Strategic Investment in Protection in Networked Systems[☆]

(Forthcoming in Network Science)

Matt V. Leduc[a,b,1], Ruslan Momot[c,2]

[a]*Stanford University, MS&E, 475 Via Ortega, Stanford, CA 94305-4121, USA*
[b]*International Institute for Applied Systems Analysis (IIASA), Schlossplatz 1, A-2361 Laxenburg, Austria*
[c]*INSEAD, Boulevard de Constance, Fontainebleau, France, 77305*

Abstract

We study the incentives that agents have to invest in costly protection against cascading failures in networked systems. Applications include vaccination, computer security and airport security. Agents are connected through a network and can fail either intrinsically or as a result of the failure of a subset of their neighbors. We characterize the equilibrium based on an agent's failure probability and derive conditions under which equilibrium strategies are monotone in degree (i.e. in how connected an agent is on the network). We show that different kinds of applications (e.g. vaccination, malware, airport/EU security) lead to very different equilibrium patterns of investments in protection, with important welfare and risk implications. Our equilibrium concept is flexible enough to allow for comparative statics in terms of network properties and we show that it is also robust to the introduction of global externalities (e.g. price feedback, congestion).

Keywords: Network Economics, Network Games, Local vs Global Externalities, Cascading Failures, Systemic Risk, Immunization, Airport Security, Computer Security
JEL Codes: D85, C72, L14, Z13

[☆] Both authors thank three anonymous reviewers for the *11th International Conference on Web and Internet Economics, WINE 2015, Amsterdam, The Netherlands, December 9-12, 2015*, where this paper was presented.

Email addresses: mattvleduc@gmail.com (Matt V. Leduc), ruslan.momot@insead.edu (Ruslan Momot)

[1] Matt V. Leduc wishes to thank Matthew O. Jackson, Ramesh Johari, Peter Glynn, Adrien Vigier, Elena Rovenskaya, Alexander Tarasyev and Stefan Thurner, Brian Fath, Yurii Yermoliev, Tatiana Ermolieva. He is grateful to the National Academy of Sciences, the U.S. National Member Organization for the International Institute for Applied Systems Analysis (IIASA), for support through a grant from the U.S. National Science Foundation during IIASA's YSSP 2013 program.

[2] Ruslan Momot thanks Ehud Lehrer (Tel-Aviv University) and Timothy Van Zandt (INSEAD) for invaluable discussions and feedback on the paper. He also wishes to thank Ekaterina Astashkina (INSEAD), Marat Salikhov (INSEAD) and Dmitriy Knyazev (University of Bonn) for their comments and suggestions.

1. Introduction

Many systems of interconnected components are exposed to the risk of cascading failures. The latter arises from interdependencies or interlinkage, where the failure of a single entity (or small set of entities) can result in a cascade of failures jeopardizing the whole system. This phenomenon occurs in various kinds of systems. Well-known examples include 'black-outs' in power grids, where overload redistribution following the failure of a single component can result in a cascade of failures that ripples through the entire grid (e.g. Rosas-Casals *et al.* (2007), Wang *et al.* (2010)). The internet and computer networks also exhibit this phenomenon—one manifestation being the spread of malware (e.g. Lelarge & Bolot (2008b), Balthrop *et al.* (2004)). Likewise, human populations are exposed to the spread of contagious diseases[3].

Studying the incentives to guard against the risk of cascading failures in such interconnected systems has received attention in recent years. In early 2015, a measles epidemic spread across the western part of the United States. It was reported that one of the causes was the unwillingness of parents to vaccinate their children (e.g. The Economist (5 February 2015), The Economist (4 February 2015), Reuters (27 August 2015)). Indeed, some people may want to avoid the perceived risks of a vaccine's side effects and free-ride on the "herd immunity" provided by the vaccination of other people. This raises the following question: what are the incentives to vaccinate against a contagious disease? The same type of question can be asked about other systems subject to the risk of cascading failures. What are the incentives to invest in computer security solutions to protect against the spread of malware? A recent wave of terror attacks within the European union also illustrates the fact that the EU is an interconnected system of many countries. Each member country is thereby exposed to the decisions of other member countries regarding investments in security and intelligence. Indeed, an attacker entering the EU area can reach any location within it. Likewise, what incentives do airports have to invest in security equipment/personnel? How does the structure of interactions between individuals, computers, airports or countries affect those incentives?

There are mainly two streams of literature studying such strategic decisions in interconnected systems. One focuses on the role played by the structure through which agents interact (e.g. a network), while the other focuses on modeling different types of attacks on the system (e.g. random attacks, targeted attacks, strategic attacks).

In the first stream of literature, early work studying games of "interdependent security" (e.g. Heal & Kunreuther (2004) and Heal *et al.* (2006)) considered a broad set of applications ranging from airline security to supply chain management, but did not yet incorporate a complex network interaction structure. More recent work has studied heterogeneous interaction structures. For example, Galeotti & Rogers (2013) consider the problem of a social planner attempting to eradicate an infection from a population. They consider a simple network consisting of two types of agents interacting with others within and across their respective social groups. They then explore the influence of assortativity on the optimal actions of a decision maker. Other papers, like ours, explore the influence of a networked interaction structure on the agents' strategic decisions in more detail. This includes Lelarge & Bolot

[3] For different applications, such as cascading risk in financial systems, see for example Acemoglu *et al.* (2015), Elliott *et al.* (2014).

(2008a) studying the case of strategic immunization and Cabrales *et al.* (2014) exploring the setting of interconnected firms choosing investments in risky projects. More recently, Cerdeiro *et al.* (2015) explored the problem of designing the network topology that provides the proper incentives to the agents.

In the second stream of literature, papers like Dziubiński & Goyal (2016) and Acemoglu *et al.* (2013) explore strategic attack models, in which a defender chooses protection levels, while an attacker chooses the targets in an attempt to maximize the number of affected agents in the network.

In this paper, we develop a framework to study the incentives that agents have to invest in protection against cascading failures in networked systems. A set of interconnected agents can each fail exogenously (fully randomly) or as a result of a cascade of failures[4] (through infected connections). Depending on the application, failure can mean a human being contracting an infectious disease, a computer being infected by a virus or an airport/country being exposed to a security event (e.g. a suspicious luggage or passenger being checked in or being in transit). Each agent must decide on whether to make a costly investment in protection against cascading failures. This investment can mean vaccination, investing in computer security solutions or airport security equipment, to name a few important examples. Strategic decisions to invest in protection are based on an agent's intrinsic failure risk as well as on his belief about his neighbors and their probability of failure. In a complex networked system, forming such a belief can be challenging. For that reason, we employ a solution concept that considerably simplifies how agents reason about the network: agents do not observe the network, but simply know the number of connections they have. This is similar to the equilibrium concept used in Galeotti *et al.* (2010), Jackson & Yariv (2007) and Leduc *et al.* (2015). This equilibrium concept allows us to preserve the heterogeneity of the networked interaction structure (each agent can have a different degree, i.e. a different number of connections) while simplifying the computation of an equilibrium. It also conveniently allows for comparative statics in terms of the network structure (as captured by the degree distribution), as well as other model parameters. This allows us to measure such things as the effect of an increase in the level of connectedness on investments in protection.

We characterize the equilibrium for three broad classes of games: (i) *games of total protection*, in which agents invest in protection against *both* their intrinsic failure risk and the failure risk of their neighbors; (ii) *games of self protection*, in which agents invest in protection *only* against their intrinsic failure risk; and (iii) *games of networked-risk protection*, in which agents invest in protection *only* against the failure risk of their neighbors. The first and third classes define games of strategic substitutes, in which some agents free-ride on the protection provided by others. Applications covered by these classes of games include vaccination and standard computer security solutions (e.g. anti-virus). The second class defines a game of strategic complements, in which agents pool their investments in protection and this can result in coordination failures. Applications covered by this class of games include airport security, border security within the European union and other types of computer security solutions (e.g. two-factor authentication (2FA)).

[4] Similar random failure mechanisms are studied in Lelarge & Bolot (2008b), Goyal & Vigier (2015), Aspnes *et al.* (2005), Blume *et al.* (2013) and Acemoglu *et al.* (2013).

Another of our contributions is to analyze the effect of the network structure on equilibrium behavior in those three classes of games. For example, in the case of vaccination, it is the agents who have *more* neighbors than a certain threshold who choose to vaccinate and the agents who are less connected who free-ride. The more connected agents thus bear the burden of vaccination, which can be seen as a positive outcome. In the case of airport security, on the other hand, it is agents who have *fewer* neighbors than a certain threshold who choose to invest in security equipment/personnel. Since the less connected airports are less likely to act as hubs that can transmit failures, this can be seen as an inefficient outcome. To our knowledge, we are the first to explicitly characterize such features, which are the consequence of network structure and can have important policy and welfare implications.

Finally, we study the case when the cost of protection is endogenized and allowed to depend on global demand. For instance, the price of vaccines or computer security solutions may increase (e.g. vaccines may be produced in limited supplies) if demand increases. It is important to understand the impact that this may have on agents' behavior as the introduction of such a *global externality* (e.g. see the global congestion case in Arribasa & Urbanoa (2014)) may conflict with the cascading failure process affecting an agent through his local connections. We characterize the equilibrium after introducing this price feedback and show that the results derived previously still hold with minor changes.

Acemoglu *et al.* (2013) and Lelarge & Bolot (2008b) are perhaps the closest work to ours. The former paper, in a setting similar to ours, shows that under random and targeted attacks both over- and underinvestment (as compared to the socially optimal level) are possible. Furthermore, the authors show that optimal investment levels are defined by network centrality measures, whereas our characterization of equilibrium investment is based on degree centrality. Additionally, we further explore the role of the network structure in defining agents' incentives to invest in protection. In particular, we study comparative statics by varying the degree distributions of the underlying network. Lelarge & Bolot (2008b) also consider different types of protection against contagion risk in trees and sparse random graphs. As compared to their probabilistic approach, the equilibrium concept we use allows for a characterization of behavior in terms of an agent's degree. We also deal with a common (possibly endogenized) cost of investment as opposed to their randomized costs. Finally, our paper contributes to the rapidly expanding stream of literature on games on networks[5].

The paper is organized as follows: Section 2 introduces the concept of cascading failures in networked systems. Section 3 develops the game theoretic framework that allows us to study the problem in a tractable way while imposing a realistic cognitive burden on agents. Section 4 characterizes the equilibrium for the three broad classes of games previously mentioned. Implications for risk and welfare are discussed. Comparative statics results in terms of the network structure (as captured by the degree distribution) and other model parameters are also presented. An extension in which the cost of protection is endogenized is also studied. Section 5 concludes with a critical evaluation of our model and a discussion of possible extensions. For clarity of exposure, all the proofs are relegated to an appendix.

[5] The reader is referred to Jackson & Zenou (2014) for a survey of the existing literature on games on networks.

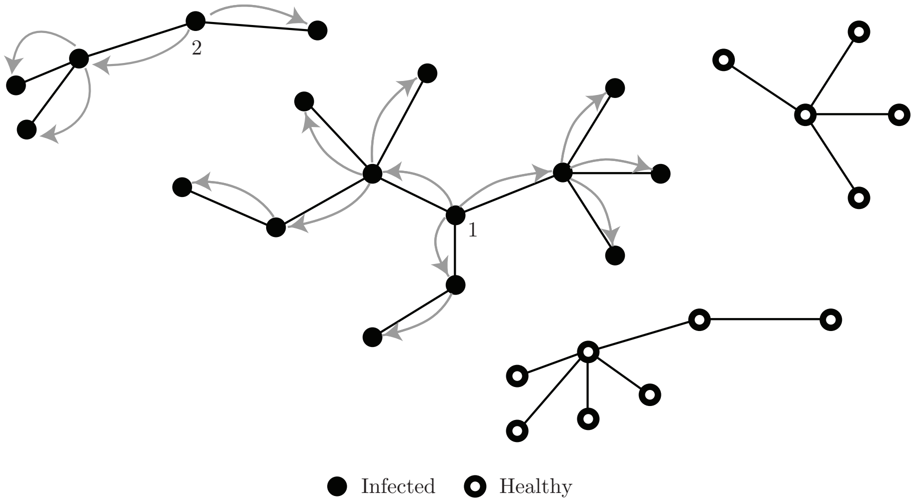


Figure 1: Example of a Contagion Cascade: individuals labeled 1 and 2 contract the disease from exogenous sources. From then on, a contagion cascade takes place in discrete steps: all their neighbors become infected. This then leads to their neighbors' neighbors to become infected and so on.

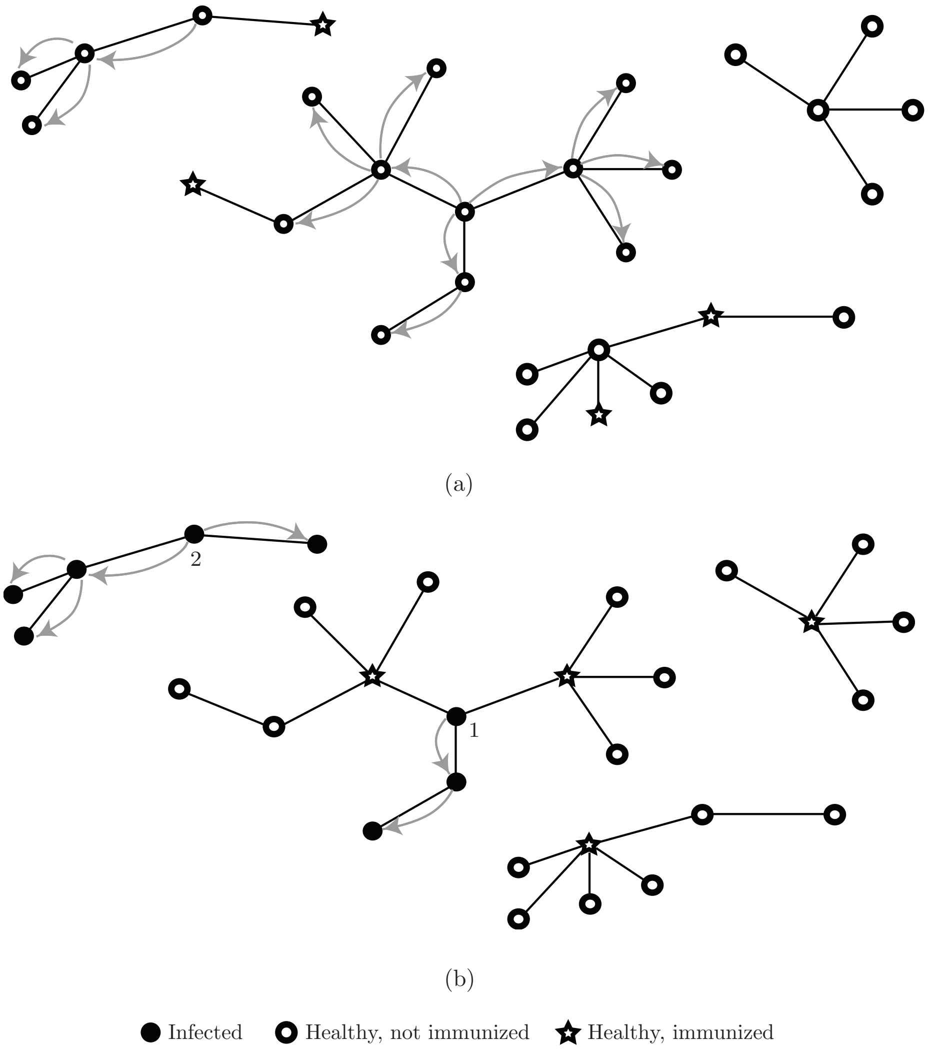
2. Cascading Failures in Networks

2.1. Overview

In this section we will discuss how cascades of failures can propagate through networks. A *cascade of failures* is defined as a process involving the subsequent failures of interconnected components. A *failure* is a general term that may represent different kinds of costly events. Let us consider, for example, the spread of a disease in a human population. Initially, some individuals get infected through exogenous sources such as livestock, mosquitos or the mutation of a pathogen. These individuals can then transmit the disease through contacts with other humans. Let us suppose that an individual is sure to catch the disease if one of his neighbors is infected. Figure 1 illustrates this process. We can see the impact of network structure on contagion. Some people lying in certain components remain healthy whereas others are infected by their neighbors. We also see that individuals with a high number of contacts tend to facilitate contagion. This is a simplified model of contagion. A more realistic model could, for example, transmit the disease only to randomly selected neighbors, depending on its virulence.

Now let us imagine that some individuals are vaccinated and therefore are not susceptible to becoming infected, neither by exogenous sources nor by contacts with other people. This will have an impact on the cascading process. Indeed, it will effectively 'cut' certain contagion channels, thereby impeding the spread of the disease. Figure 2 illustrates this. We see that the importance of the network structure becomes even more striking. In Fig. 2a), immunized individuals have been selected randomly, whereas in Fig. 2b) individuals with 4 or more contacts have been immunized. It is clear that those more connected individuals often act as hubs through which contagion can spread more easily. When these individuals are immunized, the effect of impeding the propagation of the disease tends to be much greater than when the immunized individuals are chosen at random.

In this example, the 'failure' of an individual means he becomes infected by the disease. In other applications, 'failure' can mean infection by malware. The nodes then no longer



(a)

(b)

Figure 2: Examples of Contagion Cascades in the Presence of Immunized Individuals: individuals labeled 1 and 2 contract the disease from exogenous sources. The contagion cascade then propagates. In part (a), a randomly-chosen subset of agents were vaccinated against the disease. In part (b), individuals with at least 4 contacts were vaccinated against the disease.

represent individuals but computers (or local subnetworks or autonomous systems). Antivirus software or other sorts of computer security solutions are means by which the spread of malware can be impeded.

We saw in the simple example of Fig. 2 that the configuration of the vaccinated nodes was crucial to impeding contagion. An important question is to study the incentives that an individual may have to become vaccinated. How does the network structure affect his decision to become vaccinated? What roles other individuals play in influencing that decision through their own vaccination behavior?

Given the range of applications, we will talk of an *investment in protection.* This refers to an investment made by a node in order to protect itself against the risk of failure. In the next section, we build a model of strategic investment in protection against cascading failures in networked systems. We will refer to nodes as *agents*, since they make decisions regarding this investment in protection. More generally, we will be interested in how the network structure and the failure propagation mechanism influence those decisions through the externalities that they generate.

2.2. Network

A network, as the one described previously, can be formally defined as follows. There is a set of nodes (or agents) $\mathcal{N} = \{1, 2, ..., n\}$. The connections between them are described by an *undirected* network that is represented by a symmetrical adjacency matrix $g \in \{0,1\}^{n\times n}$, with $g_{ij} = 1$ implying that i and j are connected. i can thus be affected by the failure of j and vice versa. By convention, we set $g_{ii} = 0$ for all $i \in \mathcal{N}$. The network realization g is drawn from the probability measure $P : \{0,1\}^{n\times n} \rightarrow [0,1]$ over the set of all possible networks with n nodes. We assume that P is permutation-invariant, i.e. that changing node labels does not change the measure. Each agent i has a neighborhood $N_i(g) = \{j|g_{ij} = 1\}$. The *degree* of agent i, $d_i(g)$, is the number of i's connections, i.e. $d_i(g) = |N_i(g)|$.

3. A Bayesian Network Security Game

3.1. Informational Environment

We study an informational environment similar to the one presented in Galeotti *et al.* (2010). Agents are aware of their proclivity to interact with others, but do not know who these others will be when taking actions. Formally, this means that an agent knows only his degree d_i. For example, a bank may have a good idea of the number of financial counter-parties it has but not the number of counter-parties the latter have, let alone the whole topology of the interbank system. In applications to the spread of contagious diseases, an individual may know the number of people he interacts with, but not the number of people the latter interact with. Likewise, in the case of an email network, someone may know the number of contacts he has, but not the number of contacts his contacts have.

First, since P is permutation invariant (cf. Section 2.2), we can define the *degree distribution* of P as the probability a node has degree d in a graph drawn according to P; we denote the degree distribution[6] by $f(d)$ for $d \geq 1$. Note that we are not interested in

[6] Throughout, we use the term *degree distribution* to mean *degree density.* When referring to the *cumulative distribution function (CDF)*, we will do so explicitly.

modeling agents of degree 0 (since they do not play a game) and we therefore always assume that $f(0) = 0$. We assume a countably infinite set of agents. An agent's type is his degree d and it is drawn i.i.d. according to the degree distribution $f(d)$. Likewise, the degree of each of an agent's neighbors is drawn i.i.d. according to the density function $\tilde{f}(d)$. This is the *edge-perspective degree distribution* and can be written as $\tilde{f}(d) = \frac{f(d)d}{\sum_{d' \geq 1} f(d')d'}$. This expression follows from a standard calculation in graph theory (see Jackson (2008) for more details). $\tilde{f}(d)$ is the probability that a neighbor has degree d. It therefore takes into account the fact that a higher-degree node has a higher chance of being connected to any agent and thus of being his neighbor. Thus agents reason about the graph structure in a simple way through the degree distribution.

3.2. Action Sets and Strategies

In order to protect himself against the risk of failure, we allow an agent i to make a costly investment in *protection*. This is a *one-shot investment* that can be made *in anticipation* of a cascade of failures, which may take place in the future. This investment in protection is represented by an action a_i, which is part of a binary action set $\mathcal{A} = \{0, 1\}$. The latter represents the set of possible investments in protection against failure: $a_i = 1$ means that the agent invests in protection while $a_i = 0$ means that the agent remains unprotected. In an application to computer security, a_i can represent an investment in computer security solutions or anti-virus software. In applications to disease spread, a_i can represent vaccination, whereas in the case of airport security, a_i can represent an investment in security personnel or equipment. We assume throughout that $\mathcal{A}$ is the same for all agents. The exact effect of this action on an agent's actual failure risk will be formalized later in Definition 5.

Note that all agents have access to the same information about the network (only its degree distribution $f(d)$). An agent does not know his position in the network, only the number of neighbors he has (an agent's degree is his type). An agent i's behavior is thus governed only by his degree d_i and not by his label i. We can then define a strategy in the following way.

Definition 1. *A strategy $\mu : \mathbb{N}^+ \to [0, 1]$ is a scalar-valued function that specifies, for every $d > 0$, the probability that an agent of degree d invests in protection. We denote by $\mathcal{M}$ the set of all strategies.*

Thus $\mu(d)$ is the symmetric mixed strategy played by an agent of degree d. Note that $\mathcal{M} = [0, 1]^\infty$, the space of $[0, 1]$-valued sequences. Throughout, we endow $\mathcal{M}$ with the product topology and $[0, 1]$ with the Euclidean topology.

3.3. Failure Probabilities and Utility Functions

We start with the following definition:

Definition 2. *An agent's intrinsic failure probability is denoted by $p \in [0, 1]$.*

We thus assume all agents can fail intrinsically with the same probability p. The interpretation of intrinsic failure depends on the application. In the context of malware, intrinsic failure means a computer becoming infected as a result of a direct hacking attack. In the context of the spread of contagious diseases, intrinsic failure means being infected by a virus through non-human sources, such as contact with livestock or insects. In the context of airport security, intrinsic failure can mean a suspicious luggage being checked in at the airport.

We now state a property of this network security game, which addresses how an agent reasons about the failure probability of his neighbors.

Property 1. *Each agent conjectures that each of his neighbors fails with probability $\mathcal{T}(\mu) \in [0, 1]$, independently across neighbors.*

This setting is similar to that of Jackson & Yariv (2007), where each neighbor adopts a product or an opinion with a certain probability that depends on the strategy played by the population. Note that the dependence of a neighbor's failure probability $\mathcal{T}(\mu)$ on the strategy μ played by other agents was made explicit. An agent's cascading failure probability can now be defined in terms of $\mathcal{T}(\mu)$, as seen in the following definition.

Definition 3. *For any d, let the function $q_d : [0, 1] \to [0, 1]$ denote a degree-d agent's cascading failure probability, i.e. $q_d(\mathcal{T}(\mu))$ is the probability that an agent of degree d will fail as a result of a cascade of failures, given that his neighbors each fail independently with probability $\mathcal{T}(\mu)$. For any d, $q_d(\mathcal{T}(\mu))$ is strictly increasing and continuous in $\mathcal{T}(\mu)$. Moreover, we explicitly set $q_0(\mathcal{T}(\mu)) = 0$ and thus an agent with no neighbor cannot fail as a result of a cascade of failures.*

The actual expression for $q_d(\mathcal{T}(\mu))$ depends on the type of cascade we are considering. We will consider only a situation where $\{q_d\}_d$ is an increasing sequence of functions. That is, when $d' > d$, then $q_{d'}(\mathcal{T}(\mu)) > q_d(\mathcal{T}(\mu))$ for any $\mathcal{T}(\mu) \in [0, 1]$. In other words, the cascading failure risk is higher when an agent has more connections[7]. For convenience, we will sometimes write $q_d(\mathcal{T}(\mu))$ simply as q_d.

Since an agent of degree d either fails intrinsically with probability p or in a cascade with probability q_d, we can define his *total probability of failure* as follows.

Definition 4 (Total probability of failure)**.** *The total probability of failure of an agent of degree d is*

$$\beta_d = p + (1 - p)q_d. \tag{1}$$

Thus an agent can either fail intrinsically (i.e. by himself) or as a result of the failures of a subset of his neighbors. Those neighbors who have failed may have done so intrinsically or as a result of the failure of a subset of their own neighbors.

We study a static setting, in which agents make decisions simultaneously, *in anticipation* of a cascade of failures that may happen in the future. Therefore each agent is *healthy* when he chooses an action $a \in \mathcal{A}$ representing a costly investment in protection against failure. This is a good decision model for the applications that we cover. E.g., vaccines are taken by healthy individuals in anticipation of an epidemic that may spread in the future. Likewise, investments in computer security solutions are taken for healthy computers or autonomous systems in anticipation of the spread of malware, which may take place at a later date.

[7] The reader is referred to Chapter 4 of Leduc (2014) for the case where $q_d(\mathcal{T}(\mu))$ is decreasing in d. This can model a form of diversification of failure risk across neighbors.

Similar long-term security decisions are taken in other contexts, such as airport security, for example.

We now describe how this action affects an agent's failure probability.

Definition 5. *Let the mapping $\mathcal{B} : [0,1] \times [0,1] \times \mathcal{A} \to [0,1]$ denote the effective failure probability of an agent. We assume that $\mathcal{B}(p, q_d, a)$ is continuous in all arguments, increasing in p and in q_d and that it is decreasing in a.*

Thus, $\mathcal{B}(p, q_d, a)$ is the total failure probability of an agent (defined in (1)) when he has invested a in protection against failure. Note that this definition allows this action to operate separately on p and q_d, as will be seen in Section 4. This will become useful as we study different kinds of protection. We can now state an agent's expected utility function, which will capture his decision problem.

A degree-d agent's expected utility function is given by

$$U_d(a, \mu) = -V \cdot \mathcal{B}(p, q_d(\mathcal{T}(\mu)), a) - C \cdot a. \tag{2}$$

where $C > 0$ is the cost of investing in protection, $V > 0$ is the value that is lost in the event of failure and $\mathcal{B}(\cdot, \cdot, \cdot)$ is the *effective failure probability* (cf. Definition 5).

This utility function captures the tradeoff between the expected loss $V \cdot \mathcal{B}(p, q_d(\mathcal{T}(\mu)), a)$ and the cost[8] C of investing in protection. Notice again that an agent's expected utility depends on the actions of others only through the cascading failure probability $q_d(\mathcal{T}(\mu))$, since they will affect the probability of failure $\mathcal{T}(\mu)$ of a randomly-picked neighbor. Note also that the expected utility function[9] $U_d(\cdot, \cdot)$ depends on the agent's degree d but not on his identity i. Therefore, any two agents i and j who have the same degree have the same expected utility function. From the assumptions on $\mathcal{B}$, U_d is continuous in all arguments. An agent is risk-neutral and will thus maximize this expected utility function by choosing the appropriate action a. The game thus models security decisions under contagious random attacks in a network where each agent (node) knows only his own degree and the probability that a neighbor has a certain degree.

While the cascading failure probability q_d can take many functional forms, we provide several examples which can all be modeled using the particular form $q_d(\mathcal{T}(\mu)) = 1-(1-r\mathcal{T}(\mu))^d$. This functional form results from a contact process.

Malware or Virus Spread: Let a computer be infected by a direct hacking attack with probability p. Assume that malware (i.e. computer viruses) can spread from computer to computer according to a general contact process: if a neighbor is infected, then the computer will be infected with probability r. If each neighbor is infected with probability $\mathcal{T}(\mu)$ and this infection spreads independently across each edge with probability r, then $q_d(\mathcal{T}(\mu)) = 1 - (1 - r\mathcal{T}(\mu))^d$. This contact process can also serve as a model for the spread

[8] The cost of investing in protection may represent the price of airport security equipment or computer security solutions. It may also represent the possible side-effects that may be associated with a vaccine (e.g. The Economist (4 February 2015)).

[9] Note that we could write a degree-d agent's expected utility function as $U(a, \mu, d)$. We write it with d as a subscript simply because it is a convenient notation.

of viruses among human populations. In this case, p is the probability of being infected by non-human sources (e.g. insects, livestock, etc.) and $q_d(\mathcal{T}(\mu))$ is the probability of being infected by neighbors (i.e. other persons with whom the agent interacts). The parameter r models the virulence or infectiousness of the process: given that a neighbor is infected, r is the probability[10] that he will infect the agent.

Airport and European Union Security: The contact process described above can also be applied to airport or EU security. The exogenous failure (with probability p) can mean a security event such as the failure to stop a suspicious luggage from being checked in on a flight or a terrorist entering the European union from outside through one of the EU countries with weaker border control. In these scenarios, the agents represent airports or countries, and the edges linking them represent flights and connecting routes between countries. The suspicious luggage or terrorist can then cascade, i.e. travel to one or more other airports/countries, exposing them to security risks. $q_d(\mathcal{T}(\mu)) = 1-(1-r\mathcal{T}(\mu))^d$ can then model the risk of an entity coming into contact with a security threat coming from a neighboring country or airport.

In the next two sections we develop both the optimal response of an agent to the environment described previously, as well as the consistency check that $\mathcal{T}(\mu)$ should satisfy given the strategic choices of the agents.

3.4. Consistency

We will now develop a consistency check that guarantees that a randomly-picked neighbor's failure probability $\mathcal{T}(\mu)$ is consistent with the strategy μ played by the population.

Definition 6. *Let the function $\mathcal{F} : \mathcal{M} \times [0,1] \to [0,1]$ be defined as*

$$\mathcal{F}(\mu, \alpha) = \sum_{d \geq 1} \tilde{f}(d)\mathcal{B}(p, q_{d-1}(\alpha), \mu(d)). \tag{3}$$

In the above definition[11], $\mathcal{F}(\mu, \alpha)$ is the failure probability of a randomly-picked neighbor given that agents play strategy μ and this neighbor's other neighbors fail with probability α. A fixed point $\alpha = \mathcal{F}(\mu, \alpha)$ ensures that α is the same across all agents and consistent with μ. We consider $\mathcal{F}(\mu, \alpha)$ with the following property:

Property 2. *For any $\mu \in \mathcal{M}$, $\mathcal{F}(\mu, \alpha)$ has a unique fixed point in α.*

Note that Property 2 is not particularly stringent. It is easy to verify in the contact process models of the examples described in Section 3.3.

We can now formally define $\mathcal{T}(\mu)$, the failure probability of a randomly-picked neighbor given that strategy μ is played by other agents:

[10] In Fig. 1 and Fig. 2, r was assumed to be 1 for simplicity of exposure.

[11] Note that an agent does not internalize the effect of his own failure on others when forming his belief about the failure risk of a neighbor. Hence the presence of $q_{d-1}(\alpha)$ on the right-hand side of (3) instead of $q_d(\alpha)$: the cascading failure risk of a given neighbor of degree d is only due to his $d-1$ other neighbors.

Definition 7. *Given $\mathcal{F} : \mathcal{M} \times [0,1] \to [0,1]$ satisfying Property 2, let $\mathcal{T} : \mathcal{M} \to [0,1]$ be defined as follows: for any $\mu \in \mathcal{M}$,*

$$\mathcal{T}(\mu) = \mathcal{F}(\mu, \mathcal{T}(\mu)). \tag{4}$$

3.5. Optimal Response

It is now straightforward to solve for the optimal strategy of an agent of degree d: an agent invests in protection, does not invest, or is indifferent if $U_d(1,\mu)$ is greater than, less than, or equal to $U_d(0,\mu)$, respectively. We thus have the following definition.

Definition 8. *Let $\mathcal{S}_d(\mathcal{T}(\mu)) \subset [0,1]$ denote the set of optimal responses for a degree-d agent given $\mathcal{T}(\mu)$; i.e.:*

$$\begin{aligned} U_d(1,\mu) > U_d(0,\mu) &\implies \mathcal{S}_d(\mathcal{T}(\mu)) = \{1\}; \\ U_d(1,\mu) < U_d(0,\mu) &\implies \mathcal{S}_d(\mathcal{T}(\mu)) = \{0\}; \\ U_d(1,\mu) = U_d(0,\mu) &\implies \mathcal{S}_d(\mathcal{T}(\mu)) = [0,1]. \end{aligned}$$

We can now let $\mathcal{S}(\mathcal{T}(\mu)) \subset \mathcal{M}$ denote the set of optimal strategies given $\mathcal{T}(\mu)$; i.e.,

$$\mathcal{S}(\mathcal{T}(\mu)) = \prod_{d \geq 1} \mathcal{S}_d(\mathcal{T}(\mu)).$$

Note that at least one optimal response always exists and is essentially uniquely defined, except at those degrees where an agent is indifferent.

3.6. Equilibrium

We now formally define the equilibrium concept and state our first proposition.

Definition 9 (Mean-Field Equilibrium). *A strategy μ^* constitutes a mean-field equilibrium (MFE) if $\mu^* \in \mathcal{S}(\mathcal{T}(\mu^*))$.*

This equilibrium definition ensures that both the optimality and consistency conditions are satisfied. Also note that to any equilibrium μ^*, there corresponds a unique equilibrium neighbor failure probability $\alpha^* = \mathcal{T}(\mu^*)$.

Proposition 1 (Existence). *Any network security game that satisfies Properties 1 and 2 has a mean-field equilibrium.*

An MFE is a symmetric equilibrium with the property that an agent's neighbors fail independently with the same probability $\mathcal{T}(\mu^*)$ under μ^*. An MFE is particularly easy to compute. In fact, $\alpha^* = \mathcal{T}(\mu^*)$ is obtained from a one-dimensional fixed-point equation resulting from the composition of $\mathcal{T}$ and $\mathcal{S}$, i.e. $\alpha^* = \mathcal{T}(\mathcal{S}(\alpha^*))$. μ^* is then found from the map $\mathcal{S}(\alpha^*)$ (cf. Definition 8). Allowing for correlations between the failures of neighbors would considerably complicate the analysis[12].

[12] For some work in that direction, see Chapter 3 of Leduc (2014).

4. Characterizing Equilibria

In this section, we will study three classes of games in which agents make decisions to invest in protection. We will start with games of *total protection*, in which an agent's investment decreases his total risk of failure. We will then proceed with games of *self protection*, in which an agent's investment in protection only protects him against his own intrinsic risk of failure. We will finally study an intermediate case: a game of *networked-risk protection*, in which an agent's investment in protection only protects him against the risk of failure of his neighbors

4.1. Games of Total Protection

In games of *total protection*, the investment protects both against the intrinsic failure risk and the cascading failure risk.

Examples of applications covered by this class are the spread of contagious diseases and the decision to vaccinate or malware and the investment in anti-virus or computer security solutions. Vaccination, for example, protects against both the risk of being infected by non-human (intrinsic failure risk) and human sources (cascading failure risk). It is also the case for standard anti-virus software featuring a firewall protection. This protects an agent against both direct hacking attacks (intrinsic failure risk) and malware spread through the Internet/e-mail networks (cascading failure risk).

We have the following definition:

Definition 10 (Games of total protection)**.** *In a game of total protection, the effective failure probability has the following form*

$$\mathcal{B}(p, q_d(\mathcal{T}(\mu)), a) = \Big(p + (1-p)q_d(\mathcal{T}(\mu))\Big) \cdot (1 - ka) \tag{5}$$

for some $k \in [0,1]$ *and*

$$\mathcal{F}(\mu, \alpha) = \sum_{d \geq 1} \tilde{f}(d)\Big(p + (1-p)q_{d-1}(\alpha)\Big) \cdot (1 - k\mu(d)). \tag{6}$$

In games of total protection, as can be seen in (5), an agent's investment in protection decreases his total probability of failure $p + (1-p)q_d(\mathcal{T}(\mu))$. The parameter k governs the effectiveness of the investment in protection. The higher k, the more an investment in protection reduces the failure probability.

Before stating our first theorem, we introduce the following definition.

Definition 11 (Upper-threshold strategy)**.** *A strategy* μ *is an upper-threshold strategy if there exists* $d_U \in \mathbb{N}^+ \bigcup \{\infty\}$*, such that:*

$$\begin{aligned} d < d_U &\implies \mu(d) = 0; \\ d > d_U &\implies \mu(d) = 1. \end{aligned}$$

Thus, under an upper-threshold strategy, agents with degrees *above* a certain threshold invest in protection whereas agents with degrees *below* that threshold do not invest. Note that the definition above does not place any restriction on the strategy *at* the threshold d_U itself; we allow randomization at this threshold.

Games of total protection are *submodular*. In other words, they are of *strategic substitutes*: the more other agents invest in protection (the lower $\mathcal{T}(\mu)$), the less an agent has an incentive to invest in protection. A nice property of games of total protection is that they have a unique equilibrium that is characterized by an upper-threshold strategy.

Theorem 1 (Total Protection)**.** *In a game of total protection, the equilibrium μ^* is unique. Moreover, μ^* is an upper-threshold equilibrium, i.e. μ^* is an upper-threshold strategy.*

The intuition behind this result is that, higher-degree agents are more exposed to cascading failures than lower-degree agents, thus making an investment in *total* protection relatively more rewarding. The implications of this theorem are important as higher-degree agents are more likely to act as hubs though which contagion can spread. This result can thus be seen as a satisfactory outcome since more connected agents have higher incentives to internalize the risk they impose on the system. In equilibrium, the total cost of protection is thus born by those who have a maximal effect on decreasing $\mathcal{T}(\mu)$. For example, in the case of malware, agents with a higher level of interaction (higher degree) have a higher incentive to invest in computer security (i.e. anti-virus software). The same principle applies in the case of human-born viruses: individuals who interact more have a higher incentive to get vaccinated.

Note that in spite of the above, agents tend to underinvest in equilibrium compared to the socially optimal investment level. This is the result of free-riding and is in line with classical results of moral hazard in economics and the failure of agents to take into account negative externalities.

In the next section, we study the second class of games: Games of *self protection*.

4.2. Games of Self Protection

In games of *self protection*, the investment protects only against the intrinsic failure risk.

Examples of applications covered by this class of games include airport security when luggage/passengers are only scanned at the originating airport. Airports then otherwise rely on each other's provision of security for transiting passengers/luggage. The same principle applies to security within the European Union, where travelers are only inspected at their point of entry. EU countries otherwise rely on each other's security for travelers within the EU.

Another important example is two-factor authentication (2FA) in computer networks. Consider an e-mail network and a provider such as Gmail. The latter allows its users to use such a two-factor authentication (2FA) feature. Users who take advantage of this option are asked to enter a security code sent to their mobile phone in addition to their password entered upon authentication. 2FA thus effectively protects against direct hacking attacks (a user's personal intrinsic risk). Indeed, access to the account with 2FA can only be granted conditional on the user having access to the mobile phone linked to this account. Yet, 2FA does not diminish the user's exposure to cascading failure risk (i.e. malware transmitted through the e-mail network): carelessly opening an infected e-mail attachment from a friend can fully compromise his account.

We now have the following definition:

Definition 12 (Games of self protection)**.** *In a game of self protection, the effective failure probability has the following form*

$$\mathcal{B}(p, q_d(\mathcal{T}(\mu)), a) = p \cdot (1 - ka) + (1 - p \cdot (1 - ka)) \cdot q_d(\mathcal{T}(\mu)) \tag{7}$$

for some $k \in [0, 1]$ *and*

$$\mathcal{F}(\mu, \alpha) = \sum_{d \geq 1} \tilde{f}(d) \Big(p \cdot (1 - k\mu(d)) + \big(1 - p \cdot (1 - k\mu(d))\big) \cdot q_{d-1}(\alpha) \Big). \tag{8}$$

In games of self protection, as can be seen in (7), an agent's investment in protection only decreases his intrinsic probability of failure p. It has no effect on his cascading failure probability $q_d(\mathcal{T}(\mu))$. Again, the parameter k governs the effectiveness of the investment in protection corresponding to the action a.

Before stating our second theorem, we introduce the following definition.

Definition 13 (Lower-threshold strategy)**.** *A strategy* μ *is a lower-threshold strategy if there exists* $d_L \in \mathbb{N}^+ \bigcup \{\infty\}$*, such that:*

$$\begin{aligned} d > d_L &\implies \mu(d) = 0; \\ d < d_L &\implies \mu(d) = 1. \end{aligned}$$

Under a lower-threshold strategy, agents with degrees *below* a certain threshold invest in protection whereas agents with degrees *above* that threshold do not invest. Note that the definition above does not place any restriction on the strategy *at* the threshold d_L itself; we allow randomization at this threshold.

Games of *self protection* are *supermodular*. In other words, they are of *strategic complements*: the more other agents invest in protection (the lower $\mathcal{T}(\mu)$), the more an agent has an incentive to invest in protection. Since games of self protection are effectively coordination games, there can be multiple equilibria. The next result shows that any equilibrium can be characterized by a lower-threshold strategy. In other words, the thresholds are reversed when compared to games total protection (cf. Theorem 1).

Theorem 2 (Self Protection)**.** *In a game of self protection, any equilibrium* μ^* *is a lower-threshold equilibrium. That is,* μ^* *is a lower-threshold strategy.*

Higher cascade risk thus leads to *lower* incentives to invest in protection. This is because an agent remains exposed to the failure risk of others irrespectively of whether he invests in protection. An investment in protection thus has lower returns as the cascading failure risk increases. An agent's cascading failure risk *increases* in degree, and thus higher-degree agents invest *less* in protection than lower-degree agents. The intuition is that higher-degree agents are more exposed to cascading failure risk than lower-degree agents, thus making an investment in their own *self* protection relatively less rewarding.

In the example of airport security, an airport that interacts with a high number of other airports has smaller incentives to invest in its own security, since it remains exposed to a high risk of being hit by an event coming from a connecting flight. This, as before, is assuming that the passengers/luggage are only inspected at their point of origin and not at points of transit. In the example of two-factor authentication (2FA) in an email network, it is the users with a high number of contacts who have lower incentives to enable this security feature since they remain exposed to infected email attachments from their many contacts.

The fact that, in games of self-protection, the incentives are reversed has important implications. In fact, the more connected (higher-degree) agents have a lesser incentive to invest in protection even though they are more vulnerable *and* more dangerous, i.e. they are hubs through which cascading failures can spread. More central agents thus have lower incentives to internalize the risk they impose on the system, pointing to an inefficient outcome. Moreover, in equilibrium, the total cost of protection is born by lower-degree agents: those who have the smallest effect on decreasing $\mathcal{T}(\mu)$.

4.3. Games of Networked-Risk Protection

In games of *networked-risk protection*, the investment protects only against the cascading failure risk. It does not protect against intrinsic failure risk.

Examples of applications include protection against many sexually transmitted diseases. For instance, the use of condoms protects against the transmission of HIV/AIDS through sexual partners. Nevertheless, such practices leave agents exposed to the external risk of being infected through a medical mistake in a hospital (e.g. with an infected syringe).

We have the following definition:

Definition 14 (Games of networked-risk protection)**.** *In a game of networked-risk protection, the effective failure probability has the following form*

$$\mathcal{B}(p, q_d(\mathcal{T}(\mu)), a) = p + (1-p) \cdot q_d(\mathcal{T}(\mu)) \cdot (1-ka) \tag{9}$$

for some $k \in [0,1]$ *and*

$$\mathcal{F}(\mu, \alpha) = \sum_{d \geq 1} \tilde{f}(d) \Big(p + \big(1-p\big) \cdot q_{d-1}(\alpha) \cdot (1 - k\mu(d)) \Big). \tag{10}$$

We now show that a game of networked-risk protection is structurally equivalent to a game of total protection.

Corollary 1. *A game of networked-risk protection is structurally equivalent to a game of total protection. Particularly, an equilibrium strategy* μ^* *in any game of networked-risk protection is unique and is characterized by an upper threshold.*

It is easy to see that agents have lesser incentives to invest than in the case of a game of total protection. Indeed, the marginal utility of investing in the latter case is always Vpk higher, because an investment also protects against the intrinsic failure risk. We thus conclude that $\mu^*_{tp} \succeq \mu^*_{np}$, where μ^*_{tp} and μ^*_{np} are the investment profiles in games of total and networked-risk protection, respectively. In other words, if an agent of some degree invests in the case of networked-risk protection, then he will necessarily also invest in the case of total protection. In the interest of space, we skip further in-depth discussion of the results in this section as they mainly replicate the results of Section 4.1.

4.4. Welfare, Risk and Comparative Statics

The next proposition states when the equilibrium expected utility and effective failure risk of an agent are monotone in degree.

Proposition 2 (Risk and Welfare I)**.** *Let $a_d \in \mu^*(d)$:*

- *(i) The equilibrium expected utility $U_d(a_d, \mu^*)$ is non-increasing in d.*
- *(ii) In a game of self protection, the equilibrium effective failure probability $\mathcal{B}(p, q_d(\mathcal{T}(\mu^*)), a_d)$ is non-decreasing in d.*

Note that there is no analogue to Part (ii) for games of total protection or networked-risk protection. The equilibrium effective failure probability can be non-monotone in degree. Indeed, the upper-threshold strategy means that higher-degree agents invest in protection and may thus have a lower effective failure probability than lower-degree agents.

We will now state a welfare result for games of self protection. These games are easier to analyze because they are of strategic complements. In games of self-protection, agents effectively *pool* their investments in protection and, as said earlier, there can be multiple equilibria. These equilibria can however be ordered by level of investment. Suppose there are m possible equilibria. Then, they can be ordered in the following way

$$\mu_1^* \preceq \mu_2^* \preceq ... \preceq \mu_m^*.$$

Since (8) is decreasing in μ, it follows that $\mathcal{T}(\mu_1^*) \geq \mathcal{T}(\mu_2^*) ... \geq \mathcal{T}(\mu_m^*)$.

We then have a second welfare result.

Proposition 3 (Welfare II)**.** *In a game of self protection, let $\mu_k^* \preceq \mu_l^*$ be two equilibria ordered by level of investment. Then μ_l^* weakly Pareto-dominates μ_k^*.*

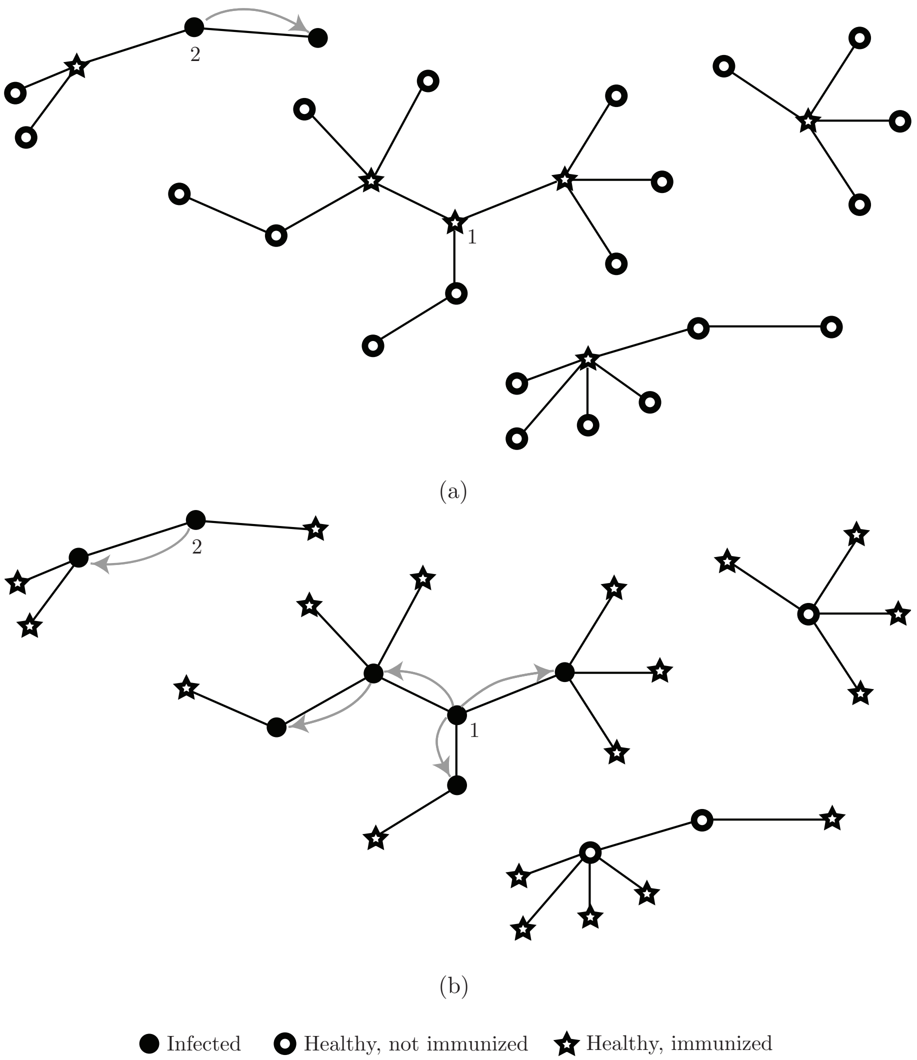
This result is not trivial. It effectively states that in the high-investment equilibrium, the decrease in risk resulting from higher investments outweighs the cost of those investments. This is due to the positive externality stemming from the effect of pooled investments in protection, which reduce all agents' failure risk.

We can focus our attention on the minimum-investment equilibrium $\underline{\mu}^*$ and the maximum-investment equilibrium $\bar{\mu}^*$. In the former, $\mathcal{T}(\underline{\mu}^*)$ is actually maximal since agents invest least, while in the latter, $\mathcal{T}(\bar{\mu}^*)$ is actually minimal since agents invest most. From Proposition 3, agents playing the minimum-investment equilibrium can be thus considered a coordination failure.

In Fig. 3, we illustrate Theorems 1 and 2 on a complex network. We see how the upper (resp. lower) threshold nature of equilibria in games of total (resp. self) protection affects the spread of cascading failures differently.

We now state a result comparing the welfare in games of total and self protection.

Proposition 4 (Welfare III)**.** *Let $W(\mu) = \sum_d f(d) U_d(\mu(d), \mu)$ be the utilitarian welfare under strategy μ. Specifically, we denote by $W^{tp}(\cdot)$ the utilitarian welfare in a game of total protection and by $W^{sp}(\cdot)$ the utilitarian welfare in a game of self protection, when all other model parameters are held fixed. Then $W^{tp}(\mu^*) > W^{sp}(\bar{\mu}^*)$, where μ^* is the unique equilibrium in a game of total protection and $\bar{\mu}^*$ be the maximum-investment equilibrium in a game of self protection.*



(a)

(b)

Figure 3: Illustration of Theorems 1 and 2 on a complex network with the cascading process of Fig. 1: possible equilibrium strategies in (a) a game of total protection and (b) a game of self protection. In (a), we see that the upper-threshold strategy insulates contagion hubs whereas in (b) we see that the lower-threshold strategy insulates periphery nodes and leaves contagion hubs vulnerable.

The above proposition states that the unique equilibrium in a game of total protection welfare-dominates the higher-investment equilibrium in a game of self protection. This result is mainly due to the fact that the return on investment in a game of total protection is higher than in a game of self protection, since it protects against the total risk of failure (not just the intrinsic risk of failure).

An advantage of our informational setting is that we can relate equilibrium behavior to network properties as captured by the edge-perspective degree distribution $\tilde{f}(d)$. We can then ask questions such as "does a higher level of connectedness[13] increase or decrease the incentives to invest in protection?" This is examined in the next proposition.

Proposition 5 (Shifting Degree Distribution). *Let $\underline{\mu}^*$ and $\bar{\mu}^*$ be the minimum- and maximum-investment equilibria in a game of self protection, when the edge-perspective degree distribution is $\tilde{f}$. Then, a first-order distributional shift[14] $\tilde{f}' \succ \tilde{f}$ results in $\underline{\mu}'^* \preceq \underline{\mu}^*$ and $\bar{\mu}'^* \preceq \bar{\mu}^*$ and thus in $\mathcal{T}'(\underline{\mu}'^*) \geq \mathcal{T}(\underline{\mu}^*)$ and $\mathcal{T}'(\bar{\mu}'^*) \geq \mathcal{T}(\bar{\mu}^*)$.*

Thus in a game of self protection, a higher level of connectedness leads to *lower* incentives to invest in protection: each of the new maximum- and minimum-investment equilibria are weakly dominated by the corresponding equilibria in the less connected network. The intuition behind this result is that an agent is more likely to be connected to a high-degree neighbor (high contagion risk and unprotected). This increases the agent's cascading failure risk and therefore lowers the incentive to invest in *self* protection. We note that in equilibrium, the corresponding neighbor failure probabilities are larger, i.e. $\mathcal{T}'(\underline{\mu}'^*) \geq \mathcal{T}(\underline{\mu}^*)$ and $\mathcal{T}'(\bar{\mu}'^*) \geq \mathcal{T}(\bar{\mu}^*)$.

Note that there is no straightforward analogue to Proposition 5 in the case of total protection or networked-risk protection. In fact shifting $\tilde{f}(d)$ may in this case increase the probability of having a protected neighbor or an unprotected one, depending on the extent of the shift in $\tilde{f}(d)$ and on the threshold d_U in the upper-threshold strategy. A shift in $\tilde{f}(d)$ could thus potentially have non-monotone effects.

When cascading failures follow a contact process as in the examples of Section 3.3, it is interesting to study the effect of a change in the infectiousness parameter r on equilibria. The following two propositions illustrate that a change in r has opposite effects, depending on whether the game is one of self protection or total protection.

Proposition 6 (Varying Infectiousness). *Suppose cascading failures follow a contact process with infectiousness parameter r, as in the examples of Section 3.3. Let $\underline{\mu}^*$ and $\bar{\mu}^*$ be the minimum- and maximum-investment equilibria in a game of self protection and let μ^* be the unique equilibrium in a game of total (or networked-risk) protection. Then, an increase $r' > r$ in infectiousness results in:*

- *(i) $\underline{\mu}'^* \preceq \underline{\mu}^*$ and $\bar{\mu}'^* \preceq \bar{\mu}^*$ and thus in $\mathcal{T}'(\underline{\mu}'^*) \geq \mathcal{T}(\underline{\mu}^*)$ and $\mathcal{T}'(\bar{\mu}'^*) \geq \mathcal{T}(\bar{\mu}^*)$.*
- *(ii) $\mu'^* \succeq \mu^*$ and $r'\mathcal{T}'(\mu'^*) \geq r\mathcal{T}(\mu^*)$.*

[13] Note that by a higher level of connectedness, we mean an edge-perspective degree distribution placing higher mass on higher-degree nodes. We do not mean the presence of short paths between any two nodes.

[14] Here $\tilde{f}' \succ \tilde{f}$ means that $\tilde{f}'$ first-order stochastically dominates $\tilde{f}$.

Part (i) says that in a game of self protection, when cascading failures follow a contact process, a higher level of infectiousness creates *lower* incentives for agents to invest in protection: the initial increase in $\mathcal{T}(\mu)$ caused by higher infectiousness causes an even greater increase in $\mathcal{T}(\mu)$ as a result of strategic interactions. The situation is very different in a game of total (or networked-risk) protection, as shown in Part (ii), where a higher level of infectiousness creates *higher* incentives for agents to invest in protection. This investment in protection is however not enough to counter the increase in $r\mathcal{T}(\mu)$ caused by a higher level of infectiousness. This is because agents free-ride on the protection provided by others and thus an increase in $r\mathcal{T}(\mu)$ cannot be completely compensated.

The next result examines the effect of an increase in the parameter k, which governs the extent of the protection resulting from an investment.

Proposition 7 (Varying the Quality of Protection). *Let $\underline{\mu}^*$ and $\bar{\mu}^*$ be the minimum- and maximum-investment equilibria in a game of self protection and let μ^* be the unique equilibrium in a game of total (or networked-risk) protection with parameter k. Then, $k' > k$ results in:*

- *(i) $\underline{\mu}'^* \succeq \underline{\mu}^*$ and $\bar{\mu}'^* \succeq \bar{\mu}^*$, and thus in $\mathcal{T}'(\underline{\mu}'^*) \leq \mathcal{T}(\underline{\mu}^*)$ and $\mathcal{T}'(\bar{\mu}'^*) \leq \mathcal{T}(\bar{\mu}^*)$.*
- *(ii) $\mu'^* \preceq \mu^*$, but $\mathcal{T}'(\mu'^*) \leq \mathcal{T}(\mu^*)$.*

Thus in a game of self protection, an increase in the protection quality results in a higher investment and a reduction in a neighbor's probability of failure. Strategic interactions thus further add to the benefits of an improvement in the protection technology. On the contrary, in a game of total (or networked-risk) protection, such an increase in the protection quality results in a lower investment. However, it still results in a reduction of a neighbor's probability of failure, which is due entirely to the increase in protection quality.

4.5. Endogenizing the Cost of Protection

So far, we have only examined network effects. That is, a utility function depends on other agents only through the failure probability of one's neighbors. In reality, global feedback effects might also influence an agent's utility. By 'global feedback effects', we mean effects that impact an agent's utility in other ways than through its neighbors on the network. For instance, prices of vaccines, computer security solutions or airport security equipment might be affected by demand (i.e. by μ). Likewise, if protection is provided under the form of insurance[15], the insurance premium might depend on the overall failure level in the population, which itself depends on the overall level of investment in protection. Such price feedback effects, in addition to network effects, are also considered in Jackson & Zenou (2014). Gagnon & Goyal (2015) also build a model in which agents' utilities are affected both by their neighbors on a social network and by effects unrelated to that network.

In this section, we introduce such global feedback effects to the model developed in the previous sections. We focus on global feedback through the cost of protection, which can take the form of a price to be paid.

[15] See, for example, Reuters (12 October 2015): "Cyber insurance premiums rocket after high-profile attacks". Oct 12, 2015. Reuters. The reader may also see Johnson *et al.* (2011) and Lelarge & Bolot (2009) for some work on insurance provision.

We will introduce the following function, which maps a strategy μ to the corresponding probability that a randomly-picked agent invests in protection:

Definition 15. *Let the function $\mathcal{G} : \mathcal{M} \to [0,1]$ be defined as:*

$$\mathcal{G}(\mu) = \sum_{d \geq 1} f(d)\mu(d). \tag{11}$$

Thus to each strategy μ corresponds a fraction $\mathcal{G}(\mu)$ of agents who invest in protection. Furthermore, it is easy to notice that this function $\mathcal{G}$ increases in μ.

We will explore a setting in which the cost of protection is influenced by global demand. Namely, when the cost of protection depends monotonically on total demand: $C_g = C \cdot g(\mathcal{G}(\mu))$, where $g(\cdot)$ is either an increasing or a decreasing continuous function of the total fraction of people $\mathcal{G}(\mu)$ willing to invest in protection. In the following examples, we outline two situations that can be modeled by the function $g(\cdot)$.

Example 1 ($g(\cdot)$ increasing). This case corresponds to the situation where the product is scarce or there are global congestion effects. For instance, a vaccine might be produced in limited quantity and thus, the more people demand it, the harder it may be to obtain it, which will have an increasing effect on price.

Example 2 ($g(\cdot)$ decreasing). This corresponds to the case of economies of scale. For instance, a new airport security technology might require significant initial R & D investments. Producing it in large numbers may thus lead to a lower cost per unit, which may lower the price.

We will slightly modify a degree-d agent's expected utility function in order to introduce the global feedback effect:

$$U_d(a, \mu) = -V \cdot \mathcal{B}(p, q_d(\mathcal{T}(\mu)), a) - C \cdot g(\mathcal{G}(\mu)) \cdot a. \tag{12}$$

Note that the cascading failure probability $q_d(\mathcal{T}(\mu))$ does not depend explicitly on the global fraction of agents who invest in protection, as it is solely driven by network effects, i.e. through an agent's neighborhood. It is also important to mention that the introduction of a global externality does not affect the definition of $\mathcal{T}(\mu)$. The latter function was defined to be the failure probability of a randomly-picked neighbor, which does not depend explicitly on the total fraction of agents investing in protection $\mathcal{G}(\mu)$.

We will now modify the optimality condition in order to ensure that this fraction $\mathcal{G}(\mu)$ arises in equilibrium. We can redefine the set of optimal responses as follows:

Definition 16. *Let $\mathcal{S}_d(\mathcal{T}(\mu), \mathcal{G}(\mu)) \subset [0,1]$ denote the set of optimal responses for a degree-d agent given $\mathcal{T}(\mu)$ and $\mathcal{G}(\mu)$; i.e.:*

$$\begin{aligned}
U_d(1,\mu) > U_d(0,\mu) &\implies \mathcal{S}_d(\mathcal{T}(\mu), \mathcal{G}(\mu)) = \{1\};\\
U_d(1,\mu) < U_d(0,\mu) &\implies \mathcal{S}_d(\mathcal{T}(\mu), \mathcal{G}(\mu)) = \{0\};\\
U_d(1,\mu) = U_d(0,\mu) &\implies \mathcal{S}_d(\mathcal{T}(\mu), \mathcal{G}(\mu)) = [0,1].
\end{aligned}$$

Let $\mathcal{S}(\mathcal{T}(\mu), \mathcal{G}(\mu)) \subset \mathcal{M}$ denote the set of optimal strategies given $\mathcal{T}(\mu)$ and $\mathcal{G}(\mu)$; i.e.,

$$\mathcal{S}(\mathcal{T}(\mu), \mathcal{G}(\mu)) = \prod_{d \geq 1} \mathcal{S}_d(\mathcal{T}(\mu), \mathcal{G}(\mu)).$$

We now only need to slightly modify the equilibrium condition:

Definition 17 (Mean-Field Equilibrium with Endogenized Cost of Protection)**.** *A strategy μ^* constitutes a mean-field equilibrium (MFE) if $\mu^* \in \mathcal{S}(\mathcal{T}(\mu^*), \mathcal{G}(\mu^*))$.*

It turns out that the main results that were stated in the previous sections of the paper are robust to the introduction of this global externality. We summarize those more general results in the following proposition.

Proposition 8 (Network Security Game with Endogenized Cost of Protection)**.**

- *(i) (Existence): There exists a mean-field equilibrium in the game with endogenized cost or protection.*
- *(ii) (Threshold Strategies): The threshold characterization of equilibria is robust to the endogenization of the cost of protection. The equilibrium is of: (1) an upper-threshold nature for a game of total protection and networked-risk protection; (2) a lower-threshold nature for a game of self protection.*
- *(iii) (Uniqueness): In a game of total protection or of networked-risk protection, the mean-field equilibrium μ^* is unique if $g(\cdot)$ is an increasing function.*

As before, there can be multiple equilibria for games of self protection.

5. Conclusion

In this paper, we developed a framework to study the strategic investment in protection against cascading failures in networked systems. Agents connected through a network can fail either intrinsically or as a result of a cascade of failures that may cause their neighbors to fail. We studied three broad classes of games covering a wide range of applications. We showed that equilibrium strategies are monotone in degree (i.e. in the number of neighbors an agent has on the network) and that this monotonicity is reversed depending on whether (i) an investment in protection insulates an agent against the risk of failure of his neighbors (games of total protection and games of networked-risk protection) or (ii) only against his own intrinsic risk of failure (games of self protection). The first case covers the important examples of vaccination, anti-virus software as well as protection against sexually-transmitted diseases. Here it is the *more* connected agents who have higher incentives to invest in protection. The second case, on the other hand, covers examples such as airport/EU security as well as other types of computer security solutions such as two-factor authentication (2FA). Here it is the *less* connected agents who have higher incentives to invest in protection. Our analysis reveals that it is the nature of strategic interactions (strategic substitutes/complements), combined with a network structure that leads to such strikingly different equilibrium behavior in each case, with important implications for the system's resilience to cascading failures.

Our model is simple and the incomplete information framework that we use allows for a tractable treatment. The unobservability of the network is a credible assumption for many applications. In the applications of vaccination or computer security, agents typically do not know the topology of the social network or email network, for example. They merely know the number of connections that they have. Our equilibrium concept then predicts the behavior of the agents based on their level of interaction with the population (their degree).

The property that neighbors fail independently imposes a realistic cognitive burden on agents and allows for a tractable way to express an agent's expected cascading failure probability. This property is similar to the local tree-like assumption used in other models such as Lelarge & Bolot (2008a) and is valid for large or relatively sparse networks. In spite of its advantages, this property may no longer be realistic for small or dense networks. However, as long as an agent's cascading failure probability is monotone in degree (which may still be approximately the case, even in some smaller/denser networks), our monotonicity results could hold, at least approximately.

In the case of airport security or EU security, the topology of the network of airports or countries can credibly be known and influence the decisions of the agents. It would be interesting to extend our analysis to the case where agents know the topology of the network. While it is likely that equilibrium behavior would still be monotone in the level of interaction of the agents with the rest of the population, degree centrality may no longer be the appropriate measure. It would be interesting to see if in this case, one could somehow relate equilibrium behavior to some other measure of network centrality as in Acemoglu *et al.* (2013). Such extensions are left for future work.

Appendix A. Proofs

Proposition 1. Note that we endow $[0,1]$ with the Euclidean topology.

For any $\alpha \in [0,1]$ define the correspondence Φ by $\Phi(\alpha) = \mathcal{T}(\mathcal{S}(\alpha))$. Any fixed point α^* of Φ, with the corresponding $\mu^* \in \mathcal{S}(\alpha^*)$ such that $\mathcal{T}(\mu^*) = \alpha^*$ constitute a MFE. We thus need to show that the correspondence Φ has a fixed point. We employ Kakutani's fixed point theorem on the composite map $\Phi(\alpha) = \mathcal{T}(\mathcal{S}(\alpha))$.

Kakutani's fixed point theorem requires that Φ have a compact domain, which is trivial since $[0,1]$ is compact. Further, $\Phi(\alpha)$ must be nonempty; again, this is straightforward, since both $\mathcal{S}$ and $\mathcal{T}$ have nonempty image.

Next, we show that $\Phi(\alpha)$ has a closed graph. We first show that $\mathcal{S}$ has a closed graph, when we endow the set of strategies with the product topology on $[0,1]^\infty$. This follows easily: if $\alpha_n \to \alpha$, and $\mu_n \to \mu$, where $\mu_n \in \mathcal{S}(\alpha_n)$ for all n, then $\mu_n(d) \to \mu(d)$ for all d. Expressing utility as a function of α, i.e. $U_d(a,\alpha) = -V \cdot \mathcal{B}(p, q_d(\alpha), a) - C \cdot a$, we see that $U_d(1,\alpha)$ and $U_d(0,\alpha)$ are continuous, and it follows that $\mu(d) \in \mathcal{S}_d(\alpha)$, so $\mathcal{S}$ has a closed graph. Note also that with the product topology on the space of strategies, $\mathcal{T}$ is continuous: if $\mu_n \to \mu$, then $\mathcal{T}(\mu_n) \to \mathcal{T}(\mu)$ by the bounded convergence theorem.

To complete the proof that Φ has a closed graph, suppose that $\alpha_n \to \alpha$, and that $\alpha'_n \to \alpha'$, where $\alpha'_n \in \Phi(\alpha_n)$ for all n. Choose $\mu_n \in \mathcal{S}(\alpha_n)$ such that $\mathcal{T}(\mu_n) = \alpha'_n$ for all n. By Tychonoff's theorem, $[0,1]^\infty$ is compact in the product topology; so taking subsequences if necessary, we can assume that μ_n converges to a limit μ. Since $\mathcal{S}$ has a closed graph, we know $\mu \in \mathcal{S}(\alpha)$. Finally, since $\mathcal{T}$ is continuous, we know that $\mathcal{T}(\mu) = \alpha'$. Thus $\alpha' \in \Phi(\alpha)$, as required.

Finally, we show that the image of Φ is convex. Let $\alpha_1, \alpha_2 \in \Phi(\alpha)$, and choose $\mu_1, \mu_2 \in \mathcal{S}(\alpha)$ such that $\alpha_1 = \mathcal{T}(\mu_1)$ and $\alpha_2 = \mathcal{T}(\mu_2)$. Since $\mathcal{F}$ is continuous in μ and since $\mathcal{T}$ is unique (this follows from Property 2), then $\mathcal{T}$ is continuous in μ. Now since $\mathcal{S}(\alpha)$ is convex, it follows that for any $\delta \in (0, 1)$,

$$\begin{aligned} \delta\mathcal{T}(\mu_1) + (1-\delta)\mathcal{T}(\mu_2) &\in [\min_{\mu\in\mathcal{S}(\alpha)} \mathcal{T}(\mu), \max_{\mu\in\mathcal{S}(\alpha)} \mathcal{T}(\mu)] \\ &= \Phi(\alpha) \end{aligned}$$

and thus $\delta\alpha_1 + (1-\delta)\alpha_2 \in \Phi(\alpha)$—as required.

By Kakutani's fixed point theorem, Φ possesses a fixed point α^*. Letting $\mu^* \in \mathcal{S}(\alpha^*)$ be such that $\mathcal{T}(\mu^*) = \alpha^*$, we conclude that μ^* is an MFE. □

Theorem 1. For convenience, since the expected utility $U_d(a, \mu)$ depends on μ only through $\alpha = \mathcal{T}(\mu)$, we may write it as a function of α as follows.

$$U_d(a, \alpha) = -V \cdot \mathcal{B}(p, q_d(\alpha), a) - C \cdot a \tag{A.1}$$

Consider the incremental expected utility for an agent of degree d, i.e.

$$\begin{aligned} \Delta U_d(\alpha) &= U_d(1, \alpha) - U_d(0, \alpha) \qquad \text{(A.2)} \\ &= -V \cdot (p + (1-p)q_d(\alpha))(1-k) - C - (-V \cdot (p + (1-p)q_d(\alpha)) \\ &= V \cdot \Big(p + (1-p)q_d(\alpha)\Big)k - C \end{aligned}$$

We will first show that any equilibrium is an upper-threshold strategy.

Consider $\Delta U_d(\alpha)$ as a function of the continuous variable d over the connected support $[1, \infty)$. From (A.2), we can write

$$\Delta U_d(\alpha) \;=\; V \cdot \Big(p + (1-p)q_d(\alpha)\Big)k - C$$

Since $q_d(\alpha)$ is non-decreasing in d, for any $\alpha \in (0, 1)$, $\Delta U_d(\alpha)$ is a non-decreasing function of d. It follows that the inverse image of $(-\infty, 0)$ is $\emptyset$ if $\Delta U_1(\alpha) > 0$ or an interval $[1, x)$ where $x \geq 1$ otherwise. The integers in such intervals (i.e. $\emptyset \bigcap \mathbb{N}^+$ or $[1, x) \bigcap \mathbb{N}^+$) represent the degrees of agents for whom not investing in protection is a strict best response, i.e. $\{d : \mathcal{S}_d(\alpha) = \{0\}\}$. It follows that the degrees of agents for whom investing in protection is a strict best response (i.e. $\{d : \mathcal{S}_d(\alpha) = \{1\}\}$) are located at the rightmost extremity of the degree support.

Thus we may write $\mu(d) = 1$, for all $d > d_U$ and $\mu(d) = 0$, for all $d < d_U$. This is valid for any best-responding strategy μ and it is therefore valid for any equilibrium strategy μ^*.

We now prove equilibrium uniqueness. We prove it in a sequence of steps:

Step 1: For all $d \geq 1$, $\Delta U_d(\alpha)$ is strictly increasing in $\alpha \in [0, 1]$. This follows directly from Definition 3.

Step 2: For all $d \geq 1$, and $\alpha' > \alpha$, $\mathcal{S}_d(\alpha') \succeq \mathcal{S}_d(\alpha)$.[16] This follows immediately from Step 1 and the definition of $\mathcal{S}_d$ in Definition 8.

Step 3: If μ', μ are strategies such that $\mu'(d) \geq \mu(d)$, then $\mathcal{T}(\mu') \leq \mathcal{T}(\mu)$. This follows from the fact that $\mathcal{F}(\mu, \alpha)$ (cf. (6) in Definition 10) is non-increasing in μ and that it is also continuous in both μ and α. Thus the unique fixed point $\bar{\alpha} = \mathcal{F}(\mu, \bar{\alpha})$ is non-increasing in μ. Therefore, $\mathcal{T}(\mu') \leq \mathcal{T}(\mu)$.

Step 4: Completing the proof. So now suppose that there are two mean-field equilibria μ^* and μ'^*, with $\mathcal{T}(\mu'^*) = \alpha'^* > \alpha^* = \mathcal{T}(\mu^*)$. By Step 2, since $\mu^* \in \mathcal{S}(\alpha^*)$ and $\mu'^* \in \mathcal{S}(\alpha'^*)$, we have $\mu'^*(d) \geq \mu^*(d)$. By Step 3, we have $\alpha^* = \mathcal{T}(\mu^*) \geq \mathcal{T}(\mu'^*) = \alpha'^*$, a contradiction. Thus the $\alpha^* = \mathcal{T}(\mu^*)$ in any MFE must be unique, as required.

It then follows from the threshold nature of the equilibrium strategy μ^* that to α^*, there corresponds a unique $\mu^* \in \mathcal{S}(\alpha^*)$ such that $\alpha^* = \mathcal{T}(\mu^*)$.

□

Theorem 2. For convenience, we do as in the proof of Theorem 1 and write the expected utility $U_d(a, \alpha)$ as a function of α.

In a game of self protection, consider now $\Delta U_d(\alpha)$ as a function of the continuous variable d over the connected support $[1, \infty)$. From (7) and (2), we can write

$$\begin{aligned} \Delta U_d(\alpha) &= U_d(1,\alpha) - U_d(0,\alpha) \\ &= -V \cdot (p(1-k) + (1-p(1-k))q_d(\alpha)) - C + V \cdot (p + (1-p)q_d(\alpha)) \\ &= V \cdot (pk - pkq_d(\alpha)) - C \end{aligned}$$

Since $q_d(\alpha)$ is non-decreasing in d, for any $\alpha \in (0,1)$, $\Delta U_d(\alpha)$ is a non-increasing function of d. It follows that the inverse image of $(-\infty, 0)$ is an interval $[1, \infty)$ if $\Delta U_1(\alpha) < 0$ or an interval (x, ∞) where $x \geq 1$ otherwise. The integers in such intervals (i.e. $[1,\infty) \bigcap \mathbb{N}^+$ or $(x, \infty) \bigcap \mathbb{N}^+$) represent the degrees of agents for whom not investing in protection is a strict best response, i.e. $\{d : \mathcal{S}_d(\alpha) = \{0\}\}$. It follows that the degrees of agents for whom investing in protection is a strict best response (i.e. $\{d : \mathcal{S}_d(\alpha) = \{1\}\}$) are located at the leftmost extremity of the degree support.

Thus we may write $\mu(d) = 1$, for all $d < d_L$ and $\mu(d) = 0$, for all $d > d_L$. This is valid for any best-responding strategy μ and it is therefore valid for the equilibrium strategy μ^*.

□

Corollary 1. The result follows from comparing the incremental expected utility of an agent of degree d in the case of networked-risk protection (see (A.3) below) with the one in the case of total protection (see (A.2) in the proof of Theorem 1):

$$\Delta U_d(\alpha) = V(1-p)q_d(\alpha)k - C \tag{A.3}$$

By comparing (A.3) to (A.2), it is easy to see that the incremental utility of investing in protection is $Vpk > 0$ higher in the case of total protection. Otherwise, $\Delta U_d(\alpha)$ is increasing in d and a similar argument as in the proof of Theorem 1 leads to the conclusion that the equilibrium strategy μ^* is of an upper-threshold nature. □

[16] Here the set relation $A \preceq B$ means that for all $x \in A$ and $y \in B$, $x \leq y$.

Proposition 2. Part (i):

Note that $q_d(\mathcal{T}(\mu^*))$ is non-decreasing in d. Thus for $d' > d$, $a_{d'} \in \mu^*(d')$ and $a_d \in \mu^*(d)$, we have

$$U_d(a_d, \mu^*) \geq U_d(a_{d'}, \mu^*) \geq U_{d'}(a_{d'}, \mu^*) \tag{A.4}$$

where the first inequality follows from $a_d \in \mu^*(d)$, while the second inequality follows from $q_d(\mathcal{T}(\mu^*))$ being non-decreasing in d. Thus $U_d(a_d, \mathcal{T}(\mu^*))$ is non-increasing in d.

Part (ii):

From Theorem 2, the equilibrium strategy μ^* is non-increasing in d. From (7), it thus follows that for $a_d \in \mu^*(d)$, $\mathcal{B}(p, q_d(\mathcal{T}(\mu^*)), a_d)$ is non-decreasing in d (since $\mathcal{B}$ is non-decreasing in $q_d(\mathcal{T}(\mu^*))$ and non-increasing in a_d).

□

Proposition 3. For convenience, we do as in the proof of Theorem 1 and write the expected utility $U_d(a, \alpha)$ as a function of α.

Let $\alpha_l^* = \mathcal{T}(\mu_l^*)$ and $\alpha_k^* = \mathcal{T}(\mu_k^*)$. We then have $\mu_l^* \in \mathcal{S}(\alpha_l^*)$ and $\mu_k^* \in \mathcal{S}(\alpha_k^*)$. Then for any d,

$$U_d(a_l, \alpha_l^*) \geq U_d(a_k, \alpha_l^*) \geq U_d(a_k, \alpha_k^*) \tag{A.5}$$

where $a_l \in \mu_l^*(d)$ and $a_k \in \mu_k^*(d)$.

The first inequality follows from a_l being a best response to α_l^* (i.e. $a_l \in \mu_l^*(d)$) for an agent of degree d. The second inequality follows from U_d being decreasing in α^*.

Since (A.5) holds for any d, all agents have expected utility that is weakly greater in the higher-investment equilibrium μ_l^*. We therefore conclude that μ_l^* weakly Pareto-dominates μ_k^*.

□

Proposition 4. First note that the expected utilities in all possible cases are

$$\begin{aligned}
U_d^{tot}(1, \mu^*) &= -V\big(p + (1-p)q_d(\mathcal{T}(\mu^*))\big)(1-k) - C \\
U_d^{tot}(0, \mu^*) &= -V\big(p + (1-p)q_d(\mathcal{T}(\mu^*))\big) \\
U_d^{s.p.}(1, \bar{\mu}^*) &= -V\big(p(1-k) + (1-p(1-k))q_d(\mathcal{T}(\bar{\mu}^*))\big) - C \\
U_d^{s.p.}(0, \bar{\mu}^*) &= -V\big(p + (1-p)q_d(\mathcal{T}(\bar{\mu}^*))\big).
\end{aligned}$$

Also note that the incremental utilities from investing in each class of games are

$$\begin{aligned}
\Delta U_d^{tot}(\mu^*) &= U_d^{tot}(1, \mu^*) - U_d^{tot}(0, \mu^*) \\
&= V\big(p + (1-p)q_d(\mathcal{T}(\mu^*))\big)k - C \qquad \text{(A.6)}
\end{aligned}$$

and that

$$\begin{aligned}
\Delta U_d^{s.p.}(\bar{\mu}^*) &= U_d^{tot}(1, \bar{\mu}^*) - U_d^{tot}(0, \bar{\mu}^*) \\
&= V\big(p - pq_d(\mathcal{T}(\bar{\mu}^*))\big)k - C. \qquad \text{(A.7)}
\end{aligned}$$

Suppose $Vpk > C$. Then, from (A.6) we see that in a game of total protection $\Delta U_d^{tot}(\mu^*) > 0$ for all d and thus $\mu^*(d) = 1$ for all d. Thus

$$W^{tot}(\mu^*) = \sum_d f(d) U_d^{tot}(1, \mu^*). \tag{A.8}$$

Also, from (A.7) we see that in a game of self protection not all agents may find it optimal to invest and thus $\bar{\mu}^*$ is some lower-threshold strategy. Therefore,

$$\begin{aligned} W^{s.p.}(\bar{\mu}^*) &= \sum_{d<d_L} f(d)U_d^{s.p.}(1,\bar{\mu}^*) + \sum_{d>d_L} f(d)U_d^{s.p.}(0,\bar{\mu}^*) \\ &\quad + f(d_L)\mu(d_L)U_d^{s.p.}(1,\bar{\mu}^*) + f(d_L)(1-\mu(d_L))U_d^{s.p.}(0,\bar{\mu}^*). \end{aligned} \tag{A.9}$$

Since $\mu^* \succeq \bar{\mu}^*$, then $\mathcal{T}(\mu^*) \leq \mathcal{T}(\bar{\mu}^*)$. Examining the expressions for the utilities allows us to conclude that $U_d^{s.p.}(1,\bar{\mu}^*) < U_d^{tot}(1,\mu^*)$ and that $U_d^{s.p.}(0,\bar{\mu}^*) < U_d^{tot}(0,\mu^*) < U_d^{tot}(1,\mu^*)$, where the last inequality follows from the fact that $\Delta U_d^{tot}(\mu^*) > 0$. Comparing the expressions for the welfare then allows us to conclude that $W^{s.p.}(\bar{\mu}^*) < W^{tot}(\mu^*)$.

Now suppose $Vpk \leq C$. Then in a game of total protection, examining (A.6) tells us that not all agents may find it optimal to invest. μ^* is thus some upper-threshold strategy and thus

$$\begin{aligned} W^{tot}(\mu^*) &= \sum_{d>d_U} f(d)U_d^{tot}(1,\mu^*) + \sum_{d<d_U} f(d)U_d^{tot}(0,\mu^*) \\ &\quad + f(d_U)\mu(d_U)U_d^{tot}(1,\mu^*) + f(d_U)(1-\mu(d_U))U_d^{tot}(0,\mu^*). \end{aligned} \tag{A.10}$$

Also, in a game of self protection, $\bar{\mu}^*(d) = 0$ for all d. Indeed $\Delta U_d^{s.p.}(\bar{\mu}^*) < 0$ (see (A.7)). Thus

$$W^{s.p.}(\bar{\mu}^*) = \sum_d f(d)U_d^{s.p.}(0,\bar{\mu}^*) \tag{A.11}$$

Since $\mu^* \succeq \bar{\mu}^*$, then $\mathcal{T}(\mu^*) \leq \mathcal{T}(\bar{\mu}^*)$. Noting that $U_d^{s.p.}(0,\bar{\mu}^*) < U_d^{tot}(0,\mu^*)$ and that $U_d^{s.p.}(0,\bar{\mu}^*) < U_d^{tot}(0,\mu^*) \leq U_d^{tot}(1,\mu^*)$ for all d such that $\mu^*(d) > 0$, we conclude by comparing the expressions for the welfare that $W^{s.p.}(\bar{\mu}^*) < W^{tot}(\mu^*)$.

Thus for all parameter ranges, $W^{s.p.}(\bar{\mu}^*) < W^{tot}(\mu^*)$. This completes the proof.

□

Proposition 5. Let $\mathcal{F}'(\mu,\alpha)$ and $\mathcal{F}(\mu,\alpha)$ denote (8) under $\tilde{f}'$ and $\tilde{f}$ respectively. $q_d(\alpha)$ is non-decreasing in d and we know from Theorem 2 that in a game of self-protection, any equilibrium strategy is a lower-threshold strategy. We therefore only need to consider such strategies. It then follows from (8) that given any lower-threshold strategy μ, $\mathcal{F}'(\mu,\alpha) \geq \mathcal{F}(\mu,\alpha)$ for all $\alpha \in [0,1]$. Since by Property 2, (8) has a single fixed point in α and we conclude that $\mathcal{T}'(\mu) \geq \mathcal{T}(\mu)$, where $\mathcal{T}'(\mu)$ and $\mathcal{T}(\mu)$ denote (4) under $\tilde{f}'$ and $\tilde{f}$ respectively.

It then follows that

$$\begin{aligned} \Phi'(\alpha) &= \mathcal{T}'(\mathcal{S}(\alpha)) \\ &\succeq \mathcal{T}(\mathcal{S}(\alpha)) \\ &= \Phi(\alpha) \end{aligned}$$

It therefore follows that $\underline{\alpha}'^* = min\{\alpha : \alpha = \Phi'(\alpha)\} \geq min\{\alpha : \alpha = \Phi(\alpha)\} = \underline{\alpha}^*$ and that $\bar{\alpha}'^* = max\{\alpha : \alpha = \Phi'(\alpha)\} \geq max\{\alpha : \alpha = \Phi(\alpha)\} = \bar{\alpha}^*$.

Thus, the equilibrium strategies are such that $\underline{\mu}'^* = \mathcal{S}(\bar{\alpha}'^*) \preceq \mathcal{S}(\bar{\alpha}^*) = \underline{\mu}^*$ and $\bar{\mu}'^* = \mathcal{S}(\underline{\alpha}'^*) \preceq \mathcal{S}(\underline{\alpha}^*) = \bar{\mu}^*$. Likewise, $\mathcal{T}'(\bar{\mu}'^*) = \underline{\alpha}'^* \geq \underline{\alpha}^* = \mathcal{T}(\bar{\mu}^*)$ and $\mathcal{T}'(\underline{\mu}'^*) = \bar{\alpha}'^* \geq \bar{\alpha}^* = \mathcal{T}(\underline{\mu}^*)$.

□

Proposition 6. Part(i):

Let $\mathcal{F}'(\mu, \alpha)$ and $\mathcal{F}(\mu, \alpha)$ denote (8) under r' and r respectively. In the case of the contact process described in the examples of Section 3.3, $q'_d(\alpha) > q_d(\alpha)$ for all $\alpha \in [0, 1]$, $d > 0$. It then follows from (8) that given any strategy μ, $\mathcal{F}'(\mu, \alpha) \geq \mathcal{F}(\mu, \alpha)$ for all $\alpha \in [0, 1]$. Since by Property 2, (8) has a single fixed point, we conclude that $\mathcal{T}'(\mu) \geq \mathcal{T}(\mu)$, where $\mathcal{T}'(\mu)$ and $\mathcal{T}(\mu)$ denote the correspondence (4) under r' and r respectively.

It then follows that

$$\begin{array}{rcl} \Phi'(\alpha) & = & \mathcal{T}'(\mathcal{S}(\alpha)) \\ & \succeq & \mathcal{T}(\mathcal{S}(\alpha)) \\ & = & \Phi(\alpha) \end{array}$$

It therefore follows that $\underline{\alpha}'^* = min\{\alpha : \alpha = \Phi'(\alpha)\} \geq min\{\alpha : \alpha = \Phi(\alpha)\} = \underline{\alpha}^*$ and that $\bar{\alpha}'^* = max\{\alpha : \alpha = \Phi'(\alpha)\} \geq max\{\alpha : \alpha = \Phi(\alpha)\} = \bar{\alpha}^*$.

Thus, the equilibrium strategies are such that $\underline{\mu}'^* = \mathcal{S}(\bar{\alpha}'^*) \preceq \mathcal{S}(\bar{\alpha}^*) = \underline{\mu}^*$ and $\bar{\mu}'^* = \mathcal{S}(\underline{\alpha}'^*) \preceq \mathcal{S}(\underline{\alpha}^*) = \bar{\mu}^*$. Likewise, $\mathcal{T}'(\bar{\mu}'^*) = \underline{\alpha}'^* \geq \underline{\alpha}^* = \mathcal{T}(\bar{\mu}^*)$ and $\mathcal{T}'(\underline{\mu}'^*) = \bar{\alpha}'^* \geq \bar{\alpha}^* = \mathcal{T}(\underline{\mu}^*)$.

Part (ii):

We prove by contradiction. Suppose $r'\mathcal{T}'(\mu'^*) < r\mathcal{T}(\mu^*)$. Then $\mathcal{S}(\mathcal{T}'(\mu'^*)) \preceq \mathcal{S}(\mathcal{T}(\mu^*))$ and thus $\mu'^* \preceq \mu^*$. Since $\mathcal{F}'(\mu, \alpha) \geq \mathcal{F}(\mu, \alpha)$ for any $\mu \in \mathcal{M}$ and $\alpha \in [0, 1]$ and since $\mathcal{F}'$ and $\mathcal{F}$ are decreasing in μ, we have that $\mathcal{F}'(\mu'^*, \alpha) \geq \mathcal{F}(\mu^*, \alpha)$ for any $\alpha \in [0, 1]$. Therefore, $\mathcal{T}'(\mu'^*) \geq \mathcal{T}(\mu^*)$ and thus, since $r' > r$, we have that $r'\mathcal{T}'(\mu'^*) > r\mathcal{T}(\mu^*)$, a contradiction. We conclude that $r'\mathcal{T}'(\mu'^*) \geq r\mathcal{T}(\mu^*)$.

It then follows that $\mathcal{S}(\mathcal{T}'(\mu'^*)) \succeq \mathcal{S}(\mathcal{T}(\mu^*))$ and thus $\mu'^* \succeq \mu^*$.

The result extends to games of networked-risk protection by their structural equivalence to games of total protection (see Corollary 1). This completes the proof.

□

Proposition 7. Part (i):

Let $\mathcal{F}'(\mu, \alpha)$ and $\mathcal{F}(\mu, \alpha)$ denote (8) under k' and k respectively. It follows from (8) that given any strategy μ, $\mathcal{F}'(\mu, \alpha) \leq \mathcal{F}(\mu, \alpha)$ for all $\alpha \in [0, 1]$. Since under by Property 2, (8) has a single fixed point, we conclude that $\mathcal{T}'(\mu) \leq \mathcal{T}(\mu)$, where $\mathcal{T}'(\mu)$ and $\mathcal{T}(\mu)$ denote the correspondence (4) under k' and k respectively.

It then follows that

$$\begin{array}{rcl} \Phi'(\alpha) & = & \mathcal{T}'(\mathcal{S}(\alpha)) \\ & \preceq & \mathcal{T}(\mathcal{S}(\alpha)) \\ & = & \Phi(\alpha) \end{array}$$

It therefore follows that $\underline{\alpha}'^* = min\{\alpha : \alpha = \Phi'(\alpha)\} \leq min\{\alpha : \alpha = \Phi(\alpha)\} = \underline{\alpha}^*$ and that $\bar{\alpha}'^* = max\{\alpha : \alpha = \Phi'(\alpha)\} \leq max\{\alpha : \alpha = \Phi(\alpha)\} = \bar{\alpha}^*$.

Thus, $\underline{\mu}'^* = \mathcal{S}(\bar{\alpha}'^*) \succeq \mathcal{S}(\bar{\alpha}^*) = \underline{\mu}^*$ and $\bar{\mu}'^* = \mathcal{S}(\underline{\alpha}'^*) \succeq \mathcal{S}(\underline{\alpha}^*) = \bar{\mu}^*$. Likewise, $\mathcal{T}'(\bar{\mu}'^*) = \underline{\alpha}'^* \leq \underline{\alpha}^* = \mathcal{T}(\bar{\mu}^*)$ and $\mathcal{T}'(\underline{\mu}'^*) = \bar{\alpha}'^* \leq \bar{\alpha}^* = \mathcal{T}(\underline{\mu}^*)$.

Part (ii):

We prove by contradiction. Suppose $\mathcal{T}'(\mu'^*) > \mathcal{T}(\mu^*)$. Then $\mathcal{S}(\mathcal{T}'(\mu'^*)) \succeq \mathcal{S}(\mathcal{T}(\mu^*))$ and thus $\mu'^* \succeq \mu^*$. Since $\mathcal{F}'(\mu, \alpha) \leq \mathcal{F}(\mu, \alpha)$ for any $\mu \in \mathcal{M}$ and $\alpha \in [0, 1]$ and since $\mathcal{F}'$ and

$\mathcal{F}$ are decreasing in μ, we have that $\mathcal{F}'(\mu'^*, \alpha) \leq \mathcal{F}(\mu^*, \alpha)$ for any $\alpha \in [0,1]$. Therefore, $\mathcal{T}'(\mu'^*) \leq \mathcal{T}(\mu^*)$, a contradiction. We conclude that $\mathcal{T}'(\mu'^*) \leq \mathcal{T}(\mu^*)$.

It then follows that $\mathcal{S}(\mathcal{T}'(\mu'^*)) \preceq \mathcal{S}(\mathcal{T}(\mu^*))$ and thus $\mu'^* \preceq \mu^*$.

The result extends to games of networked-risk protection by their structural equivalence to games of total protection (see Corollary 1). This completes the proof.

□

Proposition 8. Part (i):

The proof is analogous to that of Proposition 1, with only minor modifications.

Denote the function $\mathfrak{T}(\mu) = (\mathcal{T}(\mu), \mathcal{G}(\mu))$ and let the correspondence Ψ be such that $\Psi(\alpha, \omega) = \mathfrak{T}(\mathcal{S}(\alpha, \omega))$, with the correspondence $\mathcal{S}(\alpha, \omega)$ defined as in Definition 16.

First, note that Ψ still has a compact domain $[0,1] \times [0,1]$ and a nonempty image.

Furthermore, it is also simple to show that Ψ has a closed graph. First, note that $\mathcal{S}(\alpha, \omega)$ has a closed graph when we endow the set of strategies with the product topology on $[0,1]^\infty$. Indeed, choose any $(\alpha_n, \omega_n) \to (\alpha, \omega)$ and $\mu_n \to \mu$ such that $\mu_n \in \mathcal{S}(\alpha_n, \omega_n)$. Then $\mu_n(d) \to \mu(d)$ for any d. Expressing utility as a function of α and ω, i.e. $U_d(a, \alpha, \omega) = -V \cdot \mathcal{B}(p, q_d(\alpha), a) - C \cdot g(\omega) \cdot a$, we note that by the continuity of $U_d(1, \alpha, \omega)$ and $U_d(0, \alpha, \omega)$, it follows that $\mu(d) \in \mathcal{S}_d(\alpha, \omega)$. Thus, $\mathcal{S}$ has a closed graph. Note also that with the product topology on the space of strategies, $\mathfrak{T}$ is continuous: by the bounded convergence theorem, both $\mathcal{T}(\mu_n) \to \mathcal{T}(\mu)$ and $\mathcal{G}(\mu_n) \to \mathcal{G}(\mu)$ and therefore it is also true that $\mathfrak{T}(\mu_n) \to \mathfrak{T}(\mu)$. We now only need to consider the sequences $(\alpha_n, \omega_n) \to (\alpha, \omega)$ and $(\alpha'_n, \omega'_n) \to (\alpha', \omega')$ where $(\alpha'_n, \omega'_n) \in \Psi(\alpha_n, \omega_n)$. By choosing $\mu_n \in \mathcal{S}(\alpha_n, \omega_n)$ such that $\mathcal{T}(\mu_n) = \alpha'_n$ and $\mathcal{G}(\mu_n) = \omega'_n$, and by the same argument as in the proof of Proposition 1, we can conclude that $(\alpha', \omega') \in \Psi(\alpha, \omega)$, as desired.

Finally, the image of Ψ is convex. Indeed, $\mathfrak{T}(\mu)$ is continuous in μ. Furthermore, $\mathcal{S}(\alpha, \omega)$ is convex (which follows from convexity of $\mathcal{S}_d(\alpha, \omega)$ for any d). Convexity of the image of Ψ thus follows from an argument analogous to that presented in the proof of Proposition 1.

By Kakutani's fixed point theorem, Ψ has a fixed point (α^*, ω^*). Letting $\mu^* \in \mathcal{S}(\alpha^*, \omega^*)$ be such that $\mathfrak{T}(\mu^*) = (\alpha^*, \omega^*)$, we conclude that μ^* is an MFE.

Part (ii):

Note that the incremental expected utilities for an agent of degree d in games of total and self protection are respectively:

$$\Delta U_d(\mu) = V \cdot (p + (1-p) q_d(\mathcal{T}(\mu)))\, k - C g(\mathcal{G}(\mu)) \tag{A.12}$$

and

$$\Delta U_d(\mu) = V \cdot (pk - pk q_d(\mathcal{T}(\mu))) - C g(\mathcal{G}(\mu)) \tag{A.13}$$

It is obvious that for any given $\mathcal{T}(\mu)$ and $\mathcal{G}(\mu)$, these functions preserve the properties (i.e. monotonicity in d) that were discussed in the proofs of Theorems 1 and 2. The threshold nature of equilibria is thus maintained. By an argument analogous to that of the proof of Corollary 1, it also follows that a game of networked-risk protection is structurally equivalent to a game of total protection and thus the upper-threshold nature also follows in that case.

Part (iii):

For convenience, since the expected utility $U_d(a, \mu)$ depends on μ only through $\alpha = \mathcal{T}(\mu)$ and $\omega = \mathcal{G}(\mu)$, we may write it as a function of α and ω as follows:

$$U_d(a, \alpha, \omega) = -V \cdot \mathcal{B}(p, q_d(\alpha), a) - C \cdot g(\omega) \cdot a \tag{A.14}$$

For a game of total protection, the incremental expected utility for an agent of degree d is

$$\Delta U_d(\alpha, \omega) = V \cdot (p + (1-p)q_d(\alpha))\, k - C \cdot g(\omega) \tag{A.15}$$

We will consider the case when $g(\omega)$ is an increasing function. As in the proof of Theorem 1, we will conduct the analysis in 4 steps.

Step 1: For all $d \geq 1$, $\Delta U_d(\alpha, \omega)$ is strictly increasing in $\alpha \in [0, 1]$ and strictly decreasing in $\omega \in [0, 1]$.

Step 2: Notice that α and ω are moving $\Delta U_d(\alpha, \omega)$ in opposite directions. Hence, if both α and ω increase, we cannot conclude anything about the change in $\mathcal{S}_d(\alpha, \omega)$. However for $\alpha' > \alpha$ and $\omega' > \omega$ it holds $\mathcal{S}_d(\alpha', \omega) \succeq \mathcal{S}_d(\alpha, \omega')$.

Step 3: For any strategies μ', μ such that $\mu'(d) \geq \mu(d), \forall d \geq 1$, then $\mathcal{G}(\mu') \geq \mathcal{G}(\mu)$ and $\mathcal{T}(\mu') \leq \mathcal{T}(\mu)$. As we have noted before, the global externality does not have a direct impact on $\mathcal{F}(\mu, \alpha)$ and thus the behavior of $\mathcal{T}(\mu)$ remains as in step 3 of the proof of Theorem 1.

Step 4: Suppose that there are two mean-field equilibria $(\mu^*, \alpha^*, \omega^*)$ and $(\mu'^*, \alpha'^*, \omega'^*)$. Without loss of generality assume that $\alpha'^* > \alpha^*$. We need to consider two cases. First, if $\omega'^* \leq \omega^*$, then it is true that $\mathcal{S}(\alpha'^*, \omega'^*) \succeq \mathcal{S}(\alpha^*, \omega^*)$. As $\mu^* \in \mathcal{S}(\alpha^*, \omega^*)$ and $\mu'^* \in \mathcal{S}(\alpha'^*, \omega'^*)$, then it follows that $\mu'^* \succeq \mu^*$. However that leads to the contradiction: $\alpha^* = \mathcal{T}(\mu^*) \geq \mathcal{T}(\mu'^*) = \alpha'^*$. Finally consider the case of $\omega'^* > \omega^*$. By Proposition 8(ii), due to the threshold nature of the equilibrium, the equilibrium strategies can be ordered as either $\mu'^* \succeq \mu^*$ or $\mu'^* \preceq \mu^*$. If $\mu'^* \preceq \mu^*$ then $\omega^* = \mathcal{G}(\mu^*) \geq \mathcal{G}(\mu'^*) = \omega'^*$, which is a contradiction. If $\mu'^* \succeq \mu^*$, it follows that $\alpha^* = \mathcal{T}(\mu^*) \geq \mathcal{T}(\mu'^*) = \alpha'^*$ and we arrive at a contradiction.

Thus, we showed that in a game of total protection with both endogenized cost (with $g(\cdot)$ increasing), any MFE must be unique. Uniqueness in the case of a game of networked-risk protection follows by its structural equivalence to a game of total protection (Corollary 1). □

Acemoglu, Daron, Malekian, Azarakhsh, & Ozdaglar, Asuman. (2013). *Network security and contagion*. Tech. rept. National Bureau of Economic Research.

Acemoglu, Daron, Ozdaglar, Asuman, & Tahbaz-Salehi, Alireza. (2015). Systemic risk and stability in financial networks. *The american economic review*, **105**(2), 564–608.

Arribasa, I, & Urbanoa, A. (2014). *Local coordination and global congestion in random networks*. Tech. rept. University of Valencia, ERI-CES.

Aspnes, James, Chang, Kevin, & Yampolskiy, Aleksandr. (2005). Inoculation strategies for victims of viruses and the sum-of-squares partition problem. *Pages 43–52 of: Proceedings of the sixteenth annual acm-siam symposium on discrete algorithms*. Society for Industrial and Applied Mathematics.

Balthrop, J., Forrest, S., Newman, M.E.J., & Williamson, M.H. (2004). Technological networks and the spread of computer viruses. *Scientific reports*, **304**, 527–529.

Blume, Lawrence, Easley, David, Kleinberg, Jon, Kleinberg, Robert, & Tardos, Éva. (2013). Network formation in the presence of contagious risk. *Acm transactions on economics and computation*, **1**(2), 6.

Cabrales, Antonio, Gottardi, Piero, & Vega-Redondo, Fernando. (2014). Risk-sharing and contagion in networks. *Ssrn 2425558*.

Cerdeiro, Diego, Dziubiński, Marcin, & Goyal, Sanjeev. (2015). Contagion risk and network design. *Ssrn 2619022*.

Dziubiński, Marcin Konrad, & Goyal, Sanjeev. (2016). How do you defend a network? *Theoretical economics*.

Elliott, Matthew, Golub, Benjamin, & Jackson, Matthew O. (2014). Financial networks and contagion. *The american economic review*, **104**(10), 3115–3153.

Gagnon, Julien, & Goyal, Sanjeev. (2015). Networks, markets and inequality.

Galeotti, Andrea, & Rogers, Brian W. (2013). Strategic immunization and group structure. *American economic journal: Microeconomics*, **5**(2), 1–32.

Galeotti, Andrea, Goyal, Sanjeev, Jackson, Matthew O., Vega-Redondo, Fernando, & Yariv, Leeat. (2010). Network games. *Review of economic studies.*, **77**, 218–244.

Goyal, Sanjeev, & Vigier, Adrien. (2015). Interaction, protection and epidemics. *Journal of public economics*, **125**, 64–69.

Heal, Geoffrey, & Kunreuther, Howard. (2004). *Interdependent security: A general model*. Tech. rept. National Bureau of Economic Research.

Heal, Geoffrey, Kearns, Michael, Kleindorfer, Paul, & Kunreuther, Howard. (2006). Interdependent security in interconnected networks.

Jackson, Mathew O. (2008). Social and economic networks. *Princeton university press, nj.*

Jackson, Matthew O., & Yariv, Leeat. (2007). Diffusion of behavior and equilibrium properties in network games. *American economic review*, **97**(2), 92–98.

Jackson, Matthew Owen, & Zenou, Yves. (2014). Games on networks. *Handbook of game theory*, **4**. (Peyton Young and Shmuel Zamir, eds.).

Johnson, Benjamin, Böhme, Rainer, & Grossklags, Jens. (2011). Security games with market insurance. *Pages 117–130 of: Decision and game theory for security.* Springer.

Leduc, Mathieu V. (2014). *Mean-field models in network game theory.* Ph.D. thesis, Stanford University.

Leduc, Matt V, Jackson, Matthew O, & Johari, Ramesh. (2015). Pricing and referrals in diffusion on networks. *arxiv preprint arxiv:1509.06544.*

Lelarge, Marc, & Bolot, Jean. (2008a). A local mean field analysis of security investments in networks. *Pages 25–30 of: Proceedings of the 3rd international workshop on economics of networked systems.* ACM.

Lelarge, Marc, & Bolot, Jean. (2008b). Network externalities and the deployment of security features and protocols in the internet. *Sigmetrics'08.*

Lelarge, Marc, & Bolot, Jean. (2009). Economic incentives to increase security in the internet: The case for insurance. *Pages 1494–1502 of: Infocom 2009, ieee.* IEEE.

Reuters. (12 October 2015). *Cyber insurance premiums rocket after high-profile attacks.*

Reuters. (27 August 2015). *U.s. vaccination rates high, but pockets of unvaccinated pose risk.*

Rosas-Casals, M., Valverde, S., & Solé, R. V. (2007). Topological vulnerability of the european power grid under errors and attacks. *International journal of bifurcation and chaos*, **17**(7), 2465–2475.

The Economist. (4 February 2015). *Rand paul on vaccination: Resorting to freedom.*

The Economist. (5 February 2015). *Politics and vaccinations: What experts say, and what people hear.*

Wang, Z., Scaglione, A., & Thomas, R. J. (2010). The node degree distribution in power grid and its topology robustness under random and selective node removals. *2010 ieee international conference on communications workshops (icc)*, 1–5.