Risk-sharing and contagion in networks

Antonio Cabrales Piero Gottardi
Universidad Carlos III European University Institute

Fernando Vega-Redondo
European University Institute

March 2012

Abstract

The aim of this paper is to investigate how the capacity of an economic system to absorb shocks depends on the specific pattern of interconnections established among economic agents. The key trade-off at work is between the risk-sharing gains enjoyed by firms when they become more interconnected and the large-scale costs resulting from an increased risk exposure. We focus on two dimensions of the network structure: the size of the (disjoint) components into which the network is divided, and the “relative density” of connections within each component. Depending on the structure of the shocks (e.g. whether their distribution displays “fat” or “thin” tails), we find that the efficient configuration can vary widely in both of the aforementioned dimensions (e.g. ranging from extreme segmentation to none at all, or from complete to sparse connections). We also find that there is typically a conflict between efficiency and pairwise stability, due to a “size externality” that is not internalized by agents who belong to components that have reached an individually optimal size.

1 Introduction

Recent economic events have made it clear that looking at financial entities in isolation gives an incomplete, and possibly very misleading, impression of
the potential impact of shocks to the financial system. The aim of this paper is to investigate how the capacity of the system to absorb shocks depends on the specific pattern of interconnections established among economic agents – to fix ideas, we shall think of them as banks. The key issue we would like to analyze is the trade-off between risk-sharing and contagion. Or, to be more specific, we want to shed light on the extent to which the risk-sharing benefits to firms of becoming more highly interconnected (which provides some insurance against moderate shocks) may be offset by the large-scale costs resulting from an increased risk exposure (which, for large shocks, could entail a large wave of induced bankruptcies). Clearly, this trade-off must be at the center of any regulatory efforts of the financial world that takes a truly systemic view of the problem. This paper highlights and studies some of the considerations that should play a key role in this endeavor. In particular, by formulating the problem in a stylized and analytically tractable framework, it studies how segmentation, density of connections, or size asymmetries should be tailored to the underlying shock structure. It also sheds light on the key issue of whether the induced normative prescriptions are consistent with the individual incentives to form or remove links.

More precisely, we analyze a model in which there is a network consisting of $N$ nodes, each of them interpreted as a firm. For simplicity, in most of the paper we shall consider the case where all firms are *ex ante* symmetric and have the same level of assets and liabilities. But *ex post* they will be different since we assume that, with some probability, a shock arrives and individually hits the investment project operated by a randomly selected firm. The first *direct* effect of such a shock is to decrease the income generated by its assets, thus possibly leading to the default of the firm if the resulting income falls short of its liabilities. But if this firm has links to other firms, the latter will also be affected. To be specific, let us think of the presence of a link between two firms as reflecting an exchange of their currently held assets. Then, the overall network of connections generates patterns of mutual exposure between any pair of directly or indirectly connected firms, the magnitude of such exposure decreasing with their respective network distance. Thus, when
A shock hits a firm, all the firms which are in its own network component become affected in proportion to their exposure to that firm. So, in the end, it is the overall network structure that determines how any given shock affects different firms and what is its overall aggregate impact on the whole system.

In order to concentrate the analysis on our basic trade-off—insurance versus contagion—we focus on two dimensions of the network structure. One is the size of the (disjoint) components into which the network is divided, i.e. the degree of segmentation of the system. The other concerns the “relative density” of connections within each component, as measured by the fraction of all nodes in the component that lie at different network distances.

Network density is important because, as explained, different network distances yield different degree of exposure. In this respect, our analysis will largely focus on contrasting two polar (firm-symmetric) cases: (i) completely connected components where there is a direct link between any pair of firms in each component; (ii) minimally connected components, where each firm has the minimum number of links (i.e. two) required to obtain indirect connectivity to every other firm in the component. In the first case (complete components), the mutual exposure between any pair of firms in the same component is exactly the same. Instead, in the second case (minimally connected components), the reciprocal exposure between two firms is heterogeneous, falling with their network distance.

When firms are connected as described (i.e. through the exchange of assets), once a shock hits the project run by a particular firm, other firms connected to it are affected and may even default. Specifically, this will happen for any firm whose exposure to the shock (as measured by the fraction of its currently held assets that are affected by it) exceeds the sum of firm’s capital and liabilities. Under these conditions, a key concern is to identify the architecture of the system that minimizes the expected number of defaults in the system. This involves finding both the best degree...
of segmentation as well as the optimal link density within each separate component. As explained, the objective is to strike the best compromise between risk sharing and contagion.

Naturally, the maximum extent of risk sharing obtains when all firms form part of a single and fully connected network. But this configuration, of course, also yields the highest exposure to a large shock, which could lead to the extensive default across the whole system. There are two alternative (and in some cases complementary) ways of reducing such exposure. One is by segmentation, which isolates the firms in each component from any shock that might hit every other component. The second one is reducing the density of connections in each component, which buffers the network-mediated propagation along this component of any shock that hits one of its firms.

The paper proposes an abstract model that captures the essence of the problem and allows the study the aforementioned trade-off under a fairly general structure of shocks. Some of our results can be summarized as follows. First we find that when the probability distribution of the shocks exhibits “fat tails” (i.e. attributes a high mass to large shocks), the optimal configuration involves a maximum degree of segmentation – that is, components should be of the minimum possible size. This reflects a situation where the priority is to minimize contagion. Instead, in the opposite case where the probability distribution places high enough mass on relatively small shocks, the best configuration involves having all firms arranged in a single component. The aim in this latter case is to achieve the highest level of risk sharing. These two polar cases, however, do not exhaust all possibilities. For, quite naturally, we also find that there are suitable specifications of the shock structure (e.g. a mixture of fat and thin tails) where intermediate arrangements are optimal, i.e. the optimal degree of segmentation involves medium-sized components.

It is interesting to note that all of the previous conclusions hold irrespectively of whether components are assumed to have either of the two network structures considered, i.e. complete or lattice structures. But, in fact, an
analogous trade-off between risk-sharing and contagion can be attained by changing network density. This is indeed the route explored by our second set of results. The key conclusion there is that, under some circumstances, low density can be preferred to more segmentation as the mechanism for limiting contagion when the shocks become large. More specifically, the potential advantage of a lattice configuration is that firm exposure between firms in a component is not uniform but decays with network distance. Thus, if shocks are large (but not too much), only a fraction of the firms will default while all would if they were completely connected. Of course, an obvious consequence of this observation is that the lattice structure always induces an optimal degree of segmentation that is lower (or, equivalently, a component size that is larger) than in the case of the completely connected structure.

The results described so far consider structures where all firms have the same size. We also explore an extension with asymmetric firms, which makes it natural to study asymmetric structures. We focus, specifically, on star architectures, where large firms acts as “hubs” for small firms by providing them with their sole connections. Some of our results – like the trade-off between risk sharing and contagion – extend to this framework, but also new phenomena appear. In a star, for example, the center can become a sort of firebreak that prevents contagion from extending to all the spokes, and this can be useful for some intermediate values of the shock. We close our analysis by comparing the benefits of symmetric structures (where only firms of the same size, big or small, connect) to asymmetric ones such as the star where large firms connect to smaller firms. This comparison shows a rich interplay of size and network structure that, for both small and large shocks, opens parameter ranges where not only asymmetric but also symmetric structures are uniquely efficient.

Finally, the paper addresses the issue of whether the requirements for overall efficiency are compatible with the incentives of individuals to establish links. Formally, we model those incentives through a notion that has been formerly used in the network literature, bilateral equilibrium. For simplicity, we also restrict our attention to completely connected components,
so agents solely determine the degree of segmentation of the network. In this context, we find that there is typically a conflict between efficiency and strategic stability. This conflict derives from the fact that strategic stability typically requires significant asymmetries in component size that are inconsistent with efficiency. There are, therefore, positive externalities associated to symmetry in component size that are not internalized at a bilateral equilibrium. This, in general, forces some agents to lie in a component that is inefficiently small.

To sum up, our analysis highlights that the efficient configuration of the network – concerning, in particular, its segmentation, density, and the handling of size asymmetries – crucially depends on the nature of the shocks faced by the system. But since, as explained, one cannot generally expect that social and individual incentives be aligned, an important role for policy opens up. Our model, of course, is too stylized to allow for the formulation of concrete policy advice. It provides, however, a theoretical framework that is useful to understand the core issues and trade-offs involved, thus in turn allowing the construction of models with richer institutional detail and specific policy implications.

We end this introduction with a review of related literature. The research on financial contagion and systemic risk is quite diverse and also fast-growing. Hence we shall provide here only a brief summary of some representative instances.

Allen and Gale (2000) pioneered the study of the stability of interconnected financial systems. They propose a model in the Diamond and Dybvig (1983) tradition, with shocks affecting the aggregate preference for liquidity of depositors. Given the specific kind of shocks they consider, their main conclusion is that a complete (non-segmented) network is always the effi-

The reader is referred to Allen and Babus (2009) for a recent survey on how risk sharing in financial contexts can lead, through contagion, to large systemic effects. There is also a large body of literature that studies the general problem of risk sharing in non-financial contexts, largely motivated by its application to consumption sharing in poor economies that lack formal insurance mechanisms. Paradigmatic examples are the papers by Bramoullé and Kranton (2007), Bloch et al. (2008), and Ambrus et al. (2011)
cient structure, i.e. the one that minimizes the extent of default in every case. Our model, in contrast, shows that a richer shock structure can generate a genuine trade-off between risk-sharing and contagion in which some segmentation or/and low density are efficient. A paper where lower density of interaction is found to be efficient is Freixas et al. (2000), where again the shocks are on consumer liquidity needs but, in contrast with Allen and Gale (2000), they have ”geographical” dimension, i.e. they are related to the ”traveling flows” of agents across neighboring locations lying along a one-dimensional ring. They show that restricting links to connecting neighboring banks is beneficial because, even though it limits risk sharing, it has the positive consequence of reducing the incentives for deposit withdrawal. On the other hand, Leitmer (2005) obtains a positive role for segmentation, but in a context where agents organize in groups, which can be regarded as completely connected. Thus, in order to share risk, agents commit to transferring resources within each group, leading in the end to a situation where either the whole group succeeds or the whole group fails. No role is played, therefore, by the structure of interaction.

The recent paper by Allen, Babus, and Carletti (2011) is perhaps closest to in approach and motivation to this paper. They consider a six-firm context where each individual firm faces the need to find funds for its respective investment. Since these investments are risky, firms can gain from risk diversification, which is achieved by exchanging shares with two other firms. This gives rise to the financial network, for which two possibilities are considered: the segmented and the unsegmented structures. In each case, a signal arrives indicating whether there will be any default. This, of course, has different implications in the two structures – i.e. induces different posterior beliefs on the default probability of any given firm. Hence, it triggers different reactions by investors, which can force liquidation in one case and not in the other (depending largely on the magnitude of liquidation and bankruptcy costs).

Finally, let us refer to a complementary line of literature that, in contrast with the papers just summarized, studies the issue of contagion and systemic risk in the context of large networks (typically, randomly gener-
ated). In most of these papers, the approach is numerical, based on large-scale simulations (see e.g. Nier et al. (2007) for a good representative). Here, however, we shall focus on the recent paper by Blume et al. (2011), which integrates the mathematical theory of random networks with strategic analysis of network formation.

These authors model the financial structure as an (asymptotically large) Erdös-Rényi network with given average degree (i.e. connectivity), the key parameter. Assuming that payoffs are linear in degree and shocks (that render agents worthless) are transmitted through the links, the key issue is two-fold: what is the (socially) optimal degree, and whether this is consistent with agents’ incentives to (dis)connect. The answer is that social optimality is attained around the threshold where a large (so-called giant) component of defaults emerges, but individuals will want to connect beyond this point. So again we have a conflict between social and individual optimality, which in this case is due to the fact that agents will not internalize the effect on others of adding new channels (i.e. links) for the spread of contagion.

The rest of the paper is organized as follows. Section 2 discusses the model: first a discrete version, then its continuum idealization. Section 3 undertakes the analysis of optimal financial structures under a variety of different assumptions on the underlying structure. Section 4 addresses network formation and explores the tension between strategic stability and optimality. Finally, Section 5 concludes with a summary and an agenda for future research. For the sake of smooth discussion, all formal proofs of our results are relegated to an Appendix.

2 The model

2.1 The Environment

We consider an environment with \( N \) ex ante identical, risk-neutral financial firms and a continuum of small investors. At any given point in time, each firm has an investment opportunity - a project - which requires an initial payment equal to \( I \) and yields a random gross return \( \tilde{R}I \) at the end of
the period. The resources needed to undertake the project are obtained by
issuing liabilities (e.g. deposits or bonds) on which a deterministic rate of
return must be paid.

The random return of the project depends on whether the firm enjoys
normal returns or it is hit by a shock. In the first case, its gross returns are
deterministically equal to some normal level $R$. Instead, if the firm is hit by
a shock, we find it useful to consider two possibilities. One is that the shock
is small, which we label as $s$. In that case, the firm experiences a loss of
some fixed size $L_s$, so its gross return is $R - L_s$. The second possibility is
that the shock is large, labeled as $b$. Then, the induced loss $\bar{L}_b$ is ex ante
random, with distribution function $\Phi(L_b)$. Denote by $q$ the probability that
some shock hits the firm, and by $\pi_s$ and $\pi_b$ the conditional probabilities that
the shock be small and large, respectively. Then, the random variable $\bar{R}$ can
be summarized as follows: Summarizing, the gross unit rate of return on a
firm’s project is:

$$\bar{R} = \begin{cases} 
R & \text{with prob. } 1 - q \\
R - L_s & \text{with prob. } q \pi_s \\
R - \bar{L}_b & \text{with prob. } q \pi_b
\end{cases}$$

(1)

Since the return on a firm’s investment is subject to shocks, while the
return promised to depositors is deterministic, when hit by a shock the firm
may not be able to meet the required payments on its liabilities, in which
case it must default. Default entails two types of costs. First, there are
the liquidations costs – for simplicity, we assume that these costs leave no
resources available to make payments to creditors at the time of default.
Then, there are the additional costs derived from the loss of future earnings
possibilities for any defaulting firm. These costs are assumed to be very
substantial, so the primary objective of a firm is to minimize the probability
of default. And, from the point of view of overall social welfare, the aim
must be to minimize the expected number of defaults. As we shall explain
below, both criteria are equivalent, so that from an ex ante perspective,
individual and social incentives are fully aligned.

The supply of funds to the firms comes from the investors. We assume
these are risk neutral and require an expected gross rate of return equal
to $r$ on their investments. Since firms may default, in which case creditors receive a payment equal to zero, the gross rate of return on the deposits to the firms is set at a level $M$ that is greater or equal than $r$. Specifically, if we denote by $\varphi$ the *ex ante* probability that any given firm defaults (an endogenous variable), then the risk-arbitrage condition determining $M$ can be written as follows:

$$M = \frac{r}{1 - \varphi}. \quad (2)$$

To render the problem interesting, we assume:

**A1.** (i) $R(1 - q) > r$

(ii) $R - L_s < r$.

The first inequality ensures that a firm’s project is viable, that is, its expected return exceeds the expected return which must be paid to lenders when they anticipate no default. The second inequality implies, since $r \leq M$ and $\tilde{L}_b \geq L_s$, that if a firm can only draw on the resources generated by its project it is surely unable to pay depositors (and hence must default) when a shock $b$ hits its return.

Since, as stated above, default entails a significant cost for a firm, a firm may benefit from entering risk sharing arrangements with other firms which allow it to diversify risks. Here we consider the case where these arrangements take the form of swaps of assets between firms, that is, exchanges of claims to the yields of the firms’ investments (similarly to Allen, Babus and Carletti (2011)). These asset exchanges can be viewed as a *securitization* process of the firms’ claims, by which we indicate the (possibly iterative) procedure through which firms exchange shares on their whole array of asset holdings.

More precisely, let us posit that each firm can exchange a fraction $1 - \theta$ of its standing shares for an equal fraction of the assets held by other “neighboring” firms. Such an exchange entitles the firm to the returns earned by the new assets, as well as makes it responsible for its liabilities. Note that, due to the *ex ante* symmetry of firms, such an exchange involves trading assets of equal expected return. The specific pattern of exchange among
firms is formalized by a network, where a direct linkage between two firms reflects that they undertake a *direct* exchange of their assets. We allow for these asset swaps to occur repeatedly, once firms’ portfolios are securitized. Then, even firms that have only an *indirect* connection through intermediaries will end up having some reciprocal exposure to each other’s projects if the number of exchange rounds is high enough – in particular, as high as their network distance. In this case, therefore, the return on a firm’s assets in the end becomes a weighted average of the return on its project and the projects of the firms it traded with, directly or indirectly. A convenient way of representing these weights is through a matrix $A$ of the form

$$
A = \begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1N} \\
a_{21} & a_{22} & \cdots & a_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
a_{N1} & a_{N2} & \cdots & a_{NN}
\end{pmatrix}
$$

(3)

where for each $i, j$, $(i \neq j)$, $a_{ij} \geq 0$ denotes the fraction of shares in the investment project run by firm $i$ that is owned by firm $j$. By construction, therefore, the following adding-up constraint must be satisfied:

$$
\sum_{j=1}^{N} a_{ij} = 1 \quad (i = 1, 2, \ldots, N)
$$

(4)

On the other hand, the fact that all firms are ex-ante symmetric and the portfolio swaps are conducted one-for-one basis implies that $A$ is symmetric, i.e.

$$
a_{ij} = a_{ji} \quad \text{for all} \quad i, j = 1, 2, \ldots, N.
$$

In order to establish meaningful comparisons across different network configurations, we shall assume that, after all rounds of asset exchange have been completed, the induced network structures (as captured by the corresponding exposure matrix $A$) always involve the same “externalization of risk” by each agent. Alternative structures, therefore, differ only in terms of how such externalized risk is distributed among the rest of the agents. Formally, this amounts to imposing a common value of $a_{ii} \equiv \alpha$ for each $i = 1, 2, \ldots, N$ and every network configuration, where $\alpha$ is taken to be a
parameter of the model, bounded above zero independently of \( N \). A natural motivation for imposing such lower bound derives from considerations of moral hazard. If the effort required to operate one’s project is costly, a way to induce a high enough effort is to make sure that agents hold a certain share on the induced returns, no matter what is the underlying asset-exchange structure or the number of securitizations rounds.

For simplicity, shocks are assumed to be rare and thus taken to come at most one at a time. In the framework described at the beginning of this section (cf. (1)), this can be motivated by postulating that, even if shocks hit firms in a stochastically independent manner, the probability \( q \) that a shock hits any given firm is so low that the probability that two or more shocks arrive in a single period is of an order of magnitude that can be ignored. Or, as an extreme formalization of this idea, we could model time continuously and assume that the arrivals of small and big shocks to each firm are governed by independent Poisson processes with fixed rates \( \pi_s \) and \( \pi_b \), respectively.

Eventually, if the process runs for long enough, a shock does hit somewhere in the system with probability one. And when a shock of size \( L \) hits the return on the project run by some firm \( i \), the exposure to it of all firms in the system is given by \( A e_i L \), where \( e_i \) is the \( i \)-th unit vector \([0, \ldots, 1, \ldots, 0]^\top\). And this implies that firm \( i \) will default in response to such a shock when

\[
\alpha(R - L) + \sum_{j \neq i} a_{ij} R < M \iff \alpha L > R - M.
\]

While firm \( k \neq i \) defaults whenever

\[
\left( \alpha + \sum_{i \neq j \neq k} a_{kj} \right) R + a_{ki}(R - L) < M.
\]

We readily see, therefore, that when a firm exchanges its assets with others, it reduces its exposure to the shocks hitting its own project but also becomes

\(^3\)This requirement is analogous, for example, to a well-known provision in the recent Dodd-Frank act, recently passed in the USA to regulate further the financial system, by which, under certain circumstances, “a securitizer is required to retain not less than 5 percent of the credit risk...” (see http://www.sec.gov/about/laws/wallstreetreform-cpa.pdf).
exposed to the shocks affecting the projects of those firms with which the firm in question is directly or indirectly connected. On the nature of those shocks, we make the following key assumption

**A2.** (i) \(\pi_s > N\pi_b\).

(ii) \(\alpha(R - L) + (1 - \alpha)R \geq \frac{r}{1 - N\pi_b}\).

Part (i) in the above assumption says that small \((s)\) shocks are significantly more likely than big \((b)\) ones. It implies, in particular, that any given firm finds it more likely that it be hit individually by a small shock than by a big shock impinging anywhere in the overall system of \(N\) firms. The relevance of this condition has to be understood in combination with Part (ii) of this same assumption. The right-hand side of \(A2(ii)\) is simply an upper bound on \(M\), the (minimum) payment promised to investors in order to induce them to invest their funds in any of the banks, if it can be assured that none of them will to fail by the presence of a small shock alone. To see this, note that, under the aforementioned caveat, the worst-case scenario corresponds to a situation where a big shock hitting anywhere in the system is enough to induce the default of all firms. But if, in this scenario, the payment \(W\) that is guaranteed to investors must satisfy:

\[(1 - N\pi_b)W \geq r.\]

Since \(M\) can be no higher than any \(W\) that satisfies the above inequality, the desired upper bound follows. On the other hand, note that the left-hand side of the inequality in \(A2(ii)\) gives a lower bound on the gross return of a firm that is hit by a small shock. Thus, indeed, this inequality guarantees no default of a firm that is hit by a small shock, provided \(W\) that no other firm defaults in the same circumstances.

Consider now Assumption A2 (i)-(ii) in combination with A1(ii). They jointly guarantee that every firm will be interested in forming some links. For, on the one hand, A1(ii) implies that a firm in isolation that is hit by any shock goes into default. On the other hand, A2(ii) indicates that a firm avoids default when hit by a small shock if it enjoys any connections – i.e. if
its portfolio displays some diversification. This diversification keeps the firm safe from contagion by any small shock in the system but, in principle, does not protect it from contagion if a big shock arrives somewhere. However, the fact that, by A2(i), the latter event is less likely than being individually hit by a small shock means that any firm whose objective is to minimize the probability of default will want to form some links.

The previous argument makes clear that any pattern of asset exchange and diversification allows firms to attain full insurance against $s$ shocks. It is also clear, however, that the default performance in the event of $b$ shocks will often not be the same across different financial network structures. In general, that is, these structures will be markedly different in terms of the extent of contagion they induce when big shocks hit the system. Hence, in a nutshell, one of the primary aims of this paper is to understand what configurations minimize those detrimental side effects of risk-sharing.

To be precise, relying on the assumption that default costs are pre-eminent and agents are risk neutral, we posit that social welfare is maximized when the expected number of defaults is minimal. If agents are in an \textit{ex ante} symmetric situation, it is easy to see that this criterion is equivalent to that of minimizing the individual probability that any single agent defaults. In this case, therefore, social and individual incentives are aligned from an \textit{ex ante} viewpoint. However, as we discuss in Section 6, a wedge between them can arise if asymmetric configurations (specifically, on component size)

\footnote{Beale et al. (2011) propose a similar criterion to evaluate and compare different financial systems, based on the minimization of a “systemic cost function” defined as the expectation of a convex function of the number of defaults in the system. Thus, in this case, not only the expected number of defaults matters, but also its variability.}

\footnote{To see this, continue to denote by $\varphi$ the individual \textit{ex ante} probability of default and then write $\varphi = \sum_m \rho(m) \varpi(m)$, where $\rho(m)$ stands for the probability that $m$ defaults occur and $\varpi(m)$ for the conditional probability that any particular firm defaults when there are a total of $m$ defaults in the system. Then, since

$$
\varpi(m) = \frac{1}{N} + (1 - \frac{1}{N}) \frac{1}{N-1} + \cdots + (1 - \frac{1}{N} - \cdots - \frac{1}{N}) \frac{1}{N-m+1} = \frac{m}{M}
$$

we obtain $\sum_m \rho(m)m = \varphi N$, i.e. the expected number of defaults is proportional to the individual default probability.
are considered.

**Remark 1** We should point out that in the environment considered the mutual exposure between firms comes from the crossownership of their shares, not from mutual lending relationships. Hence the default of one firm has no direct implication for the solvency of other firms, the possibility of contagion of a large shock hitting a firm only comes from the correlation of their portfolio returns.

### 2.2 Financial Structures

We shall consider financial structures that differ along two dimensions: segmentation and network density. For the moment, we shall also restrict our analysis to symmetric configurations – which means, in particular, that all components of the network must be of identical size. This, however, implies no loss of generality for our normative analysis since, as we show below, efficient configurations are essentially symmetric. The assumption of symmetry will be then be relaxed in Sections 6 and 5, where we respectively turn to studying the implications of strategic stability (which typically requires unequally sized components) and firms that are themselves asymmetric (small and large).

By *segmentation* we mean the partition of the $N$ firms into disjoint components. Each component is formed by firms that are either directly or indirectly linked by the exchange of assets (and hence the crossownership of shares), while there is no trade across components. The measure of this segmentation is given by the number $C$ of equal-sized components in which the set of firms is divided, or equivalently by the number $K = \frac{N}{C} - 1$ of other firms to whom every firm is linked. In terms of the matrix $A$, this amounts to having a block diagonal structure with $C$ blocks along the diagonal. The larger the segmentation (the smaller $K$), the fewer the firms affected by a given shock but, *ceteris paribus*, the larger their mutual exposure and hence the probability of default if a $b$ shock hits them. At the two extremes of segmentation, we have the case $K = N - 1$, where all firms are connected (directly or indirectly), and $K = 1$, where each firm only engages in trade
with a single other firm. We shall allow for all possible values of $K$, between 1 and $N - 1$ and explore the welfare implications.

On the other hand, by network density, we refer to the proportion of direct versus indirect linkages within each component. A firm is said to be directly linked to another one if both are involved in a direct exchange of assets. In this case, we may describe the situation by saying that there is a link (or an edge) in the underlying financial network. Instead, two firms are only indirectly linked if they are connected by a multiple-link path in the financial network. In essence, what this means is that they can end up sharing some assets (and thus display some correlation of returns) if there is a repeated exchange of assets undertaken through, say, rounds of securitization. In general, of course, we can have different levels of density, ranging from 1 (all linkages are direct) to the case where the amount of direct connectivity is minimal (i.e. two links per firm, if all firms are to be in a symmetric situation within the component). For simplicity, our analysis will just compare these two extremes, which are described in more detail in what follows.

**Complete Structures**

In this case, all firms in each component have a direct link – so, in the language of networks, the component is completely connected. Hence, by symmetry, the size of each component fully characterizes the situation, up to normalization (see below). For convenience, the component size will be denoted by $K + 1$, where $K$ is the number of partner firms with which any given firm exchanges its assets. To formalize these exchanges precisely, let $\theta$ be the (common) fraction of assets each firm keeps for itself after any such round of asset exchange – hence it exchanges the fraction $1 - \theta$. Then, the pattern of exchange in a complete component of size $K + 1$ may be captured by a square matrix $B_K$ of the form:
$$B_K = \begin{pmatrix} 
\theta & (1-\theta)/K & \cdots & (1-\theta)/K \\
(1-\theta)/K & \theta & \cdots & (1-\theta)/K \\
\vdots & \vdots & \ddots & \vdots \\
(1-\theta)/K & (1-\theta)/K & \cdots & \theta 
\end{pmatrix}. \quad (5)$$

With such pattern of exchanges, if firms undergo a repeated process of securitization and exchange for some $m$ rounds, the eventual pattern of exposure in the component can be formalized by a matrix $A_K$, with the same format and interpretation as the matrix in (3), except for its has dimension $K+1$, i.e. the number of different projects run by the firms in the component. This matrix is obtained by repeated composition of $B_K$ with itself $m$ times, i.e.

$$A_K = (B_K)^m. \quad (6)$$

Clearly, the matrix $A_K$ displays the same qualitative features as the original $B_K$, namely, it is a row-stochastic matrix where all entries along the main diagonal are identical and so happens as well among all entries that are off the main diagonal. Once we then add the normalization criterion that imposes a fraction $\alpha$ of a firm’s portfolio to be invested in its own investment project, all its entries $(a_{ij})_{i,j=1}^K$ become uniquely determined as follows:

$$a_{ii} = \alpha \forall i = 1, 2, ..., K \quad (7)$$

$$a_{ij} = (1-\alpha)/K \forall i, j = 1, 2, ..., K, i \neq j. \quad (8)$$

**Minimally Connected Structures**

This corresponds to the situation where firms display the minimum number of links required to be connected, directly or indirectly (i.e. through a path of some length), to all other firms in the component. If all firms are in fully symmetric situation, it must then be that all firms have exactly two links. The architecture of every component, therefore, is that of a ring. This implies that there is a suitable labeling of the $K+1$ firms in the component such that the pattern of (direct) exchange of assets is given by the matrix
As before, we assume that the firms conduct \( m \) rounds of securitization, the matrix of final exposure \( A_K \) being obtained as in (6). In this respect, it is natural to posit that \( m \) is large enough such that all firms in the component bear some reciprocal exposure – otherwise, their lying in the same component would have no implications whatsoever for our purposes. This obviously requires that \( m \geq K/2 \). In this case, all entries in the corresponding matrix \( A_K \) will display positive entries, reflecting the aforementioned reciprocal risk exposure of every two firms in the component. Again for normalization (see the discussion above), we suppose that the main-diagonal entries of this matrix are all equal to \( \alpha \).

It is easy to see that, as \( m \to \infty \), the corresponding matrix \( A_K \) displays, in the limit, a configuration where all entries (both on- and off-the main diagonal) are equal to \( 1/(K+1) \). Indeed, the same conclusion applies to any pattern of exchange (not just those considered here), which suggests that the contrast between alternative structures is best highlighted when \( m \) is relatively low. This leads us to consider the lowest value of \( m \) that preserves the connectivity of the ring, i.e. \( m = K/2 \), where for simplicity we assume that \( K \) is even. In this case, the effect of network distance on firm exposure is highest, thus allowing for the clearest contrast between the complete and the ring architectures. In the former case, as we have explained, the exposure \( a_{ij} \) of any firm \( i \) is constant across all \( j \neq i \). In the latter, however, \( a_{ij} \) is decreasing in the ring distance between \( i \) and \( j \). It is highest at \( i = j \), with \( a_{ii} = \alpha \), and is lowest when the ring distance between \( i \) and \( j \) attains the maximum value of \( K/2 \).
The Continuum Approximation

The pattern of risk exposure induced by either the complete or ring structures can be graphically depicted through a function that, for each firm $i$, specifies the fraction of the portfolio held by each $j$, as a function of the (integer) “network distance” between both firms. In the case of the complete structure, such a function is constant for all $j \neq i$, while for the ring it is monotonically decreasing. Because of the discreteness of the domain, it becomes very involved and tedious to carry out a formal analysis of the situation – in particular, for the case of the ring structure, for which uninteresting integer considerations arise. This leads us to abstracting from these considerations by studying a continuum approximation of our model that captures the essential features of the problem. In what follows, we present and discuss the approach.

In the continuum approximation of the model, the number of firms, $N$, is taken to be a measure and the same applies to the number of firms belonging to a certain component. If the component is complete, all firms are neighbors. Instead, if the component is minimally connected, firms are taken to be placed continuously along a ring whose length equals the component size. In order to keep a formal parallelism with the discrete formulation, we shall assume that any arriving shock directly affects a unit measure of adjacent firms in one component. These firms, therefore, play the role of the single firm directly hit by a shock in the discrete context. Also to keep the formal correspondence between the two setups, the size of those components will continue to be identified with $K + 1$ so that, if a shock affects the component, there is a measure $K$ of firms indirectly affected by it.

In such a continuum setup, segmentation is modeled as in the previous section except that the size of a component, $K + 1$, can be any real positive

---

6This continuum formulation can be seen as representing a limit description of a context consisting of a large number of small firms where each of them is of relatively small size. Our preferred motivation, however, is to view it simply as a smooth approximation of a context where the number of firms is discrete but not necessarily large.
number lying between 1 and \( N \) (i.e. it need not be an integer). Differences in network density, on the other hand, are modeled as follows. If the component is complete, we continue to assume that the pattern of exposure is given by a portfolio where each firm holds a proportion \( \alpha \) of its assets on claims involving its own project and the complementary proportion \( (1 - \alpha) \) on uniform claims evenly spread over the projects run by the mass \( K \) of other firms in the component. Alternatively, if the structure of asset exchange is given by a ring, the pattern of risk exposure in a component of size \( K + 1 \) is assumed defined by a continuous function \( f(d) \) of the ring distance \( d \in [0, K/2] \) to the set of firms directly hit by the shock. This function \( f : [0, K/2] \to [0, 1] \) is taken to be positive and decreasing for all \( K/2 > d > 0 \). This simply reflects the idea that every firm in a component is affected by any shock hitting it but with an intensity that decreases with the distance to the source. In addition, we require the following three conditions:

\[
\begin{align*}
\text{(9)} \\
\quad f(0) &= \alpha \\
\quad f(K/2) &= 0 \\
\quad 2 \int_0^{K/2} f(x; K) dx &= 1 - \alpha, \quad (11)
\end{align*}
\]

The above conditions embody the three key properties satisfied by the exposure function in the discrete setup. First, (9) simply states that the firm affected directly by the shock bears a fraction \( \alpha \) of it. Second, (10) requires that the exposure becomes vanishing small when the distance to the source of the shock is maximal in the component. Finally, (11) captures the requirement that the part of the shock not absorbed by the firm hit by it must be fully absorbed by the other firms in the component.

To simplify further the analysis, we postulate that \( f \) is a two-piece linear function with the kink at the bisectrix (i.e. at a distance \( d = H \) such that \( f(H) = H \)). Thus, in view of (9)–(10), we arrive at a function of the form:

\[
\begin{align*}
\text{(12)} \\
\quad f(d; K) &= \alpha - \frac{H}{H} d \\
&= \frac{HK}{K-2H} - \frac{2H}{K-2H} d \\
&= 0 \\
\text{for } 0 \leq d \leq H \\
\text{for } H < d < K/2 \\
\text{for } d = K/2.
\end{align*}
\]
The value of $H$ (and therefore the function $f(\cdot)$) can then be characterized from the adding-up constraint (11). To see this note that

$$\int_0^{K/2} f(x; K)dx = \frac{(\alpha - H) H}{2} + H^2 + \frac{H (K/2 - H)}{2}$$

so that (11) can be rewritten as

$$2 \left( \frac{(\alpha - H) H}{2} + H^2 + \frac{H (K/2 - H)}{2} \right) = 1 - \alpha$$

hence we get:

$$H = \frac{2(1 - \alpha)}{K + 2\alpha}. \quad (14)$$

It is immediate to verify that, if $\alpha \geq 1/2$, the function $f(\cdot)$ is decreasing for all $K > 1$ (i.e. for all component sizes above the minimum). Thus, we shall assume that $\alpha \geq 1/2$, so that the simple (piece-wise linear) function given in (12) displays the essential properties of the exposure function of the original discrete setup. It can also be checked that this function is concave or convex depending on whether, respectively, $K$ is smaller or larger than $2(1 - \alpha)/\alpha$. An illustration of how it approximates the original function for the discrete setup is displayed in Figure 1.

Thus, to summarize, in the continuum version of the model, the pattern of exposure depends on network structure as follows. For a completely connected component of size $K + 1$, the exposure to a shock that hits other firms in the component is given by the function that is constantly equal to $(1 - \alpha)/K$. Instead, for a ring of the same size, the exposure is given by the function $f(d; K)$ defined in (12), which is decreasing in the distance $d$ to the unit measure of firms hit by the shock.

We are now in a position to determine the extent of default induced by any given shock of magnitude $L$ in the different cases. Recall that this

---

7Note that the function $f(\cdot)$ is decreasing if $\alpha > H$ and $K/2 > H$. Also recall that $K = 1$ induces the minimum component size $K + 1 = 2$ which is required to provide full insurance against $s$ shocks.
shock is assumed to hit a unit mass of adjacent firms in a single component. Whether such directly affected firms default or not is independent, as already argued in the previous section, of the underlying financial network structure. For they will default if, and only if,

\[ L > \frac{R - M}{\alpha}. \]

If this inequality is violated and the firms hit by the shock do not default, no other firm in the corresponding component defaults either. This simply follows from the fact that \( \alpha \geq 1/2, K \geq 1 \) and \( f(d; K) \leq \alpha \) for any \( d \geq 0 \). But if those firms directly affected by the shock do default, what happens to all the others in the component naturally depends on the size \( K + 1 \) of the component and on the pattern of the connections within it.

First, in the case where the interaction pattern is complete, the uniformity of the exposure has the following immediate implication: all firms indirectly affected (i.e. not hit by the shock but lying in the component affected) will experience the same outcome and default if, and only if, \( R - M < \frac{1 - \alpha}{K} L \), or

\[ L > \frac{K}{1 - \alpha} (R - M) \quad (15) \]

whereas none of those firms will default otherwise. Thus, if we let \( g_c(L; K) \) stand for the mass of firms that default when the shock hits some other firm in their component, that magnitude is given by the following step function:

\[
g_c(L; K) = \begin{cases} 
0 & \text{if } L \leq \frac{K}{1 - \alpha} (R - M) \\
K & \text{if } L > \frac{K}{1 - \alpha} (R - M)
\end{cases} \quad (16)
\]

In contrast, when the component is connected through a ring interaction structure (as captured by \( f(\cdot) \)), the conclusion is, in general, not so dichotomic. For, in this case, whether a firm in the component defaults or not depends on its ring distance \( d \) to those firms that have been directly affected. It defaults if, and only if,

\[ L > \frac{1}{f(d; K)} (R - M). \quad (17) \]
which is to be contrasted with \([15]\). Hence the threshold that marks the relevant “default range” is given by the distance \(\hat{d}\) such that

\[
f(\hat{d}; K) = \frac{R - M}{L}
\]

so that a firm defaults if, and only if, its distance \(d\) from the set of firms directly hit by the shock is such that

\[
d < f^{-1}((R - M) / L; K).
\]

Under a ring structure, therefore, the effect of shocks on the number of defaults is not extreme and discontinuous as under complete (direct) interaction. Rather, as the magnitude \(L\) of the shock increases, the mass of firms defaulting among those indirectly affected by it grows gradually, as determined by the function \(g_r(L; K) \equiv 2f^{-1}((R - M) / L; K)\). This function is easily seen to be as follows (see an illustration in Figure 2 for \(K = 20\) and \(\alpha = 1/2\)):

\[
g_r(L; K) = \begin{cases} 
0 & \text{for } L \leq \frac{R - M}{\alpha} \\
\frac{2H}{\alpha - H} - \frac{2H}{\alpha - H} \frac{R - M}{L} & \text{for } \frac{R - M}{\alpha} \leq L \leq \frac{R - M}{H} \\
\frac{K - 2H}{H} R - \frac{M}{L} & \text{for } L \geq \frac{R - M}{H}
\end{cases}
\]  \tag{19}

3 Optimal Financial Structures

In the environment described in the previous section – more specifically, in its continuous version – our objective is to identify the optimal financial structures that minimize the expected mass of firms defaulting. As explained, our analysis will focus on two dimensions of the “design problem”: network segmentation and network density. The first dimension is given by the number \(C\) of different components in place (and, consequently, the sizes \(K_i + 1\) of each of them). The second dimension, on the other hand, involves comparing the two structures under consideration: ring and complete components. Ring components of size \(K\) have the number of defaults
given by the function specified in (19), while complete components have the function \( g_{\nu}(L; K) \) specified in (16). Hence, formally, the aim is to study – both for the ring \((\nu = r)\) and the complete \((\nu = c)\) structures – the following optimization problem:

\[
\min_{K_i, C} \sum_{i=1}^{C} \frac{K_i + 1}{N} \sum_{i=1}^{K_i} g_{\nu}(\tilde{L}_b; K_i) \\
\text{s.t.} \sum_{i=1}^{C} \frac{K_i + 1}{N} = 1
\]  

(20)

The specification of the optimization problem (20) allows for the possibility of asymmetric structures, where the system is divided into components of different size. It will be shown, however, that the solution must be symmetric for all cases under consideration, with \( K_i = K \) for all \( i \). This implies that the problem (20) can be reduced to minimizing \( \mathbb{E} g_{\nu}(\tilde{L}_b; K) \) with respect to \( K \). And, as explained in Section 2.1, this is in turn equivalent to minimizing the individual probability \( \varphi \) that any given firm defaults (cf. Footnote 5).

We organize the analysis in three parts. First, in Subsection 3.1 we identify some clear-cut conditions under which the optimal segmentation is one of the two polar extremes – maximal or minimal – and the optimal degree of connectivity is complete. Then, in Subsection 3.2 we extend the analysis to more general specifications of the probability distribution of the \( b \) shocks, where intermediate levels of segmentation are optimal and a wider range of issues can be studied. Finally, in Subsection 3.3 we identify scenarios where not only intermediate levels of segmentation are optimal but also a low density of connections, as embodied by the ring structure.

### 3.1 Polarized segmentation

In order to get a clear understanding of the forces at work, we shall start by examining the case where the probability distribution of the \( b \) shocks is of the Pareto family with support \([1, \infty)\) and density \( \gamma/L^{\gamma+1}_b \). By modulating the decay parameter \( \gamma \), this formulation already allows the discussion of many questions of interest such as the contrast between fat and thin tails in the shock distribution (i.e. between scenarios where large shocks are relatively frequent or not). Our analysis will be carried in two steps. Firstly, we shall
focus on understanding how $\gamma$ affects the optimal degree of segmentation (as described by $K$) for each of the two structures considered: the ring and the completely connected components. Secondly, we shall compare these two structures.

Let $D_r(\alpha, K, \gamma) = \mathbb{E}_\gamma g_r(\tilde{L}_b; K)$, i.e. the expected mass of firms in a ring of size $K + 1$ who default when a $b$ shock hits some other firms in the component. We have:

$$D_r(\alpha, K, \gamma) = \int_1^\infty g_r(L; K) \frac{\gamma}{L^{\gamma+1}} \, dL$$

$$= \int_{\frac{R-M}{\alpha}}^\infty \left( K - \frac{K - 2H}{\alpha - H} \right) \frac{\gamma}{L^{\gamma+1}} \, dL + \int_{\frac{R-M}{\alpha}}^{\frac{R-M}{\alpha}} \left( \frac{2\alpha H}{\alpha - H} - \frac{2H}{\alpha - H} \right) \frac{\gamma}{L^{\gamma+1}} \, dL$$

$$= \gamma \left[ -K \frac{1}{\gamma y^\gamma} + \frac{K - 2H}{H} \frac{1}{(\gamma + 1) y^{\gamma+1}} \right]_{\frac{R-M}{\alpha}}^{\infty} + 2\gamma \left[ -\frac{\alpha H}{\alpha - H} \frac{1}{\gamma y^\gamma} + \frac{H}{\alpha - H} \frac{1}{(\gamma + 1) y^{\gamma+1}} \right]_{\frac{R-M}{\alpha}}^{\frac{R-M}{\alpha}}.$$ 

To simplify matters, we shall make hereafter the following simplifications. First, since the specific value of $\alpha$ has little interesting bearing on the analysis, the concrete one of $\alpha = 1/2$ will be chosen throughout\textsuperscript{8}. Second, we shall normalize $R - M$ to unity. In general, of course, the value of $M$, as determined by (2), should depend on the underlying network structure, which affects the default probability $\varphi$ faced by any given firm. For the sake of identifying optimal structures, however, we can take $M$ to be constant. This follows from the fact that any structure that qualifies as optimal under the assumption of a constant $M$, will continue being so if this value is adjusted in the correct direction. Assuming, therefore, that $M$ is constant we can then choose to have $R - M = 1$ without loss of generality. And then,

\textsuperscript{8}A convenient consequence of this choice for $\alpha$ is that, just as for the discrete version of the model, the pattern of exposure is exactly the same for the complete and the ring structure if the component size is the smallest (i.e. $K = 1$). The difference between the two structures then grows wider the higher is $K$. 

25
with these simplifications, we can write

\[
D_r(1/2, K, \gamma) = \gamma \left[ \frac{KH^\gamma}{\gamma} - \frac{K - 2H}{H} \frac{H^{\gamma+1}}{(\gamma + 1)} \right] + 2\gamma \left[ \frac{H^\gamma}{1 - 2H} \frac{1}{\gamma} + \frac{2H}{1 - 2H} \frac{H^{\gamma+1}}{(\gamma + 1)} + \frac{H}{1 - 2H} \frac{1}{\gamma^2} - \frac{H}{1 - 2H} \frac{1}{(\gamma + 1)^2} \right].
\]

(21)

Studying how the above expression behaves in terms of \( K \), we shall first find that whenever \( \gamma > 1 \) (i.e. the distribution function \( \tilde{L}_b \) does not have fat tails) the function attains a minimum at the highest value of \( \hat{K} = N - 1 \). (Recall that the minimum admissible value of \( K \) is 1 and its maximal value is \( N - 1 \).) This readily tells us that the optimal segmentation structure is to have a single component with \( N \) firms. On the other hand, when \( \gamma < 1 \) (i.e. the distribution has fat tails) the function will be seen to attain a minimum at \( \hat{K} = 1 \). Assuming that \( N \) is even (which will be assumed throughout, for simplicity), the latter conclusion implies that the optimal component size is the minimal one of 2, which in turn implies a maximum degree of segmentation. These conclusions are stated in the following result, for which a detailed proof can be found in the Appendix.

\textbf{Proposition 1} When the shock \( \tilde{L}_b \) has a Pareto distribution, the optimal segmentation pattern for the ring structure is always symmetric and the degree of segmentation is minimal (one single component) if \( \gamma > 1 \), and maximal (i.e. \( N/2 \) components) if \( \gamma < 1 \).

The previous result shows that there is indeed a trade-off between risk sharing and contagion. On the one hand, when the distribution of the shocks exhibits fat tails (hence large shocks are relatively likely), the predominant consideration is to control contagion rather than achieving risk sharing. This leads to minimizing the expected number of defaults by breaking the network

\[9\text{If } N \text{ were not an even number but large, it can be shown that the optimal segmentation structure would still be almost symmetric consisting of } \frac{N}{2} - 1 \text{ components of size } 2 \text{ and a residual component of a larger size. This is a direct implication of the fact that the function } D_r \text{ is concave in } K \text{ for } \gamma < 1. \text{ For completeness, this result is stated and proved as Lemma } 3 \text{ in the Appendix.}\]
into disjoint components of minimal size, which limits the extent to which a shock may spread its consequences far into the system. Instead, when the distribution has no fat tails, the most important consideration becomes risk-sharing, which is maximized by placing all firms in a single component.

Next, we turn to studying the analogous question for the case where the components are completely connected. In this case, the expected mass of defaults in a completely connected component of size $K + 1$ when this component is hit by a $b$ shock is:

$$D_c(1/2, K, \gamma) = \mathbb{E}_\gamma g_c(\bar{L}_b; K) = K \Pr (L \geq 2K) = K \left( \frac{1}{2K} \right)^\gamma.$$  \hfill (22)

Hence

$$\frac{\partial D_c}{\partial K} = - (\gamma - 1) \left( \frac{1}{2K} \right)^\gamma \geq 0 \iff \gamma \geq 1,$$

which again implies that segmentation is maximal (component size is minimum) when $\gamma < 1$ (the shock distribution has fat tails), while is is minimal in the case where $\gamma > 1$.\footnote{As a minor side-point, also note that $D_c(\cdot)$ is a convex function of $K$ when $\gamma < 1$. This implies that if $N$ were not even and the system could not be divided into $N/2$ components of minimum size 2, the optimal arrangement would still be symmetric consisting of $N/2$ components of uniform size.}

Hence, concerning optimal segmentation, we arrive at the same conclusion as for the ring which is formally stated in the following result.

**Proposition 2** When the shock $\bar{L}_b$ has a Pareto distribution, the optimal segmentation pattern for the completely connected structure is always symmetric and the degree of segmentation is minimal (one single component) if $\gamma > 1$, and maximal (i.e. $N/2$ components) if $\gamma < 1$.

Finally, we want to compare the two network structures (ring and complete components) in order to identify which of the two is optimal when not only segmentation but also network density can be decided. In view of Propositions 1 and 2, it is enough to compare the expected number of defaults under either maximal or minimum segmentation for each of the
structures when, respectively, $\gamma$ is lower or higher than unity. This leads to the following result.

**Proposition 3** If the shock $\tilde{L}_b$ has a Pareto distribution and $N$ is large enough, the completely connected structure weakly dominates the ring structure for all values of $\gamma$. More specifically, both are equivalent when $\gamma < 1$ and the complete network is uniquely optimal when $\gamma > 1$.

Combining the results obtained in this subsection, we conclude that if shocks are Pareto distributed, the optimal network must display extreme density and polarized (maximum or minimum) segmentation. In this case, therefore, all adjustment to the underlying risk conditions (i.e. fat or thin tails) is implemented through varying segmentation. As our subsequent analysis will show, however, neither the polarized segmentation nor the complete component connectivity are features maintained if shocks are distributed in different manner. In particular, intermediate levels of segmentation and sparse connectivity can be optimal under certain circumstances.

### 3.2 Intermediate Degrees of Segmentation

Propositions 1 and 2 establish that, when the distribution of the shocks has a simple Pareto structure and thus it either has, or does not have, fat tails, the optimal degree of segmentation is always extreme, i.e. maximal or minimal. We show next that this is no longer true when the distribution of the shocks is more complex, as for instance when it is given by the mixture of two Pareto distributions.

**Proposition 4** Suppose that the shock $\tilde{L}_b$ is distributed as a mixture of a Pareto distribution with parameter $\gamma > 1$ and another Pareto distribution with parameter $\gamma' < 1$, with respective weights $p$ and $1 - p$. Then, if $N > 4$, there exist $p_0$, $p_1$ (0 < $p_0$ < $p_1$ < 1) such that, if $p \in (p_0, p_1)$, the optimal segmentation pattern for the completely connected structure is symmetric with an intermediate component size $K^* + 1$ such that $1 < K^* < N - 1$. 

28
The previous result establishes that an intermediate level of segmentation is optimal among completely connected structures when the shock distribution involves a mixture of Pareto distributions displaying with both fat and thin tails. A similar conclusion arises when the components display a ring structure, although an analytic result in this case is hard to obtain. We illustrate matters, therefore, through the following example.

**Example 1** Set $\gamma = 2$, $\gamma' = 0.5$ and $p = 0.95$. For these values we find that the value of $K$ which minimizes $D_c(1/2, K, 2, 0.5, 0.95)$ is $\hat{K}_c = 5.65$, and at this value the expected mass of defaults (when a $b$ shock hits some other firms in the component) is 0.13. The value of $K$ which minimizes the corresponding expression for the ring structure, $D_r(1/2, K, 2, 0.5, 0.95)$, is higher, $\hat{K}_r = 8.02$ and the expected mass of defaults in this case is also higher, equal to 0.145. The fact that $\hat{K}_r > \hat{K}_c$ can be heuristically understood as a reflection of the fact that, when arranged optimally, components with a ring structure compensate for their lower density of connections with a larger size. See Figure 1 for a graphical description of the situation.

To find the optimal financial structure for the whole system we also need to specify the value of $N$. Let $N = 10$. In this case it is clear that a symmetric structure with equal components of size $\hat{K}_c + 1$ or $\hat{K}_r + 1$ is not feasible in either case. We find, however, that both for the complete and ring structures the optimal structure is given by two equal-sized components of size $K + 1 = 5$. Moreover, in line with the indicated conclusions at $\hat{K}_r$ and $\hat{K}_c$, we obtain that the optimal complete structure still dominates the optimal ring structure since the former yields an expected number of defaults equal to 0.26 in contrast with 0.32 induced by the latter. This conclusion actually proves to be robust with respect to other possible specifications of the parameter values of the environment. A graphical account of the situation is provided in Figure 2, where we focus on the relevant case where the total population $s$ divided in at most two components of varying size.
Figure 1: The expected number of defaults in a given component as a function of its size $K$ for both complete components ($D_c(\cdot)$) and ring components ($D_r(\cdot)$). The shock $b$ is distributed according to a mixture of two Pareto distributions with exponents $\gamma = 2$ and $\gamma = 0.5$ and respective weights $p = 0.95$ and $1 - p = 0.05$. The minimum of the functions $D_c(\cdot)$ and $D_r(\cdot)$ is denoted by $\hat{K}_c(=5.65)$ and $\hat{K}_r(=8.02)$, respectively.

3.3 Sparse Connections

Let us consider now the case where the probability distribution of the $b$ shocks is not smooth because it has some atoms. More precisely, let $\Phi(L_b)$ be the mixture of a Pareto distribution with $\gamma > 1$ and a Dirac distribution putting all probability mass on a shock of magnitude $\bar{L} > 2(N - 1)$. On the one hand, the Pareto distribution considered has no fat tails, which as we saw in Section 3.1 not only favors minimal segmentation ($K = N - 1$) but also singles out the complete structure over the ring. But, on the other hand, the shock $\bar{L}$ induced through the Dirac distribution is such that, if firms are arranged in a single component, then all firms default when they are completely connected but some do survive if arranged in a ring. Because of these conflicting considerations, we show below that there is an
Figure 2: The expected total number of defaults in a system consisting of $N = 10$ firms that is divided in two components of size $K$ and $N - K$ for both complete components ($W_c(\cdot)$) and ring components ($W_r(\cdot)$). The shock $b$ is distributed according to a mixture of two Pareto distributions with exponents $\gamma = 2$ and $\gamma = 0.5$ and respective weights $p = 0.95$ and $1 - p = 0.05$. The socially optimal value that minimizes the expected number of defaults is $K^* = 5$ for both the complete and ring structures.

open region of parameter values for which the second effect prevails and hence the optimal financial structure is a ring – that is, sparser connections are optimal.

**Proposition 5** Assume that, with probability $p$, the shock follows a Pareto distribution with parameter $2 > \gamma > 1$ and with probability $1 - p$ it equals $\bar{L} = 2(N - 1) + 1$. Then, for all values of $N$ such that

\[
N > 1 + \left( \frac{1}{2^{\gamma-1}} + \frac{1}{\gamma - 1} \left( \frac{1}{2^{\gamma+1}} + \frac{1}{\gamma - \frac{1}{\gamma + 1}} \right) \right)^{\frac{1}{2 - \gamma}}
\]

and $p$ such that

\[
\frac{1 - p}{p} < (\gamma - 1) \left( \frac{1}{2(N - 1)} \right)^{\gamma}.
\]

31
the optimal financial structure is a single ring component.

Condition (24) says that the weight \( p \) on the Pareto distribution is sufficiently high so that the optimal segmentation structure is determined by it and a single component ensues, both for the ring and the complete structures. But, given that there is also a significant probability that a large shock arrives that cannot be completely shared, some attempt at “controlling the induced damage” may be in order. And this is indeed what the ring achieves – a suitable compromise between the extent of risk sharing allowed by extensive indirect connectivity (i.e. minimal segmentation) and the limits to wide risk contagion resulting from sparse direct connections.

4 Stability and optimality

We now examine the relationship between the optimal pattern of linkages derived in the previous section and the individual incentives to form those linkages. We explore, in other words, whether social welfare is aligned with individual payoffs. To model the strategic considerations involved in the creation and destruction of links, the network-formation game is assumed to be conducted as follows:

- Agents independently submit their proposals concerning the set of agents in the whole population \( P \) each one of them wants to connect to. Formally, a strategy of each agent \( i \) is a subset \( L_i \subset P \).

- Links are formed (only) between the agents who reciprocally list each other at the proposal-submission stage. Formally, given profile of strategies \( L \equiv (L_i)_{i\in P} \) a link between any two agents \( j \) and \( k \) is established iff \( j \in L_k \) and \( k \in L_j \).

Given the above network-formation rules, a specific network \( \Gamma(L) \) is induced by each strategy profile \( L \). Then, in line with our motivation of the model, we posit that the resulting payoff of each agent \( i \) is simply decreasing with \( \varphi_i(\Gamma(L)) \), its own default probability resulting from the network induced by \( L \).
In such a network-formation game, an undesirable feature of the standard concept of Nash Equilibrium is that it leads to a vast multiplicity of equilibrium networks, a consequence of the fact that the formation of any link induces a coordination problem between the two agents involved. (As an extreme illustration, note that the empty network can always be supported by a Nash equilibrium where every agent proposes nobody to link with.)

To address this issue, it is common in the literature to consider a strengthening of the Nash equilibrium notion that reduces miscoordination by allowing sets of agents to deviate jointly (see e.g. Goyal and Vega Redondo (2007) or Calvó-Armengol and Iklíç (2009)). In the continuum framework considered here, we shall capture this idea by means of the concept we shall label **Coalition-Proof Equilibrium** (CPE), where any group (i.e. coalition) of agents can coordinate their deviations.

**Definition 1** A strategy profile \( \mathbf{L} \equiv (L_i)_{i \in P} \) of the network-formation game defines a Coalition-Proof equilibrium (CPE) if there is no subset of agents \( W \) of positive measure and a substrategy profile \( (L'_j)_{j \in W} \) for all of them such that\(^{11}\)

\[
\forall i \in W, \quad \varphi_i \left[ \Gamma \left( (L'_j)_{j \in W}, (L_k)_{k \in P \setminus W} \right) \right] < \varphi_i \left[ \Gamma (\mathbf{L}) \right].
\]  

The CPE notion precludes unilateral profitable deviations by any (positive-measure) set of agents, so it is a obviously a refinement of the standard notion of Nash Equilibrium. It is a very strong refinement in that it embodies no limit on the amount/measure of agents that may coordinate their deviations. But, in general, it would be natural to impose some such limit since large-scale deviations may be very difficult to implement. In our case, however, introducing some such limit would have no significant implications on the analysis, so we choose to avoid it for expositional simplicity\(^{12}\).

\(^{11}\)In line with what was postulated in Section 2, we shall continue to assume that, after any change in connections has been implemented, the agents involved in the change continue to distribute the fraction \( 1 - \alpha \) of their own assets among all their neighbors (old and new) in a uniform manner.

\(^{12}\)More specifically, associated to any \( \eta > 0 \), one could define the notion of \( \eta \)-CPE, such that only joint deviations by at a set of agents of measure no higher than \( \eta > 0 \) are to be
For the sake of focus, we analyze CPE for the class of distributions considered in Section 3.2 and show that, under certain generic conditions, there is a conflict between group optimality and individual incentives. As it will be recalled, in that setup Proposition 4 established that, for an open set of the parameter space, the optimal configuration involves a symmetric segmentation of the whole population in several (equal-sized) components that are completely connected. Building upon that result, Proposition 6 below shows that, under analogous conditions, one can be sure that the uniform socially optimal component size \( K^* \) is lower than the size \( \hat{K} \) that the individuals of any given component would prefer. This brings about the conflict between individual and social incentives that gives rise to the inefficiency of CPE-configurations. In a CPE displaying complete components, individual incentives (supported by coalitional deviations) lead to at least one component of individually optimal size \( \hat{K} \) and at least one other component of a size smaller than \( K^* \). This is not socially optimal because, in essence, convexity considerations favor a more balanced configuration where the sizes of those two components be closer.

**Proposition 6** Suppose that the shock \( \tilde{L}_b \) is distributed as a mixture of a Pareto distribution with parameter \( \gamma > 1 \) and another Pareto distribution with parameter \( \gamma' < 1 \), with respective weights \( p \) and \( 1 - p \). Then, if \( N > 4 \), there exist \( p_0, p_1 \) \((0 < p_0 < p_1 < 1)\) such that, if \( p \in (p_0, p_1) \), the socially optimal configuration cannot be supported at a CPE of the network formation game. Among the completely connected structures, the only CPE configuration is asymmetric with all but one component displaying the size \( \hat{K} \) that minimizes \( D_c(1/2, K, \gamma, \gamma', p_0) \) and one component of a lower size.

(see Remark 2 in the Appendix) that Proposition 6 applies to such notion of equilibrium for any value of \( \eta > 0 \).

Another requirement typically demanded in Game Theory from coalition-based notions of equilibrium is that the coalitional deviations considered should be robust, in the sense of being themselves immune to a subcoalition profitably deviating from it. Again, this requirement has no bite for our analysis since, as we also explain in Remark 2, all the profitable deviations that need to be allowed are robust in the aforementioned sense.
The source of the conflict between social and individual optimality lies in the fact that, in having each separate component strive for the size that is individually (or component-wise) optimal, it may impose an externality on outside individuals who are consequently forced to remain in a component that is too small to share risk efficiently. The network-formation game allows for this situation to be supported by a CPE, and hence the inefficiency that results under certain circumstances. As a simple illustration, refer back to Example 1 where the efficient configuration involved two completely connected components of common size equal to $K^* = 5$. We found, however, that the optimal size for each separate component is $\hat{K} = 5.65$, the value that minimizes $D_c(1/2, K, 2, 0.5, 0.95)$. This entails an asymmetric CPE configuration with two components, one displaying a size of 5.65 and the other a smaller “residual size” of 4.35.

The existence of a conflict between efficiency and strategic stability is of course hardly novel nor surprising in the field of social networks (see e.g. Jackson and Wolinsky (1996) for an early instance of it). For, typically, the creation or destruction of any link between two agents must impose externalities on others that not internalized by the two agents involved in the linking decision. In the context of risk-sharing, this tension has been studied in a recent paper by Bramoullé and Kranton (2007) – hereafter labelled BK – and it is interesting to understand the differences with our approach. We close, therefore, this section with a brief comparison of the two models.

BK consider a theoretical framework consisting of a finite number agents affected by i.i.d. income shocks. Linkages generate risk sharing in a way similar to that of our model, except in two important respects: (a) risk-sharing is complete (i.e. uniform) across all members in a component; (b) there is no risk of contagion, so the size of optimal components is just limited by the fact that links are assumed to be costly. Focusing on the notion of strategic stability (which is weaker than ours),\footnote{Strategic stability allows only for coalitions of at most two players and rules out as well the simultaneous creation and destruction of links. See Jackson and Wolinsky (1996) for details.} BK are also interested in
comparing efficient and equilibrium configurations. They find that whenever equilibrium structures exist (not always), there are at most two asymmetric components, with sizes smaller than the optimal one. The contrast between BK’s conclusions and ours derives from the nature of the externality in each case. In BK, given that the cost of any new link is borne alone by the two agents involved, at equilibrium there is underinvestment in link formation (which is a “public good”). Instead, in our case there are no linking costs and the nature of the uninternalized externality has to do with the fact that having some agents abandon an optimally sized component in order to reach a socially beneficial homogenization of component sizes is not in the interest of those who would be involved in the move. Thus, in the end, it is the need to meet an overall feasibility constraint induced by overall population size that typically generates inefficiencies.

5 Asymmetric structures

So far we have concentrated the discussion on groups of firms where all of them are exactly identical. Although this allows us to obtain analytical results and gather intuition, the real world contains firms of very different sizes, so it is useful to see how our framework extends when firms are different in size.

For simplicity, consider a situation with two types of firms. One type involves firms that are relatively “small”, their size being normalized to unity. Thus, in effect, they are exactly as we have been considering so far. The other type involves firms of a larger size $\beta > 1$. Such a larger size has two implications. First, their returns when no shock affects them (directly or indirectly) are scaled-up compared to those of smaller size, i.e. they are multiplied by $\beta$ and, naturally, the same factor applies to the value of its assets. Second, they face a probability of being directly hit by a shock that is also $\beta$ times larger, even though we assume that the size distribution of this shock is equal to that of small firms. although the the probability of a shock. In a sense, one can view a large firm as analogous to a completely connected component of smaller firms of the sort we have been considering.
so far. The only key difference is that, since they all belong to the same entity, there is no longer the requirement that each “unit” of the composite group should remain responsible of a proportion $\alpha$ of its own assets and corresponding liabilities.

Having different kinds of firms allows us to study new types of structures in which large and small firms play asymmetric roles. Our analysis will focus, specifically, on comparing two cases: (a) complete (symmetric) structures involving firms of identical size, and (b) “star” structures consisting of large firms that act as central hubs. We shall illustrate the main ideas in a simple scenario where there are only two big firms and $2\beta$ small ones. In this context, the complete structures involve two symmetric components: one has the two big firms and another with $2\beta$ firms that are completely connected. In contrast, the asymmetric situation has two stars, each consisting of $\beta$ small firms that are solely connected to a common big firm.

For the purposes of our present discussion, we do not gain much by representing the situation through a continuum formulation, so we return to the basic set-up where each firm is conceived as a discrete entity (small or large, of sizes one and $\beta$). Our first step is to specify the pattern of exchange and induced risk exposure prevailing in each case. As explained in Subsection 2.2, this exposure can be represented through a corresponding matrix $B_K$ for a component of size $K + 1$.

When firms are arranged in complete structures, the matrices $B_K$ are of the form specified in (5), its entries given by (7)-(8) if we impose the condition that each firm keeps a fraction $\alpha$ of its own assets. In our present case, it particularizes to two component sizes, $K = 1$ and $K = 2\beta - 1$, for the complete components consisting of big and small firms, respectively.

On the other hand, when firms are arranged in star structures consisting of $\beta + 1$ firms, the corresponding exposure matrix is given by:

$$
\tilde{B} = \begin{pmatrix}
\theta & (1 - \theta)/\beta & (1 - \theta)/\beta & \cdots & (1 - \theta)/\beta \\
(1 - \theta) & \theta & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
(1 - \theta) & 0 & 0 & \cdots & \theta
\end{pmatrix}
$$
where, for notational simplicity, we dispense with any reference to the component size \((K = \beta + 1)\).

Note that the entries of the matrix \(\tilde{B}\) reflect the fact that a big firm (indexed by \(i = 1\)) must offer only a share \((1 - \theta)/\beta\) of its assets for a larger share \((1 - \theta)\) in the assets of small firms (those indexed by \(i = 2, 3, \ldots, \beta + 1\)). But if we maintain the assumption that firms in a component undergo exactly as many rounds of securitization as needed in order to have some amount of risk sharing between any two firms in the component, the matrix \(\tilde{B}\) will not account for all asset exchange. Instead, there must be one additional round of it, the resulting matrix \(A\) of effective exposure given by a two-fold composition of \(\tilde{B}\), i.e. \(\tilde{A} = \tilde{B}^2\), or

\[
\tilde{A} = \begin{pmatrix}
\theta^2 + (1 - \theta)^2 & 2\theta(1 - \theta)/\beta & (1 - \theta)/\beta & \cdots & (1 - \theta)/\beta \\
2\theta(1 - \theta) & \theta^2 + (1 - \theta)^2/\beta & (1 - \theta)^2/\beta & \cdots & (1 - \theta)^2/\beta \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
2\theta(1 - \theta) & (1 - \theta)^2/\beta & (1 - \theta)^2/\beta & \cdots & \theta^2 + (1 - \theta)^2/\beta
\end{pmatrix}.
\]

Finally, if we continue to suppose that firms attain the maximum degree of diversification that is consistent with all of them keeping at least an \(\alpha\) share in its own project, the value of \(\theta\) can be pinned down from the following condition:

\[
\alpha = \theta^2 + (1 - \theta)^2/\beta \tag{26}
\]

The above expression reflects the fact that it is small firms that will be able to attain the largest degree of risk “externalization” \((1 - \alpha)\), since the value of their assets is lowest. Instead, the larger firms, whose assets are worth \(\beta\) more than those of smaller ones, cannot possibly attain the same level under asset exchanges with smaller firms that trade equal value. This implies that, in a star network, the large firms are forced to keep a larger share \(\alpha'\) on their

---

\(^{14}\)In this section, we maintain a general value of \(\alpha\) (rather than the specific value of 1/2 hitherto considered) in order to highlight that, depending on their size and network position, agents may face different constraints on the extent to which they can “externalize risk.”
own project, where

\[ \alpha' = \theta^2 + (1 - \theta)^2 > \alpha \]  \hspace{1cm} (27)

Noting that \( \alpha' - \alpha = (1 - \theta)^2 (\beta - 1) / \beta \), this means we can rewrite the exposure matrix as follows:

\[
\begin{pmatrix}
\alpha' & (1 - \alpha') / \beta & (1 - \alpha') / \beta & \cdots & (1 - \alpha') / \beta \\
1 - \alpha' & \alpha & (\alpha' - \alpha) / (\beta - 1) & \cdots & (\alpha' - \alpha) / (\beta - 1) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 - \alpha' & (\alpha' - \alpha) / (\beta - 1) & (\alpha' - \alpha) / (\beta - 1) & \cdots & \alpha
\end{pmatrix}
\]

Based on the above exposure matrix, the criterion for default is just the natural extension of the one that we have been postulating throughout. That is, a firm \( i \) defaults when a shock hits a firm \( j \) in the same component (possibly \( i = j \)) and \( \tilde{a}_{ij}L \) exceeds their gross returns (where \( \tilde{a}_{ij} \) denotes the \( ij \) entry of \( \tilde{A} \)). This, of course, amounts to a different condition if firm is big or small: if \( i \) is small, default occurs when \( \tilde{a}_{ij}L > 1 \), while if it is large the condition is \( \tilde{a}_{ij}L > \beta \).

As advanced, our aim in this section is to compare two types of arrangements:

(i) a star structure composed of two stars, each of them consisting of a \( \beta \) firm acting as a hub and \( \beta \) spokes of unit size.

(ii) a symmetric structure composed of two complete components: one consisting of two firms of size \( \beta \), and the other of \( 2\beta \) firms of unit size.

To get a sharp understanding of the situation, we explore how different values \( L \) for the magnitude of the big shock (which hits a randomly selected firm) affect the expected number of defaults. The contrasting implications induced for each structure are detailed in the following result:

**Proposition 7** There are two different scenarios, depending on the value of \( \theta \):
1) If $\theta > 1/3$, the star structure has a higher expected number of defaults than the symmetric structures whenever the magnitude $L$ of the shock hitting a randomly selected firm satisfies:

$$\frac{\beta}{\alpha'} < L \leq \frac{\beta}{\alpha'} \frac{\beta - 1}{1 - \alpha'} < L \leq \frac{2\beta - 1}{1 - \alpha}$$

while the star structure has a lower number expected number of defaults than the symmetric structures whenever

$$\frac{\beta}{1 - \alpha} < L \leq \frac{\beta}{1 - \alpha'} \frac{2\beta - 1}{1 - \alpha} < L \leq \frac{\beta - 1}{\alpha' - \alpha}.$$

2) If $\theta < 1/3$, the star structure has a higher expected number of defaults than the symmetric structures whenever

$$\frac{\beta}{\alpha'} < L \leq \frac{\beta}{\alpha'} \frac{\beta - 1}{1 - \alpha'} < L \leq \frac{2\beta - 1}{1 - \alpha}$$

while the star structure has a lower expected number of defaults than the symmetric structures whenever

$$\frac{\beta}{1 - \alpha} < L \leq \frac{\beta - 1}{\alpha' - \alpha} \frac{2\beta - 1}{1 - \alpha} < L \leq \frac{\beta}{1 - \alpha'}.$$

The conclusions of Proposition 7 are illustrated in Figure 3 for the case where $\theta > 1/3$. We focus our ensuing discussion on this case, since the alternative one of $\theta < 1/3$ is qualitatively analogous.

First, the relative dominance of the symmetric structure in the range

$$\frac{\beta}{\alpha'} \leq L \leq \frac{\beta}{\alpha}$$

is easy to explain. The star structure forces the larger firm in the hub to hold more of its own assets than in the symmetric one ($\alpha'$ versus $\alpha$), in order to match the lower value of the assets owned by the smaller firms located the spokes. This, in turn, decreases the extent to which it can share risk, a limitation that is detrimental against relatively small shocks. But, on the other hand, it is precisely this limitation that protects the hub against large shocks, which is why the start structure dominates in the following range:

$$\frac{\beta}{1 - \alpha} \leq L \leq \frac{\beta}{1 - \alpha'}.$$
\[ \theta > 1/3 \]

The parameter ranges for the magnitude \( L \) of the shock where either the complete or the star structures dominate (i.e. yield a lower expected number of defaults) if \( \theta > 1/3 \). The non-shaded areas correspond to parameter ranges where both structures yield the same results.

Finally, there are two additional ranges involving relatively high values of \( L \), where the considerations at work are somewhat more subtle. These ranges are:

\[
\frac{\beta}{1 - \alpha'} \leq L \leq \frac{2\beta - 1}{1 - \alpha}, \quad \frac{2\beta - 1}{1 - \alpha} \leq L \leq \frac{\beta - 1}{\alpha' - \alpha}
\]

When the shock lies in the first range, the complete structure dominates because complete homogeneous components do provide enough risk sharing to avert default by most of the firms if the shock hits the component where there are more of them – i.e. the component consisting of only small firms. In the second range, however, the shock is so large that all firms in an homogeneous component default if affected by the shock, directly or indirectly. The situation, however, is somewhat better in a star component, but only if the shock hits one of the spoke firms. In this case, the small firm thus hit defaults, as well as the large hub. But the latter firm acts a sort of buffer, preventing a sizable fraction of the shock to spread and cause further defaults in the component (in contrast with the situation just described for an homogeneous component).
6 Conclusion

We have proposed a stylized model to study the problem that arises when firms need to share resources to weather shocks that can threaten their survival, but then are exposed to the risk coming from those same connections that help them in the time of need. Depending on the characteristics of the shock distribution, a wide variety of different possibilities can be optimal. For example, maximal segmentation in small groups is optimal if big shocks are likely, while very large groups are optimal when most shocks are of moderate magnitude. There are also conditions, however, when an intermediate group size is optimal or when groups should be large but display some internal “detachment” (i.e. sparse connectivity).

The former consideration pertain to social optimality, i.e. to the minimization of defaults. We have explored whether such overall objective is aligned with individual optimality. And we have seen that, in general, there is a conflict between strategic incentives and social welfare. This tension arises from the fact that when a component attains the size that minimizes the default probability of their members, it will block admitting new members from a smaller components, ignoring the negative externality of their behavior. Finally, we have also studied asymmetric structures and found that certain asymmetries (e.g. a “central” agent acting as a hub) can have useful properties as a firebreak in certain cases.

There are many issues that this paper did not study in depth. Although we have identified conditions under which sparse internal connections and asymmetries are beneficial, we do not have a theorem providing general conditions under which alternative topologies are optimal. Since our work has highlighted a disparity between efficiency and equilibrium outcomes, it would be important to extend our normative analysis in this manner. Only then would it be possible to understand better what are the options and consequences of alternative policy measures impinging on the incentives of agents (say, banks) to connect for the purpose of sharing risk. Another important extension would be to integrate such risk-management decisions with other considerations (e.g. cooperation, exploitation of synergies) that
also underlie economic connections in the real world.

Appendix

Proof of Proposition 1. Introducing in (21) the value of $H$ obtained in (14) we get:

$$D_r(1/2, K, \gamma) = \left( K \left( \frac{1}{\gamma + 1} \right) - \frac{2}{K - 1} \frac{1}{\gamma + 1} \right) \left( \frac{1}{K + 1} \right)^\gamma + \frac{1}{K - 1} \frac{1}{\gamma + 1} \left( \frac{1}{2} \right)^{\gamma - 1},$$

and hence

$$\frac{\partial D_r}{\partial K}(1/2, K, \gamma) = -\frac{1}{(K - 1)^2} \frac{1}{(\gamma + 1)} \left( \frac{1}{2} \right)^{\gamma - 1}$$

$$+ \left( K \left( \frac{-\gamma}{\gamma + 1} \right) \frac{1}{K + 1} + \frac{2}{K - 1} \frac{1}{\gamma + 1} \right) \left( \frac{1}{K + 1} \right)^\gamma \tag{28}$$

Now note that the inequality $\partial D_r(1/2, K, \gamma)/\partial K > 0$ is equivalent to:

$$\frac{2(K - 1)}{K + 1} \gamma + (K - 1)^2 + 2 > (K + 1) \left( \frac{K + 1}{2} \right)^{\gamma - 1} + \gamma (K - 1)^2 \frac{1}{K + 1}$$

or

$$(K - 1)^2 \left( 1 - \frac{\gamma K}{K + 1} \right) + 2 \left( 1 + \gamma \frac{K - 1}{K + 1} \right) > (K + 1)^\gamma \frac{1}{2^{\gamma - 1}}$$

which can be rewritten as

$$(K - 1)^2 + 2 + \gamma (K - 1) (2 - K) > (K + 1)^\gamma \frac{1}{2^{\gamma - 1}}. \tag{30}$$

So, using the identities

$$(K - 1)^2 + 2 + (K - 1) (2 - K) = K + 1$$

and

$$(K + 1)^\gamma \frac{1}{2^{\gamma - 1}} = 2 \left( \frac{K + 1}{2} \right)^\gamma$$

we can equivalently write (30) as follows:

$$\frac{K + 1}{2} - \left( \frac{K + 1}{2} \right)^\gamma < \frac{1}{2} (1 - \gamma) (K - 1) (2 - K) > 0.$$
Denote by $\Xi(K, \gamma)$ the left hand side of the previous inequality, conceived as a function of $K$ and $\gamma$. Then, to complete the proof, we establish the following claim:

$$\forall K > 1, \quad \Xi(K, \gamma) \geq 0 \iff \gamma \leq 1.$$ (31)

To show the claim, first note that $\Xi(K, 1) = 0$ for all $K$. Hence we conclude that $\frac{\partial D}{\partial K}(1/2, K, \gamma) = 0$ for $\gamma = 1$ and all $K$. On the other hand, examining the partial derivative of $\Xi$ w.r.t. $\gamma$, we find

$$\frac{\partial \Xi}{\partial \gamma}(K, \gamma) = -\left(\frac{K + 1}{2}\right)^\gamma \ln \frac{K + 1}{2} + \frac{1}{2} (K - 1) (2 - K)$$

$$\leq -\ln \frac{K + 1}{2} + \frac{1}{2} (K - 1) (2 - K),$$

the inequality being strict for all $K > 1$.

Having established (31) we conclude, as desired, that the minimum of $D_r(1/2, K, \gamma)$ is attained at the maximum value of $K$ (i.e. $N - 1$) when $\gamma > 1$, while it is attained at the lowest value of $K$ (i.e. $K = 1$, since $N$ is even) when $\gamma < 1$. The proof is thus complete. ■

For simplicity, Proposition 1 presumes that $N$ is even, which renders feasible an exact partition of the system in components of size 2. As indicated, however (see Footnote 9), the gist of the result extends naturally to cases where $N$ is any large real number, in which case the maximum segmentation involves $N/2$ components of size 2 and a residual. This conclusion readily follows from the following lemma.

**Lemma 1** If $\gamma < 1$, the function $D_r(1/2, K, \gamma)$ is strictly concave in $K$ for all $K > 1$.

**Proof.** From the expression for $D_r(1/2, K, \gamma)$ given in (28), denote by $A(K)$ and $B(K)$ its first and second term respectively. That is,

$$A(K) = \left(K \left(\frac{1}{\gamma + 1}\right) - \frac{2}{K - 1} \left(\frac{1}{2}\right)^{\gamma-1} \left(\frac{1}{\gamma + 1}\right) \left(\frac{1}{K + 1}\right)^{\gamma}\right)$$

$$B(K) = \frac{1}{K - 1} \left(\frac{1}{\gamma + 1}\right).$$
hence we have:

\[
\frac{dA(K)}{dK} = \left( \frac{-K^2 + K + 2}{K - 1} \right) \frac{1}{K + 1} \cdot \frac{\gamma + 1}{\gamma + 1} + \left( \frac{1}{\gamma + 1} + \frac{2}{(K - 1)^2} \right) \frac{1}{\gamma + 1} \left( \frac{1}{K + 1} \right) ^\gamma \\
= \frac{1}{\gamma + 1} \left( \left( \frac{-K^2 + K + 2}{K - 1} \right) \frac{\gamma}{\gamma + 1} + 1 + \frac{2}{(K - 1)^2} \right) \left( \frac{1}{K + 1} \right) ^\gamma \\
= \frac{1}{\gamma + 1} \left( \frac{K^2 - 2K + 3}{K - 1} - \frac{\gamma}{K + 1} \left( K^2 - K - 2 \right) \right) \left( \frac{1}{K - 1} \right) \left( \frac{1}{K + 1} \right) ^\gamma \\
= \frac{1}{\gamma + 1} \left( \frac{K^2 - 2K + 3}{(K - 1)^2} \right) \left( \frac{1}{K + 1} \right) ^\gamma - \frac{\gamma}{K + 1} \left( K^2 - K - 2 \right) \left( \frac{1}{K - 1} \right) \left( \frac{1}{K + 1} \right) ^\gamma \\

\frac{dB(K)}{dK} = -\frac{1}{(K - 1)^2} \left( \frac{1}{\gamma + 1} \right) ^\gamma
\]

So that taking another derivative for each of the previous expressions we obtain:

\[
\frac{d^2 A(K)}{dK^2} = -\frac{1}{\gamma + 1} \left( \frac{1}{K + 1} \right) ^\gamma \times \\
\left( -K^3 \gamma^2 + K^3 \gamma + 4K^2 \gamma^2 - 2K^2 \gamma - 5K \gamma^2 + 5K \gamma + 4K + 2 \gamma^2 - 4 \gamma + 4 \right) \\
= -\frac{1}{\gamma + 1} \left( \frac{1}{K + 1} \right) ^\gamma \times \\
\left( K^3 \left( -\gamma^2 + \gamma \right) + 2K^2 \left( 2 \gamma^2 - \gamma \right) + 5K \left( -\gamma^2 + \gamma \right) + 4K + 2 \left( \gamma - 1 \right)^2 + 2 \right) \\
\frac{d^2 B(K)}{dK^2} = 2 \left( \frac{1}{\gamma + 1} \right) ^\gamma \left( \frac{1}{K - 1} \right) ^3
\]
and therefore:

\[
\frac{\partial^2 D_r(1/2, K, \gamma)}{\partial K^2} = \frac{d^2 A(K)}{dK^2} + \frac{d^2 B(K)}{dK^2} = \\
= 2 \left( \frac{1}{2} \right)^{\gamma-1} \frac{1}{(\gamma + 1)(K - 1)^3} \left( K^3 (\gamma^2 + \gamma) + 2K^2 (2\gamma^2 - \gamma) + 5K (-\gamma^2 + \gamma) + 4K + 2 (\gamma - 1)^2 + 2 \right) \\
= \frac{1}{2\gamma (\gamma + 1)(K - 1)^2} \left( 4 \left(\frac{2\gamma}{K + 1}\right)^{\gamma+1} K^3 (-\gamma^2 + \gamma) + 2K^2 (\gamma^2 - \gamma) + K (-\gamma^2 + \gamma) + 2K^2 \gamma^2 + 4K (-\gamma^2 + \gamma) + 4K + 2 (\gamma - 1)^2 + 2 \right) \\
= \frac{1}{2\gamma (\gamma + 1)(K - 1)^2} \left( (K - 1)^2 K (-\gamma^2 + \gamma) + 2K^2 \gamma^2 + 4K (-\gamma^2) + 2\gamma^2 - 2\gamma^2 + 4K \gamma + 4K + 2 (\gamma - 1)^2 + 2 \right) \\
= \frac{1}{2\gamma (\gamma + 1)(K - 1)^2} \left( (K - 1)^2 K (-\gamma^2) + 2\gamma^2 (K - 1)^2 + 4\gamma (K - 1) + 4 (K + 1) \right)
\]

Hence

\[
\frac{\partial^2 D_r(1/2, K, \gamma)}{\partial K^2} < 0
\]

if and only if

\[
G(K) \equiv \frac{(K + 1)^{\gamma+1}}{(K - 1)^2 (K (\gamma - \gamma^2) + 2\gamma^2) + \gamma (K - 1) + \frac{K + 1}{2}} < 1 \quad (32)
\]

First, we observe that \( G(1) = 1 \). Thus, to confirm (32), it is enough to show that the derivative of \( G \) is negative. To this end note that, letting \( x \equiv K - 1 \) for notational simplicity, we have that \( \frac{d}{dx} G(K) < 0 \) if, and only if,

\[
\frac{d}{dx} \left( \left( \frac{x}{2} + 1 \right)^{\gamma+1} \right) < \frac{d}{dx} \left( \frac{\gamma x \left( \frac{x}{2} (x (1 - \gamma) + 1 + \gamma) + \gamma + \frac{1}{2} \right) + \frac{x}{2} + 1}{\left( \gamma x \left( \frac{x}{2} (x (1 - \gamma) + 1 + \gamma) + \gamma + \frac{1}{2} \right) + \frac{x}{2} + 1 \right)^{\gamma+1}} \right)
\]
In view of the fact that
\[
\frac{d}{dx} \left( \left( x + \frac{1}{2} \right)^{\gamma+1} \right) = \frac{(\gamma+1) \left( x + \frac{1}{2} \right)^\gamma}{\left( x + \frac{1}{2} \right)^{\gamma+1}} = \frac{1}{x + \frac{1}{2}}
\]
one can write (32) as follows:
\[
\frac{1}{x + \frac{1}{2}} < \frac{\frac{1}{2} \gamma + \frac{1}{4} x \gamma + \frac{1}{4} x^2 \gamma + \frac{3}{8} x^2 \gamma - \frac{3}{8} x^2 \gamma^2 + 1}{x + \frac{1}{2} x \gamma + \frac{1}{8} x^2 \gamma + \frac{1}{8} x^3 \gamma + \frac{1}{8} x^2 \gamma^2 - \frac{1}{8} x^3 \gamma^2 + \frac{1}{2}}
\]
which is equivalent to the following inequalities:
\[
\left( \frac{1}{2} \gamma + \frac{1}{4} x \gamma + \frac{1}{4} x^2 \gamma + \frac{3}{8} x^2 \gamma - \frac{3}{8} x^2 \gamma^2 + 1 \right) (x + 2) >
\]
\[
\frac{x}{2} + \frac{1}{2} x^2 \gamma + \frac{1}{8} x^3 \gamma + \frac{1}{8} x^2 \gamma^2 - \frac{1}{8} x^3 \gamma^2 + 1
\]
\[
x + \frac{1}{4} \gamma + \frac{5}{8} x \gamma + \frac{1}{8} x^2 \gamma + \frac{7}{16} x^2 \gamma + \frac{3}{8} x^3 \gamma + \frac{1}{16} x^2 \gamma^2 - \frac{3}{8} x^3 \gamma^2 + \frac{1}{2} >
\]
\[
x + \frac{1}{2} x \gamma + \frac{1}{8} x^2 \gamma + \frac{1}{8} x^3 \gamma + \frac{1}{8} x^2 \gamma^2 - \frac{1}{8} x^3 \gamma^2 + \frac{1}{2}
\]
\[
\frac{1}{4} \gamma + \frac{1}{8} x \gamma + \frac{1}{8} x^2 \gamma + \frac{5}{16} x^2 \gamma + \frac{2}{8} x^3 \gamma - \frac{1}{16} x^2 \gamma^2 - \frac{2}{8} x^3 \gamma^2 > 0
\]
the above inequality being always true if \( \gamma < 1 \), which completes the proof.

**Proof of Proposition 3** From (22) and (28) we get:
\[
D_r(1/2, K, \gamma) - D_r(1/2, K, \gamma)
\]
\[
= \left( \frac{1}{2} \right)^\gamma \left( \frac{1}{K} \right)^{\gamma-1} - K \left( \frac{1}{\gamma + 1} \right) + \frac{2}{K-1} \left( \frac{1}{1} \right)^{\gamma-1} - \left( \frac{1}{2} \right)^{\gamma-1}
\]
Hence, using l'Hôpital,
\[
\lim_{K \to 1} (D_r(1/2, K, \gamma) - D_r(1/2, K, \gamma))
\]
\[
= \left( \frac{1}{2} \right)^\gamma \left[ 1 - \left( \frac{1}{\gamma + 1} \right) \right] + \lim_{K \to 1} \frac{1}{K-1} \left( \frac{1}{\gamma + 1} \right)^\gamma - \left( \frac{1}{2} \right)^\gamma
\]
\[
= \left( \frac{1}{2} \right)^\gamma \frac{1}{(\gamma + 1)^{\gamma}} - \frac{-\gamma 2^{\gamma}}{(\gamma + 1)^{\gamma}} = 0.
\]
This implies that the two structures are equally optimal when $\gamma < 1$ since in this case (under the maintained assumption that $N$ is even) a maximal segmentation in components of size $K = 1 = 2$ minimizes the expected number of defaults under both the ring and the completely connected structures.

Consider now the case of $\gamma > 1$, for which $K = N - 1$ (i.e. minimal segmentation) is optimal for both structures. If we evaluate (33) at such common optimal value for both structures we find:

$$D_c(1/2, K, \gamma) - D_r(1/2, K, \gamma) = \left(\frac{1}{2}\right)^\gamma \left(\frac{1}{K}\right)^{\gamma-1} - K \left(\frac{1}{\gamma+1}\right)^\gamma \frac{1}{K+1} + \frac{2}{K-1} \left(\frac{1}{K+1}\right)^\gamma - \left(\frac{1}{2}\right)^\gamma.$$

Hence for all $K \geq 1 + \gamma$

$$D_c(1/2, K, \gamma) - D_r(1/2, K, \gamma) < 0,$$

which implies that if $\gamma > 1$ and $N \geq 2 + \gamma$ the desired conclusion follows. This completes the proof. □

**Proof of Proposition 4** Abusing slightly previous notation, denote by $D_c(1/2, K, \gamma, \gamma', p) = \mathbb{E}_{\gamma, \gamma' ; \rho \gamma_{\rho b} (\tilde{L}_b ; K)}$ the expected mass of defaults in a complete component of size $K + 1$ when a $b$ shock hits some other firm in the component, and this shock is distributed as a mixture of a Pareto distribution with parameter $\gamma$ and a Pareto distribution with parameter $\gamma'$, with respective weights $p$ and $1 - p$. From (22), we can write:

$$D_c(1/2, K, \gamma, \gamma', p) = p K \left(\frac{1}{2K}\right)^\gamma + (1 - p) K \left(\frac{1}{2K}\right)^\gamma'.$$

Hence

$$\frac{\partial D_c}{\partial K}(1/2, K, \gamma, \gamma', p) = -p (\gamma - 1) \left(\frac{1}{2K}\right)^\gamma - (1 - p) (\gamma' - 1) \left(\frac{1}{2K}\right)^\gamma'.$$

(34)
Since $\gamma > 1$ and $\gamma' < 1$ we have that $\frac{\partial D_c}{\partial K} > 0$ is equivalent to
\[
(1 - p) (1 - \gamma') \left( \frac{1}{2K} \right)^{\gamma'} > p (\gamma - 1) \left( \frac{1}{2K} \right)^{\gamma},
\]
or
\[
K > \frac{1}{2} \left( \frac{p (\gamma - 1)}{(1 - p) (1 - \gamma')} \right)^{\frac{1}{1 - \gamma'}}
\]
which implies that the function is minimized at the point
\[
\hat{K}(p) = \frac{1}{2} \left( \frac{p (\gamma - 1)}{(1 - p) (1 - \gamma')} \right)^{\frac{1}{1 - \gamma'}}
\]
provided this point is admissible, i.e. $\hat{K}(p) \in [1, N - 1]$.

We now compute the second derivative of $D_c$ and find:
\[
\frac{\partial^2 D_c}{\partial K^2} (1/2, K, \gamma, \gamma', p) = p (\gamma - 1) \frac{\gamma}{K} \left( \frac{1}{2K} \right)^{\gamma} + (1 - p) (\gamma' - 1) \frac{\gamma'}{K} \left( \frac{1}{2K} \right)^{\gamma'}
\]
\[
\leq p (\gamma - 1) \frac{\gamma}{K} \left( \frac{1}{2K} \right)^{\gamma} + (1 - p) (\gamma' - 1) \frac{\gamma'}{K} \left( \frac{1}{2K} \right)^{\gamma'}
\]
\[
= -\frac{\gamma}{K} \frac{\partial D_c}{\partial K} (1/2, K, \gamma, \gamma', p)
\]
and therefore $\frac{\partial^2 D_c}{\partial K^2} (1/2, K, \gamma, \gamma', p) > 0$ for all feasible $K < \hat{K}(p)$, i.e. the function $D_c(\cdot)$ is convex in this range.

The socially optimal structure solves the following problem:
\[
\min_{K_i, C} \sum_{i=1}^C \frac{K_{i+1}}{N} D_c(\alpha, K_i, \gamma, \gamma', p)
\mathrm{s.t.} \sum_{i=1}^C \frac{K_{i+1}}{N} = 1 \quad K_i \geq 1
\]
Denote by $(K^*_i)_{i=1}^C$ a vector of component sizes that solves this optimization problem. We need to show that there exists some appropriate range $[p_0, p_1]$ such that if $p \in [p_0, p_1]$, $K^*_i = K^*_j = K^*$ for all $i, j = 1, 2, \ldots, C$ and some common $K^*$ with $2 \leq K^* \leq N - 2$.

Choose $p_0$ such that $\hat{K}(p_0) = \frac{N}{2} - 1$. Such a choice is feasible and unique since $\hat{K}(\cdot)$ is increasing in $p$, $\hat{K}(0) = 0$, and $\hat{K}(p) \to \infty$ as $p \to 1$. Suppose
now that the weight $p$ involved in the mixture underlying $\tilde{L}_b$ satisfies $p \geq p_0$.

Next we show that, in this case, the vector $(K^*_i)_{i=1}^C$ solving (35) must satisfy:

$$\forall i, j = 1, 2, \ldots, C, \quad K^*_i = K^*_j \leq \hat{K}(p)$$  \hspace{1cm} (36)

Let $K^*_i$ and $K^*_j$ stand for any two component sizes that are part of the solution to the optimization problem. First we note that, since $\hat{K}(p) \geq N/2 - 1$, if $K^*_i > \hat{K}(p)$ then we must have that $K^*_j < \hat{K}(p)$. But such asymmetric arrangement cannot be part of a solution to (35) because $D_c(1/2, \cdot, \gamma, \gamma', p)$ is increasing at $K^*_i$ and decreasing at $K^*_j$. Hence a small transfer of firms from component $i$ to component $j$ would decrease the expected number of defaults. The only possibility, therefore, is that $K^*_j \leq \hat{K}(p)$ and $K^*_j \leq \hat{K}(p)$.

To complete the argument for (36), we finally show that $K^*_i = K^*_j$. Suppose otherwise, and then consider substituting both (dissimilar) components for two components of equal size $\frac{1}{2}(K^*_i + K^*_j)$. Since neither $K^*_i$ or $K^*_j$ exceed $\hat{K}(p)$, both values lie in the convex part of the function $D_c(1/2, \cdot, \gamma, \gamma', p)$. It follows, therefore, that the indicated substitution would reduce the overall expected number of defaults, which contradicts that the original heterogeneous configuration is optimal.

We have thus shown that if $p \geq p_0$, the unique optimal configuration involves a uniform segmentation in components of common size $K^*(p) \leq \hat{K}(p)$. Since making $p = p_0$ we have $K^*(p_0) = \hat{K}(p_0) = N/2 - 1$ and $\hat{K}(p)$ is increasing and continuous in $p$, it is clear that we can choose some $p_1$ with $p_0 < p_1 < 1$ such that

$$\forall p \in [p_0, p_1], K^*(p) = N/2 - 1.$$  \hspace{1cm} (37)

For $N > 4$, this is consistent with the requirement that the optimal component size $K^* + 1$ be “intermediate,” i.e. satisfy $1 < K^* < N - 1$, and thus the proof of the proposition is complete. ■

**Proof of Proposition 5.** The expected mass of firms not directly hit by a shock who default in a completely connected component of size $K$ hit by a $b$ shock when the probability distribution of the $b$ shock as stated in the claim is:

$$D_c(1/2, K, \gamma, p) = (1 - p)K + pK \left( \frac{1}{2K} \right) \gamma.$$  \hspace{1cm} (38)
Differentiating the above expression with respect to \( K \) yields:

\[
\frac{\partial D_c(1/2, K, \gamma, p)}{\partial K} = (1 - p) - (\gamma - 1) p \left( \frac{1}{2K} \right)^\gamma,
\]

which is negative for all \( K \) as long as (24) is satisfied. This establishes that the optimal degree of segmentation for the complete structure is minimal, that is, \( K = N - 1 \).

Next, using (19) and (28), and noting that \( \bar{L} = 2(N - 1) + 1 > \frac{1}{\alpha} = K + 1 \) (where we use the simplifying assumptions that \( \alpha = 1/2 \) and \( R - M = 1 \)), we obtain the following expression for the expected mass of defaults in the case of the ring structure:

\[
D_r(1/2, K, \gamma, p) = (1 - p) \left( K - \left( K - \frac{2}{K + 1} \right) \frac{K + 1}{L} \right) + p \left[ \left( K \left( \frac{1}{\gamma + 1} \right) - \frac{2}{K - 1} \right) \left( \frac{1}{K + 1} \right)^\gamma \right] + p \left[ \frac{1}{K - 1} \right] \left( \frac{1}{\gamma + 1} \right)^\gamma.
\]

It suffices then to show that the expected mass of defaults is smaller for the ring than for the completely connected structure when \( K = N - 1 \). This comparison yields:

\[
D_c(1/2, N - 1, \gamma, p) > D_r(1/2, N - 1, \gamma, p) \iff (1 - p) (N - 1) + p (N - 1) \left( \frac{1}{2(N - 1)} \right)^\gamma > (1 - p) \left( N - 1 - \frac{N^2 - N - 2}{2N - 1} \right) + p \left( \frac{N - 1}{\gamma + 1} - \frac{2}{N - 2} \frac{1}{\gamma + 1} \right) \left( \frac{1}{N} \right)^\gamma + p \left[ \frac{1}{N - 2} \frac{1}{2(N - 1)} \frac{1}{\gamma + 1} \right]
\]

or

\[
\left( \frac{N - 1}{\gamma + 1} - \frac{2}{N - 2} \frac{1}{\gamma + 1} \right) \left( \frac{1}{N} \right)^\gamma + \frac{1}{N - 2} \frac{1}{2(N - 1)} \frac{1}{\gamma + 1} - (N - 1) \left( \frac{1}{2(N - 1)} \right)^\gamma.
\]

Using (24) the above inequality holds if

\[
(\gamma - 1) \left( \frac{1}{2(N - 1)} \right)^\gamma \frac{N^2 - N - 2}{2N - 1} > \left( \frac{N - 1}{\gamma + 1} - \frac{2}{N - 2} \frac{1}{\gamma + 1} \right) \left( \frac{1}{N} \right)^\gamma + \frac{1}{N - 2} \frac{1}{2(N - 1)} \frac{1}{\gamma + 1} - (N - 1) \left( \frac{1}{2(N - 1)} \right)^\gamma
\]

or

\[
\left( \frac{1}{2(N - 1)} \right)^\gamma \left[ (\gamma - 1) \frac{N^2 - N - 2}{2N - 1} + N - 1 \right] - \frac{N^2 - 3N - 2}{N - 2} \frac{1}{\gamma + 1} \left( \frac{1}{N} \right)^\gamma - \frac{1}{N - 2} \frac{1}{2(N - 1)} \frac{1}{\gamma + 1} > 0,
\]

51
or
\[
\left(\frac{1}{2(N-1)}\right)^\gamma (\gamma - 1) \frac{N^2-N-2}{2(N-1)} \frac{N^2-3N+2}{(N-2)\gamma+1} \frac{(\gamma + 1) N^\gamma - (2(N - 1))^{\gamma}}{4(N-1)(N-1)N}\]

Notice that, if \(N \geq 5\) (which is implied by (23)),
\[
\left((N - 1)^2 + (N - 3)\right) 4(N - 1) >\]

or
\[
\frac{(N - 1)^2 + N - 3}{4(N - 1) + 2} > \frac{(N - 1)^2}{4(N - 1)}.
\]

Hence
\[
(\gamma - 1) \left(\frac{1}{2(N-1)}\right)^\gamma 2 \frac{(N-1)^2 + N - 3}{2(N-1) + 1} + (N - 1) \left(\frac{1}{2(N-1)}\right)^\gamma
\]
\[
-2 \left(\frac{(N-1)}{2}\right) \left(\frac{1}{\gamma + 1}\right) - \frac{1}{\gamma + 1} \left(\frac{1}{N - 2}\right) \left(\frac{1}{N}\right)^\gamma - \frac{1}{N - 2} \frac{1}{2\gamma + 1} \left(\frac{1}{\gamma + 1}\right)
\]
\[
> (\gamma - 1) \left(\frac{1}{2(N-1)}\right)^\gamma 2 \left(\frac{(N-1)^2}{4(N-1)}\right)
\]
\[
+ (N - 1) \left(\frac{1}{2(N-1)}\right)^\gamma - \frac{N - 1}{\gamma + 1} \left(\frac{1}{N - 1}\right) \left(\frac{1}{N-1}\right)^\gamma - \frac{1}{N - 2} \frac{1}{2\gamma + 1} \left(\frac{1}{\gamma + 1}\right)
\]
\[
= \frac{1}{(N - 1)^{\gamma-1}} (\gamma - 1) \frac{1}{2\gamma + 1} + \frac{1}{2\gamma - \frac{1}{\gamma + 1}} - \frac{1}{N - 2} \frac{1}{2\gamma + 1} \left(\frac{1}{\gamma + 1}\right),
\]

where the latter expression is positive if
\[
\frac{N - 2}{(N - 1)^{\gamma-1}} = (N - 1)^{2-\gamma} - \frac{1}{(N - 1)^{\gamma-1}} > \frac{1}{2\gamma + 1} \frac{1}{(N - 1)^{\gamma-1}} + \frac{1}{2\gamma - \frac{1}{\gamma + 1}}
\]

which is implied by the inequality
\[
(N - 1)^{2-\gamma} - \frac{1}{2\gamma + 1} > \frac{1}{(N - 1)^{\gamma-1}} \frac{1}{2\gamma + 1} + \frac{1}{2\gamma - \frac{1}{\gamma + 1}}
\]

that is in turn equivalent to (23). This completes the proof of the proposition. □
Proof of Proposition 6. Let $L \equiv (L_i)_{i \in P}$ be a CPE of the network-formation game and denote by $\{C_q\}_{q=1}^r$ the list of components of the CPE network $\Gamma(L)$, with respective sizes $|C_q| = K_q + 1$ for each $q = 1, 2, ..., r$. Denote by $\hat{K}$ the unique value that minimizes $D_c(1/2, K, \gamma, \gamma', p)$ (c.f. the proof of Proposition 4). The proof of the Proposition can be decomposed in the following two claims.

Claim 1 $\forall q = 1, 2, ..., s$, $K_q \leq \hat{K}$.

Claim 2 Let $K_{\tilde{q}} < \hat{K}$ for some $\tilde{q} \in 1, 2, ..., r$. Then, $K_q = \hat{K}$ for all $q \neq \tilde{q}$.

To establish Claim 1, suppose that if $K_q > \hat{K}$ for some $q$. Then, choose a subset $D$ of the agents in the subpopulation $P_q$ of component $C_q$ whose measure satisfies $K_q - \hat{K} > |D| > 0$. Then, since the function $D_c(1/2, \cdot, \gamma, \gamma', p)$ is increasing beyond $\hat{K}$, we have that

$$D_c(1/2, K_q - |D|, \gamma, \gamma', p) < D_c(1/2, K_q, \gamma, \gamma', p)$$

which implies that the agents in $C_q \setminus D$ could enjoy a lower default probability if they could form a separate component. But this, of course, they can always do by jointly deleting the links to the agents in $D$, which establishes the desired conclusion.

Next, we proof Claim 2. Suppose, for the sake of contradiction, that there are two components of $\Gamma(L)$, $C_{q'}$ and $C_{q''}$, such that $0 < K_{q'} < K_{q''} < \hat{K}$. Consider then two subsets of each component: $D' \subset K_{q'}$, $D'' \subset K_{q''}$, both of which are of equal measure $\varepsilon$ with $0 < \varepsilon < \hat{K} - K_{q''}$. Let the individuals in $D' \cup D''$ contemplate a joint deviation from their strategies in $L$ whereby each of the agents in $D'$ establishes additional links with all of the individuals in $D''$, while keeping all their previous links intact. Then, every individual in $D'$ ends up having a total measure of links equal to $K_{q'} + \varepsilon$ and every individual in $D''$ a total measure of $K_{q''} + \varepsilon$. Denote by $\hat{\Gamma}$ the network induced by these additional links. Then, maintaining the assumption that the deviating agents continuing to exchange $1 - \alpha$ of their own assets with all their neighbors (old and new) in a uniform manner, the following inequalities apply:

$$\forall i \in D', \varphi_i(\hat{\Gamma}) = D_c(1/2, K_{q'} + \varepsilon, \gamma, \gamma', p) < D_c(1/2, K_{q'}, \gamma, \gamma', p) = \varphi_i(\Gamma(L))$$
∀i ∈ D′′ ϕ_i(Γ) = D_c(1/2, K_{q''} + ε, γ, γ', p) < D_c(1/2, K_{q'} + , γ, γ', p) = ϕ_i(Γ(L))

Hence all agents involved benefit from the deviation, which contradicts the hypothesis that the original profile L ≡ (L_i)_{i∈P} defined a CPE. This contradiction establishes the Claim and proves the Proposition. ■

Remark 2 It is immediate to check that the line of argument used above in proving Proposition 6 still applies if the notion of CPE were redefined to limit the size/measure of deviating coalitions to be no higher than some pre-established value η, arbitrarily small but positive. For, in the proof of both Claims 7 and 8, the only condition required is that some positive measure of agents be involved, which is obviously consistent with any positive value for the aforementioned η.

An additional observation worth making is that the deviations contemplated in the proof of Proposition 6 are “internally consistent” in the following sense. Given that any coalition of the required composition is set to deviate, this deviation is itself in the interest of any subset of the coalition that might reconsider the situation. This requirement (which is commonly demanded in the game-theoretic literature for coalition-based notions of equilibrium), is clearly satisfied in our case since, by refusing to follow suit with the deviation, any subcoalition can only loose.

Proof of Proposition 7 To prove the result, we compute the expected losses associated to the two structures (star and complete) for a full partition of the range of variation of the shock L. Two scenarios need to be treated separately in terms of the value of the parameter θ: θ > 1/3 and θ < 1/3. Both are qualitatively similar, but they differ in the specific subranges for L that must be considered. This pertains, specifically, on how two relevant thresholds compare, which is as stated by the following auxiliary lemma.

Lemma 2 Given β, let α and α' be determined by (26)-(27). Then,

\[
\frac{2\beta - 1}{1 - \alpha} > \frac{\beta}{1 - \alpha'} \iff \frac{\beta - 1}{\alpha' - \alpha} > \frac{\beta}{1 - \alpha'} \iff \theta > 1/3. \tag{39}
\]
Proof of Lemma 2:

To establish the last equivalence in (39), note that the following list of expressions are all equivalent:

\[
\frac{\beta - 1}{\alpha' - \alpha} > \frac{\beta}{1 - \alpha'} \\
(\beta - 1)(1 - \alpha') > \beta(\alpha' - \alpha) = \beta \frac{\beta - 1}{\beta} (1 - \theta)^2 \\
2(1 - \theta)\theta > (1 - \theta)^2 \\
2\theta > 1 - \theta
\]

and the last inequality can be simply rewritten as \( \theta > 1/3 \).

The first equivalence in (39), on the other hand, is an immediate consequence of the fact that the following expressions are all equivalent as well:

\[
\frac{2\beta - 1}{1 - \alpha} > \frac{\beta}{1 - \alpha'} \\
(\beta - 1)(1 - \alpha') + \beta(1 - \alpha') > \beta(1 - \alpha) \\
(\beta - 1)(1 - \alpha') > \beta(\alpha' - \alpha) \\
\frac{\beta - 1}{\alpha' - \alpha} > \frac{\beta}{1 - \alpha'}
\]

hence the desired conclusion simply follows from the last equivalence in (39).

\( \Box \)

The proof of the proposition entails computing and comparing the expected number of defaults that result under each of the two configurations considered: star and symmetric structures. Denote by \( \Psi_{\text{star}}(L) \) and \( \Psi_{\text{sym}}(L) \) the functions that specify those defaults as a function of the magnitude \( L \) of the shock. In what follows, we construct these functions, in a way that Lemma 39 guarantees is well defined.

We consider first the case of \( \theta > 1/3 \). Ex ante, a shock hitting a component of the network has the same probability of striking a large firm or a small firm. So let us begin by determining the expected number of defaults if the shock hits a small firm and the structure of the component to which it belongs is a star. For any given value of \( L \), it is given by the following
function:

\[
\Psi_{\text{star}}(L) = \begin{cases} 
0 & \text{for } L \leq \frac{1}{\alpha} \\
1 & \text{for } \frac{1}{\alpha} < L \leq \frac{\beta}{1-\alpha} \\
1 + \beta & \text{for } \frac{\beta}{1-\alpha} < L \leq \frac{\beta-1}{\alpha-\alpha} \\
2\beta & \text{for } L > \frac{\beta-1}{\alpha-\alpha}
\end{cases}
\]

Instead, if the structure is still a star but the shock hits a large firm (i.e. the hub), the expected number of defaults is given by the following function:

\[
\Psi_{\text{star}}^\ell(L) = \begin{cases} 
0 & \text{for } L \leq \frac{\beta}{1-\alpha} \\
\beta & \text{for } \frac{\beta}{1-\alpha} < L \leq \frac{\beta}{1-\alpha} \\
2\beta & \text{for } L > \frac{\beta}{1-\alpha}
\end{cases}
\]

Hence, overall, when both of the former possibilities (small and large firm being hit by the shock) have the same ex ante probability, the corresponding function is given by:

\[
\Psi_{\text{star}}(L) = \begin{cases} 
0 & \text{for } L \leq \frac{1}{\alpha} \\
\frac{1}{2} \times 1 + \frac{1}{2} \times 0 & \text{for } \frac{1}{\alpha} < L \leq \frac{\beta}{1-\alpha} \\
\frac{1}{2} \times 1 + \frac{1}{2} \times \beta & \text{for } \frac{\beta}{1-\alpha} < L \leq \frac{\beta-1}{\alpha-\alpha} \\
\frac{1}{2} \times (1 + \beta) + \frac{1}{2} \times 2\beta & \text{for } \frac{\beta-1}{\alpha-\alpha} < L \leq \frac{2\beta-1}{1-\alpha} \\
2\beta & \text{for } L > \frac{2\beta-1}{1-\alpha}
\end{cases}
\]

(40)

The alternative to the star arrangement is to place two large firms (of size \(\beta\)) in one component, and 2\(\beta\) small firms in the other, completely connected. In this case, since there is no asymmetry within each component, every firm keeps \(\alpha\) of its own project. And since the shock reaches with equal probability each of the two components, the expected losses are given by the following function:

\[
\Psi_{\text{sym}}(L) = \begin{cases} 
0 & \text{for } L \leq \frac{1}{\alpha} \\
\frac{1}{2} \times 1 + \frac{1}{2} \times 0 & \text{for } \frac{1}{\alpha} \leq L \leq \frac{\beta}{1-\alpha} \\
\frac{1}{2} \times 1 + \frac{1}{2} \times \beta & \text{for } \frac{\beta}{1-\alpha} \leq L \leq \frac{2\beta-1}{1-\alpha} \\
\frac{1}{2} \times (1 + \beta) + \frac{1}{2} \times 2\beta & \text{for } \frac{2\beta-1}{1-\alpha} \leq L \leq \frac{2\beta-1}{1-\alpha} \\
2\beta & \text{for } L \geq \frac{2\beta-1}{1-\alpha}
\end{cases}
\]

(41)
A straightforward comparison of the functions $\Psi_{\text{star}}(L)$ and $\Psi_{\text{sym}}(L)$ given in (40) and (41), in combination with Lemma 39, yields the part of the Proposition that is associated to the case where $\theta > 1/3$.

For the case with $\theta < 1/3$, an analogous procedure must be applied, now giving rise to corresponding functions that, for comparison, we denote by $\hat{\Psi}_{\text{star}}(L)$, $\hat{\Psi}_{\text{star}}^\ell(L)$, $\hat{\Psi}_{\text{sym}}(L)$, and $\hat{\Psi}_{\text{sym}}^\ell(L)$. Two of them are as before, i.e. $\hat{\Psi}_{\text{star}}^\ell(L) = \Psi_{\text{star}}^\ell(L)$ and $\hat{\Psi}_{\text{sym}}^\ell(L) = \Psi_{\text{sym}}^\ell(L)$. However, the function specifying the expected number of defaults if the shock hits a small firm in a star is now different and given by:

$$
\hat{\Psi}_{\text{star}}^s(L) = \begin{cases} 
0 & \text{for } L \leq \frac{1}{\alpha} \\
1 & \text{for } \frac{1}{\alpha} < L \leq \frac{\beta-1}{\alpha-\alpha'} \\
\beta & \text{for } \frac{\beta-1}{\alpha-\alpha'} < L \leq \frac{\beta}{1-\alpha'} \\
2\beta & \text{for } L > \frac{\beta}{1-\alpha'}
\end{cases}
$$

which combined with $\hat{\Psi}_{\text{star}}^\ell(L)$ (equal to $\Psi_{\text{star}}^\ell(L)$) gives rise to the overall function corresponding to the the star structure in this case:

$$
\hat{\Psi}_{\text{star}}(L) = \begin{cases} 
0 & \text{for } L \leq \frac{1}{\alpha} \\
\frac{1}{2} \times 1 + \frac{1}{2} \times 0 & \text{for } \frac{1}{\alpha} < L \leq \frac{\beta}{\alpha'} \\
\frac{1}{2} \times 1 + \frac{1}{2} \times \beta & \text{for } \frac{\beta}{\alpha'} < L \leq \frac{\beta-1}{\alpha-\alpha'} \\
\frac{1}{2} \times \beta + \frac{1}{2} \times \beta & \text{for } \frac{\beta-1}{\alpha-\alpha'} < L \leq \frac{\beta}{1-\alpha'} \\
2\beta & \text{for } L > \frac{\beta}{1-\alpha'}
\end{cases}
$$

Finally, comparing the function $\hat{\Psi}_{\text{star}}(L)$ specified above with the function $\hat{\Psi}_{\text{star}}(L)$ (equal to $\Psi_{\text{star}}(L)$) given in (41), the statement of the Proposition for the case of $\theta < 1/3$ can be readily verified. This completes the proof of the result.

References


