$ decreases. In particular, $\overline{\lambda}$ is increasing in v, decreasing in c and increasing in v/c. Now, when $\lambda < \overline{\lambda}$, a change in the parameters will either:

1. Have no effect if it is still true that $\lambda < \overline{\lambda}$, or

2. Cause the unique viral equilibrium to be a mixed strategy equilibrium if it becomes the case that $\overline{\lambda} < \lambda$.

In the second case, diffusion decreases (by Proposition 5), and the result follows. The analysis is virtually identical for public bads except that the equilibrium can jump from no adoption to full adoption rather than from full adoption to a viral mix.

6.10 Proof of Lemma 2

Suppose agents are playing the strategy σ. Let $F_{1,\sigma}(z)$ be the generating function for the distribution over finite "forward component sizes" in the potential adopter network (*not* necessarily the whole network) from following a randomly chosen edge. Then

$$F_{1,\sigma}(z) = zG_1[1 - \sigma + \sigma F_{1,\sigma}(z)], \tag{25}$$

from which it follows that

$$F_{1,\sigma}(1) = G_1[1 - \sigma + \sigma F_{1,\sigma}(1)] = \rho_\sigma.$$

We can find the expected finite forward component size in the potential adopter network by taking the derivative of (25) at 1 and substituting in the expression for $F_{1,\sigma}(1)$. I.e.

$$F'_{1,\sigma}(1) = G_1[1-\sigma+\sigma F_{1,\sigma}(1)] + \sigma F'_{1,\sigma}(1) G'_1[1-\sigma+F_{1,\sigma}(1)] = \frac{\rho_\sigma}{1-\sigma G'_1(1-\sigma+\sigma\rho_\sigma)}. \quad (26)$$

Since, in the limit, the graph is locally tree-like, each neighbor of an agent can be considered as an independent cluster of potential influence (we discuss this in more depth in Appendix B). If σ is nonviral, then $\rho_\sigma = 1$, and exposure gives no information about the degree distribution. Hence the total expected influence of an agent i is $F'_{1,\sigma}(1)$ multiplied by i's expected number of potential-adopter neighbors: $\mathbb{E}[A_i] = \sigma G'_1(1)$. That is,

$$\mathcal{I}_\theta(\sigma) = \frac{1}{1-\sigma G'_1(1)} \times \mathbb{E}[A_i] = \frac{\sigma G'_1(1)}{1-\sigma G'_1(1)}.$$

When σ is viral, there is a giant component of potential adopters. Here there are two cases we need to consider. The first is that the seed is in the giant forward component (which happens with probability $\to 1$ in the limit of large networks). In this case, expected influence is the expected finite forward component through each neighbor; this is simply $F'_{1,\sigma}(1)$ for each neighbor. The second case is that the seed is in a finite forward component, in which case an agent is pivotal for the information reaching the giant component. We return to this second case at the end, for now we calculate the expected degree conditional on exposure.

Let $\mathcal{S}_i$ denote the event that at least one neighbor of a randomly chosen vertex i is in the giant component of potential adopters. Let A_i denote the number of neighbors of i who are potential adopters. Let d_i denote i's degree. The quantity we need to find is $\mathbb{E}[A_i \mid \mathcal{S}_i]$. By the law of total expectation,

$$\mathbb{E}[A_i \mid \mathcal{S}_i] = \sum_{d_i=1}^{\infty} \mathbb{E}[A_i \mid \mathcal{S}_i, d_i] \mathbb{P}(d_i \mid \mathcal{S}_i).$$

By Sadler (2020, Theorem 3),

$$\mathbb{P}(A_i = k \mid \mathcal{S}_i, d_i) = \frac{1-\rho_\sigma^k}{1-(1-\sigma+\sigma\rho_\sigma)^{d_i}} \mathbb{P}(\text{Bin}(d,\sigma) = k).$$

So we have

$$\mathbb{E}[A_i \mid \mathcal{S}_i, d_i] = \sum_{k=1}^{d_i} k \left(\frac{1-\rho_\sigma^k}{1-(1-\sigma+\sigma\rho_\sigma)^{d_i}} \mathbb{P}(\text{Bin}(d_i,\sigma) = k) \right).$$

Now notice that (dropping the i subscripts)

$$\mathbb{E}[A \mid \mathcal{S}, d]\mathbb{P}(d \mid \mathcal{S}) = \left[\sum_{k=1}^{d} k \left(\frac{1-\rho_\sigma^k}{1-(1-\sigma+\sigma\rho_\sigma)^d}\mathbb{P}(\text{Bin}(d,\sigma)=k)\right)\right]\left[\frac{p_d\left(1-(1-\sigma+\sigma\rho_\sigma)^d\right)}{1-\sum_{k\geq 0}p_k(1-\sigma+\sigma\rho_\sigma)^k}.\right]$$

$$= \frac{1}{1-G_0(1-\sigma+\sigma\rho_\sigma)}\sum_{k=1}^{d} k p_d\left(1-\rho_\sigma^k\right)\mathbb{P}(\text{Bin}(d,\sigma)=k).$$

By Tonelli's theorem, the double sum $\sum_{d=1}^{\infty}\sum_{k=1}^{n}$ can instead be computed as

$$\begin{aligned}\sum_{k=1}^{\infty}\sum_{d=k}^{\infty} k p_d\left(1-\rho_\sigma^k\right)\mathbb{P}(\text{Bin}(d,\sigma)=k) &= \sum_{k=1}^{\infty} k(1-\rho_\sigma^k)\sum_{d=k}^{\infty} p_d\mathbb{P}(\text{Bin}(d,\sigma)=k)\\ &= \sum_{k=1}^{\infty} k(1-\rho_\sigma^k)\frac{\sigma^k}{k!}G_0^{(k)}(1-\sigma)\end{aligned}$$

Subtracting 1 from the above expression (for the link along which exposure occurred) gives us the numerator for expected influence under viral strategies. Hence if the seed is in the giant forward component then expected influence is

$$F'_{1,\sigma}(1)\mathbb{E}[A_i \mid \mathcal{S}_i] = \frac{\rho_\sigma}{1-\sigma G_1'(1-\sigma+\sigma\rho_\sigma}\left(\frac{\sum_{k=1}^{\infty} k(1-\rho_\sigma^k)\frac{\sigma^k}{k!}G_0^{(k)}(1-\sigma)}{1-G_0(1-\sigma+\sigma\rho_\sigma)}-1\right).$$

Finally, returning to the "second case" mentioned earlier in the proof, it follows from the above analysis that the seed is in a finite forward component with probability $\frac{F'_{1,\sigma}(1)\mathbb{E}[A_i|\mathcal{S}_i]}{cn}$, where cn is the size of the giant component. But in this case, influence is equal to the entire giant component, i.e. cn, so on net the expected influence is simply $F'_{1,\sigma}(1)\mathbb{E}[A_i \mid \mathcal{S}_i]$. Adding this to expected influence gives us a factor of 2, which proves Lemma 2. We show the explicit case of Lemma 1 in the Online Appendix.

Online Appendix

A Poisson Model

A.1 Branching Processes and Duality

There are several texts one can consult for a detailed treatment of branching processes, e.g. Van der Hofstad (2023a); Athreya and Ney (1972). The most important results on Galton-Watson branching processes are also outlined in Appendix A of Sadler (2020). As such, we confine ourselves here to presenting results on the duality properties of supercritical branching processes, and for this we follow Van der Hofstad (2023a).

Let $\mathcal{D} \equiv \{p_k\}_{k\geq 0}$ be a probability distribution over the nonnegative integers, and suppose that $\mathcal{D}$ is the offspring distribution of a branching process. The branching process is said to be *subcritical* if $\mathbb{E}[\mathcal{D}] < 1$, and *supercritical* if $\mathbb{E}[\mathcal{D}] > 1$. This is because the *extinction probability,* ρ i.e. the probability that the branching process eventually dies out, is 1 if $\mathbb{E}[\mathcal{D}] < 1$, and strictly less than 1 if $\mathbb{E}[\mathcal{D}] > 1$.

Let $G_{\mathcal{D}}(z) = \sum_k p_k z^k$ denote the generating function for the distribution $\mathcal{D}$. Then the extinction probability ρ is defined by the smallest solution in $[0, 1]$ to the equation $\rho = G_{\mathcal{D}}(\rho)$.

Call the distributions $\{p_k\}_{k\geq 0}$ and $\{p_k^{'}\}_{k\geq 0}$ a *conjugate pair* if

$$p'_k = \rho^{k-1} p_k.$$

It is easy to check that $\{p'_k\}_{k\geq 0}$ is a probability distribution, since

$$\begin{aligned}\sum_{k=0}^{\infty} p'_k &= \rho^{-1} \sum_{k=0}^{\infty} p_k \rho^k \\ &= \rho^{-1} G_{\mathcal{D}}(\rho) \\ &= \rho^{-1} \rho = 1.\end{aligned}$$

It turns out that the distribution for a supercritical branching process conditioned on extinction, is precisely equal to the conjugate distribution defined above. This is stated formally in the following theorem.

Theorem A1 (Van der Hofstad (2023a, Theorem 3.7))**.** *Let $\{p_k\}_{k\geq 0}$ and $\{p'_k\}_{k\geq 0}$ be a conjugate pair of offspring distributions. The branching process with distribution $\{p_k\}_{k\geq 0}$ conditioned on*

extinction has the same distribution as the branching process with offspring distribution $\{p'_k\}_{k\geq 0}$.

The proof follows directly from Bayes' rule. Theorem A1 takes on a particularly nice form for Poisson branching processes. Let $\{p_k\}_{k\geq 0}$ be a Poisson distribution with mean λ. Then the generating function is given by

$$G_0(z) = e^{\lambda(z-1)},$$

and therefore the extinction probability $\rho = \rho_\lambda$ satisfies

$$\rho_\lambda = e^{\lambda(\rho_\lambda - 1)}. \tag{A1}$$

If $\lambda > 1$, then a branching process with offspring distribution $\{p_k\}_{k\geq 0}$ is supercritical, and therefore by Theorem A1 the distribution conditional on extinction is

$$p'_k = \rho_\lambda^{k-1} p_k = \frac{\rho_\lambda^k}{e^{\lambda(\rho_\lambda - 1)}} \cdot \frac{e^{-\lambda}\lambda^k}{k!} = \frac{e^{-\lambda\rho_\lambda}(\lambda\rho_\lambda)^k}{k!},$$

where the second equality follows from Eqn. (A1). But this distribution is again Poisson, with mean

$$\mu_\lambda \equiv \lambda\rho_\lambda < 1.$$

It follows that a branching process with offspring distribution $\mathrm{Po}(\lambda)$ (where $\lambda > 1$) conditioned on extinction, has the same distribution as a branching process with offspring distribution $\mathrm{Po}(\lambda\rho_\lambda)$. We call a branching process with offspring distribution $\mathrm{Po}(\lambda\rho_\lambda)$ the *subcritical dual* of the supercritical branching process with offspring distribution $\mathrm{Po}(\lambda)$.

Once we introduce percolation, the branching process has offspring distribution $\mathrm{Po}(\lambda\sigma)$, and so the subcritical dual has offspring distribution $\mathrm{Po}(\lambda\sigma\rho_\sigma)$, where we omit the dependence of ρ on λ for readability. Finally, although we formally prove Section 6.2, this duality gives us another way of seeing why the influence function takes on the specific form that it does. In particular consider the following theorem.

Theorem A2 (Van der Hofstad (2023a, Theorem 3.5))**.** *Let T denote the total offspring of a branching process with i.i.d. offspring X, having mean offspring $\mu < 1$, then*

$$\mathbb{E}[T] = \frac{1}{1-\mu}.$$

It follows from Theorem A2 that the expected offspring of the subcritical dual branching process in our setting is

$$\mathbb{E}[T] = \frac{1}{1-\lambda\sigma\rho_\sigma}.$$

This is remarkably close to our expression for the expected influence, in fact the expected influence is $2\lambda\sigma\rho_\sigma\mathbb{E}[T] - 2\rho_\sigma$. The reason for the additional factor $\lambda\sigma\rho_\sigma$ is that when we choose an edge at random in the network and follow it to a vertex incident with it, the vertex has, in expectation, $\lambda\sigma\rho_\sigma$ "forward" neighbors who are potential adopters. So the expected influence is just $\mathbb{E}[T]$ for each of these neighbors, minus an adjustment factor ρ_σ due to viral inference. The "2" comes from the fact that we have to count both the cases where the seed is in the "giant forward component" or the "finite forward component" (see Section 6.10 for more details).

The reason we can obtain this result via branching processes is that the Erdös-Renyi random graph is intimately related to a branching process with Poisson offspring distribution. We omit the details here, but one can show that $G(n, \lambda/n)$ converges "locally in probability" to a branching process with offspring distribution $\text{Po}(\lambda)$ (see Van der Hofstad, 2023b, Theorem 2.18). In other words, our analysis of $G(n, \lambda/n)$ in terms of Poisson branching processes can be made precise. We take this approach in the following proof, but present it in a slightly more flexible way that also works for graphs with an arbitrary degree distribution.

B Influence

This section provides a rigorous foundation for our definition of influence in finite graphs and its behavior in the limit as $n \to \infty$. The following explanation can be easily adapted to a graph with an arbitrary degree sequence, and even any distribution over values v. In our case, the degree sequence is Poisson and the distribution over values is a point mass.

Fix a graph $G = (V, E)$ on $V = \{1, 2, \ldots, n\}$ vertices with edge set E, and fix a strategy σ which induces a subnetwork of potential adopters $G_a = (V_a, E_a) \subseteq G$. Let $s \in \{1, 2, \ldots, n\}$ be the seed agent, and for each vertex j define $\mathcal{P}_{G_a}(j, s)$ to be the set of all paths in the potential adopter network G_a from j to s. That is,

$$\mathcal{P}_{G_a}(j, s) = \{v_0 v_1, \ldots, v_k \colon k \in \mathbb{N},\, v_i v_{i+1} \in E_a \text{ for all } i, \text{ and } v_i \text{ are distinct}\}.$$

Write $G_a - \{i\}$ for the graph obtained after removing from G_a the vertex i and any edges incident with it. Write $G_a + \{i\}$ for the graph obtained by adding $i \in V$ to the subgraph $G_a \subseteq G$. If $i \in V_a$, then $G_a + \{i\} = G_a$. The influence $\text{Inf}(i) = \text{Inf}_{\{G, G_a, s\}}(i)$ of i in G is defined by[1]

$$\text{Inf}(i) = \{j \in V \colon P_{G_a + \{i\}}(j, s) \neq \emptyset, \text{ and } P_{G_a - \{i\}}(j, s) = \emptyset\}.$$

[1]Influence is defined for any vertex in G and not just in G_a, because agents outside the potential adopter network still consider what their influence *would* be if they were to join the potential adopter network.

That is, the set of all agents $j \in V_a$ for whom every path to s contains i, if i were in the potential adopter network. Now, if $\text{Inf}(i) \neq \emptyset$, then at least one of i's neighbors is in $\text{Inf}(i)$. To see this, note that if every $j \in \text{Inf}(i)$ is not a neighbor of i in $G_a + \{i\}$, then any path from j to s must pass through at least one of i's neighbors. If we let $P = v_0 v_1, \ldots, v_k \in \mathcal{P}(j, s)$ with $v_0 = j$ and $v_k = s$, and $k > 2$, (by the assumption that j is not a neighbor of i) then there is some $\ell < k$ with $v_\ell = i$, and where $v_{\ell-1} \neq j$ is a neighbor of i. But if $v_{\ell-1} \notin \text{Inf}(i)$, then there exists some other path from $v_{\ell-1}$ to s not passing through i. Call this path P_ℓ. Then $P v_{\ell-1} P_\ell$[2] is a path from j to s not containing i, a contradiction of the fact that $j \in \text{Inf}(i)$.

This gives us a useful way to reformulate i's influence. We have $j \in \text{Inf}(i)$, if and only if i is a cutvertex separating a component of agents containing j from a component of agents containing s. This is useful because if i's influence is nonempty, then as we have seen, i has at least one neighbor over whom he has influence, and every agent over which i has influence is connected to one of i's neighbors. Hence for any $j \in \text{Inf}(i)$, we know that j is in a component with at least one of i's neighbors. If follows that if i is connected to the seed, then i's influence is the sum over the component sizes in $G_a - \{i\}$ of each of his neighbors over whom he has influence (being careful not to double count if two of i's neighbors are in the same component). Formally, letting $C_G(j)$ denote the component in G containing j, and letting $N_G(i)$ denote the set of neighbors if i in G, we have

$$\text{Inf}(i) = \begin{cases} \bigcup_{j:\, s \notin C_{G_a - \{i\}}(j)} C_{G_a - \{i\}}(j)(s), & s \in C_{G_a + \{i\}}(i), \\ \emptyset, & s \notin C_{G_a + \{i\}}(i). \end{cases} \tag{B2}$$

In words, the influence set of i is the union over all components containing i's neighbors who are not on some alternate path back to the seed.

We claim that for any fixed n, the expected size of the influence $\text{Inf}(i)$ of any agent i is a polynomial in the adoption probability σ. For any fixed n, if an agent knew the structure of the graph G and the potential adopter subnetwork G_a, he could calculate his influence. In our model, G and G_a are both random. This is because nature draws a graph $G = \mathbb{G}(n, p)$ rather than the graph being fixed. Note that the realization of G does not depend at all on strategies.

Fixing a realization of the graph G, the potential adopter network is determined by n independent Bernoulli experiments, each with success probability σ, and so (as claimed) the probability that any particular potential adopter subgraph is realised is determined by a polynomial in σ. Since agents are aware upon exposure that at least one of their neighbors has adopted, they condition the probability of realising any particular potential adopter network on the knowl-

[2]This notation is standard in Diestel (2000).

edge that at least one of their neighbors is a potential adopter. It follows that the conditional probability is a ratio of positive polynomials in the adoption probabilities σ.

Finally, the seed agent is chosen uniformly at random. Hence for a seed s, a realised graph G and a potential adopter subnetwork G_a, an agent i can calculate their influence. That is, the expected size of the influence of i is well defined given G and G_a. But in our model, strategies are conditional upon exposure, i.e. i has the opportunity to act only if they are in the same component as s. Agents don't know the time period t at which they are exposed, nor do they know which of their neighbors exposed them, as such, upon exposure they condition on the fact that they are connected to the seed in the potential adopter network, and on nothing else. So the quantity we are really interested in is the expected influence conditional on the event $\mathcal{S}_i \equiv \{s \in C_{G_a+\{i\}}(i)\}$. Note that since $\text{Inf}(i) = \emptyset$ on the complement of $\mathcal{S}_i$, the expected influence $\mathbb{E}_\sigma[|\text{Inf}(i)|]$ is identical to $\mathbb{E}_\sigma[|\text{Inf}(i)|\mathbb{1}_{\mathcal{S}_i}]$. Hence for any agent i, ex-ante the expected size of i's influence under the strategy $\sigma \in [0,1]$ and conditional on $\mathcal{S}_i$ is

$$\begin{aligned}\mathbb{E}_\sigma[|\text{Inf}(i)| \mid \mathcal{S}_i] &= \frac{\mathbb{E}_\sigma[|\text{Inf}(i)|\mathbb{1}_{\mathcal{S}_i}]}{\mathbb{P}(\mathcal{S}_i)} \\ &= \frac{\mathbb{E}_\sigma[|\text{Inf}(i)|]}{\mathbb{P}(\mathcal{S}_i)} \\ &= \frac{1}{\mathbb{P}(\mathcal{S}_i)} \sum_G \sum_{G_a} \sum_s |\text{Inf}_{G_a,G,s}(i)| \mathbb{P}(G)\mathbb{P}(G_a \mid G, \sigma)\mathbb{P}(s \mid G_a, G, \sigma), \end{aligned} \tag{B3}$$

where $\mathbb{P}(s \mid G_a, G, \sigma) = \frac{1}{n-1}$ is fixed and independent of the choice of G and s, while each $\mathbb{P}(G_a \mid G, \sigma)$ is polynomial in σ, and each $\mathbb{P}(G)$ is a number in $[0,1]$ determined by the degree distribution. We have written $\text{Inf}_{G,G_a,s}$ to emphasize that the influence of i depends on the graph G, the potential adopter network $G_a \subseteq G$ and the seed s.

The most important consequence of this is that the expected influence upon exposure is a continuous function of σ, since it is a ratio of polynomials in σ with a denominator that has no roots in $[0,1]$ (except at $\sigma = 0$ which is trivial). In the limit as $n \to \infty$ we have to take some care because expected influence can become infinite and continuity in this context just means that around some "critical threshold" we can make influence arbitrarily large.

We now take a closer look at what happens as $n \to \infty$. We have proved that i's influence can be obtained by taking the union over the components containing i's "forward neighbors", that is, those neighbors who are not in the same forward component as the seed. We claim that as $n \to \infty$, in the absence of a giant component every agent i of degree d has influence over $d-1$ of his neighbors with probability $1 - o(1)$.

To prove the above claim, we use a Lemma from from Bollobás & Riordan (2015) which es-

tablishes that the configuration model produces a locally tree-like graph. Given a graph G, let $G_{\leq t}(v)$ denote the subgraph of G induced by the vertices within distance t of v; that is, up to the "t-th neighbours" of v. Let $\mathcal{T} = \mathcal{T}_{\mathcal{D}}$ be a branching process (or a "random rooted tree") on $X_1, X_2, \ldots$ with $X_i \sim \mathcal{D}'$, the forward degree distribution (independently for all i). As with G, let $\mathcal{T}_{\mathcal{D}}|_t$ be the subtree of $\mathcal{T}_{\mathcal{D}}$ induced by the vertices within distance t of the root (that is, the first t generations of the process).

Lemma B3 (Bollobás and Riordan (2015))**.** *Let v be a vertex of $G = \mathbb{G}^{\star}_{\mathbf{d}}$ chosen uniformly at random. Then we may couple the random graphs $G_{\leq t}(v)$ and $T_{\mathcal{D}}|_t$ so that they are isomorphic as rooted graphs with probability $1 - o(1)$ as $n \to \infty$.*

This gives us the following corollary.

Corollary B3. *Let v be a vertex of $G = \mathbb{G}^{\star}_{\mathbf{d}}$ chosen uniformly at random. If $t \geq 1$ is a constant, then w.h.p. the neighbourhood $G_{\leq t}(v)$ of v in G is a tree.*

Suppose that the graph has no giant component of potential adopters, so that all component sizes are finite. Consider choosing a vertex i and random, and suppose i has degree d. Corollary B3 implies that if i is exposed then with probability $1 - o(1)$, i is a cut vertex separating $d - 1$ components– one for each of his neighbors– from the component containing s. This is because by assumption all of i's neighbors are contained in finite components, and so for any $\epsilon > 0$ we can choose a finite size t such that the fraction of vertices in $\text{Inf}(i)$ which are also contained in the tree $G_{\leq t}(v)$ (minus the component containing i and s) is $1 - \epsilon$. It follows that i's expected influence is simply the expected "forward component size" his neighbors who are potential adopters. When agents do not know their own degrees, they compute the expected influence through each of their forward neighbors. In expectation there are $\sigma\mathbb{E}[\mathcal{D}'] = \sigma\lambda$ forward neighbors who are potential adopters. Since each neighbor is (w.h.p.) contained in a separate forward component, i's expected influence under our simplifying assumption is $\sigma\lambda\mathcal{J}_\lambda(\sigma)$ where we define $\mathcal{J}_\lambda(\sigma)$ to be the number of agents contained in the component of *one* of i's potential adopting neighbors, under the strategy σ. Equivalently, we can define $\mathcal{J}_\lambda(\sigma)$ as the expected influence of an agent with one forward neighbor in the potential adopter network. In the limit as $n \to \infty$ we have an explicit expression for the component sizes and forward component sizes (See Callaway et al. (2000) and Newman et al. (2001)), namely

$$\mathcal{I}_\lambda(\sigma) = \lambda\sigma\mathcal{J}(\sigma) = \frac{\lambda\sigma}{1 - \lambda\sigma}. \tag{B4}$$

It remains to consider what happens when there are components of infinite size in the potential adopter network. We would like do define influence analogously to the above– as the expected "forward component size". If the seed is connected to the giant component only on paths passing

through i, then i's influence is infinite—it is the entire giant component. On the other hand, if the seed is connected to the giant component on other paths not passing through i, then i's influence is the *finite* forward component size. This is because all agents not on finite forward components will be exposed independently of the action of i (a different chain of people from the giant component will reach them). Hence when the seed is in the giant forward component, the influence of an agent i of degree d is, in expectation, their number of potential adopting neighbors multiplied by the expected "forward component size" through each of these neighbors, conditional on those components being finite (i.e. the paths dying out). This turns out to be precisely the expectation of the generating function $F_{1,\sigma}(z)$, which we define in Section 6.10 and we show its expectation to be

$$F'_{1,\sigma}(1) = \frac{\rho_\sigma}{1 - G'_{1,\sigma}(1 - \sigma + \sigma\rho_\sigma)}, \tag{B5}$$

where ρ_σ is the forward extinction probability. When $\rho_\sigma = 1$, (B5) coincides with $\mathcal{J}(\sigma)$ from (B4). However, due to viral inference the expected number of potential adopting neighbors is different when strategies are viral. Let $\mathbb{E}[A_i \mid \mathcal{S}_i]$ be the expected degree of a randomly chosen agent in the giant component. Let $\mathcal{F}$ be the event that the seed is in a finite forward component, and let cn be the size of the giant component. Then total expected influence for large n under viral strategies is

$$\begin{aligned}
\mathcal{I}_\theta(\sigma) &= F'_{1,\sigma}(1)\mathbb{E}[A_i \mid \mathcal{S}_i]\,(1 - \mathbb{P}(\mathcal{F})) + cn\mathbb{P}(\mathcal{F}) \\
&= F'_{1,\sigma}(1)\mathbb{E}[A_i \mid \mathcal{S}_i]\left(1 - \frac{F'_{1,\sigma}(1)\mathbb{E}[A_i \mid \mathcal{S}_i]}{cn}\right) + cn\frac{F'_{1,\sigma}(1)\mathbb{E}[A_i \mid \mathcal{S}_i]}{cn} \\
&= F'_{1,\sigma}(1)\mathbb{E}[A_i \mid \mathcal{S}_i]\left(1 - o(\tfrac{1}{n})\right) + F'_{1,\sigma}(1)\mathbb{E}[A_i \mid \mathcal{S}_i] \\
&\to 2F'_{1,\sigma}(1)\mathbb{E}[A_i \mid \mathcal{S}_i],
\end{aligned}$$

as $n \to \infty$. As we show in Section 6.10, the correct expression for expected influence is

$$\mathcal{I}_\theta(\sigma) = 2F'_{1,\sigma}(1)\left(\frac{1}{1 - G_0(1 - \sigma + \sigma\rho_\sigma)}\sum_{k=1}^{\infty} k(1 - \rho_\sigma^k)\frac{\sigma^k}{k!}G_0^{(k)}(1 - \sigma) - 1\right).$$

This concludes our formal discussion of the notion of influence.

C The Degree Distribution

In order for $\lim_{n\to\infty} \mathbf{d}^{(n)}$ to be well-behaved, we assume—as is standard—that there exists a distribution $\mathcal{D}$ with finite expectation and with p.m.f. $\{p_k\}_{k\in\mathbb{N}}$, such that for each $k \in \mathbb{N}$,

$$\lim_{n\to\infty} \frac{n_k(\mathbf{d}^{(n)})}{n} = p_k, \text{ and} \tag{C6}$$

$$\lim_{n\to\infty} \frac{m(\mathbf{d}^{(n)})}{n} = \frac{\mathbb{E}(\mathcal{D})}{2}. \tag{C7}$$

The function $n_k(\mathbf{d}^{(n)})$ is the number of vertices of degree k in $\mathbf{d}^{(n)}$, while $m(\mathbf{d}^{(n)})$ is the number of edges in the graph with degree sequence $\mathbf{d}^{(n)}$.

Under these conditions, all of our quantities of interest (the extinction probability, size of the largest component, etc.) are analytic except at the critical threshold (see Janson, 2009, Theorem 3.11 for details). This justifies (among other things) our implicit differentiation of ρ_σ in the proof of Proposition 2.

D Equilibrium

Although we analyze equilibria in the "limit-game", our pure strategy equilibria exist in any game with sufficiently large n. Our mixed strategy equilibria are ϵ-equilibrium for sufficiently large n. To see this, we focus on the case where a public good diffuses on the network. Note first that for all n, we have $\mathcal{I}_\lambda(0) = 0$, so consider the case where $\sigma = 1$ is a strict pure strategy equilibrium in the limit-game.

Since $\sigma = 1$ is a strict pure strategy equilibrium, it must be that $\mathcal{I}_\lambda(1) > \frac{c}{v} - 1$. But since $\mathcal{I}_\lambda^{(n)}(\sigma)$ converges to $\mathcal{I}_\lambda(\sigma)$ as $n \to \infty$, there must be some N for which $\mathcal{I}_\lambda^{(n)}(1) > \frac{c}{v} - 1$ for all $n \geq N$. Hence $\sigma = 1$ is an equilibrium in all sufficiently large games.

Finally, suppose $\sigma \in (0, 1)$ is a nonviral mixed strategy equilibrium of the limit-game. Then $\mathcal{I}_\lambda(\sigma) = \frac{c}{v} - 1$. Now fix $\epsilon > 0$. As above, for sufficiently large n we must have that $\mathcal{I}_\lambda^{(n)}(\sigma + \epsilon) > \frac{c}{v} - 1$ and $\mathcal{I}_\lambda^{(n)}(\sigma - \epsilon) < \frac{c}{v} - 1$ (alternatively, for viral equilibria the inequalities are reversed). Hence by continuity of $\mathcal{I}_\lambda^{(n)}(\sigma)$ (established in Appendix B), there is a mixed strategy equilibrium in $(\sigma - \epsilon, \sigma + \epsilon)$, i.e., within distance ϵ of σ.

E Proofs: Additional Details

E.1 Lemma 1 as a Corollary of Lemma 2

In general, the expression given in Lemma 2 for the expected number of potential adopting neighbors the simplest we can give because we don't know anything specific about $G_0^{(k)}$. However, for the Poisson distribution we have

$$G_0^{(k)}(z) = \lambda^k G_0(z).$$

Plugging this into the above expression gives

$$\begin{aligned}\sum_{k=1}^{\infty} k(1-\rho_\sigma^k)\frac{\sigma^k}{k!}G_0^{(d)}(1-\sigma) &= G_0(1-\sigma)\sum_{k=1}^{\infty}\left(\frac{\sigma^k}{(k-1)!} - \frac{(\sigma\rho_\sigma)^k}{(k-1)!}\right)\lambda^k \\ &= G_0(1-\sigma)\left(\lambda\sigma e^{\lambda\sigma} - \lambda\sigma\rho_\sigma e^{\lambda\sigma\rho_\sigma}\right)\end{aligned}$$

But now recall that $G_0(z) = e^{\lambda(z-1)}$, so

$$\begin{aligned}\lambda\sigma G_0(1-\sigma)\left(e^{\lambda\sigma} - \rho_\sigma e^{\lambda\sigma}\right) &= \lambda\sigma e^{-\lambda(1-\sigma-1)}\left(e^{\lambda\sigma} - \rho_\sigma e^{\lambda\sigma}\right) \\ &= \lambda\sigma\left(1 - \rho_\sigma e^{\lambda\sigma(\rho_\sigma-1)}\right) \\ &= \lambda\sigma(1-\rho_\sigma^2)\end{aligned}$$

where the last equality comes from the fact that $\rho_\sigma = e^{\lambda\sigma(\rho_\sigma-1)}$. So for the Poisson distribution we arrive at

$$\mathbb{E}[A \mid \mathcal{S}, d] = \frac{\lambda\sigma\left(1-\rho_\sigma^2\right)}{1 - G_0(1-\sigma+\sigma\rho_\sigma)},$$

and finally since the denominator is equal to $1-\rho_\sigma$ we get $\mathbb{E}[A \mid \mathcal{S}, d] = \lambda\sigma(1+\rho_\sigma)$, which proves Lemma 1. In general if $\sigma = 1$, the expression can be simplified to

$$\mathbb{E}[A \mid \mathcal{S}] = \frac{1}{1-G_0(\rho_1)}\sum_{k=1}^{\infty} k(1-\rho_1^k)p_k = \frac{G_0'(1)(1-\rho_1)(1+\rho_1)}{1-G_0(\rho_1)}.$$

E.2 Remainder of Proof of Proposition 2

Consider the limit

$$L \equiv \lim_{\lambda\to 1^+}\frac{\lambda-1}{1-\rho_1}$$

Since the numerator and denominator both approach 0, by L'Hôpital's rule, we have

$$L = \lim_{\lambda \to 1^+} \frac{1}{-\frac{d\rho_1}{d\lambda}} = \frac{1 - \lambda\rho_1}{\rho_1(1-\rho_1)}.$$

A second application of L'Hôpital's rule gives

$$\begin{aligned} L = \lim_{\lambda \to 1^+} \frac{-\rho_1 - \lambda\frac{d\rho_1}{d\lambda}}{\frac{d\rho_1}{d\lambda}(1-2\rho_1)} &= \left(\lim_{\lambda \to 1^+} \frac{1}{-\frac{d\rho_1}{d\lambda}}\right)\left(\lim_{\lambda \to 1^+} \frac{\rho_1}{1-2\rho_1}\right) + \left(\lim_{\lambda \to 1^+} \frac{-\lambda}{1-2\rho_1}\right) \\ L &= L(-1) + 1 \\ \implies L &= \frac{1}{2}. \end{aligned}$$

On the other hand, we have

$$\lim_{\lambda \to 1^+} \lambda\rho_1 = 1 \implies \lim_{\lambda \to 1^+} (1 - \lambda\rho_1)^2 = 0 < \frac{1}{2} = L.$$

This shows that (24) holds as $\lambda \to 1^+$. Next we show that both sides of (24) are strictly increasing. The RHS is immediate, since

$$\frac{d}{d\lambda}(1-\lambda\rho_1)^2 = -2\underbrace{\frac{d(\lambda\rho_1)}{d\lambda}}_{<0}\underbrace{(1-\lambda\rho_1)}_{>0} > 0.$$

For the LHS, we have

$$\frac{d}{d\lambda}\left(\frac{\lambda-1}{1-\rho_1}\right) = \frac{1-\rho_1+\frac{d\rho_1}{d\lambda}(\lambda-1)}{(1-\rho_1)^2},$$

which is positive iff

$$\begin{aligned} & 1 - \rho_1 + (\lambda-1)\left(\frac{\rho_1(\rho_1-1)}{1-\lambda\rho_1}\right) > 0 \\ \iff & 1 > \frac{(\lambda-1)\rho_1}{1-\lambda\rho_1} \\ \iff & \rho_1 < \frac{1}{2\lambda-1}. \end{aligned}$$

To prove the last inequality above it suffices to show that

$$e^{\lambda\left(\frac{1}{2\lambda-1}-1\right)} \leq \frac{1}{2\lambda-1}, \tag{E8}$$

since this would imply that the smallest solution in $[0,1]$ to the equation $\rho_1 = e^{\lambda(\rho_1 - 1)}$ must be at some $\rho_1 < \frac{1}{2\lambda - 1}$. By using the fact that

$$\begin{aligned}\lambda\left(\frac{1}{2\lambda-1}-1\right) = \frac{2\lambda(1-\lambda)}{2\lambda-1} &= \frac{1}{2}\left[\frac{1-(2\lambda-1)^2}{2\lambda-1}\right] \\ &= \frac{1}{2}\left[\frac{1}{2\lambda-1} - (2\lambda-1)\right] \\ &= \frac{1}{2(2\lambda-1)} + \frac{1}{2} - \lambda,\end{aligned}$$

so we can rewrite (E8) as

$$e^{-\lambda}\left(e^{\frac{1}{2(2\lambda-1)}+\frac{1}{2}}(2\lambda-1) - e^{\lambda}\right) \leq 0.$$

At $\lambda = 1$ the LHS of the above inequality is equal to 0, so it suffices to prove that the function

$$f(\lambda) \equiv e^{\frac{1}{2(2\lambda-1)}+\frac{1}{2}}(2\lambda-1) - e^{\lambda}$$

is strictly increasing. To show this, first note that for $\lambda > 1$,

$$2(\lambda-1)^2 > 0 \implies (2\lambda-1)\lambda - 3\lambda + 2 > 0 \implies \lambda > \frac{3\lambda+2}{2\lambda-1}.$$

Next recall that $\log(1+x) < x$ for all $x \neq 0$, so we have

$$\frac{\lambda}{2\lambda-1} + \log\left(1 + \left(1 - \frac{1}{2\lambda-1}\right)\right) < \frac{\lambda}{2\lambda-1} + 1 - \frac{1}{2\lambda-1} = \frac{3\lambda-2}{2\lambda-1} < \lambda.$$

Exponentiating both sides gives

$$e^{\frac{\lambda}{2\lambda-1}}\frac{4\lambda-3}{2\lambda-1} < e^{\lambda}.$$

Finally, some algebraic manipulation shows that

$$\frac{df}{d\lambda} = e^{\lambda} - e^{\frac{\lambda}{2\lambda-1}}\frac{4\lambda-3}{2\lambda-1}.$$

So we see that $\frac{df}{d\lambda} > 0$ iff $e^{\lambda} > e^{\frac{\lambda}{2\lambda-1}}\frac{4\lambda-3}{2\lambda-1}$, which we have proved! This proves (E8) which in turn proves that $\rho_1 < \frac{1}{2\lambda-1}$. Hence

$$\frac{d}{d\lambda}\left(\frac{\lambda-1}{1-\rho_1}\right) > 0$$

for all $\lambda > 1$. To summarise, we have thus far shown that

$$\frac{\lambda - 1}{1 - \rho_1} > (1 - \lambda\rho_1)^2 \text{ for all } \lambda \geq 2 \tag{E9}$$

$$\lim_{\lambda \to 1^+} \left(\frac{\lambda - 1}{1 - \rho_1} \right) = \frac{1}{2} > 0 = \lim_{\lambda \to 1^+} (1 - \lambda\rho_1)^2, \tag{E10}$$

$$\frac{d}{d\lambda} \left(\frac{\lambda - 1}{1 - \rho_1} \right) > 0 \tag{E11}$$

$$\frac{d}{d\lambda} (1 - \lambda\rho_1)^2 > 0. \tag{E12}$$

To complete the proof, we show that $(1 - \lambda\rho_1)^2|_{\lambda=2} < \frac{1}{2}$, which shows that (E9) also holds when $1 < \lambda < 2$. It suffices to prove that at $\lambda = 2$,

$$\rho_1 \geq \frac{\sqrt{2} - 1}{2\sqrt{2}},$$

since this would imply

$$(1 - 2\rho_1)^2 \leq \left(1 - \frac{\sqrt{2} - 1}{\sqrt{2}}, \right) = \frac{1}{2}.$$

But for this it suffices to show that

$$e^{2\left(\frac{\sqrt{2}-1}{2\sqrt{2}} - 1\right)} \geq \frac{\sqrt{2} - 1}{2\sqrt{2}},$$

since by definition of ρ_1 this means $\rho_1 \geq \frac{\sqrt{2}-1}{2\sqrt{2}}$. It is easily verified that

$$e^{2\left(\frac{\sqrt{2}-1}{2\sqrt{2}} - 1\right)} \approx 0.181 > 0.146 \approx \frac{\sqrt{2} - 1}{2\sqrt{2}},$$

and this proves that (E9) for all $\lambda > 1$, which in turn proves that

$$\frac{d\mathcal{I}_\lambda(1)}{d\lambda} < 0$$

for all $\lambda > 1$, as desired. It is also worth noting that as $\lambda \to \infty$, the above argument shows $\lambda\rho_1 \to 0$ and therefore $\mathcal{I}_\lambda(\sigma) \to 0$ for any fixed σ.

E.3 Note on Proposition 5

We used the notation $\frac{C_{1,\pi}(\lambda)}{n}$ in the text to refer to the size of the largest component in the potential adopter network in the largest equilibrium. This is different to the ex-ante expected

size of diffusion, but the comparative statics are the same for both. In particular,

$$\frac{C_{1,\pi}(\lambda)}{n} = \sigma^*(1-\rho_{\sigma^*}),$$

where σ^* is the largest equilibrium under π and given λ, while the ex-ante expected size of diffusion is

$$\frac{C_{1,\pi}(\lambda)}{n} \times \mathbb{P}(\text{seed triggers a large cascade}) = \sigma^*(1-\rho_{\sigma^*})^2,$$

so we see that the relevant comparative statics are identical from the ex-ante perspective.

F Extensions

F.1 A Note on Condition 1

Condition 1 can be weakened to the following.

Condition 2. Let $\sigma \in [0,1]$ be any strategy.

(i) If σ is viral, then $\mathcal{I}_\theta(\sigma)$ is strictly decreasing in σ.

(ii) $\mathcal{I}_\theta(1)$ is strictly decreasing in θ for all $\theta > \theta^{\text{crit}}$.

In Condition 2 we have removed case (i) of Condition 1 since we can show it always holds, and we have replaced the condition that $\mathcal{I}_\theta(\sigma)$ is strictly decreasing in θ for all viral σ, with the condition that this is true at $\sigma = 1$. This amended condition is sufficient for virtually the same equilibrium analysis as we did in the Poisson model, because full adoption of public goods is always an equilibrium around the critical threshold, and (ii) guarantees that as θ get large there is a point where full adoption is no longer an equilibrium.

F.2 Examples

F.2.1 Zipf Distribution

We do everything here in terms of α, but it is straightforward to translate all our results in terms of $\theta = (1-e^\alpha)^{-1}$. In general the generating function for a Zipf distribution is given by

$$G_0(z) = z^{\underline{k}}\Phi(ze^{-\alpha}, 1, \underline{k}).$$

So in the special case where $\underline{k} = 1$, we get

$$G_0(z) = \frac{\ln(1-e^{-\alpha}z)}{\ln(1-e^{-\alpha})}.$$

It follows that

$$G_1(z) = \frac{G_0'(z)}{G_0'(1)} = \frac{1-e^{-\alpha}}{1-e^{-\alpha}z},$$

as claimed. The extinction probability ρ_σ under the strategy σ must satisfy

$$\rho_\sigma = \frac{1-e^{-\alpha}}{1-e^{-\alpha}(1-\sigma+\sigma\rho_\sigma)}.$$

We claim that $\rho_\sigma = \frac{e^\alpha-1}{\sigma}$ solves this. To see this, observe that

$$\rho_\sigma\left(1-e^{-\alpha}(1-\sigma+\sigma\rho_\sigma)\right) = -\rho_\sigma^2\sigma e^{-\alpha} + \left(1-e^{-\alpha}(1-\sigma)\right)\rho_\sigma,$$

and so substituting $\rho_\sigma = \frac{e^\alpha-1}{\sigma}$ we get

$$\begin{aligned}
&-\left(\frac{e^\alpha-1}{\sigma}\right)^2(\sigma e^{-\alpha} + \left(1-e^{-\alpha}(1-\sigma)\right)\frac{e^\alpha-1}{\sigma} \\
&\qquad = \frac{1}{\sigma}\left[-(e^{2\alpha}-2e^\alpha+1)e^{-\alpha} + (e^\alpha-1-1+e^{-\alpha}) + \sigma(1-e^{-\alpha})\right] \\
&\qquad = \frac{1}{\sigma}\sigma(1-e^{-\alpha}) \\
&\qquad = 1-e^{-\alpha},
\end{aligned}$$

and so $\frac{e^\alpha-1}{\sigma}$ is a fixed point of the equation $\rho_\sigma = G_1(1-\sigma+\sigma\rho_\sigma)$. It is in fact the smallest solution and therefore the extinction probability.[3]

Importantly, it follows from our analysis that

$$\sigma\rho_\sigma = e^\alpha - 1.$$

To calculate the expected influence, we first calculate

$$G_1'(z) = \frac{e^{-\alpha}(1-e^{-\alpha})}{(1-e^{-\alpha}z)^2} = \frac{e^\alpha-1}{(e^\alpha-z)^2}.$$

[3]We omit the details here but this follows from the fact that it is the smallest root of the quadratic which solves $\rho_\sigma = G_1(1-\sigma+\sigma\rho_\sigma)$.

So at $z = 1 - \sigma + \sigma\rho_\sigma$,

$$\begin{aligned} G_1'(1-\sigma+\sigma\rho_\sigma) &= G_1'(1-\sigma+(e^\alpha-1)) \\ &= G_1'(e^\alpha-\sigma) \\ &= \frac{e^\alpha-1}{(e^\alpha-e^\alpha+\sigma)^2} \\ &= \frac{e^\alpha-1}{\sigma^2} \\ &= \frac{\rho_\sigma}{\sigma}. \end{aligned}$$

Hence $\sigma G_1'(1-\sigma+\sigma\rho_\sigma) = \rho_\sigma$. The last step is to calculate the expected number of neighbors who are potential adopters. We have

$$G_0^{(k)}(1-\sigma) = G_1^{(k-1)}(1-\sigma)G_0'(1) = -\frac{(k-1)!(e^\alpha-1)}{(e^\alpha-(1-\sigma))^k} \times \frac{1}{(e^\alpha-1)\ln(1-e^{-\alpha})},$$

and so

$$\sum_{k=1}^{\infty} k(1-\rho_\sigma^k)\frac{\sigma^k}{k!}G_0^{(k)}(1-\sigma) = -\frac{1}{\ln(1-e^{-\alpha})}\sum_{k=1}^{\infty} k\left[\left(\frac{\sigma}{e^\alpha-(1-\sigma)}\right)^k - \left(\frac{\rho_\sigma\sigma}{e^\alpha-(1-\sigma)}\right)^k\right],$$

Now,

$$\sum_{k=1}^{\infty}\left(\frac{\sigma}{e^\alpha-(1-\sigma)}\right)^k = \frac{\sigma/(e^\alpha-(1-\sigma))}{\left(\frac{\sigma}{e^\alpha-(1-\sigma)}-1\right)^2} = \frac{1}{1-\frac{\sigma}{e^\alpha-(1-\sigma)}} - 1 = \frac{\sigma}{e^\alpha-1}$$

and similarly,

$$\sum_{k=1}^{\infty}\left(\frac{\sigma\rho_\sigma}{e^\alpha-(1-\sigma)}\right)^k = \frac{\sigma/(e^\alpha-(1-\sigma))}{\left(\frac{\sigma\rho_\sigma}{e^\alpha-(1-\sigma)}-1\right)^2} = \frac{1}{1-\frac{\sigma\rho_\sigma}{e^\alpha-(1-\sigma)}} - 1 = \frac{e^\alpha-1}{\sigma}.$$

Finally, we have

$$\frac{1}{1-G_0(1-\sigma+\sigma\rho_\sigma)} = \frac{1}{1-\frac{\ln(e^{-\alpha}\sigma)}{\ln(1-e^{-\alpha})}} = \frac{\ln(1-e^{-\alpha})}{\ln(e^\alpha-1)-\ln(\sigma)}.$$

Making the substitution $\rho_\sigma = \frac{e^\alpha - 1}{\sigma}$ and putting everything together, we have

$$\begin{aligned}\frac{1}{1 - G_0(1-\sigma+\sigma\rho_\sigma)} \sum_{k=1}^{\infty} k(1-\rho_\sigma^k)\frac{\sigma^k}{k!} G_0^{(k)}(1-\sigma) &= -\frac{1}{\ln(\rho_\sigma)}\left(\rho_\sigma^{-1} - \rho_\sigma\right) \\ &= \frac{1-\rho_\sigma^2}{-\rho_\sigma \ln(\rho_\sigma)}.\end{aligned}$$

Hence the expected influence when σ is viral (making the substitution $\theta = (e^\alpha - 1)^{-1} = (\sigma\rho_\sigma)^{-1}$, and multiplying by 2) is

$$\begin{aligned}\mathcal{I}_\theta(\sigma) = \frac{2\rho_\sigma}{1 - \sigma G_1'(1-\sigma+\sigma\rho_\sigma)}\left[\frac{1-\rho_\sigma^2}{-\rho_\sigma\ln(\rho_\sigma)} - 1\right] &= \frac{2\rho_\sigma}{1-\rho_\sigma}\left[\frac{1-\rho_\sigma^2}{-\rho_\sigma\ln(\rho_\sigma)} - 1\right] \\ &= \frac{2(1+\rho_\sigma)}{-\ln(\rho_\sigma)} - \frac{2\rho_\sigma}{1-\rho_\sigma} \\ &= \frac{2(\theta\sigma+1)}{\theta\sigma\ln(\theta\sigma)} - \frac{2}{\theta\sigma - 1},\end{aligned}$$

which is precisely the expression given in Section 5.2.1. On the other hand if σ is nonviral, then

$$\mathcal{I}_\theta(\sigma) = \frac{\sigma G_1'(\sigma)}{1 - \sigma G_1'(\sigma)} = \frac{\frac{\sigma}{e^\alpha - 1}}{1 - \frac{\sigma}{e^\alpha - 1}} = \frac{1}{\frac{e^\alpha - 1}{\sigma} - 1}.$$

We do not attempt to show that Condition 1 holds but it is evidently true from Figure 7.

F.2.2 Exponential Distribution

The extinction probability is the smallest solution in $[0, 1]$ to the equation

$$\rho_\sigma = \frac{1}{[c - (1-\sigma+\sigma\rho_\sigma)(c-1)]^2}.$$

Using computational software (or otherwise), we find that when $\rho_\sigma < 1$, it is given by

$$\rho_\sigma = \frac{2 + \sigma(c-1)}{2\sigma(c-1)} - \frac{\sqrt{4 + \sigma(c-1)}}{2\sqrt{\sigma(c-1)}}.$$

Notice that the key parameter here is $\sigma(c-1)$. In fact we can rewrite everything in terms of $\mu = \sigma(c-1)$ so that

$$\rho_\sigma = \frac{2+\mu}{2\mu} - \frac{\sqrt{4}+\mu}{2\sqrt{\mu}}.$$

The expected forward degree is given by

$$G_1'(1) = 2(c-1) > 1 \iff c > \frac{3}{2}.$$

Or in terms of α,

$$\frac{1}{1-e^{-\alpha}} > \frac{3}{2} \iff 0 < \alpha < \ln(3).$$

So here, $\alpha = \ln(3)$ is the critical threshold for the emergence of the giant component. Some work shows

$$\sigma G_1'(1-\sigma+\sigma z) = \frac{2(c-1)\sigma}{(c-(c-1)(1-\sigma+\sigma z))^3}.$$

and substituting ρ_σ into this function gives

$$\sigma G_1'(1-\sigma+\sigma\rho_\sigma) = \frac{16}{\sqrt{(c-1)\sigma}(\sqrt{(c-1)\sigma}+\sqrt{4+(c-1)\sigma})^3}.$$

Putting this all together and making the substitution substituting $\mu = \sigma(c-1)$, we have

$$\frac{\rho_\sigma}{1-\sigma G_1'(1-\sigma+\sigma\rho_\sigma)} = \frac{2}{\mu^2+4\mu+\sqrt{\mu}\sqrt{4+\mu}(\mu-2)}.$$

Next we calculate the expected forward degree conditional on exposure under a viral strategy. We have

$$G_0^{(k)}(1-\sigma) = \frac{k!(\theta-1)^{k-1}\theta}{(1+\sigma(\theta-1))^{k+1}}$$

So

$$\begin{aligned}
\sum_k k(1-\rho_\sigma^k)\frac{\sigma^k}{k!}\left[\frac{k!(\theta-1)^{k-1}}{1+\sigma(\theta-1))^{k-1}}\right] &= \frac{\theta\sigma}{(1+\sigma(\theta-1)}\sum_k k(1-\rho_\sigma^k)\left[\frac{\sigma(\theta-1)}{1+\sigma(\theta-1)}\right]^{k-1} \\
&= \frac{\theta\sigma}{(1+\sigma(\theta-1)}\left[\frac{1}{\left(1-\frac{\sigma(\theta-1)}{1+\sigma(\theta-1)}\right)^2} - \rho_\sigma\frac{1}{\left(1-\frac{\sigma(\theta-1)}{1+\sigma\rho_\sigma(\theta-1)}\right)^2}\right] \\
&= \theta\sigma\left[1-\frac{\rho_\sigma}{1+\sigma(\theta-1)(1-\rho_\sigma)}\right].
\end{aligned}$$

Next, we have

$$\frac{1}{1-G_0(1-\sigma+\sigma\rho_\sigma)} = \frac{1+\sigma(1-\rho_\sigma)(\theta-1)}{1+\sigma(1-\rho_\sigma)(\theta-1)-(1-\sigma+\sigma\rho_\sigma)} = \frac{1+\sigma(1-\rho_\sigma)(\theta-1)}{\theta\sigma(1-\rho_\sigma)}.$$

Putting everything together, we have

$$\begin{aligned}\frac{1}{1-G_0(1-\sigma+\sigma\rho_\sigma)}\sum_{k=1}^{\infty}k(1-\rho_\sigma^k)\frac{\sigma^k}{k!}G_0^{(k)}(1-\sigma) &= \frac{1+\sigma(1-\rho_\sigma)(\theta-1)}{\theta\sigma(1-\rho_\sigma)}\cdot\theta\sigma\left[1-\frac{\rho_\sigma}{1+\sigma(\theta-1)(1-\rho_\sigma)}\right] \\ &= \frac{1}{1-\rho_\sigma}\left[1-\sigma(1-\rho_\sigma)(\theta-1)-\rho_\sigma\right] \\ &= 1+\sigma(\theta-1).\end{aligned}$$

Hence subtracting 1 for the neighbor in the giant component, the expected number of potential adopting neighbors under an exponential degree distribution is simply $\sigma(\theta-1)$ as claimed.

F.2.3 Power Law

The expected degree for a power law is

$$\mathbb{E}[\{p_k\}_{k\geq 0}] = \frac{\zeta(\gamma-1)}{\zeta(\gamma)},$$

and the second moment is

$$\frac{\zeta(\gamma-2)}{\zeta(\gamma)}.$$

So the critical percolation threshold is

$$\frac{\frac{\zeta(\gamma-1)}{\zeta(\gamma)}}{\frac{\zeta(\gamma-2)}{\zeta(\gamma)}-\frac{\zeta(\gamma-1)}{\zeta(\gamma)}} = \frac{\zeta(\gamma-1)}{\zeta(\gamma-2)-\zeta(\gamma-1)}.$$

We evaluate the forward distribution and extinction probability numerically using Mathematica given the above and given the closed form expression for the generating function $G_1(z)$ in Section 5.2.3.