Submitted to *Operations Research*
manuscript (Please, provide the manuscript number!)

Robust Risk Quantification via Shock Propagation in Financial Networks

Dohyun Ahn, Nan Chen
Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, dohyun.ahn@cuhk.edu.hk, nchen@se.cuhk.edu.hk

Kyoung-Kuk Kim
College of Business, Korea Advanced Institute of Science and Technology, Seoul, Republic of Korea, kkim@kaist.ac.kr

Given limited network information, we consider robust risk quantification under the Eisenberg-Noe model for financial networks. To be more specific, motivated by the fact that the structure of the interbank network is not completely known in practice, we propose a robust optimization approach to obtain worst-case default probabilities and associated capital requirements for a specific group of banks (e.g., SIFIs) under network information uncertainty. With this analytical tool, we observe the effects of various incomplete network information on these worst-case quantities and provide insights into the collection of network information from the perspective of financial regulators. All claims are numerically illustrated using data from the European banking system.

Key words: Risk quantification; Financial network; Robust optimization; Information uncertainty

1. Introduction

Along with the evolution of interconnectedness in financial systems, the importance of financial institutions' exposure to losses arising from other institutions' defaults has increased significantly in recent years. The more banks are exposed to such losses, the more likely they are to suffer from the domino effects of financial failures. As seen in the financial crisis of 2008, large financial institutions may face severe financial stress in quick succession. For another example, since the onset of the Greek sovereign debt crisis in 2010, some European countries have failed to make

scheduled debt payments to their creditors such as the International Monetary Fund, resulting in significant stress on many European banks (Guerrieri et al. 2012).

Such repeating financial crises have provoked heated academic debates on how to measure, mitigate, and manage the aforementioned risk with a focus on the impact of the financial system's network topology. Above all, Eisenberg and Noe (2001) have played a key role in those discussions. Specifically, the authors provide a clearing mechanism that settles payment obligations of financial institutions based on a fixed-point characterization. This mechanism effectively describes default cascades triggered by interbank liabilities in financial networks, leading to many subsequent studies. To name a few, Elsinger et al. (2006) develop an empirical approach to assess systemic risk, Capponi et al. (2016) analyze the effect of liability concentration on systemic losses, and Cifuentes et al. (2005) and Rogers and Veraart (2013) propose extended models incorporating asset fire sales.[1] Liu and Staum (2010) and Feinstein et al. (2018) conduct sensitivity analyses of clearing payments, and Barucca et al. (2020) study a network-based valuation model for interbank claims. See Birge et al. (2018) for a good review of this strand of the literature. Interested readers can also consult Gai and Kapadia (2010) and Elliott et al. (2014) for other models of default contagion in the networks.

Nonetheless, only a few papers use the Eisenberg-Noe model to investigate the impact of random losses in banks' assets on their solvency, which is of practical importance to regulators. For example, Chen et al. (2016) find a lower bound of the probability that a shock to a single bank leads to other banks' default. Khabazian and Peng (2019) provide a lower bound of the probability of bankruptcy in the system when all banks receive normally distributed shocks. However, these works are not applicable to the case of multivariate shocks following a general class of probability distributions.

Another limitation of the model is that it requires full information on interbank exposures in the network, which is rarely available in practice as commonly noted in three well-known surveys in this context (Glasserman and Young 2016, Capponi 2016, Benoit et al. 2017). Instead, various types of partial information can be collected. For example, each bank's aggregate interbank exposure can be generally known from the balance sheet. The Basel III framework requires banks to report large

interbank exposures, defined as exposures greater than or equal to 10% of their Tier 1 capital (BCBS 2020a,c). Also, the Basel Committee annually investigates the interbank assets and liabilities of the so-called systemically important financial institutions (SIFIs), but not all banks are examined.

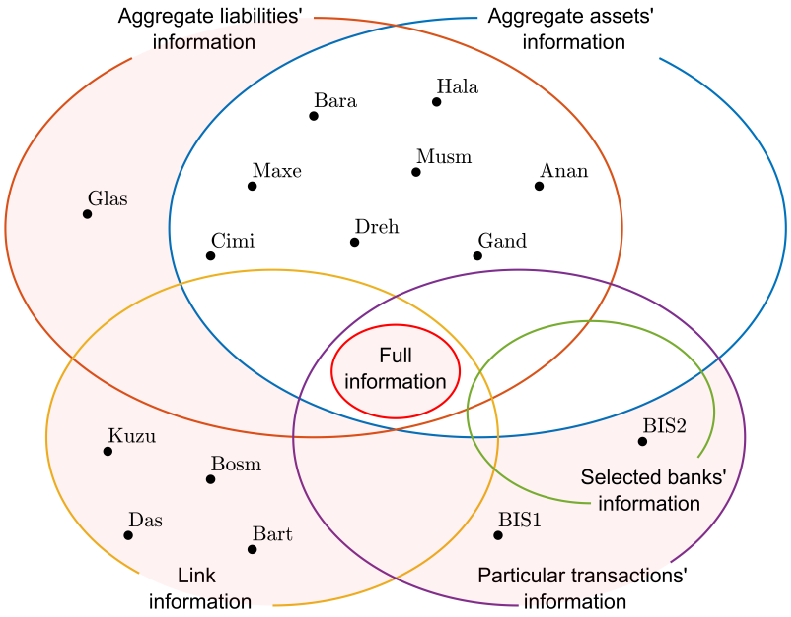
To address the issue of incomplete network information, many studies have developed network reconstruction methods using the aggregate information of interbank assets and liabilities (e.g. Anand et al. 2015, Upper and Worms 2004, Baral and Fique 2012, Cimini et al. 2015, Drehmann and Tarashev 2013, Halaj and Kok 2013, Musmeci et al. 2013, Gandy and Veraart 2017). However, since these methods do not consider random shocks to financial institutions, they do not provide a clear answer to risk quantification (in particular, to the estimation of risk capital). Moreover, Anand et al. (2018) find that none of those methods is generally superior to the other methods and that the results highly depend on the jurisdiction of the data and the performance measure. See Squartini et al. (2018) for a systematic review of recent network reconstruction methods.

Several other works focus on connectedness between banks, which we call link information, to understand each bank's risk contribution. For example, Kuzubaş et al. (2014), Das (2016), Bosma et al. (2019) and Bartesaghi et al. (2020) use the link information to rank the banks according to their systemic importance and identify SIFIs through the notion of network centrality. Still, the effects of random shocks are not discussed in those works.

To our knowledge, Glasserman and Young (2015) is so far the only paper that considers both random shocks and partial network information. They use the aggregate liabilities' information to obtain upper bounds of banks' default probability and expected loss, assuming that random shocks follow a particular type of distribution. However, little is known of a risk quantification method that can accommodate various types of incomplete network information and random shocks.

Motivated by these limitations, this paper studies robust risk quantification based on multivariate random shocks to financial institutions and incomplete network information under the Eisenberg-Noe model. In particular, we provide a comprehensive framework for estimating the worst-case default probabilities under different kinds of network information. In Figure 1, we illustrate the

Figure 1 The network information sets that this work covers.



Note. The shaded area represents the network information that we address. The existing literature indicated here includes Anand et al. (2015) (Anan), Upper and Worms (2004) (Maxe), Baral and Fique (2012) (Bara), Cimini et al. (2015) (Cimi), Drehmann and Tarashev (2013) (Dreh), Halaj and Kok (2013) (Hala), Musmeci et al. (2013) (Musm), Gandy and Veraart (2017) (Gand), Glasserman and Young (2015) (Glas), Kuzubaş et al. (2014) (Kuzu), Bosma et al. (2019) (Bosm), Bartesaghi et al. (2020) (Bart), Das (2016) (Das), BCBS (2020a) (BIS1), and BCBS (2020c) (BIS2).

network information applicable to our framework and classify the aforementioned studies based on the partial information they consider. We particularly focus on the default event of a specific set of banks inspired by the fact that regulators pay close attention to the solvency of SIFIs. Indeed, since the collapse of SIFIs could lead to a financial crisis, the Basel Committee developed higher loss absorbency requirements for those institutions to secure the financial system by preventing their bankruptcy (BCBS 2020b). The detailed contributions of this work are as follows.

1. **MILP formulation for worst-case default probabilities.** We first characterize a bank's default event with respect to a shock vector, assuming that full network information is available. This is then leveraged to formulate a mixed-integer linear program that identifies the worst-case default event of one or more banks in a specific set, taking into account all possible network configurations under limited network information. Our approach facilitates the unbiased estima-

tion of the worst-case default probability and outperforms existing robust optimization methods that provide more conservative solutions or require computationally intractable procedures.

2. **Worst-case systemic risk capital.** We compute the minimum capital amount required to keep all SIFIs solvent with a certain probability to protect the financial system from falling into a crisis, which we call *worst-case risk capital*.[2] Specifically, we formulate a chance-constrained optimization problem in which the worst-case default probability of SIFIs is controlled by a cash injection strategy. The difficulty of solving the problem is alleviated by the shock propagation mechanism and sample average approximation.
3. **Information sensitivity.** We proceed to apply our results to various types of limited network information based on real-world data. For information to be gathered efficiently, we make the following suggestions on regulatory policies:
   (a) When it comes to regulating some selected banks (e.g. SIFIs), gathering only those banks' information would be one of the best options for regulators to take. Information from the other banks is found to have little impact on improving the worst-case quantities associated with the target banks. Given that data collection can be costly, this result helps effectively reduce the amount of information regulators need to collect.
   (b) Another effective option is to collect information on large exposures since they are likely to be the main source of potential shock contagion. Our numerical analysis indicates that this information leads to small gaps between the associated worst-case quantities and the true quantities. This alternative, however, calls for a judicious choice of the threshold for large exposures to balance the cost of gathering information with its effectiveness.
   (c) In the case it is difficult to require banks to report their full data, collecting only link information without exact amounts is a sensible alternative, provided that the network has low density. Our study demonstrates that the worst-case quantities under this kind of information are close to the true quantities for sparse networks.[3]
   (d) It is unlikely that relying on only aggregate information or a part of target banks' information will help. We observe a significant gap between the true quantities and the worst-case quantities under such information.

We note that the observations in contributions 3-(a) and 3-(b) share a common thread with Musmeci et al. (2013) and Cimini et al. (2015) since they numerically show the robustness of systemic risk estimates to partial network information. However, in contrast to those works, this paper considers random shocks for risk quantification and investigates how the difference between true and worst-case risk quantities varies according to various types of limited network information.

The remainder of the paper is organized as follows. Section 2 introduces the underlying model and describes the problem formulation. In Section 3, we characterize the region for a shock vector to drive a specific bank to default using the notion of shock propagation. In Section 4, we obtain the worst-case default probabilities, which change according to the level of available information. In the section that follows, we apply our results to the computation of the worst-case risk capital under incomplete network information. In Section 6, we verify the practical applicability of our results via numerical experiments using real-world data. Finally, Section 7 concludes the paper. All proofs can be found in the online supplement.

2. Problem Formulation

We begin this section by introducing the basic notations in this paper. All vectors are column vectors, denoted by bold symbols; e.g., $\mathbf{u} = (u_1, \dots, u_d)^\top \in \mathbb{R}^d$. The Euclidean norm of a vector $\mathbf{u}$ is denoted by $\|\mathbf{u}\|$. We use $\mathbf{1}$, $\mathbf{0}$, and $\mathbf{I}$ for vectors of ones and zeros and the identity matrix in a suitable dimension, respectively. For any two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^d$, $\mathbf{u} \le \mathbf{v}$ means an entry-wise inequality, $[\mathbf{u}, \mathbf{v}] = \{\mathbf{x} \in \mathbb{R}^d \mid \mathbf{u} \le \mathbf{x} \le \mathbf{v}\}$, $\mathbf{u} \wedge \mathbf{v} = (\min\{u_1, v_1\}, \dots, \min\{u_d, v_d\})^\top$, and $\mathbf{u}^+ = (\max\{u_1, 0\}, \dots, \max\{u_d, 0\})^\top$. For any matrix $\mathbf{M}$, we denote by $\mathbf{M}_{-i}$ the matrix obtained by eliminating its i-th column and row (similarly defined for vectors). For any two index sets $\mathcal{I}, \mathcal{J} \subset \{1, \dots, d\}$, $\mathbf{v}_\mathcal{I}$ is the vector obtained by restricting the entries of the vector $\mathbf{v} \in \mathbb{R}^d$ to $\mathcal{I}$, and $\mathbf{M}_{\mathcal{I},\mathcal{J}}$ is the matrix obtained by restricting the components of the $d \times d$ matrix $\mathbf{M}$ to $\mathcal{I} \times \mathcal{J}$. If $\mathcal{I} = \mathcal{J}$, we simply write $\mathbf{M}_\mathcal{I}$.

2.1. The Model and Basic Assumptions

In this paper, we consider an n-bank financial system, where banks are indexed by $1, \dots, n$, and adopt an extended framework of Eisenberg and Noe (2001). The extension comes from, first, the

existence of liabilities to entities outside the financial network, and second, the inclusion of random shocks to the external asset values and bankruptcy costs. This modeling framework is proposed by Glasserman and Young (2015) and it consists of the following ingredients:

- $\bar{p}_{ij}$ is the payment obligation from bank i to bank j, with $\bar{p}_{ii} = 0$;
- for each i, $c_i, b_i \geq 0$ are the external assets and the external liabilities of bank i, respectively;
- each bank's balance sheet is given by
  1. the asset side: $c_i + \sum_{j \neq i} \bar{p}_{ji}$;
  2. the liability side: $\bar{p}_i := b_i + \sum_{j \neq i} \bar{p}_{ij}$;
  3. the initial net worth (book value): $w_i := c_i + \sum_{j \neq i} \bar{p}_{ji} - \bar{p}_i$;
- the proportion of bank i's obligation to bank j is defined by $a_{ij} := (\bar{p}_{ij}/\bar{p}_i)\mathbf{1}_{\{\bar{p}_i > 0\}}$;
- a random variable X_i is a shock to the external asset c_i;
- $\eta \geq 0$ is a multiplier for bankruptcy costs which will be introduced below.

After the shock, bank i's external asset and its net worth become $c_i - X_i$ and $w_i - X_i$, respectively. Here, $\mathbf{1}_{\{\bar{p}_i > 0\}}$ yields 1 if $\bar{p}_i > 0$ and 0 otherwise. We denote the matrix of the relative liabilities by $\mathbf{A} = (a_{ij})$ and the sum of the i-th row of $\mathbf{A}$ by β_i. The quantity β_i is the so-called *financial connectivity* of bank i, measuring its reliance on funding sources inside the financial system. In this paper, we denote by $\mathbf{x} = (x_1, \ldots, x_n)$ a realization of the random shock vector $\mathbf{X} = (X_1, \ldots, X_n)$.

In this framework, bank i defaults if bank i fails to pay its full liabilities $\bar{p}_{ij}$'s and b_i, i.e., if the asset side of its balance sheet is not large enough to keep it solvent. Then its assets are further reduced by *bankruptcy costs* proportional to its shortfall in payments, given by

$$\eta \left\{ \bar{p}_i - \left(c_i - x_i + \sum_{j \neq i} p_j a_{ji} \right) \right\}, \tag{1}$$

where p_j denotes the total debt payment of bank j. After deducting the bankruptcy costs, its residual assets are distributed to its creditors according to the pro rata allocation rule, where interbank liabilities $\bar{p}_{ij}$'s and external liabilities b_i have the equal priority.

Based on the above-mentioned features of this framework, the celebrated notion of the clearing payment vector $\mathbf{p}(\mathbf{x}) \in \mathbb{R}_+^n$ is defined as a solution to the following implicit equation:

$$p_i = \bar{p}_i \wedge \left\{ c_i - x_i + \sum_{j \neq i} p_j a_{ji} - \eta \left(\bar{p}_i - c_i + x_i - \sum_{j \neq i} p_j a_{ji} \right) \right\}^+, \quad i = 1, \ldots, n. \tag{2}$$

Note that the term in the braces represents the amount of remaining assets after accounting for bankruptcy costs. Thus, (2) implies that each bank either meets its payment obligation or distributes all its remaining assets to the creditors. We note that the solvency condition for bank i is equivalent to $p_i(\mathbf{x}) = \bar{p}_i$; the default of bank i happens only when $p_i(\mathbf{x}) < \bar{p}_i$.

To facilitate our analysis, we impose the following modeling assumptions in the paper.

ASSUMPTION 1. (a) *The initial net worth is positive and bounded by the external asset, i.e.,* $0 < w_i \le c_i$ *for each* i.

(b) *There does not exist any subset* $\mathcal{S}$ *of banks such that* $a_{ij} = a_{ji} = 0$ *for all* $(i,j) \in \mathcal{S} \times \mathcal{S}^{\mathrm{c}}$.

(c) $0 \le \eta < (\max_i \beta_i)^{-1} - 1$.

(d) *For each* i, $x_i \le c_i - \eta\bar{p}_i/(1+\eta)$.

Item (a) allows us to focus on the impact of a shock vector on the system stability, and it also indicates that interbank assets are smaller than total liabilities. Assumption 1-(b) implies that the financial network is irreducible because if such set $\mathcal{S}$ exists, then banks in $\mathcal{S}$ can be regarded as being outside the system and considered separately. Condition (c) makes the function $\mathbf{p}(\mathbf{x})$ unique (Glasserman and Young 2015), and item (d) prevents bankruptcy costs in (1) from exceeding the total assets so that taking the positive part in (2) becomes unnecessary.[4]

2.2. The Main Problem

In this paper, under the framework in Section 2.1, we focus on the probability that at least one bank in a specific set $\mathcal{T}$ defaults, i.e.,

$$\mathsf{P}\left(\mathbf{X} \in \bigcup_{i \in \mathcal{T}} \mathsf{D}_i\right) \tag{3}$$

where D_i is a set of shock vectors that make bank i default. Due to the shock propagation, for each $i \in \mathcal{T}$, the default event D_i of bank i results not only from the loss in the bank's external assets but also from its exposure to the loss in interbank transactions.

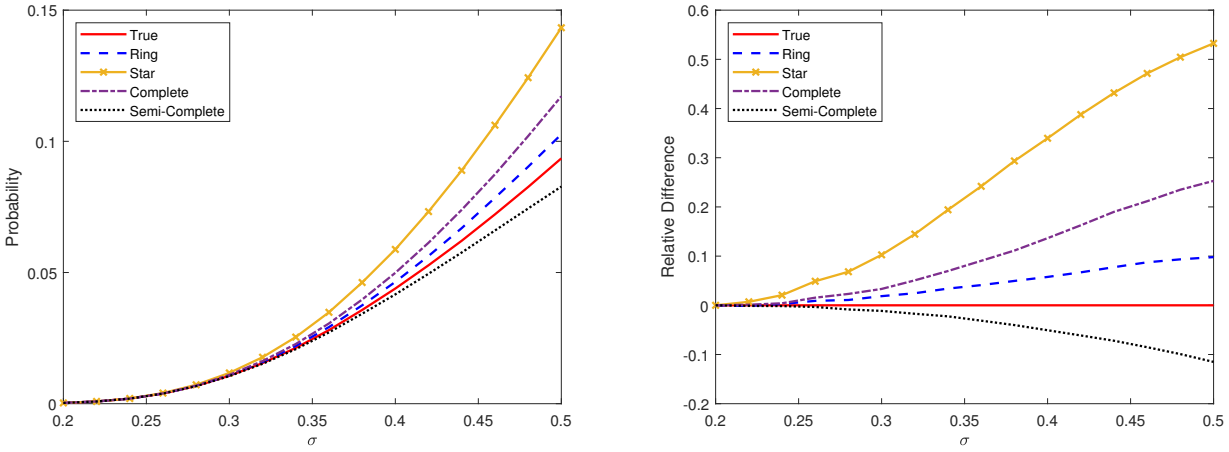
If we know full network information $\mathbf{A}$ and the distribution of $\mathbf{X}$, then the default probability (3) can be easily estimated via Monte Carlo simulation: for l simulated shocks $\mathbf{x}^1, \ldots, \mathbf{x}^l$, by solving (2),

Table 1 Possible examples of misspecified networks in Example 1.

Network Type	Ring	Star	Complete	Semi-Complete
Relative Liability Matrix	$\begin{bmatrix} 0 & .8 & 0 & 0 & 0 \\ 0 & 0 & .8 & 0 & 0 \\ 0 & 0 & 0 & .8 & 0 \\ 0 & 0 & 0 & 0 & .8 \\ .8 & 0 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 & 0 & .8 \\ 0 & 0 & 0 & 0 & .8 \\ 0 & 0 & 0 & 0 & .8 \\ 0 & 0 & 0 & 0 & .8 \\ .2 & .2 & .2 & .2 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & .1 & .1 & .1 & .5 \\ .1 & 0 & .1 & .1 & .5 \\ .1 & .1 & 0 & .1 & .5 \\ .1 & .1 & .1 & 0 & .5 \\ .2 & .2 & .2 & .2 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & .4 & .2 & .2 & 0 \\ .2 & 0 & .4 & .2 & 0 \\ .2 & .2 & 0 & .4 & 0 \\ .4 & .2 & .2 & 0 & 0 \\ .2 & .2 & .2 & .2 & 0 \end{bmatrix}$

The leading zeros are omitted in the matrices for the ease of exposition. The naming convention of the networks is consistent with the literature, except for the semi-complete network whose name stems from the fact that its subnetwork consisting of banks 1 to 4 is a complete network.

Figure 2 Bank 5's default probabilities under different network structures in Example 1.



Note. The left panel illustrates the bank's default probabilities with different σ under the true network and the misspecified networks in Table 1. The right panel describes the relative differences between the default probabilities under the misspecified networks and under the true network. The Monte Carlo method with 10^5 replications is used for the estimation of the probabilities.

one can find clearing payments $\mathbf{p}(\mathbf{x}^1), \ldots, \mathbf{p}(\mathbf{x}^l)$, based on which the frequency of the target event is measured. Alternatively, Ahn and Kim (2018) introduce an efficient computational method for the probability (3) using conditional Monte Carlo and importance sampling. However, the public, banks, or regulators often face a lack of full network information and have only partial information about financial networks. This poses a huge challenge to accurately assess the default probability (3) due to possible misspecification of the target financial network. Below we provide a simple example showing that the network misspecification leads to a misestimation of the default probability.

EXAMPLE 1. Consider a 5-bank financial network where $w_i = 2$, $\bar{p}_i = 8$ for $i = 1, \ldots, 5$, $\eta = 0.1$ and the matrix $\mathbf{A}$ of relative liabilities is given by $a_{ij} = 0.2$ for all $i \neq j$. Note that $\beta_1 = \ldots = \beta_5 = 0.8$. However, if only the aggregate information $(\beta_1, \ldots, \beta_5)$ is available, one may incorrectly specify the network structure as, for example, a ring network or a star network. Table 1 provides four possible misspecified networks given the aggregate information. Assuming that the random shocks $X_1, \ldots, X_5$ are i.i.d. and follow a lognormal distribution with parameters $\mu = 0$ and $\sigma \in [0.2, 0.5]$, we demonstrate in Figure 2 that the default probabilities of bank 5 under the misspecified networks could deviate significantly from its true default probability. The star network and semi-complete network in Table 1 result in the smallest and largest default probabilities for bank 5, respectively, and the other networks in the table lead to values in between. The case of multiple target banks will be discussed in Section 4.

Given such network uncertainties that make the exact computation of (3) intractable, one would need to explore its worst-case version while utilizing all incomplete but available network information. Hence, given each bank's total liabilities $(\bar{p}_i)$, external liabilities (b_i), equities (w_i) and partial network information, we aim at investigating the *worst-case default probability*:

$$\mathsf{P}\big(\mathbf{X} \in \mathsf{D}_{\mathcal{A},\mathcal{T}}\big). \tag{4}$$

Here, $\mathcal{A}$ is the set of all possible matrices of relative liabilities under given the partial network information, which we call the *network uncertainty set* hereafter. The set $\mathsf{D}_{\mathcal{A},\mathcal{T}} := \bigcup_{\mathbf{A}\in\mathcal{A}} \bigcup_{i\in\mathcal{T}} \mathsf{D}_i$ is a collection of shock vectors that might cause at least one bank in $\mathcal{T}$ to default in the worst case. We often suppress the subscript $\mathcal{T}$ if it is clear from the context. By definition, (4) bounds (3) from above, and they are equal when the full network information is known, i.e., when $\mathcal{A}$ is a singleton.

REMARK 1. Given a lack of knowledge on the true network structure, for practical purposes, one might consider estimating the probability (3) with a reduced-form approach such as a one-factor Gaussian copula model. This relies on each bank's marginal default probability and the correlation coefficients between the common market factor and idiosyncratic factors without considering network effects. However, if the network effects are not involved, each bank's marginal probability will depend only on losses in external assets, leading to the underestimation of the target probability.[5]

3. Shock Propagation

To facilitate the derivation of the set $\mathsf{D}_{\mathcal{A},\mathcal{T}}$, we characterize the solvency condition via the analysis of shock propagation in the following lemma.

LEMMA 1. *For each shock realization* $\mathbf{x}$*, bank* i *is solvent if and only if* $\Phi_i(\mathbf{x}) \le w_i$*, where*

$$\Phi_i(\mathbf{x}) := x_i + \max_{\boldsymbol{\zeta} \in \mathsf{Q}_i} \boldsymbol{\zeta}^\top (\mathbf{x}_{-i} - \mathbf{w}_{-i}), \tag{5}$$

$\mathsf{Q}_i := \left\{ \mathbf{v}_{-i} \in \mathbb{R}_+^{n-1} | \mathbf{v}_{\mathcal{I}} = (1+\eta)(\mathbf{I} - (1+\eta)\mathbf{A}_{\mathcal{I}})^{-1} \mathbf{a}^i_{\mathcal{I}},\ \mathbf{v}_{\mathcal{I}^c} = \mathbf{0},\ \mathcal{I} \subset \{1,2,\ldots,n\} \setminus \{i\} \right\}$*,* $\mathbf{a}^i$ *is the* i*-th column of* $\mathbf{A}$*, and* w_i *is bank* i*'s initial net worth defined in Section 2.1 for each* $i \in \{1,2,\ldots,n\}$*.*

For each shock realization $\mathbf{x}$, $\Phi_i(\mathbf{x})$ in (5) can be viewed as the total shock to bank i, which aggregates the *direct* shock to bank i (i.e., the first term in the right-hand side of (5)) and the *indirect* shock (i.e., the second term in the right-hand side of (5)) that indicates shock propagation from other banks to bank i. Based on the above lemma, the set D_i defined in Section 2.2 can be represented as $\mathsf{D}_i = \{\mathbf{x} \in [\mathbf{0}, \mathbf{c}] \,|\, \Phi_i(\mathbf{x}) > w_i\}$, which is the set of shock vectors that cause the total shock to bank i to exceed its net worth.[6]

To understand the implication of the vector $\boldsymbol{\zeta}$ in (5), we observe that for each shock realization $\mathbf{x}$ and for each i, there exists some $\mathcal{I}$ in $\{1,2,\ldots,n\} \setminus \{i\}$ such that

$$\Phi_i(\mathbf{x}) = x_i + \left(\mathbf{a}^i_{\mathcal{I}}\right)^\top \left((\mathbf{I} - (1+\eta)\mathbf{A}_{\mathcal{I}})^{-1} \right)^\top \left((1+\eta)(\mathbf{x}_{\mathcal{I}} - \mathbf{w}_{\mathcal{I}}) \right), \tag{6}$$

where the set $\mathcal{I}$ represents the set of insolvent banks. This implies that excess shocks of banks in $\mathcal{I}$, i.e., $\mathbf{x}_{\mathcal{I}} - \mathbf{w}_{\mathcal{I}}$, are increased by a factor $(1+\eta)$ and further amplified while circulating within the set $\mathcal{I}$ since $(\mathbf{I} - (1+\eta)\mathbf{A}_{\mathcal{I}})^{-1} = \mathbf{I} + (1+\eta)\mathbf{A}_{\mathcal{I}} + (1+\eta)^2\mathbf{A}^2_{\mathcal{I}} + \cdots$. Then, the resulting shocks are transferred to bank i through $\mathbf{a}^i_{\mathcal{I}}$. Note that the banks in the set $\mathcal{I}^c$ do not affect bank i. Accordingly, the vector $\boldsymbol{\zeta}$ in (5) captures the banks involved in shock propagation to bank i and the extent of the propagated shocks, and thus, it can be interpreted as the impact of other banks' excess shocks on bank i.[7]

Before concluding this section, we highlight that for each shock realization $\mathbf{x}$, the condition that no bank in $\mathcal{T}$ defaults is given by $\Phi_i(\mathbf{x}) \le w_i$ for each $i \in \mathcal{T}$, where $\Phi_i(\mathbf{x})$ can be rewritten as

$$\Phi_i(\mathbf{x}) = x_i + \max\left\{ \sum_{j\in\mathcal{T}^c} \zeta_j(x_j - w_j) \,\middle|\, 0 \le \zeta_j \le (1+\eta)\left(a_{ji} + \sum_{k\in\mathcal{T}^c} a_{jk}\zeta_k\right) \ \forall j \in \mathcal{T}^c \right\}, \tag{7}$$

and for each $j \in \mathcal{T}^c$, the optimal solution ζ_j^* of (7) coincides with that of (5). Equation (7) holds because bank i in $\mathcal{T}$ is not affected by the other banks in $\mathcal{T}$ (i.e., $\zeta_j^* = 0$ for all $j \in \mathcal{T} \setminus \{i\}$ due to their solvency) and the vectors in Q_i are the extreme points of the feasible set of the above linear program. This result will be particularly useful for characterizing the worst-case total shock to multiple target banks under network information uncertainty.

4. Robust Quantification of Default Probabilities

In this section, we discuss tractable quantification of (4) under incomplete network information. Since $\mathsf{D}_i = \{\mathbf{x} \in [\mathbf{0}, \mathbf{c}] \,|\, \Phi_i(\mathbf{x}) > w_i\}$ by Section 3, for fixed sets $\mathcal{A}$ and $\mathcal{T}$, one can derive that

$$\mathsf{D}_{\mathcal{A},\mathcal{T}} = \left\{ \mathbf{x} \in [\mathbf{0}, \mathbf{c}] \,\middle|\, \max_{i\in\mathcal{T}} \left\{ \overline{\Phi}_i(\mathbf{x}) - w_i \right\} > 0 \right\},$$

where $\overline{\Phi}_i(\mathbf{x}) := \max_{\mathbf{A}\in\mathcal{A}} \Phi_i(\mathbf{x})$ means the worst-case total shock to bank i for each shock realization $\mathbf{x}$. A naive Monte Carlo estimation of the worst-case default probability (4) requires solving a bilevel optimization problem for each shock realization $\mathbf{x}$: the inner layer solves (7) to find $\Phi_i(\mathbf{x})$ given $\mathbf{A}$, and the outer layer maximizes $\Phi_i(\mathbf{x})$ over the set $\mathcal{A}$. Such an optimization problem is, however, difficult to solve in general. Hence, we first propose a tractable formulation of $\max_{\mathbf{A}\in\mathcal{A}} \Phi_i(\mathbf{x})$ that facilitates the computation of (4) and then apply it to various examples of partial information.

4.1. Main Result

We assume that total liabilities $\bar{p}_1, \ldots, \bar{p}_n$ and external liabilities $b_1, \ldots, b_n$ are given since their information is often available to the public in practice, which implies that $\beta_j = (\bar{p}_j - b_j)/\bar{p}_j$ is known for all j. Also, individual interbank transactions are assumed to be partially (or not) observable. In particular, the amount of the liabilities $\bar{p}_{jk}$ of bank j to bank k may be exactly known for some j, k. In some other cases, their positive lower bounds may be known instead of exact amounts; for

example, if the liabilities between banks j and k are only known to exist, they would be at least as large as a very small amount, e.g., a dollar.

Let $\mathcal{K}$ denote the set of indices (j,k) such that the amount of $\bar{p}_{jk}$ is exactly known. Similarly, we define $\widetilde{\mathcal{K}}$ as the set of indices (j,k) such that only a positive lower bound of $\bar{p}_{jk}$ is known. Then, considering all aforementioned cases, we construct the following network uncertainty set:

$$\mathcal{A} = \left\{ \widetilde{\mathbf{A}} = (\tilde{a}_{jk}) \in \mathbb{R}_+^{n\times n} \;\middle|\; \sum_{k=1}^{n} \tilde{a}_{jk} = \beta_j \;\forall j,\; \tilde{a}_{jk} = a_{jk} \;\forall (j,k) \in \mathcal{K},\; \tilde{a}_{jk} \geq a_{jk} \;\forall (j,k) \in \widetilde{\mathcal{K}} \right\}, \tag{8}$$

where $\{a_{jk}\}_{(j,k)\in\mathcal{K}}$ are the known relative liabilities, and $\{a_{jk}\}_{(j,k)\in\widetilde{\mathcal{K}}}$ are positive constants representing the known lower bounds of the relative liabilities corresponding to $\widetilde{\mathcal{K}}$. As we shall see in Section 4.2, given the available network information, $\mathcal{K}$, $\widetilde{\mathcal{K}}$, and $\{a_{jk}\}_{(j,k)\in\mathcal{K}\cup\widetilde{\mathcal{K}}}$ can be properly determined so that the network uncertainty set $\mathcal{A}$ in (8) represents the set of all possible network configurations under that network information.

To facilitate discussions, we classify banks in $\mathcal{T}^{\mathrm{c}}$ according to the availability of their network information. For fixed $i \in \mathcal{T}$, we denote by $\mathcal{G}_1^i := \{j \in \mathcal{T}^{\mathrm{c}} | (j,k) \in \mathcal{K} \;\forall k \in \{i\} \cup \mathcal{T}^{\mathrm{c}}\}$ the set of banks whose liabilities to bank i and banks in $\mathcal{T}^{\mathrm{c}}$ are *all* known, by $\mathcal{G}_2^i := \{j \in \mathcal{T}^{\mathrm{c}} | (j,i) \notin \mathcal{K}\}$ the set of banks whose liabilities to bank i are *not* known, and by $\mathcal{G}_3^i := \mathcal{T}^{\mathrm{c}} \setminus (\mathcal{G}_1^i \cup \mathcal{G}_2^i)$ the set of banks whose liabilities to bank i are known but whose liabilities to some banks in $\mathcal{T}^{\mathrm{c}}$ are unknown. Clearly, the three sets are disjoint, there is no priority among them, and $\mathcal{T}^{\mathrm{c}} = \mathcal{G}_1^i \cup \mathcal{G}_2^i \cup \mathcal{G}_3^i$. Based on these sets, we establish our main result on $\overline{\Phi}_i(\mathbf{x})$ by applying the worst-case scenario to the potential shock propagation from bank j in $\mathcal{T}^{\mathrm{c}}$ to bank i, i.e., the right-hand side of the constraints in the linear program (7), considering the network uncertainty set $\mathcal{A}$ of relative liabilities.

Recall that ζ_j in (5) and (7) measures the impact of bank j's excess shocks on bank i. Observe that for each $j \in \mathcal{G}_1^i$, there is no uncertainty in its liabilities a_{ji} and $\{a_{jk}\}_{k\in\mathcal{T}^{\mathrm{c}}}$ (see Figure 3-(a) in a four-bank system with $\mathcal{T} = \{4\}$ and $\mathcal{G}_1^4 = \{1\}$). Thus, the corresponding constraints in (7) remain unchanged as follows:

$$0 \leq \zeta_j \leq (1+\eta)\left(a_{ji} + \sum_{k\in\mathcal{T}^{\mathrm{c}}} a_{jk}\zeta_k\right) \text{ for } j \in \mathcal{G}_1^i. \tag{9}$$

By contrast, if the direct link from bank j to bank i is not known (i.e., $j \in \mathcal{G}_2^i$), one may intuit that the worst-possible network would concentrate all unknown liabilities of bank j on that direct link; Figure 3-(c) shows the worst-case allocation of bank 2's liabilities when we consider a four-bank system with $\mathcal{T} = \{4\}$ and $\mathcal{G}_2^4 = \{2\}$ (Figure 3-(b)). Let $\tilde{\beta}_j := \beta_j - \sum_{\{k|(j,k)\in\mathcal{K}\cup\widetilde{\mathcal{K}}\}} a_{jk}$, which represents the sum of unknown relative liabilities of bank j. Then, the above intuition implies that the worst-case version of the constraints in (7) could be written as

$$0 \le \zeta_j \le (1+\eta)\left(\tilde{\beta}_j + \sum_{k\in\mathcal{K}_j\cup\widetilde{\mathcal{K}}_j} a_{jk}\zeta_k\right) \text{ for } j \in \mathcal{G}_2^i, \tag{10}$$

where $\mathcal{K}_j := \{k \in \mathcal{T}^c | (j,k) \in \mathcal{K}\}$ and $\widetilde{\mathcal{K}}_j := \{k \in \mathcal{T}^c | (j,k) \in \widetilde{\mathcal{K}}\}$.

On the other hand, if $j \in \mathcal{G}_3^i$, since the direct link a_{ji} is known, the worst-possible network would allocate all unknown liabilities of bank j to the link to a bank in $\mathcal{T}^c \setminus \mathcal{K}_j$ that has the greatest impact on bank i, i.e., $\arg\max_{l\in\mathcal{T}^c\setminus\mathcal{K}_j} \zeta_l$. For example, in a four-bank system with $\mathcal{T} = \{4\}$ and $\mathcal{G}_3^4 = \{3\}$ (see Figure 3-(d)), bank 3's total unknown liabilities are allocated to its link to bank 1 if the impact of bank 1 on bank 4 is greater than that of bank 2 on bank 4, i.e., $\zeta_1 > \zeta_2$ (see Figure 3-(e)); otherwise, they are allocated to its link to bank 2 (see Figure 3-(f)). Hence, the associated worst-case constraints may correspond to

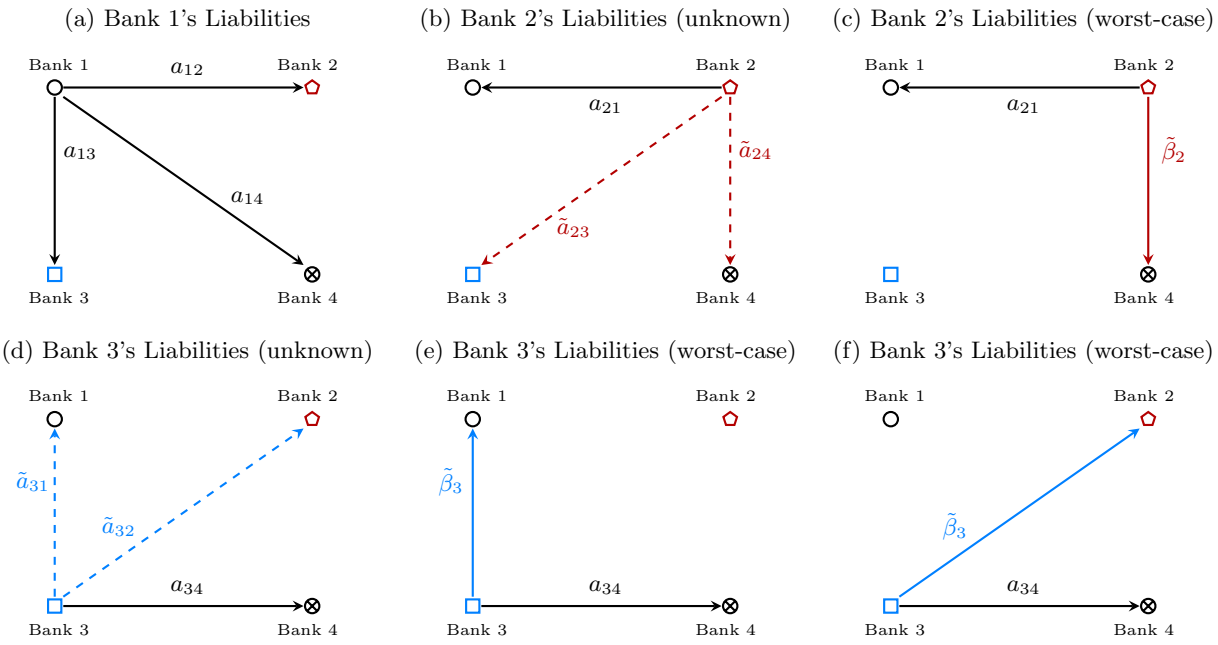
$$0 \le \zeta_j \le (1+\eta)\left(a_{ji} + \sum_{k\in\mathcal{K}_j\cup\widetilde{\mathcal{K}}_j} a_{jk}\zeta_k + \tilde{\beta}_j\left(\max_{l\in\mathcal{T}^c\setminus\mathcal{K}_j}\zeta_l\right)\right) \text{ for } j \in \mathcal{G}_3^i. \tag{11}$$

The right-hand sides in (10) and (11) represent the worst-possible shock propagation from bank j to bank i. It is worth noting that banks in $\mathcal{T}$ are minimally liable since no unknown liabilities are allocated to the links between them.

These considerations of worst-possible network structures help develop the tractable formulation of $\overline{\Phi}_i(\mathbf{x})$ by maximizing the total shock to bank i, i.e., $x_i + \sum_{j\in\mathcal{T}^c}\zeta_j(x_j - w_j)$, subject to the above constraints on $\{\zeta_j\}_{j\in\mathcal{T}^c}$, which is formalized in the following theorem.

THEOREM 1. *Assume that every bank in $\mathcal{T}$ is solvent and the network uncertainty set $\mathcal{A}$ satisfies* (8)*. Then, for all $\mathbf{x} \in [\mathbf{0}, \mathbf{c}]$ and for each $i \in \mathcal{T}$, $\overline{\Phi}_i(\mathbf{x})$ is the maximum of $x_i + \sum_{j\in\mathcal{T}^c}\zeta_j(x_j - w_j)$*

Figure 3 An illustration of worst-possible networks with a four-bank system when $\mathcal{T} = \{4\}$.



Note. We assume that all relative liabilities are known except a_{23}, a_{24}, a_{31}, and a_{32}. Black solid arrows represent known liabilities (a_{ij}), and dashed arrows denote unknown liabilities ($\tilde{a}_{ij}$). Red and blue solid arrows are the liabilities of banks 2 and 3, respectively, in the worst-possible network.

among all impacts $\{\zeta_j\}_{j\in\mathcal{T}^c}$ *of banks in* $\mathcal{T}^c$ *on bank* i *satisfying the worst-case shock propagation constraints* (9), (10), *and* (11). *Further, the nonlinear constraints* (11) *can be written as the following equivalent linear constraints:*

$$\begin{cases} 0 \le \zeta_j \le (1+\eta)\left(a_{ji} + \sum_{k\in\mathcal{K}_j\cup\tilde{\mathcal{K}}_j} a_{jk}\zeta_k + \tilde{\beta}_j\zeta_l + 1 - z_{jl}\right) & \text{for } j \in \mathcal{G}_3^i,\ l \in \mathcal{T}^c \setminus \mathcal{K}_j; \\ \sum_{l\in\mathcal{T}^c\setminus\mathcal{K}_j} z_{jl} = 1 & \text{for } j \in \mathcal{G}_3^i; \\ z_{jl} \in \{0,1\} & \text{for } j \in \mathcal{G}_3^i,\ l \in \mathcal{T}^c \setminus \mathcal{K}_j, \end{cases} \tag{12}$$

where for each $j \in \mathcal{T}^c$, *the binary variable* z_{jl} *determines whether bank* l *has the greatest impact on bank* i *among banks in* $\mathcal{T}^c \setminus \mathcal{K}_j$.

The maximization problem in Theorem 1, constructed by replacing (11) with (12), is a mixed-integer linear program (MILP) that is much easier to solve than the original formulation $\max_{\mathbf{A}\in\mathcal{A}} \Phi_i(\mathbf{x})$. Thus, it greatly facilitates the computation of $\mathsf{P}(\mathbf{X} \in \mathsf{D}_{\mathcal{A}})$. Specifically, the Monte Carlo estimate for $\mathsf{P}(\mathbf{X} \in \mathsf{D}_{\mathcal{A}})$ can be computed by solving the MILP for each $i \in \mathcal{T}$ and checking

if $\max_{i\in\mathcal{T}}\{\overline{\Phi}_i(\mathbf{x})-w_i\}>0$. While solving the MILP $|\mathcal{T}|$ times may seem time-consuming, this is not a critical issue since the set $\mathcal{T}$ would often be a collection of SIFIs and the number of SIFIs is limited in practice. Also, we can relieve the computational burden by not solving the MILP for some shock realizations. Since $\max_{i\in\mathcal{T}}\{x_i-w_i\}>0$ implies $\max_{i\in\mathcal{T}}\{\overline{\Phi}_i(\mathbf{x})-w_i\}>0$, it is enough to solve the MILP only when $x_i\le w_i$ for all $i\in\mathcal{T}$. Later, we will see some cases that further improve the time-efficiency of the algorithm by transforming it into linear programs, or by obtaining $\overline{\Phi}_i(\mathbf{x})$ even without solving an optimization problem.

Remark 2. For the robust quantification of (3), one may consider two alternatives. The first method is to use the robust counterpart of the linear programming version of (2): max $\mathbf{1}^\top\mathbf{p}$ s.t. $(\mathbf{I}-(1+\eta)\mathbf{A}^\top)\mathbf{p}\le(1+\eta)(\mathbf{c}-\mathbf{x})-\eta\bar{\mathbf{p}}$, $\mathbf{0}\le\mathbf{p}\le\bar{\mathbf{p}}$, $\forall\mathbf{A}\in\mathcal{A}$, where the result of Soyster (1976) can be applied. However, the solution turns out to be more conservative than ours. The second approach is to identify the worst possible network configuration in $\mathcal{A}$ that yields the largest default probability, i.e., $\max_{\mathbf{A}\in\mathcal{A}}\mathsf{P}\left(\mathbf{X}\in\bigcup_{i\in\mathcal{T}}\mathsf{D}_i\right)$, using sample average approximation. However, this requires solving a large-scale bilevel nonlinear optimization problem that is computationally intractable. In contrast, our method relies on solving a single-level MILP in Theorem 1, which is smaller in size and easier to solve. A numerical performance comparison is presented in Appendix A.

Remark 3. Given the network uncertainty set $\mathcal{A}$, one might alternatively consider the worst-case "systemic default" probability (WSDP). While there is no formal definition of systemic default, it is often understood as the default of all banks in the system (Tasca et al. 2014, Battiston et al. 2016, Roukny et al. 2018). In this case, the WSDP can be easily characterized under our model as long as $\mathbf{X}$ is a continuous random vector. Firstly, a simple calculation shows that for each $\mathbf{x}$ and $\mathbf{A}$, systemic default occurs if and only if $\mathbf{s}(\mathbf{x};\mathbf{A})>\mathbf{0}$, where $\mathbf{s}(\mathbf{x};\mathbf{A}):=(1+\eta)(\mathbf{I}-(1+\eta)\mathbf{A}^\top)^{-1}(\mathbf{x}-\mathbf{w})$ denotes the vector of each bank's payment shortfall. Hence, similar to (4), the WSDP can be defined by $\mathsf{P}(\mathbf{X}\in\widetilde{\mathsf{D}}_{\mathcal{A}})$, where $\widetilde{\mathsf{D}}_{\mathcal{A}}:=\bigcup_{\mathbf{A}\in\mathcal{A}}\{\mathbf{x}\,|\,\mathbf{s}(\mathbf{x};\mathbf{A})>\mathbf{0}\}$ is the set of shock vectors that could cause all banks to default in the worst case. Then, by Farkas' lemma, it is not difficult to show that $\mathsf{P}(\mathbf{X}\in\widetilde{\mathsf{D}}_{\mathcal{A}})=\mathsf{P}(\Psi(\mathbf{X})=0)$, where $\Psi(\mathbf{x}):=\max\{(\mathbf{w}-\mathbf{x})^\top\boldsymbol{\xi}\,|\,(\mathbf{I}-(1+\eta)\mathbf{A})\boldsymbol{\xi}\ge\mathbf{0}\text{ for all }\mathbf{A}\in\mathcal{A}\}$. Note

that the optimization problem for $\Psi(\mathbf{x})$ can be simply converted into a tractable linear program via a standard approach to robust linear optimization with polyhedral uncertainty (Bertsimas et al. 2011). Accordingly, the WSDP can be estimated in a manner similar to the estimation of $\mathsf{P}(\mathbf{X} \in \mathsf{D}_{\mathcal{A}})$ discussed above. See Section EC.1.3 of the online supplement for technical details.

4.2. Applications to Various Network Information

This subsection provides several examples of partial network information to illustrate the usefulness of the MILP formulation in Theorem 1 from the perspectives of individual banks and regulators.

Individual bank's information. Suppose that bank 1 wants to estimate the worst-case default probability (4) of its counterparty bank n, i.e., $\mathsf{P}(\mathbf{X} \in \mathsf{D}_{\mathcal{A},\{n\}}) = \mathsf{P}(\overline{\Phi}_n(\mathbf{X}) > w_n)$, using all the information bank 1 has, i.e., $\mathcal{K} = \bigcup_{j=1}^{n}\{(1,j),(j,1),(j,j)\}$ and $\widetilde{\mathcal{K}} = \emptyset$. In practice, bank 1 would compute such a probability conditional on itself remaining solvent prior to bank n's bankruptcy. In this case, solving the MILP results in

$$\overline{\Phi}_n(\mathbf{x}) = x_n + (1+\eta)\sum_{j=2}^{n-1}(\beta_j - a_{j1})(x_j - w_j)^+ \tag{13}$$

since $\mathcal{G}_1^n = \mathcal{G}_3^n = \emptyset$ and $\mathcal{G}_2^n = \{2,\ldots,n-1\}$. This corresponds to the fact that, from bank 1's perspective, the worst-case shock propagation to bank n occurs when all unknown liabilities of banks 2 to $n-1$ are associated with bank n. In this worst-case network structure, the condition that bank 1 is solvent before bank n's default is given by $x_1 + (1+\eta)\sum_{j=2}^{n-1} a_{j1}(x_j - w_j)^+ \le w_1$, and hence, we can estimate the said probability by sampling $\mathbf{X}$ from the distribution satisfying this condition. Note that the increase or decrease in the amount of bank 1's lending to bank n does not affect (4) since $\overline{\Phi}_n(\mathbf{x})$ and w_n remain unchanged irrespective of these actions. However, if bank 1 allows bank n to roll over the loan, the probability will be reduced because it has the same effect as temporarily increasing w_n while keeping $\overline{\Phi}_n(\mathbf{x})$ unchanged.

Multiple banks' information. Suppose that regulators want to estimate banks' default probabilities but have limited information obtained only from some of the banks. Let $\mathcal{S}$ denote the set of banks that provide their own information to the regulators. Then, we have $\mathcal{K} =$

$(\bigcup_{i\in\mathcal{S},j=1,\dots,n}\{(i,j),(j,i)\})\cup(\bigcup_{j\in\mathcal{S}^{\rm c}}\{(j,j)\})$ and $\widetilde{\mathcal{K}}=\emptyset$. Given this situation, the MILP in Theorem 1 can be used for any $\mathcal{S},\mathcal{T}\subset\{1,\dots,n\}$. Below we present two particular cases of $\mathcal{S}$ and $\mathcal{T}$ that would be of interest to regulators. We find that in those cases, the MILP can be greatly simplified.

- The regulators might want to estimate the worst-case default probability (4) of the banks whose information is not available. Thus, let us assume that the target banks' information is not available, i.e., $\mathcal{T}\subset\mathcal{S}^{\rm c}$. Then, we have $\mathcal{G}_1^i=\mathcal{S}$, $\mathcal{G}_2^i=\mathcal{T}^{\rm c}\setminus\mathcal{S}$, and $\mathcal{G}_3^i=\emptyset$ for each $i\in\mathcal{T}$, and thus, one can show that the MILP can be recast as the following linear program:

$$\max\ x_i+\mathbf{u}^\top(\mathbf{x}_{\mathcal{T}^{\rm c}}-\mathbf{w}_{\mathcal{T}^{\rm c}})\ \text{s.t.}\ \left((1+\eta)^{-1}\mathbf{I}-\widetilde{\mathbf{A}}\right)\mathbf{u}\le\mathbf{q},\ \mathbf{u}\in\mathbb{R}_+^{|\mathcal{T}^{\rm c}|},\tag{14}$$

  where

$$\widetilde{\mathbf{A}}=\begin{bmatrix}\mathbf{A}_{\mathcal{S}} & \mathbf{A}_{\mathcal{S},\mathcal{T}^{\rm c}\setminus\mathcal{S}}\\ \mathbf{A}_{\mathcal{T}^{\rm c}\setminus\mathcal{S},\mathcal{S}} & \mathbf{0}\end{bmatrix}\ \text{and}\ \mathbf{q}=\begin{bmatrix}\mathbf{a}_{\mathcal{S}}^i\\ \boldsymbol{\beta}_{\mathcal{T}^{\rm c}\setminus\mathcal{S}}-\mathbf{A}_{\mathcal{T}^{\rm c}\setminus\mathcal{S},\mathcal{S}}\mathbf{1}\end{bmatrix}.$$

- Given incomplete network information, even if the regulators observe the target banks' information, i.e., $\mathcal{T}=\mathcal{S}$, the exact computation of their default probability (3) may not be feasible. In this case, we have $\mathcal{G}_1^i=\mathcal{G}_2^i=\emptyset$ and $\mathcal{G}_3^i=\mathcal{T}^{\rm c}$ for each $i\in\mathcal{T}$, and hence, the MILP is equivalent to the following simple MILP formulation for $i\in\mathcal{T}$:

$$\begin{aligned}\max\quad & x_i+\mathbf{u}^\top(\mathbf{x}_{\mathcal{T}^{\rm c}}-\mathbf{w}_{\mathcal{T}^{\rm c}})\\ \text{s.t.}\quad & (1+\eta)^{-1}\mathbf{u}\mathbf{1}^\top-\widetilde{\boldsymbol{\beta}}_{\mathcal{T}^{\rm c}}\mathbf{u}^\top+\mathbf{Z}\le\left(\mathbf{a}_{\mathcal{T}^{\rm c}}^i+\mathbf{1}\right)\mathbf{1}^\top,\\ & \mathbf{Z}\mathbf{1}=\mathbf{1},\ \text{diag}\mathbf{Z}=\mathbf{0},\\ & \mathbf{Z}\in\{0,1\}^{|\mathcal{T}^{\rm c}|\times|\mathcal{T}^{\rm c}|},\ \mathbf{u}\in\mathbb{R}_+^{|\mathcal{T}^{\rm c}|},\end{aligned}$$

  where we simplify the notation by adding a redundant constraint $\text{diag}\mathbf{Z}=\mathbf{0}$.

Large exposures' information. As discussed in the introduction, according to the Basel III framework (BCBS 2020a), a bank's large exposures greater than or equal to 10% of the bank's capital are likely to be reported to regulators. We can deal with such a case in the MILP formulation by letting $\mathcal{K}$ be the set of interbank transactions corresponding to large exposures and $\widetilde{\mathcal{K}}$ be an empty set. We will see in Sections 6 and EC.4 that this information contains significant interbank

liabilities that could largely affect network stability in practice, and thus, the worst-case default probabilities under this information are likely to be close to the true values.

Link information. Recall that link information refers to the information on whether each interbank link exists or not. We assume that such link information is available but the exact amount of interbank liabilities is not observed. Let $\mathcal{K}$ and $\widetilde{\mathcal{K}} = \mathcal{K}^{\mathrm{c}}$ denote the set of non-existing links and the set of existing links, respectively. Then, for some $\epsilon > 0$, the uncertainty set $\mathcal{A}$ can be given by

$$\mathcal{A} = \left\{ \widetilde{\mathbf{A}} \in \mathbb{R}_+^{n\times n} \;\middle|\; \sum_{k=1}^{n} \tilde{a}_{jk} = \beta_j \text{ for all } j,\ \tilde{a}_{jk} = 0 \text{ for all } (j,k) \in \mathcal{K},\ \tilde{a}_{jk} \geq \epsilon \text{ for all } (j,k) \notin \mathcal{K} \right\}. \quad (15)$$

Clearly, the uncertainty set $\mathcal{A}$ in (15) gets smaller as the set $\mathcal{K}$ becomes larger. This implies that the lower the network density, the closer the worst-case probability (4) under the link information is to the true probability (3). Note that the corresponding MILP problem is not as simple as the equivalent formulations in the previous cases, but it is still easy to formulate and solve.

In the following theorems, we further observe the closeness between the probabilities (3) and (4) given the link information (15) in the presence of regularly varying shocks and lognormal shocks.[8]

THEOREM 2. *Suppose that $\mathcal{A}$ satisfies (15) for fixed $\mathcal{K}$ and $\epsilon > 0$ and that X_i be a nonnegative random variable having a regularly varying distribution with index $\rho_i > 1$ for each i, i.e., for all $t > 0$, $\lim_{x\to\infty} f_i(tx)/f_i(x) = t^{-\rho_i}$ where $f_i(\cdot)$ is the density function of X_i. Let $\mathbf{X}^m = \mathbf{X}/m$ for $m = 1, 2, \ldots$, and $\rho_* = \min_{i\in\mathcal{H}\cup\mathcal{T}} \rho_i$, where $\mathcal{H}$ is the set of banks that have a directed path to a bank in $\mathcal{T}$. Assume that for large m, $X_1^m, \ldots, X_n^m$ are independent and constrained to be in $[\mathbf{0}, \mathbf{c}]$ almost surely. Then, for any $\mathbf{A} \in \mathcal{A}$,*

$$\lim_{m\to\infty} \frac{1}{\log m} \log \mathsf{P}\left(\mathbf{X}^m \in \mathsf{D}_{\mathcal{A}}\right) = \lim_{m\to\infty} \frac{1}{\log m} \log \mathsf{P}\left(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) = -\rho_* + 1. \quad (16)$$

THEOREM 3. *Suppose that $\mathcal{A}$ satisfies (15) for fixed $\mathcal{K}$ and $\epsilon > 0$ and that X_i follow a lognormal distribution with parameters μ_i and σ_i for each i. Let $\mathbf{X}^m = \mathbf{X}/m$ for $m = 1, 2, \ldots$, and $\sigma_* = \max_{i\in\mathcal{H}\cup\mathcal{T}} \sigma_i$, where the set $\mathcal{H}$ is defined as in Theorem 2. Assume that for large m, $X_1^m, \ldots, X_n^m$ are independent and constrained to be in $[\mathbf{0}, \mathbf{c}]$ almost surely. Then, for any $\mathbf{A} \in \mathcal{A}$,*

$$\lim_{m\to\infty} \frac{1}{(\log m)^2} \log \mathsf{P}\left(\mathbf{X}^m \in \mathsf{D}_{\mathcal{A}}\right) = \lim_{m\to\infty} \frac{1}{(\log m)^2} \log \mathsf{P}\left(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) = -\frac{1}{2\sigma_*^2}.$$

The above theorems commonly indicate that under the link information (15), the probabilities (3) and (4) are asymptotically equivalent as the shock size gets smaller in the sense that for any $\mathbf{A} \in \mathcal{A}$,

$$\lim_{m\to\infty} \frac{\log \mathsf{P}\left(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right)}{\log \mathsf{P}\left(\mathbf{X}^m \in \mathsf{D}_{\mathcal{A}}\right)} = 1.$$

This not only highlights the importance of the link information but also allows us to use (4) as a proxy of (3) when it comes to small but heavy-tailed shocks. The set $\mathcal{H}$ can be identified using the link information (15) and represents the set of banks that may affect the banks in $\mathcal{T}$.[9] Also, $(-\rho_i)$ and σ_i indicate the heavy-tailedness of the shock X_i. Thus, among the shocks that could affect the banks in $\mathcal{T}$, the one with the heaviest tail has the most powerful influence on the probabilities (3) and (4). See Section EC.2 of the online supplement for more discussions on the two theorems.

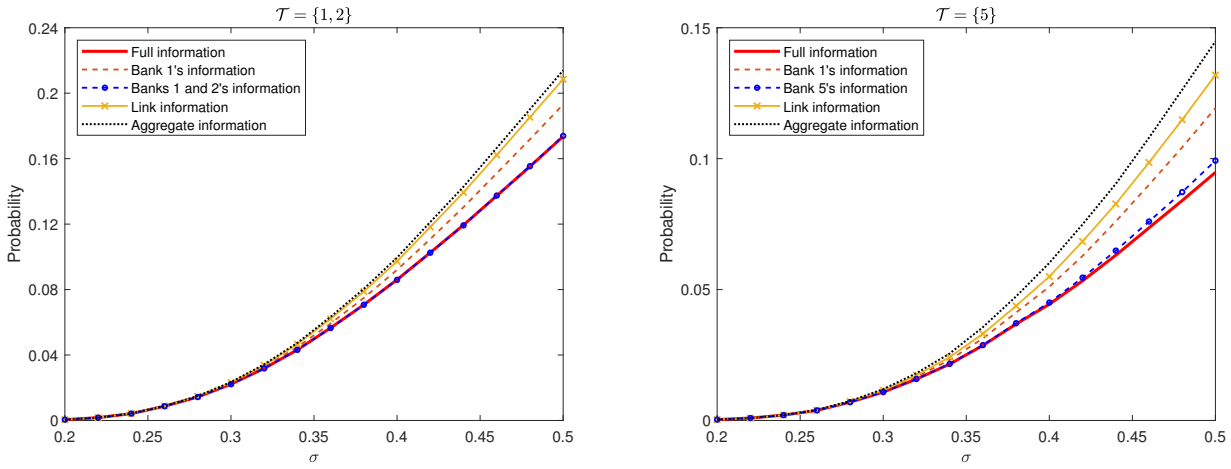
Aggregate information. Suppose that individual interbank liabilities are not observed at all and only the aggregate information $(\beta_1, \ldots, \beta_n)$ is available, which is the case in Glasserman and Young (2015, 2016). Then, $\mathcal{K} = \{(1,1),(2,2),\ldots,(n,n)\}$ and $\widetilde{\mathcal{K}} = \emptyset$. In this case, it is easy to see that $\overline{\Phi}_i(\mathbf{x}) = x_i + (1+\eta)\boldsymbol{\beta}_{\mathcal{T}^c}^\top(\mathbf{x}_{\mathcal{T}^c} - \mathbf{w}_{\mathcal{T}^c})^+$ for $i \in \mathcal{T}$, and the worst-case default probability becomes

$$\mathsf{P}(\mathbf{X} \in \mathsf{D}_{\mathcal{A}}) = \mathsf{P}\left(\max_{i\in\mathcal{T}} \left\{x_i - w_i + (1+\eta)\boldsymbol{\beta}_{\mathcal{T}^c}^\top(\mathbf{x}_{\mathcal{T}^c} - \mathbf{w}_{\mathcal{T}^c})^+\right\} > 0\right), \tag{17}$$

Thus, we do not solve an optimization problem for estimating this probability, which helps us reduce the computation time greatly. Note that (17) is analogous to the result of Glasserman and Young (2015); in fact, both coincide given a single shock and a single target bank. However, we consider a multivariate shock vector $\mathbf{X}$ and the target event $\bigcup_{i\in\mathcal{T}} \mathsf{D}_i$ in (17), whereas their result is based on a single shock to a specific bank and the target event $\bigcap_{i\in\mathcal{T}} \mathsf{D}_i$.

Comparison of network information effects. We revisit the underlying network in Example 1 to illustrate how the worst-case default probability (4) changes according to the network information. In Figure 4, we present the estimated values of (4) with the target sets $\{1,2\}$ and $\{5\}$ under five different types of network information: full network information, bank 1's information, target banks' information, link information, and aggregate information. The same shock distribution in Example 1 is assumed. For the link information, we arbitrarily set $\epsilon = 0.05$. Note that all

Figure 4 The worst-case default probabilities under different types of network information.



Note. The figure exhibits the worst-case default probabilities (4) for $\mathcal{T} = \{1, 2\}$ (left) and $\mathcal{T} = \{5\}$ (right) with different σ under different types of network information. The Monte Carlo method with 10^5 replications is used for the estimation of the probabilities.

interbank liabilities in this example are large exposures, equal to 80% of banks' capital, meaning that the large exposures' information is exactly the same as the full information.

In the left panel of Figure 4, we observe that the estimate of (4) given banks 1 and 2's information is closer to the true value (3) than the estimate of (4) given bank 1's information. This corresponds to our intuition that more information would lead to a better estimate. Also, in both panels, the information of target banks in $\mathcal{T}$ gives the best approximation, implying that the firsthand shock propagation is likely to dominate the whole shock propagation in financial networks. We shall see in Section 6 that these observations are consistent even in the case of real-world financial networks.

The magnitude of the differences between the true and worst-case default probabilities (up to 50% in this example) greatly depends on several inputs, including partial network information, shock distributions, bankruptcy costs, the size of interbank liabilities, and the number of target banks. However, no matter how large or small the gap is, our methodology allows us to compare the effects of different types of network information on the proximity between the true and worst-case default probabilities, which eventually gives us practical insights into information collection.

5. Worst-Case Risk Capital

The higher loss absorbency requirements in the Basel III framework ask global systemically important banks (G-SIBs) to hold additional capital to reduce their default probabilities. The task of gauging this risk capital poses a challenge when full network information is not available. In this section, we apply the main results in Section 4 so as to estimate the minimum additional capital amount, which we call *worst-case risk capital*, required to keep all SIFIs solvent with probability α.

Let $\mathcal{T}$ be the set of SIFIs, which is assumed to be given, and fix $\alpha \in [0,1)$. Then, using the probability (4), our problem can be mathematically formulated as follows:

$$\min \quad \sum_{i\in\mathcal{T}} \nu_i \quad \text{s.t.} \quad \mathsf{P}\left(\max_{i\in\mathcal{T}} \left\{\overline{\Phi}_i(\mathbf{X}) - w_i - \nu_i\right\} > 0\right) \le 1-\alpha, \ \nu_i \ge 0, \ i \in \mathcal{T}. \tag{18}$$

Note that the left-hand side of the first constraint of (18) is the probability (4) with $w_i + \nu_i$ instead of w_i for $i \in \mathcal{T}$. The optimal solution to (18), say $\boldsymbol{\nu}^\alpha$, limits the probability (4) to $1-\alpha$, and the total amount $\nu_\circ^\alpha := \sum_{i\in\mathcal{T}} \nu_i^\alpha$ is the worst-case risk capital. However, (18) is a joint chance-constrained optimization problem that is known to be difficult to solve; since the probability does not have a closed form in general, it is hard to check the feasibility of the probabilistic constraint, and even if it is feasible, the feasible region might not be convex.

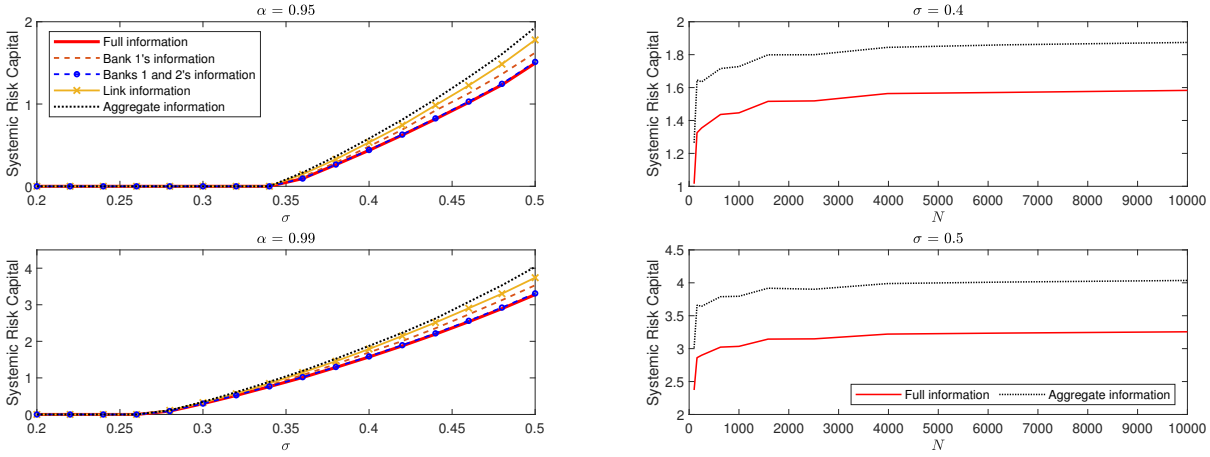
To tackle this issue, in the following theorem, we propose a sample average approximation (SAA) of the problem (18) assuming that the random shock vector $\mathbf{X}$ can be sampled.

THEOREM 4. *Assume that $\mathbf{X}$ is a continuous random shock vector and $\mathbf{x}^1, \ldots, \mathbf{x}^N$ are its independent and identically distributed samples. Let V be the set of optimal solutions of (18) and M be a sufficiently large constant such that $\overline{\Phi}_i(\mathbf{X}) < M$ almost surely for all $i \in \mathcal{T}$. If $(\boldsymbol{\nu}^N, \mathbf{z}^N)$ solves the following MILP:*

$$\begin{aligned}
\min \quad & \sum_{i\in\mathcal{T}} \nu_i \\
\text{s.t.} \quad & \sum_{j=1}^{N} z_j \le N(1-\alpha), \\
& w_i + \nu_i + M z_j \ge \overline{\Phi}_i(\mathbf{x}^j), \quad i \in \mathcal{T}, \ j = 1, \ldots, N, \\
& 0 \le \nu_i \le M, \quad i \in \mathcal{T}, \\
& z_j \in \{0,1\}, \quad j = 1, \ldots, N,
\end{aligned} \tag{19}$$

then $\nu_\circ^N := \sum_{i\in\mathcal{T}} \nu_i^N \to \nu_\circ^\alpha$ and $\inf_{\boldsymbol{\nu}^\alpha \in \mathsf{V}} \|\boldsymbol{\nu}^N - \boldsymbol{\nu}^\alpha\| \to 0$ with probability 1 as $N \to \infty$.

Figure 5 The worst-case risk capital under different types of network information.



Note. The left panel exhibits the worst-case risk capital $\nu_\circ^N$ for $\mathcal{T}=\{1,2\}$ with $N=10,000$, $\alpha=0.95,0.99$, and different σ under five types of network information in Figure 4. The right panel shows the worst-case risk capital $\nu_\circ^N$ for $\mathcal{T}=\{1,2\}$ with $\sigma=0.4,0.5$, $\alpha=0.99$, and different N under full information and aggregate information.

Theorem 4 provides significant computational benefits because the new problem (19) is much easier to solve than the original problem (18), and the approximation quality improves as the sample size increases. Furthermore, the random vector $\mathbf{X}$ can be a general continuous probability distribution, and hence, shocks to banks are possibly correlated.

Ahn and Kim (2019) and Demange (2018) study tractable mathematical programs for the optimal capital injection under the Eisenberg-Noe framework. However, those approaches essentially differ from ours. Firstly, their policies are scenario-based solutions and do not adopt any probabilistic approach, i.e., each deterministic shock scenario may lead to a different capital allocation. Secondly, they do not take incomplete network information and SIFIs' default into special consideration.

In Figure 5, using Example 1 again, we describe the worst-case risk capital under different types of network information, different σ, and different N, for the set $\mathcal{T}=\{1,2\}$. In the left panel of the figure, we observe the impact of network information on estimating the risk capital, which is similar to the results in Figure 4. The right panel, on the other hand, numerically shows that the risk capital is underestimated when the sample size is small, but converges rapidly as the sample size increases. In general, the small sample size would lead to the underestimation of the (worst-case) default probability and risk capital because default events are rare.

REMARK 4. One may also consider convex approximation (Hong et al. 2011), robust approximation (Yuan et al. 2017), and scenario approaches (Calafiore and Campi 2005, 2006) to overcome the intractability of the problem (18). However, those methods either require more distributional assumptions or lead to conservative bounds only. As long as we can sample the random vector, it is relatively easier to apply the SAA than the other methods. See Luedtke and Ahmed (2008) and Pagnoncelli et al. (2009) for more details on the SAA for joint chance-constrained optimization.

Before we end this section, we briefly discuss the applicability of our SAA approach to the measurement of external creditors' risk exposure to the financial system. If we regard the set of external creditors as node 0 and denote the associated relative liabilities of bank i by a_{i0} for each i, then a slight modification to Lemma 1 suffices to show that the creditors' worst-case losses are equal to $\overline{\Phi}_0(\mathbf{x}) \coloneqq \max_{\mathbf{A}\in\mathcal{A}} \Phi_0(\mathbf{x})$ where $\Phi_0(\mathbf{x}) \coloneqq \max\left\{\boldsymbol{\zeta}^\top(\mathbf{x}-\mathbf{w}) \,|\, (\mathbf{I}-(1+\eta)\mathbf{A})\boldsymbol{\zeta} \le (1+\eta)\mathbf{a}^0,\, \boldsymbol{\zeta}\in\mathbb{R}^n_+\right\}$ and $\mathbf{a}^0 = (a_{10}, a_{20}, \ldots, a_{n0})^\top$, which can be computed using an MILP similar to the one in Theorem 1. Thus, the associated value-at-risk, $\mathrm{VaR}_\alpha\big(\overline{\Phi}_0(\mathbf{X})\big) = \min\left\{y\in\mathbb{R}_+ \,|\, \mathsf{P}\left(\overline{\Phi}_0(\mathbf{X}) > y\right) \le 1-\alpha\right\}$, can be estimated via Theorem 4. According to Rockafellar and Uryasev (2000), given the samples $\overline{\Phi}_0(\mathbf{x}^1), \ldots, \overline{\Phi}_0(\mathbf{x}^N)$, the conditional value-at-risk of the worst-case losses can be approximated as $\mathrm{CVaR}_\alpha(\overline{\Phi}_0(\mathbf{X})) \approx \min\{y\in\mathbb{R}_+ \,|\, \overline{\Phi}_0(\mathbf{x}^j) + \{N(1-\alpha)\}^{-1}\sum_{j=1}^N z_j \le z_j + y,\, z_j \ge 0,\, j=1,\ldots,N\}$.

6. Numerical Experiments

To demonstrate how our theoretical results in Sections 4 and 5 can be applied in practice, we use data from the 2011 EU-wide stress test conducted by the European Banking Authority (EBA). For illustration, we restrict our focus on 11 German banks that participated in the test as in Chen et al. (2016) and Gandy and Veraart (2017). Our numerical results based on the full dataset can be found in Section EC.4 of the online supplement. Note that both datasets lead to the same conclusions.

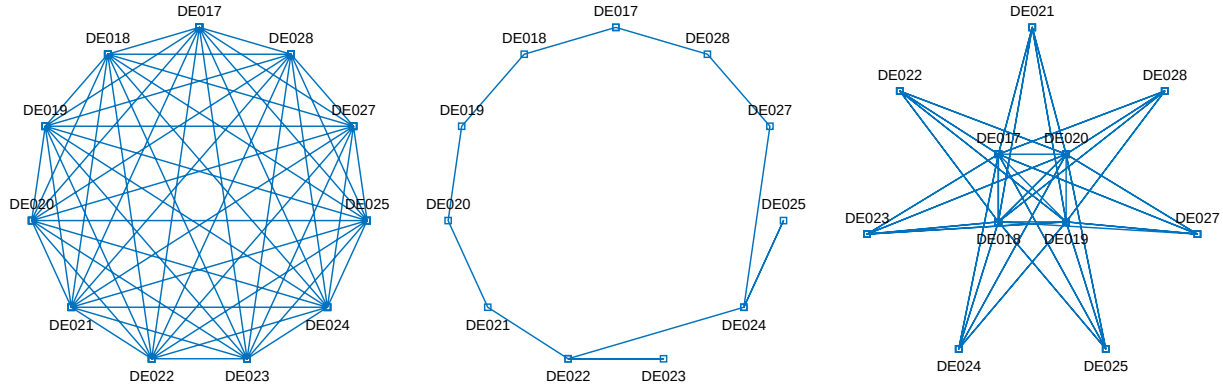
The dataset of the German banks is provided in Table 2. We use the numbers in the first column of Table 2 for the subscripts of variables; e.g., DE017's net worth is denoted by $w_1 = 30,361$. Since this dataset does not report each bank's bilateral interbank exposures, we adopt the three types of reconstructed networks used in Chen et al. (2016): complete, ring-like, and core-periphery networks

Table 2 Data of German banks from the 2011 EBA EU-wide stress test.

No.	Code	Bank Name	Total Assets	EAD	Equity ($\mathbf{w}$)	External Assets ($\mathbf{c}$)
1	DE017	Deutsche Bank AG	1,905,630	47,102	30,361	1,858,528
2	DE018	Commerzbank AG	771,201	49,871	26,728	721,330
3	DE019	Landesbank B-W	374,413	91,201	9,838	283,212
4	DE020	DZ Bank AG	323,578	100,099	7,299	223,479
5	DE021	Bayerische Landesbank	316,354	66,535	11,501	249,819
6	DE022	Norddeutsche Landesbank	228,586	54,921	3,974	173,665
7	DE023	Hypo Real Estate Holding AG	328,119	7,956	5,539	320,163
8	DE024	WestLB AG Dusseldorf	191,523	24,007	4,218	167,516
9	DE025	HSH Nordbank AG Hamburg	150,930	4,645	4,434	146,285
10	DE027	Landesbank Berlin AG	133,861	27,707	5,162	106,154
11	DE028	DekaBank Deutsche Girozentrale	130,304	30,937	3,359	99,367

All quantities are exhibited in millions of euros.

Figure 6 The recovered interbank network structures from the EBA stress test data.

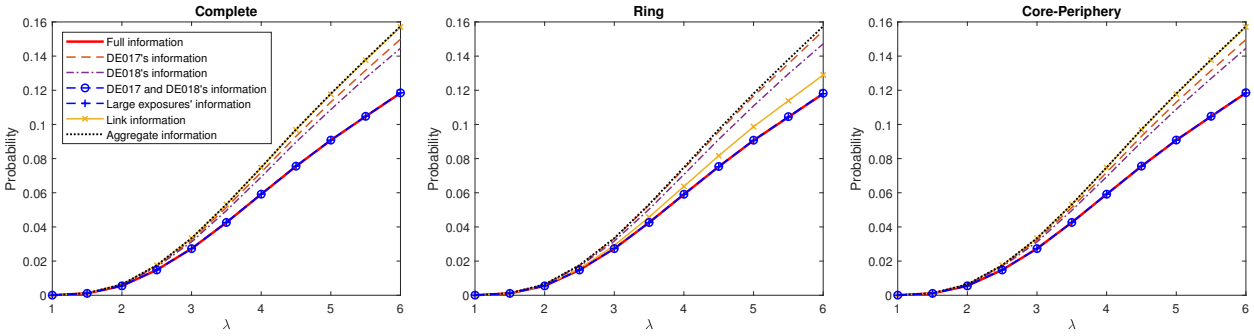


Note. The three structures above represent complete, ring-like, and core-periphery networks, respectively. For each structure, nodes represent individual banks, and edges stand for their interbank exposures.

(see Figure 6). This reconstruction is based on an entropy-minimization method in Upper and Worms (2004), assuming that each bank's interbank assets and liabilities are equal to its exposure at default (EAD).[10] We refer the reader to Chen et al. (2016) for more details. For the core-periphery network, DE017, DE018, DE019, and DE020 are selected as the core banks according to the size of the total assets. The matrices of the relative liabilities for these networks are presented in Section EC.4.3 of the online supplement.

We assume that the random shocks $X_1, \ldots, X_{11}$ follow Pareto distributions, taking into account that heavy-tailed shocks could lead to large shock propagation. Note that if only a little shock propagation occurs, it is obvious that the worst-case default probability and risk capital under partial information are almost identical to the true quantities. For each i, the probability density f_i of X_i

Figure 7 The worst-case default probabilities for $\mathcal{T} = \{\text{DE017}, \text{DE018}\}$ under different network information and different network structures.



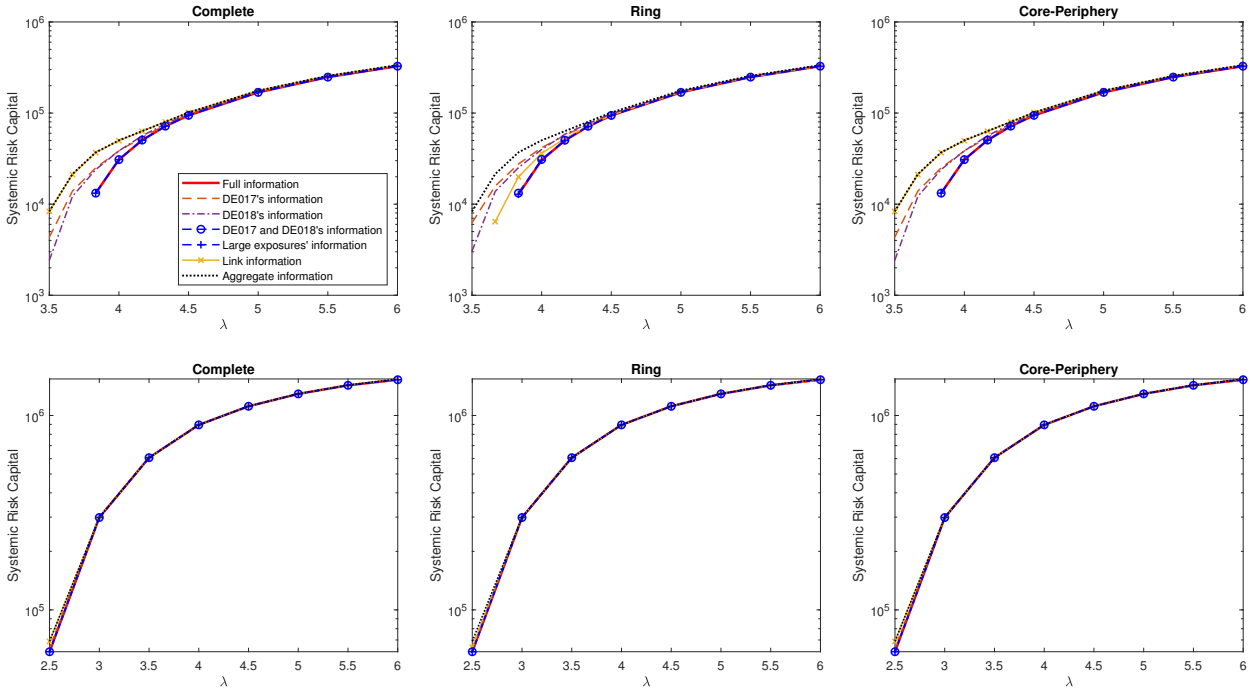
Note. The Monte Carlo method with 10^5 replications is used for the estimation of the probabilities. The legend in the left subfigure applies to all the subfigures.

is assumed to depend on the amount of external assets as follows: $f_i(x) = \theta_i^{-1}(1 + \lambda x/\theta_i)^{-(1/\lambda+1)}$, where $\theta_i = c_i/c_1$ and $\lambda \in [1, 6]$. It is also assumed that the shocks are independent and constrained to be in $[\mathbf{0}, \mathbf{c}]$ to satisfy the modeling assumption in Section 2.

We set Deutsche Bank (DE017) and Commerzbank (DE018) as target banks, i.e., $\mathcal{T} = \{\text{DE017, DE018}\}$, since they were the two largest banks in terms of asset size and were identified as G-SIBs by the Financial Stability Board in 2011. Further, we consider seven types of network information: full network information, DE017's information, DE018's information, DE017 and DE018's information, large exposures' information, link information, and aggregate information. Recall that the worst-case quantities under full information correspond to the true quantities and those under partial information should be larger than the true values.

In Figures 7 and 8, we observe the impact of network information on the worst-case default probabilities (4) and risk capitals (18) as in Figures 4 and 5, respectively. We arbitrarily set $\epsilon = 0.002$ for the link information and $\eta = 0$ for all cases. We find that most interbank exposures are large exposures due to the small network size, and hence, the worst-case analysis under large exposures' information is hardly different from the quantities under full information. While this observation does not extend to the case of larger networks, the experiments with the full EBA dataset show that large exposures' information remains useful in larger networks; see Section EC.4.1 of the online supplement. This result is associated with contribution 3-(b) in the introduction.

Figure 8 The worst-case risk capitals for $\mathcal{T} = \{\text{DE017}, \text{DE018}\}$ under different network information and different network structures.



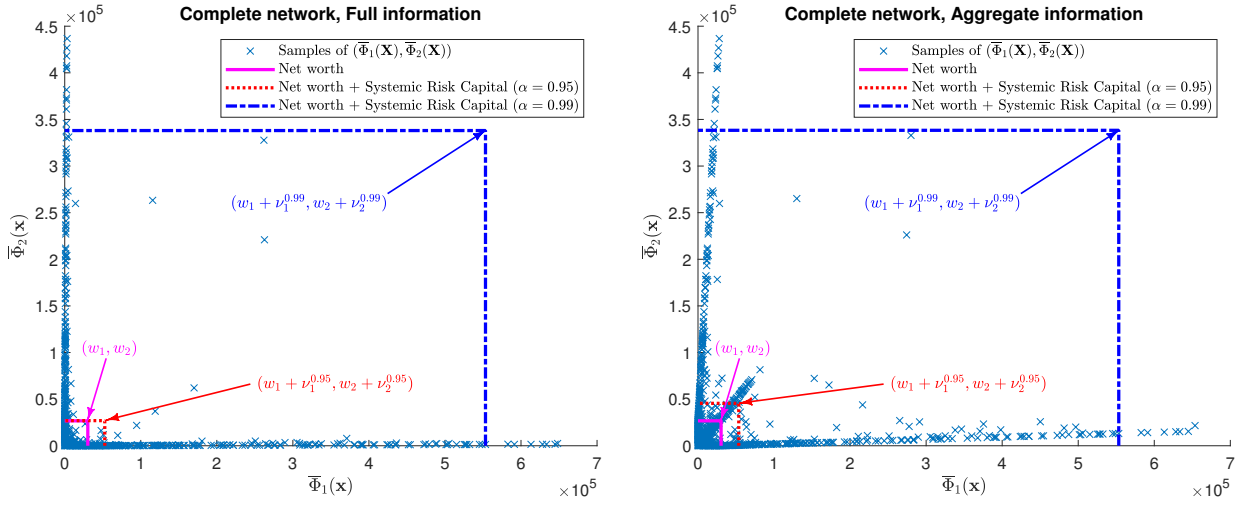
Note. The upper subfigures are when $\alpha = 0.95$, and the lower subfigures are when $\alpha = 0.99$. The number of samples (N) is 10,000. All subfigures share the same legend in the upper left subfigure.

Notably, the impact of link information on approximation quality is highly sensitive to the network structure. While it is most effective under the ring network, the performance hardly improves under the complete network. This confirms that the value of the link information is greater for a sparse network than for a dense network, as previously inferred from the corresponding uncertainty set (15). This corresponds to contribution 3-(c) in the introduction.

More importantly, it is consistently observed that the target banks' information results in the best approximations while the aggregate information and the information of a single target bank do not help much, which pertains to contributions 3-(a) and 3-(d) in Section 1. The figures nevertheless show that the information of DE018 is more useful than that of DE017 regardless of the network structure, which is because DE018 has greater financial connectivity than DE017 (see Table 2).

In contrast to the result in Figure 5, Figure 8 shows that the worst-case risk capital is close to the true risk capital in most cases and the degree of proximity increases as α increases. This seemingly

Figure 9 The distributions of $(\overline{\Phi}_1(\mathbf{X}), \overline{\Phi}_2(\mathbf{X}))$ with full information (left) and aggregate information (right).



Note. The complete network is used for the illustration. We set $\lambda = 4$. Each rectangle represents the region in which the total shocks to DE017 and DE018 are completely covered by their net worths and the corresponding risk capital allocated to each of them.

conflicting phenomenon stems from the nature of factors that affect shock propagation and that of the data we use. Figure 9 provides two scatter plots of the samples of $(\overline{\Phi}_1(\mathbf{X}), \overline{\Phi}_2(\mathbf{X}))$ under full information and aggregate information, respectively. The figure tells us that large $\overline{\Phi}_i$'s are driven by direct shocks but small $\overline{\Phi}_i$'s by indirect shocks. Hence, the shock propagation has more effects on computing the risk capital for smaller α. Note that in the figure, the default probability estimate counts the number of samples outside the region $[0, w_1] \times [0, w_2]$, whereas for risk capital, $(\nu_1^\alpha, \nu_2^\alpha)$ is determined to limit the number of samples outside the region $[0, w_1 + \nu_1^\alpha] \times [0, w_2 + \nu_2^\alpha]$ to at most $N(1-\alpha)$. In addition, compared to the case in Figure 5, the financial connectivities of the target banks are relatively small (see Table 2), which strengthens the impact of direct shocks.

7. Conclusion

In this paper, we addressed robust risk quantification under the Eisenberg-Noe model with incomplete network information. Particularly, we provided an MILP problem to identify worst-case shock propagation to a specific group of banks from other banks, based on which worst-case default probabilities are quantified. In response to recent changes in financial regulations, we also expanded our approach to the problem of computing risk capital which secures SIFIs against the worst-case shock

propagation. We formulated the problem using chance-constrained optimization and suggested a sample average approximation scheme for computational tractability. Our numerical observations revealed the impact of partial information on the worst-case default probability and risk capital. They were found to be potentially useful in estimating the true quantities in the presence of certain network data such as target banks' information, large exposures' information, or link information.

This work opens up several interesting directions for further investigation. Firstly, analyzing the difference between true and worst-case default probabilities and its sensitivity to information availability could be insightful. Secondly, relaxing specific modeling features, such as extending to the Eisenberg-Noe model with fire sales, would be interesting. Further, based on the mapping from the shock vector $\mathbf{x}$ to the loss $\Phi_i(\mathbf{x})$, one may explore the problem of selecting the most likely shock scenario given the loss outcome, which is called reverse stress testing. Lastly, assuming that unknown interbank liabilities are random, one might tackle the issue of incomplete network information differently by finding a confidence interval of the total loss given a shock realization.

Appendix A: Comparison with Different Approaches

In Table 3, we compare the performance of our approach with that of the two alternative methods for the robust quantification of (3) in Remark 2, denoted by Robust LP and Bilevel NLP, respectively. We consider two different networks: the core-periphery network constructed by the data of 11 German banks in Section 6 and that constructed by the full EBA data with 80 European banks in Section EC.4.1 of the online supplement. We use the same target banks as in those sections. We assume Pareto distributions for random shocks, where $\lambda = 4$ for the small network case and $\lambda = 2$ for the large network case. Based on 1,000 presampled shock realizations, we apply different approaches to estimate the worst-case default probabilities under SIFI information and aggregate information, and thus, the computation times in Table 3 do not include sampling times. We use the DELL PowerEdge R630 server with Dual Intel Xeon E5-2697 2.6GHz CPU and 128 GB RAM.

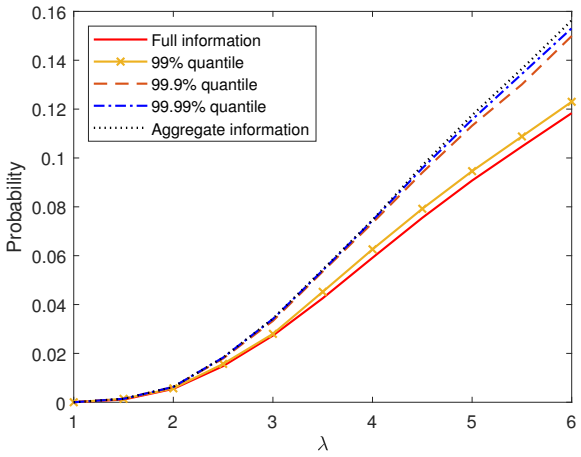
Our approach shows outstanding performance compared to the two alternative methods. In particular, Robust LP is computationally fast, but it can be easily seen in Table 3 that the estimates are too conservative to be used in practice. Bilevel NLP turns out to be extremely slower than our method. More importantly, we find that this method may often produce unreliable estimates despite such high computational costs and fail to distinguish SIFI and aggregate information as seen in their identical estimates. Note that in the case of aggregate information, our approach is fast and accurate since it does not require solving an MILP; see (17).

Table 3 Estimates and estimation times under different approaches for the robust quantification of (3) in Remark 2

Small network				
	SIFI information		Aggregate information	
Methods	Estimate	Time (sec)	Estimate	Time (sec)
Our approach	0.053	4.902	0.071	0.003
Robust LP	0.062	3.498	1	0.346
Bilevel NLP	0.053	3.038×10^2	0.053	4.634×10^2
Large network				
	SIFI information		Aggregate information	
Methods	Estimate	Time (sec)	Estimate	Time (sec)
Our approach	0.083	31.139	0.097	0.009
Robust LP	0.120	4.203	1	0.415
Bilevel NLP	0.083	1.985×10^4	0.083	3.004×10^4

Robust LP means the method of using the robust counterpart of (2), and Bilevel NLP represents the method of using a bilevel nonlinear program to find the worst possible network configuration that yields the greatest default probability.

Figure 10 The worst-case default probabilities and three different quantiles of the default probabilities.



Note. The Monte Carlo method with 10^5 replications is used for the estimation of the probabilities.

Furthermore, we numerically examine the practical validity of our worst-case default probabilities. We revisit the example of 11 German banks in Section 6 and assume that only aggregate information is available. We first generate 10^6 network structures that satisfy the condition in the set $\mathcal{A}$ with $\mathcal{K}=\{(1,1),(2,2)\ldots,(n,n)\}$ and $\widetilde{\mathcal{K}}=\emptyset$ based on a homogeneous random graph model, in which every possible edge occurs independently with probability 1/2. Given 10^5 shock realizations, we estimate the default probabilities for each network structure and obtain their 99%, 99.9%, and 99.99% quantiles. In other words, we have an empirical distribution of the target default probability based on a million different network structures. Figure 10 compares our worst-case default probabilities with those quantiles when $\mathcal{T}=\{\text{DE017, DE018}\}$. We use the core-periphery network in Section 6 for the full information benchmark. We observe that the 99.9% and 99.99% quantiles

are close to our worst-case default probabilities with aggregate information. This implies that our worst-case default probabilities are not overly conservative. We also note that the more information is available (e.g., SIFI information), the smaller the gap between our worst-case default probabilities and the quantiles, which further highlights the value of information on interbank liabilities.

Endnotes

1. Capponi and Larsson (2015) and Capponi and Weber (2022) discuss fire-sale spillover effects between banks with applications to capital constraints and portfolio diversification, respectively.

2. Although the concept of the risk capital we consider is analogous to systemic risk measures that address capital injection to each bank (Biagini et al. 2019, Feinstein et al. 2017), our approach differs in that we consider the worst-case capital requirement against incomplete network information.

3. Since interbank networks typically have a core-periphery structure (in 't Veld et al. 2020), the entire interbank network is generally sparse, but the network of core banks could be relatively dense in practice. In this case, link information may not be useful for approximating the true quantities.

4. Based on the empirical data, Ahn (2020) points out that conditions (c) and (d) are not restrictive from the practical point of view.

5. See Figure EC.9 of the online supplement for the 5-bank financial network in Example 1.

6. This lemma also allows us to visualize the default region and to understand an explicit form of the total shock; see Sections EC.3 and EC.4.2, respectively, of the online supplement. Such analyses are hardly possible if the set D_i is represented by $\mathbf{p}(\mathbf{x})$ in which $\mathbf{x}$ is implicitly entangled.

7. The vector $\boldsymbol{\zeta}$ can be viewed as a dual of the weighted Bonacich centrality in network analysis (see, e.g., Candogan et al. 2012). This centrality is used as a measure of how influential each single node is in a network, while $\boldsymbol{\zeta}$ captures the influence of other nodes on a particular node.

8. Based on empirical observations, it is widely accepted that financial shocks have heavy tails in practice; see, e.g., McNeil et al. (2015) and Bradley and Taqqu (2003).

9. See Section EC.2 of the online supplement for its mathematical definition.

10. The EAD quantifies a bank's total claims on all other banks, and hence, is considered as the size of its interbank assets. As interbank liabilities are not reported in the EBA data, their size is

roughly assumed to be equal to the EAD or its perturbed value in the literature (Glasserman and Young 2015, Chen et al. 2016, Gandy and Veraart 2017, Veraart 2020, Amini and Feinstein 2023).

Acknowledgments

The authors thank Paul Glasserman for his constructive comments on an earlier version of this paper. The authors also appreciate the area editor (John Birge), the anonymous associate editor, and two anonymous reviewers for their valuable and insightful feedback that improved the paper. The work of D. Ahn was supported by the Early Career Scheme from the Research Grants Council of Hong Kong (Project No. 24210420). N. Chen acknowledges the funding supports from Research Grants Council, Hong Kong, via GRF grants No. 14207918 and No. 14208620. Parts of this research of the second author has been done during his visit to AI-powered Financial Technology (AIFT) Ltd., Hong Kong. The work by K. Kim was supported by the National Research Foundation of Korea funded by the Ministry of Science and ICT (NRF-2019R1A2C1003144).

References

Acemoglu D, Ozdaglar A, Tahbaz-Salehi A (2015) Systemic risk and stability in financial networks. *American Economic Review* 105(2):564–608.

Ahn D (2020) Shock amplification in financial networks with applications to the CCP feasibility. *Quantitative Finance* 20(7):1045–1056.

Ahn D, Kim KK (2018) Efficient simulation for expectations over the union of half-spaces. *ACM Transactions on Modeling and Computer Simulation* 28(3):Article 23.

Ahn D, Kim KK (2019) Optimal intervention under stress scenarios: A case of the Korean financial system. *Operations Research Letters* 47(4):257–263.

Amini H, Cont R, Minca A (2016) Resilience to contagion in financial networks. *Mathematical Finance* 26(2):329–365.

Amini H, Feinstein Z (2023) Optimal network compression. *European Journal of Operational Research* 306(3):1439–1455.

Amini H, Minca A (2016) Inhomogeneous financial networks and contagious links. *Operations Research* 64(5):1109–1120.

Anand K, Craig B, von Peter G (2015) Filling in the blanks: Network structure and interbank contagion. *Quantitative Finance* 15(4):625–636.

Anand K, van Lelyveld I, Ádám Banai, Friedrich S, Garratt R, Hałaj G, Fique J, Hansen I, Jaramillo SM, Lee H, Molina-Borboa JL, Nobili S, Rajan S, Salakhova D, Silva TC, Silvestri L, de Souza SRS (2018) The missing links: A global study on uncovering financial network structures from partial data. *Journal of Financial Stability* 35:107–119.

Baral P, Fique JP (2012) Estimation of bilateral connections in a network: Copula vs. maximum entropy. *Mimeo* 1–10.

Bartesaghi P, Benzi M, Clemente GP, Grassi R, Estrada E (2020) Risk-dependent centrality in economic and financial networks. *SIAM Journal on Financial Mathematics* 11(2):526–565.

Barucca P, Bardoscia M, Caccioli F, D'Errico M, Visentin G, Caldarelli G, Battiston S (2020) Network valuation in financial systems. *Mathematical Finance* 30(4):1181–1204.

Battiston S, Caldarelli G, May RM, Roukny T, Stiglitz JE (2016) The price of complexity in financial networks. *Proceedings of the National Academy of Sciences* 113(36):10031–10036.

BCBS (2020a) Large exposures. *The Basel Framework* (Bank for International Settlement).

BCBS (2020b) Risk-based capital requirements. *The Basel Framework* (Bank for International Settlement).

BCBS (2020c) Scope and definitions. *The Basel Framework* (Bank for International Settlement).

Benoit S, Colliard JE, Hurlin C, Pérignon C (2017) Where the risks lie: A survey on systemic risk. *Review of Finance* 21(1):109–152.

Bertsimas D, Brown DB, Caramanis C (2011) Theory and applications of robust optimization. *SIAM Review* 53(3):464–501.

Biagini F, Fouque J, Frittelli M, Meyer-Brandis T (2019) A unified approach to systemic risk measures via acceptance sets. *Mathematical Finance* 29(1):329–367.

Birge JR, Khabazian A, Peng J (2018) Optimization modeling and techniques for systemic risk assessment and control in financial networks. *INFORMS TutORials in Operations Research* 64–84.

Bosma JJ, Koetter M, Wedow M (2019) Too connected to fail? inferring network ties from price co-movements. *Journal of Business & Economic Statistics* 37(1):67–80.

Bradley BO, Taqqu MS (2003) Financial risk and heavy tails. Rachev ST, ed., *Handbook of Heavy Tailed Distributions in Finance* (Elsevier).

Calafiore G, Campi MC (2005) Uncertain convex programs: randomized solutions and confidence levels. *Mathematical Programming* 102(1):25–46.

Calafiore G, Campi MC (2006) The scenario approach to robust control design. *IEEE Transactions on Automatic Control* 51(5):742–753.

Candogan O, Bimpikis K, Ozdaglar A (2012) Optimal pricing in networks with externalities. *Operations Research* 60(4):883–905.

Capponi A (2016) Systemic risk, policies, and data needs. *INFORMS TutORials in Operations Research* 185–206.

Capponi A, Chen PC, Yao DD (2016) Liability concentration and systemic losses in financial networks. *Operations Research* 64(5):1121–1134.

Capponi A, Larsson M (2015) Price contagion through balance sheet linkages. *The Review of Asset Pricing Studies* 5(2):227–253.

Capponi A, Weber M (2022) Systemic portfolio diversification. Available at SSRN: https://ssrn.com/abstract=3345399.

Chen N, Liu X, Yao DD (2016) An optimization view of financial systemic risk modeling: The network effect and the market liquidity effect. *Operations Research* 64(5):1089–1108.

Cifuentes R, Ferrucci G, Shin HS (2005) Liquidity risk and contagion. *Journal of the European Economic Association* 3(2/3):556–566.

Cimini G, Squartini T, Garlaschelli D, Gabrielli A (2015) Systemic risk analysis on reconstructed economic and financial networks. *Scientific Reports* 5:15758.

Das SR (2016) Matrix metrics: Network-based systemic risk scoring. *Journal of Alternative Investments* 18(4):33–51.

Demange G (2018) Contagion in financial networks: A threat index. *Management Science* 64(2):955–970.

Drehmann M, Tarashev N (2013) Measuring the systemic importance of interconnected banks. *Journal of Financial Intermediation* 22(4):586–607.

Eisenberg L, Noe TH (2001) Systemic risk in financial systems. *Management Science* 47(2):236–249.

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

Elsinger H, Lehar A, Summer M (2006) Risk assessment for banking systems. *Management Science* 52(9):1301–1314.

Feinstein Z, Pang W, Rudloff B, Schaanning E, Sturm S, Wildman M (2018) Sensitivity of the Eisenberg-Noe clearing vector to individual interbank liabilities. *SIAM Journal on Financial Mathematics* 9(4):1286–1325.

Feinstein Z, Rudloff B, Weber S (2017) Measures of systemic risk. *SIAM Journal on Financial Mathematics* 8(1):672–708.

Gai P, Kapadia S (2010) Contagion in financial networks. *Proceedings of the Royal Society A* 466(2120):2401–2423.

Gandy A, Veraart LAM (2017) A Bayesian methodology for systemic risk assessment in financial networks. *Management Science* 63(12):4428–4446.

Glasserman P, Heidelberger P, Shahabuddin P (2000a) Variance reduction techniques for estimating value-at-risk. *Management Science* 46(10):1349–1364.

Glasserman P, Heidelberger P, Shahabuddin P (2000b) Variance reduction techniques for value-at-risk with heavy-tailed risk factors. Joines JA, Barton RR, Kang K, Fishwick PA, eds., *Proceedings of the 2000 Winter Simulation Conference*, 604–609 (IEEE).

Glasserman P, Heidelberger P, Shahabuddin P (2002) Portfolio value-at-risk with heavy-tailed risk factors. *Mathematical Finance* 12(3):239–269.

Glasserman P, Young HP (2015) How likely is contagion in financial networks? *Journal of Banking & Finance* 50:383–399.

Glasserman P, Young HP (2016) Contagion in financial networks. *Journal of Economic Literature* 54(3):779–831.

Guerrieri L, Iacoviello M, Minetti R (2012) Banks, sovereign debt, and the international transmission of business cycles. *NBER International Seminar on Macroeconomics* 9(1):181–213.

Halaj G, Kok C (2013) Assessing interbank contagion using simulated networks. *Computational Management Science* 10:157–186.

Hong LJ, Yang Y, Zhang L (2011) Sequential convex approximations to joint chance constrained programs: A Monte Carlo approach. *Operations Research* 59(3):617–630.

in 't Veld D, van der Leij M, Hommes C (2020) The formation of a core-periphery structure in heterogeneous financial networks. *Journal of Economic Dynamics and Control* 119:103972.

Khabazian A, Peng J (2019) Vulnerability analysis of the financial network. *Management Science* 65(7):3302–3321.

Kuzubaş TU, Ömercikoğlu I, Saltoğlu B (2014) Network centrality measures and systemic risk: An application to the Turkish financial crisis. *Physica A: Statistical Mechanics and its Applications* 405:203–215.

Liu M, Staum J (2010) Sensitivity analysis of the Eisenberg-Noe model of contagion. *Operations Research Letters* 38(5):489–491.

Luedtke J, Ahmed S (2008) A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization* 19(2):674–699.

McNeil AJ, Frey R, Embrechts P (2015) *Quantitative Risk Management: Concepts, Techniques and Tools* (Princeton, NJ: Princeton University Press).

Musmeci N, Battiston S, Caldarelli G, Puliga M, Gabrielli A (2013) Bootstrapping topological properties and systemic risk of complex networks using the fitness model. *Journal of Statistical Physics* 151:720–734.

Pagnoncelli BK, Ahmed S, Shapiro A (2009) Sample average approximation method for chance constrained programming: Theory and applications. *Journal of Optimization Theory and Applications* 142(2):399–416.

Rockafellar RT, Uryasev S (2000) Optimization of conditional Value-at-Risk. *Journal of Risk* 2(3):21–41.

Rogers LCG, Veraart LAM (2013) Failure and rescue in an interbank network. *Management Science* 59(4):882–898.

Roukny T, Battiston S, Stiglitz JE (2018) Interconnectedness as a source of uncertainty in systemic risk. *Journal of Financial Stability* 35:93–106.

Soyster AL (1976) Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research* 21(5):1154–1157.

Squartini T, Caldarelli G, Cimini G, Gabrielli A, Garlaschelli D (2018) Reconstruction methods for networks: The case of economic and financial systems. *Physics Reports* 757:1–47.

Tasca P, Mavrodiev P, Schweitzer F (2014) Quantifying the impact of leveraging and diversification on systemic risk. *Journal of Financial Stability* 15:43–52.

Upper C, Worms A (2004) Estimating bilateral exposures in the German interbank market: Is there a danger of contagion? *European Economic Review* 48(4):827–849.

Veraart LAM (2020) Distress and default contagion in financial networks. *Mathematical Finance* 30:705–737.

Yuan Y, Li Z, Huang B (2017) Robust optimization approximation for joint chance constrained optimization problem. *Journal of Global Optimization* 67(4):805–827.

Online Supplement

Appendix EC.1: Proofs of Theoretical Results

In this section, we denote by $\mathbf{e}^i$ the i-th column of the identity matrix in a suitable dimension, and for any two vectors $\mathbf{u}$ and $\mathbf{v}$, we define $\mathbf{u} \vee \mathbf{v} = (\max\{u_1, v_1\}, \ldots, \max\{u_d, v_d\})^\top$.

EC.1.1. Proof of Lemma 1

In this proof, we assume that $\eta = 0$. The proof for the case of $\eta > 0$ can be found in Ahn (2020).

LEMMA EC.1. *For each* $\mathbf{x} \in [\mathbf{0}, \mathbf{c}]$, $\mathbf{p}(\mathbf{x})$ *is a solution to the following linear program:*

$$\max \mathbf{1}^\top \mathbf{p} \quad s.t. \ (\mathbf{I} - \mathbf{A}^\top)\mathbf{p} \le \mathbf{c} - \mathbf{x}, \mathbf{0} \le \mathbf{p} \le \bar{\mathbf{p}}, \tag{EC.1}$$

The above lemma is proved in Eisenberg and Noe (2001). In what follows, based on this lemma, we prove the statement in four steps. Without loss of generality, we set bank n as the target.

Step 1. Let $\tilde{\mathsf{Q}}_n \coloneqq \{\mathbf{u} \in \mathbb{R}_+^{n-1} | (\mathbf{I} - \mathbf{A}_{-n})\mathbf{u} \le \mathbf{a}^n_{-n}\}$. We want to show that Q_n is the set of extreme points of $\tilde{\mathsf{Q}}_n$. To find extreme points, all we need to do is to find $n-1$ linearly independent and active constraints from the conditions $\mathbf{u} \ge \mathbf{0}$, $(\mathbf{I} - \mathbf{A}_{-n})\mathbf{u} \le \mathbf{a}^n_{-n}$. Take any $\mathbf{u} \in \tilde{\mathsf{Q}}_n$ and set $\mathcal{J} = \{i | u_i = 0\}$ with $\mathcal{I} = \{1, \ldots, n-1\} \backslash \mathcal{J}$. There must be an additional $(n - 1 - |\mathcal{J}|)$ number of linearly independent constraints for u to be an extreme point of $\tilde{\mathsf{Q}}_n$. Suppose that u is an extreme point and it has

$$[(\mathbf{I} - \mathbf{A}_{-n})\mathbf{u}]_j = a_{jn} \tag{EC.2}$$

for some $j \in \mathcal{J}$ as one of such constraints. Then, we get

$$[(\mathbf{I} - \mathbf{A}_{-n})\mathbf{u}]_j = u_j - \sum_{i=1}^{n-1} a_{ji} u_i = -\sum_{i \in \mathcal{I}} a_{ji} u_i = a_{jn}.$$

Since all quantities are nonnegative in the above equation and $u_i > 0$ for $i \in \mathcal{I}$, we have $a_{ji} = 0$ for all $i \in \mathcal{I}$ and $a_{jn} = 0$. That means, these conditions must be satisfied for (EC.2) to be one of the active constraints that determine u.

As a result, if those conditions are met, then the constraint (EC.2) can be re-written as

$$[(\mathbf{I}-\mathbf{A}_{-n})\mathbf{u}]_j = u_j - \sum_{i\in\mathcal{J}} a_{ji}u_i = 0.$$

This, however, is clearly not independent of the constraints $\mathbf{u}_{\mathcal{J}} = \mathbf{0}$. Hence, the remaining $n-1-|\mathcal{J}|$ number of constraints, i.e. $((\mathbf{I}-\mathbf{A}_{-n})\mathbf{u})_i = a_{in}$ for $i\in\mathcal{I}$, must be active. Those $n-1$ constraints yield $\mathbf{u}_{\mathcal{J}} = \mathbf{0}$, $\mathbf{u}_{\mathcal{I}} = (\mathbf{I}-\mathbf{A}_{\mathcal{I}})^{-1}\mathbf{a}^n_{\mathcal{I}}$.

For the converse, take $\mathbf{u}$ such that $\mathbf{u}_{\mathcal{J}} = \mathbf{0}$ and $\mathbf{u}_{\mathcal{I}} = (\mathbf{I}-\mathbf{A}_{\mathcal{I}})^{-1}\mathbf{a}^n_{\mathcal{I}}$ for any fixed $\mathcal{I}\subset\{1,\dots,n-1\}$. We need to show that $\mathbf{u}\in\mathsf{Q}_n$. By the assumption on $\mathbf{A}$, $\mathbf{I}-\mathbf{A}_{\mathcal{I}}$ is also a nonsingular M-matrix and thus its inverse is nonnegative, which implies $\mathbf{u}_{\mathcal{I}}\geq\mathbf{0}$. Hence, for $j\in\mathcal{J}$,

$$[(\mathbf{I}-\mathbf{A}_{-n})\mathbf{u}]_j = -\sum_{k\in\mathcal{I}} a_{jk}u_k \leq 0,$$

whereas, for $i\in\mathcal{I}$, from $(\mathbf{I}-\mathbf{A}_{\mathcal{I}})\mathbf{u}_{\mathcal{I}} = \mathbf{a}^n_{\mathcal{I}}$

$$[(\mathbf{I}-\mathbf{A}_{-n})\mathbf{u}]_i = u_i - \sum_{k\in\mathcal{I}} a_{ik}u_k = a_{in}.$$

This is simply $(\mathbf{I}-\mathbf{A}_{-n})\mathbf{u}\leq\mathbf{a}^n$ and thus $\mathbf{u}\in\tilde{\mathsf{Q}}_n$. The linear independency of the constraints is clear from the invertibility of $\mathbf{I}-\mathbf{A}_{\mathcal{I}}$.

Step 2. Recall that $\mathbf{p}(\mathbf{x})$ is the vector of clearing payments between banks in the network. The solvency condition for bank n is equivalent to $p_n(\mathbf{x}) = \bar{p}_n$; the default of bank n happens only when $p_n(\mathbf{x}) < \bar{p}_n$. We first show that this condition $p_n(\mathbf{x}) = \bar{p}_n$ is equivalent to that the following set is nonempty:

$$\mathsf{R} = \left\{\mathbf{p}\in\mathbb{R}^n \;\middle|\; \mathbf{p}\in[\mathbf{0},\bar{\mathbf{p}}], p_n = \bar{p}_n, (\mathbf{I}-\mathbf{A}^\top)\mathbf{p}\leq\mathbf{c}-\mathbf{x}\right\}.$$

The sufficiency of $p_n(\mathbf{x}) = \bar{p}_n$ is trivial. For the converse, we note that if $\mathbf{p}^1,\mathbf{p}^2\in[\mathbf{0},\bar{\mathbf{p}}]$ satisfy $(\mathbf{I}-\mathbf{A}^\top)\mathbf{p}^i\leq\mathbf{c}-\mathbf{x}$ for $i=1,2$, then $\tilde{\mathbf{p}} = \mathbf{p}^1\vee\mathbf{p}^2$ also satisfies $\tilde{\mathbf{p}}\in[\mathbf{0},\bar{\mathbf{p}}]$ and $(\mathbf{I}-\mathbf{A}^\top)\tilde{\mathbf{p}}\leq\mathbf{c}-\mathbf{x}$. This is because

$$\mathbf{p}^i\leq\mathbf{c}-\mathbf{x}+\mathbf{A}^\top\mathbf{p}^i\leq\mathbf{c}-\mathbf{x}+\mathbf{A}^\top\tilde{\mathbf{p}}\Rightarrow\tilde{\mathbf{p}}\leq\mathbf{c}-\mathbf{x}+\mathbf{A}^\top\tilde{\mathbf{p}}.$$

Let $\mathbf{p}(\mathbf{x})$ be the clearing vector with the maximal $\mathbf{1}^\top\mathbf{p}$. If $\mathbf{p}^o$ is a feasible vector in R, then $\mathbf{p}(\mathbf{x})\vee\mathbf{p}^o$ should also have the maximal $\mathbf{1}^\top\mathbf{p}$. Since the maximizer is unique (from Assumption 1-(d)) and $p_n^o=\bar{p}_n$, we have $p_n(\mathbf{x})=\bar{p}_n$.

Step 3. We then claim that R being nonempty is in turn equivalent to

$$\sup_{\mathbf{p}\in\tilde{\mathsf{R}}}\sum_{i=1}^n a_{in}p_i\geq\bar{p}_n-(c_n-x_n) \tag{EC.3}$$

where

$$\tilde{\mathsf{R}}=\Big\{\mathbf{p}\;\Big|\;\mathbf{p}\leq\bar{\mathbf{p}},\,p_n=\bar{p}_n,\,((\mathbf{I}-\mathbf{A}^\top)\mathbf{p})_i\leq c_i-x_i,,\,\forall i\neq n\Big\}.$$

Note first that $(\mathbf{0},\bar{p}_n)^\top\in\tilde{\mathsf{R}}$ because, for each $i\neq n$,

$$\left((\mathbf{I}-\mathbf{A}^\top)\begin{pmatrix}\mathbf{0}\\ \bar{p}_n\end{pmatrix}\right)_i=-a_{ni}\bar{p}_n\leq 0\leq c_i-x_i.$$

Hence the left-hand side of (EC.3) is well defined.

To show equivalence, let us take $\mathbf{p}^o\in\mathsf{R}$. Then, obviously $\mathbf{p}^o\in\tilde{\mathsf{R}}$ and

$$\big((\mathbf{I}-\mathbf{A}^\top)\mathbf{p}^o\big)_n\leq c_n-x_n\Rightarrow\bar{p}_n-(c_n-x_n)\leq\sum_{i=1}^n a_{in}p_i^o\leq\sup_{\mathbf{p}\in\tilde{\mathsf{R}}}\sum_{i=1}^n a_{in}p_i.$$

For the converse, first check that, if $\mathbf{p}^1,\mathbf{p}^2\in\tilde{\mathsf{R}}$, then $\mathbf{p}^1\vee\mathbf{p}^2\in\tilde{\mathsf{R}}$ thanks to $\mathbf{A}\geq 0$. This implies that, for any feasible $\mathbf{p}\in\tilde{\mathsf{R}}$, $\mathbf{p}\vee(\mathbf{0},\bar{p}_n)^\top\in\tilde{\mathsf{R}}$, and that this $\mathbf{p}\vee(\mathbf{0},\bar{p}_n)^\top$ has a greater, if not equal, value than $\mathbf{p}$ or $(\mathbf{0},\bar{p}_n)^\top$ for the objective function value of the left side of (EC.3). As a result, $\sum_{i=1}^n a_{in}p_i$ takes its supremum on $\tilde{\mathsf{R}}\cap[\mathbf{0},\bar{p}]$. Therefore, its maximizer exists and this maximizer becomes a feasible vector in R.

Step 4. The left-hand side of (EC.3) can be rewritten as follows:

$$\max_{\mathbf{p}'\in\mathbb{R}^{n-1}}\ {\mathbf{a}_{-n}^n}^\top\mathbf{p}'\quad\text{such that }\ (\mathbf{I}-\mathbf{A}_{-n}^\top)\mathbf{p}'\leq\mathbf{d},\ \mathbf{p}'\leq\bar{\mathbf{p}}_{-n},$$

where $\mathbf{d}\in\mathbb{R}^{n-1}$ is given by $d_i=c_i-x_i+a_{ni}\bar{p}_n$ for $i=1,\ldots,n-1$. Then, its dual program is

$$\min_{\mathbf{u},\mathbf{v}\in\mathbb{R}_+^{n-1}}\ \mathbf{d}^\top\mathbf{u}+(\bar{\mathbf{p}}_{-n})^\top\mathbf{v}\quad\text{such that }\ (\mathbf{I}-\mathbf{A}_{-n})\mathbf{u}+\mathbf{v}=\mathbf{a}_{-n}^n,$$

which can be rewritten as follows:

$$\min_{\mathbf{u}\in\tilde{\mathcal{Q}}_n} \left(\mathbf{d}-(\mathbf{I}-\mathbf{A}_{-n}^\top)\bar{\mathbf{p}}_{-n}\right)^\top \mathbf{u}+(\bar{\mathbf{p}}_{-n})^\top \mathbf{a}_{-n}^n.$$

Since the primal has an optimal solution, by the strong duality theorem, (EC.3) holds if and only if

$$\begin{aligned}
0 \le & \min_{\mathbf{u}\in\tilde{\mathcal{Q}}_n} \left(\mathbf{d}-(\mathbf{I}-\mathbf{A}_{-n}^\top)\bar{\mathbf{p}}_{-n}\right)^\top \mathbf{u}+(\bar{\mathbf{p}}_{-n})^\top \mathbf{a}_{-n}^n + c_n - x_n - \bar{p}_n \\
= & \min_{\mathbf{u}\in\tilde{\mathcal{Q}}_n} \sum_{i=1}^{n-1}\left(c_i - x_i + \sum_{j\neq i}\bar{p}_j a_{ji} - \bar{p}_i\right)u_i + c_n - x_n + \sum_{i=1}^{n-1}\bar{p}_i a_{in} - \bar{p}_n \\
= & \min_{\mathbf{u}\in\tilde{\mathcal{Q}}_n} \sum_{i=1}^{n-1}\left(c_i + \sum_{j\neq i}\bar{p}_{ji} - \bar{p}_i - x_i\right)u_i + c_n + \sum_{i=1}^{n-1}\bar{p}_{in} - \bar{p}_n - x_n \\
= & \min_{\mathbf{u}\in\tilde{\mathcal{Q}}_n} \sum_{i=1}^{n-1}(w_i - x_i)u_i + w_n - x_n.
\end{aligned}$$

As a consequence, $p_n(\mathbf{x}) = \bar{p}_n$ is equivalent to $\Phi_n(\mathbf{x}) = x_n + \max_{\mathbf{u}\in\tilde{\mathcal{Q}}_n}(\mathbf{x}_{-n}-\mathbf{w}_{-n})^\top\mathbf{u} \le w_n$. □

EC.1.2. Proof of Theorem 1

The linear program (7) implies that under the assumption of Theorem 1, for $i\in\mathcal{T}$ and $\mathbf{x}\in[\mathbf{0},\mathbf{c}]$, $\overline{\Phi}_i(\mathbf{x})$ is the optimal value of the following problem:

$$\max\ x_i + (\mathbf{x}_{\mathcal{T}^c}-\mathbf{w}_{\mathcal{T}^c})^\top\mathbf{u} \quad \text{s.t.} \quad \left((1+\eta)^{-1}\mathbf{I}-\widetilde{\mathbf{A}}_{\mathcal{T}^c}\right)\mathbf{u} \le \tilde{\mathbf{a}}^i_{\mathcal{T}^c},\ \widetilde{\mathbf{A}}\in\mathcal{A},\ \mathbf{u}\in\mathbb{R}_+^{|\mathcal{T}^c|}. \tag{EC.4}$$

Fix $\widetilde{\mathbf{A}}\in\mathcal{A}$ and assume $\mathcal{T}^c=\{1,\dots,d\}$ with $d<n$ without loss of generality. Note that by Assumption 1–(c), the matrix $((1+\eta)^{-1}\mathbf{I}-\widetilde{\mathbf{A}}_{\mathcal{T}^c})^{-1}$ is a nonnegative matrix and $(1+\eta)\beta_j<1$ for each j. Then, a feasible solution $\mathbf{u}$ of (EC.4) satisfies $\mathbf{u}\le((1+\eta)^{-1}\mathbf{I}-\widetilde{\mathbf{A}}_{\mathcal{T}^c})^{-1}\tilde{\mathbf{a}}^i_{\mathcal{T}^c}$. In addition, consider a Markov chain with $2n$ states and transition probabilities $\{q_{ij}\}_{i,j=1,\dots,2n}$ defined by

$$q_{ij} = \begin{cases} (1+\eta)\tilde{a}_{ij}, & \text{if } i,j=1,\dots,n; \\ 1-\sum_{j=1}^n(1+\eta)\tilde{a}_{ij}, & \text{if } i=1,\dots,n \text{ and } j=n+i; \\ \mathbf{1}_{\{i=j\}}, & \text{otherwise.} \end{cases}$$

States $n+1$ to $2n$ are absorbing states. If the chain is currently in state $j\in\mathcal{T}^c$, the probability that it never reaches state i without visiting another state in $\mathcal{T}$ is given by $1-(((1+\eta)^{-1}\mathbf{I}-\widetilde{\mathbf{A}}_{\mathcal{T}^c})^{-1}\tilde{\mathbf{a}}^i_{\mathcal{T}^c})_j$,

and the probability that the chain will be in state $n+j$ at the next step is given by $1-(1+\eta)\beta_j$. Thus for all $j \in \mathcal{T}^{\mathrm{c}}$, we observe $1-(((1+\eta)^{-1}\mathbf{I}-\widetilde{\mathbf{A}}_{\mathcal{T}^{\mathrm{c}}})^{-1}\tilde{\mathbf{a}}^i_{\mathcal{T}^{\mathrm{c}}})_j \geq 1-(1+\eta)\beta_j$, and hence $u_j < 1$.

Also, the problem (EC.4) is equivalent to

$$\begin{aligned}
\max \quad & x_i + \sum_{j\in\mathcal{T}^{\mathrm{c}}} (x_j - w_j)u_j \\
\text{s.t.} \quad & (1+\eta)^{-1}u_j - \sum_{k\in\mathcal{T}^{\mathrm{c}}} \tilde{a}_{jk}u_k \leq \tilde{a}_{ji}, \quad j \in \mathcal{T}^{\mathrm{c}} \\
& \sum_{k=1}^{n} \tilde{a}_{jk} = \beta_j, \quad j \in \mathcal{T}^{\mathrm{c}} \\
& \tilde{a}_{jk} = a_{jk}, \quad j \in \mathcal{T}^{\mathrm{c}},\ (j,k) \in \mathcal{K} \\
& \tilde{a}_{jk} \geq a_{jk}, \quad j \in \mathcal{T}^{\mathrm{c}},\ (j,k) \in \widetilde{\mathcal{K}} \\
& \tilde{a}_{jk} \geq 0, \quad j \in \mathcal{T}^{\mathrm{c}},\ k = 1,\ldots,n \\
& u_j \geq 0, \quad j \in \mathcal{T}^{\mathrm{c}}.
\end{aligned} \tag{EC.5}$$

Let $\Delta^j = \{(\tilde{a}_{j1},\ldots,\tilde{a}_{jn}) \in \mathbb{R}^n_+ \mid \sum_{l=1}^n \tilde{a}_{jl} = \beta_j,\ \tilde{a}_{jk} = a_{jk} \text{ if } (j,k) \in \mathcal{K},\ \tilde{a}_{jk} \geq a_{jk} \text{ if } (j,k) \in \widetilde{\mathcal{K}}\}$ for $j \in \mathcal{T}^{\mathrm{c}}$. Then, the first five constraints of the above problem can be combined to a single constraint

$$\min_{(\tilde{a}_{j1},\ldots,\tilde{a}_{jn})\in\Delta^j} \left((1+\eta)^{-1}u_j - \sum_{k\in\mathcal{T}^{\mathrm{c}}} \tilde{a}_{jk}u_k - \tilde{a}_{ji} \right) \leq 0, \quad j \in \mathcal{T}^{\mathrm{c}}. \tag{EC.6}$$

Recall that $\mathcal{K}_j = \{k \in \mathcal{T}^{\mathrm{c}} | (j,k) \in \mathcal{K}\}$, $\widetilde{\mathcal{K}}_j = \{k \in \mathcal{T}^{\mathrm{c}} | (j,k) \in \widetilde{\mathcal{K}}\}$, $\mathcal{G}^i_1 = \{j \in \mathcal{T}^{\mathrm{c}} | (j,i) \in \mathcal{K}, \mathcal{K}_j = \mathcal{T}^{\mathrm{c}}\}$, $\mathcal{G}^i_2 = \{j \in \mathcal{T}^{\mathrm{c}} | (j,i) \notin \mathcal{K}\}$, and $\mathcal{G}^i_3 = \{j \in \mathcal{T}^{\mathrm{c}} | (j,i) \in \mathcal{K}, \mathcal{K}_j \neq \mathcal{T}^{\mathrm{c}}\}$. The left-hand side of (EC.6) is another linear programming problem whose objective value is equal to

$$\begin{cases}
(1+\eta)^{-1}u_j - \sum_{k\in\mathcal{T}^{\mathrm{c}}} a_{jk}u_k - a_{ji}, & \text{if } j \in \mathcal{G}^i_1; \\
(1+\eta)^{-1}u_j - \sum_{k\in\mathcal{K}_j\cup\widetilde{\mathcal{K}}_j} a_{jk}u_k - \tilde{\beta}_j, & \text{if } j \in \mathcal{G}^i_2; \\
\min_{l\in\mathcal{T}^{\mathrm{c}}\setminus\mathcal{K}_j} \left((1+\eta)^{-1}u_j - \tilde{\beta}_j u_l - \sum_{k\in\mathcal{K}_j\cup\widetilde{\mathcal{K}}_j} a_{jk}u_k - a_{ji} \right), & \text{if } j \in \mathcal{G}^i_3.
\end{cases}$$

since Δ^j is a convex polytope with extreme points $\{\tilde{\beta}_j \mathbf{e}^l + \sum_{\{k|(j,k)\in\mathcal{K}\cup\widetilde{\mathcal{K}}\}} a_{jk}\mathbf{e}^k\}_{\{l|(j,l)\notin\mathcal{K}\}}$. Then, the optimization problem (EC.5) can be rewritten as follows:

$$\begin{aligned}
\max \quad & x_i + \sum_{j\in\mathcal{T}^{\mathrm{c}}} (x_j - w_j)u_j \\
\text{s.t.} \quad & (1+\eta)^{-1}u_j - \sum_{k\in\mathcal{T}^{\mathrm{c}}} a_{jk}u_k \le a_{ji}, \quad j \in \mathcal{G}_1^i \\
& (1+\eta)^{-1}u_j - \sum_{k\in\mathcal{K}_j\cup\widetilde{\mathcal{K}}_j} a_{jk}u_k \le \tilde{\beta}_j, \quad j \in \mathcal{G}_2^i \\
& \min_{l\in\mathcal{T}^{\mathrm{c}}\setminus\mathcal{K}_j} \left((1+\eta)^{-1}u_j - \tilde{\beta}_j u_l - \sum_{k\in\mathcal{K}_j\cup\widetilde{\mathcal{K}}_j} a_{jk}u_k \right) \le a_{ji}, \quad j \in \mathcal{G}_3^i \\
& u_j \ge 0, \quad j \in \mathcal{T}^{\mathrm{c}}.
\end{aligned} \tag{EC.7}$$

Since the left-hand side of the third constraint in (EC.7) is less than 1, by introducing binary integer variables $\{z_{jl}\}_{j\in\mathcal{G}_3^i, l\in\mathcal{T}^{\mathrm{c}}\setminus\mathcal{K}_j\neq\emptyset}$ satisfying $\sum_{l\in\mathcal{T}^{\mathrm{c}}\setminus\mathcal{K}_j} z_{jl} = 1$ for each $j \in \mathcal{G}_3^i$, the second constraint becomes

$$(1+\eta)^{-1}u_j - \tilde{\beta}_j u_l - \sum_{k\in\mathcal{K}_j} a_{jk}u_k \le a_{ji} + 1 - z_{jl}, \quad j \in \mathcal{G}_3^i, l \in \mathcal{T}^{\mathrm{c}} \setminus \mathcal{K}_j.$$

Note that it is redundant when $z_{jl} = 0$. Consequently, the result follows. □

EC.1.3. Technical Details on Remark 3

We first claim that all banks in the system default if and only if $\mathbf{s}(\mathbf{x};\mathbf{A}) > \mathbf{0}$, where $\mathbf{s}(\mathbf{x};\mathbf{A})$ is defined as in Remark 3. From Assumption 1–(d), by subtracting $\bar{p}_i$ from both sides of (2), the equation becomes

$$\bar{\mathbf{p}} - \mathbf{p} = (1+\eta)\left(\mathbf{x} - \mathbf{w} + \mathbf{A}^\top(\bar{\mathbf{p}} - \mathbf{p})\right)^+, \tag{EC.8}$$

whose unique solution $\mathbf{p} = \mathbf{p}(\mathbf{x})$ exists by Assumption 1–(c). If all banks in the system default, i.e., $\mathbf{p}(\mathbf{x}) < \bar{\mathbf{p}}$, then since the superscript '+' in the right-hand side of (EC.8) can be omitted, we have $\bar{\mathbf{p}} - \mathbf{p}(\mathbf{x}) = (1+\eta)(\mathbf{I} - (1+\eta)\mathbf{A}^\top)^{-1}(\mathbf{x} - \mathbf{w}) = \mathbf{s}(\mathbf{x};\mathbf{A}) > \mathbf{0}$. Conversely, if $\mathbf{s}(\mathbf{x};\mathbf{A}) > \mathbf{0}$, then $\mathbf{s}(\mathbf{x};\mathbf{A}) = (1+\eta)\left(\mathbf{x} - \mathbf{w} + \mathbf{A}^\top\mathbf{s}(\mathbf{x};\mathbf{A})\right) > \mathbf{0}$, and thus, $\mathbf{p} = \bar{\mathbf{p}} - \mathbf{s}(\mathbf{x};\mathbf{A})$ is a solution to (EC.8). Hence, we have $\mathbf{s}(\mathbf{x};\mathbf{A}) = \bar{\mathbf{p}} - \mathbf{p}(\mathbf{x}) > \mathbf{0}$ due to the uniqueness of the solution to (EC.8), leading to the default of all banks. This proves the claim.

Suppose that $\mathbf{X}$ is a continuous random vector. Then, since $\mathsf{P}(\mathbf{s}(\mathbf{X};\mathbf{A})=\mathbf{0})=0$, it suffices to show that $\mathbf{s}(\mathbf{x};\mathbf{A})\geq\mathbf{0}$ for some $\mathbf{A}\in\mathcal{A}$ if and only if $\Psi(\mathbf{x})=0$, where $\Psi(\cdot)$ is defined as in Remark 3. To that end, we observe the following relationship: for fixed $\mathbf{A}\in\mathcal{A}$,

$$\begin{aligned}\mathbf{s}(\mathbf{x};\mathbf{A})\geq\mathbf{0} \quad &\Leftrightarrow \quad \left\{\mathbf{s}\geq\mathbf{0}:(\mathbf{I}-(1+\eta)\mathbf{A})^\top\mathbf{s}=\mathbf{x}-\mathbf{w}\right\}\neq\emptyset\\ &\Leftrightarrow \quad \left\{\boldsymbol{\xi}:(\mathbf{I}-(1+\eta)\mathbf{A})\boldsymbol{\xi}\geq\mathbf{0},\ (\mathbf{w}-\mathbf{x})^\top\boldsymbol{\xi}>\mathbf{0}\right\}=\emptyset,\end{aligned}$$

where the second equivalence holds by Farkas' lemma. Thus, $\mathbf{s}(\mathbf{x};\mathbf{A})\geq\mathbf{0}$ for some $\mathbf{A}\in\mathcal{A}$ implies

$$\mathsf{C}\cap\left\{\boldsymbol{\xi}:(\mathbf{w}-\mathbf{x})^\top\boldsymbol{\xi}>\mathbf{0}\right\}=\emptyset, \tag{EC.9}$$

where $\mathsf{C}:=\bigcap_{\mathbf{A}\in\mathcal{A}}\{\boldsymbol{\xi}:(\mathbf{I}-(1+\eta)\mathbf{A})\boldsymbol{\xi}\geq\mathbf{0}\}=\{\boldsymbol{\xi}:(1+\eta)\max_{\tilde{\mathbf{A}}\in\mathcal{A}}(\sum_k\tilde{a}_{jk}\xi_k)\leq\xi_j,\ \ j=1,\ldots,n\}$. Note that (EC.9) can be rewritten as $\Psi(\mathbf{x})=0$ since by definition $\Psi(\mathbf{x})$ is nonnegative for all $\mathbf{x}$.

It remains to prove that (EC.9) implies $(\mathbf{I}-(1+\eta)\mathbf{A})^\top\mathbf{s}=\mathbf{x}-\mathbf{w}$ for some $\mathbf{s}\geq\mathbf{0}$ and $\mathbf{A}\in\mathcal{A}$. For ease of exposition, we only consider the case where for each j, there exists k such that $(j,k)\notin\mathcal{K}$; the proof for the other case is similar and hence is omitted. For fixed $j\in\{1,\ldots,n\}$, it is easy to check that $\max_{\tilde{\mathbf{A}}\in\mathcal{A}}(\sum_k\tilde{a}_{jk}\xi_k)$ can be recast as

$$\sum_{\{k:(j,k)\in\mathcal{K}\cup\widetilde{\mathcal{K}}\}}a_{jk}\xi_k+\max\left\{\sum_{\{k:(j,k)\notin\mathcal{K}\}}\tilde{a}_{jk}\xi_k:\sum_{\{k:(j,k)\notin\mathcal{K}\}}\tilde{a}_{jk}=\tilde{\beta}_j,\ \ \tilde{a}_{jk}\geq0\ \forall k\text{ s.t. }(j,k)\notin\mathcal{K}\right\},$$

where the second term is equal to $\max_{\{k:(j,k)\notin\mathcal{K}\}}(\tilde{\beta}_j\xi_k)$ by the strong duality theorem. This implies that $\mathsf{C}=\widetilde{\mathsf{C}}:=\bigcap_{j=1}^n\{\boldsymbol{\xi}:(1+\eta)(\sum_{\{k:(j,k)\in\mathcal{K}\cup\widetilde{\mathcal{K}}\}}a_{jk}\xi_k+\tilde{\beta}_j\xi_l)\leq\xi_j\text{ for all }l\text{ s.t. }(j,l)\notin\mathcal{K}\}$. Accordingly, if (EC.9) is true, $\widetilde{\mathsf{C}}\cap\left\{\boldsymbol{\xi}:(\mathbf{w}-\mathbf{x})^\top\boldsymbol{\xi}>\mathbf{0}\right\}=\emptyset$, and hence, by Farkas' lemma, there exist nonnegative constants $(\gamma_{jk})_{j\in\{1,\ldots,n\},(j,k)\notin\mathcal{K}}$ such that for all j,

$$\sum_{\{k:(j,k)\notin\mathcal{K}\}}\gamma_{jk}-(1+\eta)\left(\sum_{\{i:(i,j)\in\mathcal{K}\cup\widetilde{\mathcal{K}}\}}a_{ij}\sum_{\{k:(i,k)\notin\mathcal{K}\}}\gamma_{ik}+\sum_{\{i:(i,j)\notin\mathcal{K}\}}\tilde{\beta}_i\gamma_{ij}\right)=x_j-w_j. \tag{EC.10}$$

Let $s_j=\sum_{\{k:(j,k)\notin\mathcal{K}\}}\gamma_{jk}\geq0$ for $j=1,\ldots,n$, $\bar{a}_{ij}=a_{ij}$ for $(i,j)\in\mathcal{K}$, $\bar{a}_{ij}=a_{ij}+\tilde{\beta}_i\gamma_{ij}/s_i$ for $(i,j)\in\widetilde{\mathcal{K}}$, and $\bar{a}_{ij}=\tilde{\beta}_i\gamma_{ij}/s_i$ for $(i,j)\notin\mathcal{K}\cup\widetilde{\mathcal{K}}$. Then, $\overline{\mathbf{A}}=(\bar{a}_{ij})\in\mathcal{A}$, and the equation (EC.10) becomes $s_j-(1+\eta)\sum_{i=1}^n\bar{a}_{ij}s_i=x_j-w_j$ for all j, i.e., $(\mathbf{I}-(1+\eta)\overline{\mathbf{A}})^\top\mathbf{s}=\mathbf{x}-\mathbf{w}$, which completes the proof.

Finally, the above analysis shows that $\Psi(\mathbf{x})$ is equal to the optimal value of the linear program $\max_{\boldsymbol{\xi}\in\widetilde{\mathsf{C}}}(\mathbf{w}-\mathbf{x})\boldsymbol{\xi}$, which can also be derived by a standard approach to robust linear optimization with polyhedral uncertainty in Bertsimas et al. (2011).

EC.1.4. Proof of Theorem 2

We assume $\eta = 0$ without loss of generality. Let $\widetilde{\mathbf{X}}$ be an unconstrained version of $\mathbf{X}$. Then, we have

$$\mathsf{P}\left(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) = \frac{\mathsf{P}\left(\widetilde{\mathbf{X}}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right)}{\mathsf{P}\left(\widetilde{\mathbf{X}}^m \in [\mathbf{0},\mathbf{c}]\right)},$$

and hence $\log \mathsf{P}(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i) \sim \log \mathsf{P}(\widetilde{\mathbf{X}}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i)$ as m increases since $\mathsf{P}(\widetilde{\mathbf{X}}^m \in [\mathbf{0},\mathbf{c}]) \to \mathsf{P}(\widetilde{\mathbf{X}} \geq \mathbf{0}) > 0$ and $\mathsf{P}(\widetilde{\mathbf{X}}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i) \to 0$ as $m \to \infty$. This is true for $\mathsf{P}(\mathbf{X}^m \in \mathsf{D}_{\mathcal{A}})$ as well. Thus, it suffices to show that

$$\lim_{m\to\infty} \frac{1}{\log m} \log \mathsf{P}\left(\widetilde{\mathbf{X}}^m \in \mathsf{D}_{\mathcal{A}}\right) = \lim_{m\to\infty} \frac{1}{\log m} \log \mathsf{P}\left(\widetilde{\mathbf{X}}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) = -\rho_* + 1.$$

For simplicity, abusing notation, we write $\mathbf{X}$ as the unconstrained version in the rest of the proof. Note that the complementary cumulative distribution function $\bar{F}_i(x) \coloneqq \mathsf{P}(X_i > x)$ is also regularly varying with index $\rho_i - 1$ by Karamata's theorem, and it is well known that for any $\mathcal{I} \in \{1, 2, \ldots, n\}$,

$$\mathsf{P}\left(\sum_{i\in\mathcal{I}} X_i > x\right) \sim \sum_{i\in\mathcal{I}} \mathsf{P}(X_i > x) \text{ as } x \to \infty. \tag{EC.11}$$

Let $\mathsf{Q}^{\mathcal{T}} \coloneqq \bigcup_{i\in\mathcal{T}} \{\boldsymbol{\xi} \in \mathbb{R}^n_+ \,|\, \xi_i = 1, \xi_j = 0 \text{ for } j \in \mathcal{T}\setminus\{i\}, \xi_k = u_k \text{ for } k \in \mathcal{T}^{\mathrm{c}} \text{ and } \mathbf{u} \in \mathsf{Q}_i^{\mathcal{T}}\}$, where $\mathsf{Q}_i^{\mathcal{T}}$ is the collection of extreme points of the feasible set in (7). Then, for all $\boldsymbol{\xi} \in \mathsf{Q}^{\mathcal{T}}$, $\boldsymbol{\xi}^\top \mathbf{w} \geq \min_{i\in\mathcal{T}} w_i$ and $\boldsymbol{\xi} \leq \mathbf{1}$ from the proof of Theorem 1. Also, it is easy to see that for any $\mathbf{A} \in \mathcal{A}$, $\{i | \xi_i > 0,\ \boldsymbol{\xi} \in \mathsf{Q}^{\mathcal{T}}\} \subset \widetilde{\mathcal{H}} \coloneqq \mathcal{H} \cup \mathcal{T}$. Thus, the following relationship can be found for large m:

$$\begin{aligned}
\mathsf{P}\left(\mathbf{X}^m \in \mathsf{D}_{\mathcal{A}}\right) &= \mathsf{P}\left(\bigcup_{i\in\mathcal{T}} \left\{\overline{\Phi}_i(\mathbf{X}^m) > w_i, \mathbf{X}^m \leq \mathbf{c}\right\}\right) \\
&\leq \mathsf{P}\left(\bigcup_{i\in\mathcal{T}} \left\{\overline{\Phi}_i(\mathbf{X}^m) > w_i\right\}\right) \\
&= \mathsf{P}\left(\bigcup_{i\in\mathcal{T}} \left\{X_i^m + \max_{\mathbf{A}\in\mathcal{A}} \max_{\mathbf{u}\in\mathsf{Q}_i^{\mathcal{T}}} \mathbf{u}^\top \left(\mathbf{X}^m_{\mathcal{T}^{\mathrm{c}}} - \mathbf{w}_{\mathcal{T}^{\mathrm{c}}}\right) > w_i\right\}\right) \\
&= \mathsf{P}\left(\bigcup_{\mathbf{A}\in\mathcal{A}} \bigcup_{\boldsymbol{\xi}\in\mathsf{Q}^{\mathcal{T}}} \left\{\boldsymbol{\xi}^\top \mathbf{X}^m > \boldsymbol{\xi}^\top \mathbf{w}\right\}\right) \\
&\leq \mathsf{P}\left(\bigcup_{\mathbf{A}\in\mathcal{A}} \bigcup_{\boldsymbol{\xi}\in\mathsf{Q}^{\mathcal{T}}} \left\{\mathbf{1}^\top \mathbf{X}^m_{\widetilde{\mathcal{H}}} > \min_{i\in\mathcal{T}} w_i\right\}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \mathsf{P}\Big(\mathbf{1}^\top \mathbf{X}^m_{\widetilde{\mathcal{H}}} > \min_{i\in\mathcal{T}} w_i\Big) \\
&\sim \sum_{i\in\widetilde{\mathcal{H}}} \mathsf{P}\Big(X_i > m\Big(\min_{i\in\mathcal{T}} w_i\Big)\Big) && \text{(EC.12)} \\
&\sim \sum_{i\in\widetilde{\mathcal{H}}} \mathsf{P}\big(X_i > m\big)\Big(\min_{i\in\mathcal{T}} w_i\Big)^{-\rho_i+1} && \text{(EC.13)} \\
&\le \mathsf{P}\big(X_* > m\big) \sum_{i\in\widetilde{\mathcal{H}}} \Big(\min_{i\in\mathcal{T}} w_i\Big)^{-\rho_i+1}
\end{aligned}
$$

where $\mathsf{P}(X_* > m) \coloneqq \max_{i\in\widetilde{\mathcal{H}}} \mathsf{P}(X_i > m)$ for large m. The asymptotic equivalence (EC.12) follows from (EC.11), and the asymptotic equivalence (EC.13) holds since $f_1(x) \sim g_1(x)$ and $f_2(x) \sim g_2(x)$ implies $f_1(x) + f_2(x) \sim g_1(x) + g_2(x)$ if $f_1(x)f_2(x) > 0$. Thus, we have

$$
\lim_{m\to\infty} \frac{\log \mathsf{P}\big(\mathbf{X}^m \in \mathsf{D}_{\mathcal{A}}\big)}{\log \mathsf{P}\big(X_* > m\big)} \ge 1.
$$

On the other hand, we define $\tilde{\mathbf{e}} \coloneqq \sum_{i\in\widetilde{\mathcal{H}}} \mathbf{e}^i$ and

$$
\tilde{c}_i = \begin{cases} 1 + \tilde{\mathbf{e}}^\top \mathbf{w}, & i \in \widetilde{\mathcal{H}}; \\ c_i, & \text{otherwise.} \end{cases}
$$

Then, since $\mathbf{z} \in \mathbb{R}_+^{|\widetilde{\mathcal{H}}|} \cap [\mathbf{0}, \tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}}]^{\mathrm{c}}$ implies $\mathbf{1}^\top \mathbf{z} \ge \tilde{\mathbf{e}}^\top \mathbf{w}$, we similarly make the following observation for large m:

$$
\begin{aligned}
\mathsf{P}\big(\tilde{\mathbf{e}}^\top \mathbf{X}^m > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m \in [\mathbf{0}, \tilde{\mathbf{c}}]\big) &= \mathsf{P}\big(\mathbf{1}^\top \mathbf{X}^m_{\widetilde{\mathcal{H}}} > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m_{\widetilde{\mathcal{H}}} \in [\mathbf{0}, \tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}}]\big)\mathsf{P}\big(\mathbf{X}^m_{\widetilde{\mathcal{H}}^{\mathrm{c}}} \in [\mathbf{0}, \tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}^{\mathrm{c}}}]\big) \\
&\sim \mathsf{P}\big(\mathbf{1}^\top \mathbf{X}^m_{\widetilde{\mathcal{H}}} > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m_{\widetilde{\mathcal{H}}} \in [\mathbf{0}, \tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}}]\big) \\
&= \mathsf{P}\big(\mathbf{1}^\top \mathbf{X}^m_{\widetilde{\mathcal{H}}} > \tilde{\mathbf{e}}^\top \mathbf{w}\big) - \mathsf{P}\big(\mathbf{X}^m_{\widetilde{\mathcal{H}}} \in [\mathbf{0}, \tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}}]^{\mathrm{c}}\big) \\
&\ge \mathsf{P}\big(\mathbf{1}^\top \mathbf{X}^m_{\widetilde{\mathcal{H}}} > \tilde{\mathbf{e}}^\top \mathbf{w}\big) - \mathsf{P}\big(\mathbf{1}^\top \mathbf{X}^m_{\widetilde{\mathcal{H}}} > 1 + \tilde{\mathbf{e}}^\top \mathbf{w}\big) \\
&\sim \sum_{i\in\widetilde{\mathcal{H}}} \mathsf{P}\big(X_i > m\tilde{\mathbf{e}}^\top \mathbf{w}\big) - \sum_{i\in\widetilde{\mathcal{H}}} \mathsf{P}\big(X_i > m\big(1 + \tilde{\mathbf{e}}^\top \mathbf{w}\big)\big) \\
&\sim \sum_{i\in\widetilde{\mathcal{H}}} \mathsf{P}\big(X_i > m\big)\Big\{\big(\tilde{\mathbf{e}}^\top \mathbf{w}\big)^{-\rho_i+1} - \big(1 + \tilde{\mathbf{e}}^\top \mathbf{w}\big)^{-\rho_i+1}\Big\} \\
&\ge \mathsf{P}\big(X_* > m\big)\Big\{\big(\tilde{\mathbf{e}}^\top \mathbf{w}\big)^{-\rho_*+1} - \big(1 + \tilde{\mathbf{e}}^\top \mathbf{w}\big)^{-\rho_*+1}\Big\}.
\end{aligned}
$$

Hence, we have

$$\lim_{m\to\infty} \frac{\log \mathsf{P}\big(\tilde{\mathbf{e}}^\top \mathbf{X}^m > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m \in [\mathbf{0}, \tilde{\mathbf{c}}]\big)}{\log \mathsf{P}\big(X_* > m\big)} \le 1.$$

Next, fix $\varepsilon > 0$ small enough. Then, for large m and for any $i \in \mathcal{T}$,

$$\begin{aligned}
\log \mathsf{P}\left(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) &\ge \log \mathsf{P}\big(\mathbf{X}^m \in \mathsf{D}_i\big) \\
&\ge \log \mathsf{P}\big(\mathbf{X}^m \in \overline{\mathsf{D}}_i \cap [\varepsilon\mathbf{1}, \mathbf{c}]\big) \\
&= \log \int \mathbf{1}_{\left\{\mathbf{x}/m \in \overline{\mathsf{D}}_i \cap [\varepsilon\mathbf{1}, \mathbf{c}]\right\}} f_1(x_1)\cdots f_n(x_n) d\mathbf{x} \\
&= \log \int \mathbf{1}_{\left\{\mathbf{z} \in \overline{\mathsf{D}}_i \cap [\varepsilon\mathbf{1}, \mathbf{c}]\right\}} m^n f_1(mz_1)\cdots f_n(mz_n) d\mathbf{z} \\
&\sim \log\big(m^n f_1(m)\cdots f_n(m)\big) + \log \int \mathbf{1}_{\left\{\mathbf{z} \in \overline{\mathsf{D}}_i \cap [\varepsilon\mathbf{1}, \mathbf{c}]\right\}} z_1^{-\rho_1}\cdots z_n^{-\rho_n} d\mathbf{z} \qquad \text{(EC.14)} \\
&\sim \log\big(m^n f_1(m)\cdots f_n(m)\big) + \log \int \mathbf{1}_{\left\{\tilde{\mathbf{e}}^\top \mathbf{z} \ge \tilde{\mathbf{e}}^\top \mathbf{w},\, \mathbf{z} \in [\varepsilon\mathbf{1}, \tilde{\mathbf{c}}]\right\}} z_1^{-\rho_1}\cdots z_n^{-\rho_n} d\mathbf{z} \\
&\sim \log \int \mathbf{1}_{\left\{\tilde{\mathbf{e}}^\top \mathbf{z} \ge \tilde{\mathbf{e}}^\top \mathbf{w},\, \mathbf{z} \in [\varepsilon\mathbf{1}, \tilde{\mathbf{c}}]\right\}} m^n f_1(mz_1)\cdots f_n(mz_n) d\mathbf{z} \qquad \text{(EC.15)} \\
&= \log \int \mathbf{1}_{\left\{\tilde{\mathbf{e}}^\top \mathbf{x}/m \ge \tilde{\mathbf{e}}^\top \mathbf{w},\, \mathbf{x}/m \in [\varepsilon\mathbf{1}, \tilde{\mathbf{c}}]\right\}} f_1(x_1)\cdots f_n(x_n) d\mathbf{x} \\
&= \log \mathsf{P}\big(\tilde{\mathbf{e}}^\top \mathbf{X}^m > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m \in [\varepsilon\mathbf{1}, \tilde{\mathbf{c}}]\big)
\end{aligned}$$

where $\overline{\mathsf{D}}_i$ is the closure of D_i. It is easy to see that as m increases, $f_1(mz_1)\cdots f_n(mz_n) \sim f_1(m)\cdots f_n(m) z_1^{-\rho_1}\cdots z_n^{-\rho_n}$ uniformly in $\mathbf{z}$ on the compact sets $\overline{\mathsf{D}}_i \cap [\varepsilon\mathbf{1}, \mathbf{c}]$ and $\{\mathbf{z} \in [\varepsilon\mathbf{1}, \tilde{\mathbf{c}}] | \tilde{\mathbf{e}}^\top \mathbf{z} \ge \tilde{\mathbf{e}}^\top \mathbf{w}\}$ since $\lim_{m\to\infty} f_j(mx)/f_j(m) = x^{-\rho_j}$ locally uniformly in x on $(0,\infty)$ for each j by Proposition 2.4 of Resnick (2007). Thus, the asymptotic equivalences (EC.14) and (EC.15) hold. Since ε is arbitrary, the above relationship implies

$$\lim_{m\to\infty} \frac{\log \mathsf{P}\big(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\big)}{\log \mathsf{P}\big(\tilde{\mathbf{e}}^\top \mathbf{X}^m > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m \in [\mathbf{0}, \tilde{\mathbf{c}}]\big)} \le 1.$$

Therefore, we finally arrive at

$$\lim_{m\to\infty} \frac{\log \mathsf{P}\big(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\big)}{\log \mathsf{P}\big(X_* > m\big)} \le 1.$$

According to Proposition 2.6 of Resnick (2007), $\lim_{m\to\infty} \log \mathsf{P}\big(X_i > m\big)/\log m = -\rho_i + 1$ for all i. Hence, the result follows. $\square$

EC.1.5. Proof of Theorem 3

Indeed, we can establish the following theorem which provides a more general result about the asymptotic default probability in the case of lognormal shocks than Theorem 3. This theorem covers the case when the shocks are possibly correlated.

THEOREM EC.1. *Suppose that $\mathcal{A}$ satisfies* (15) *for fixed $\mathcal{K}$ and $\epsilon > 0$. We denote the set $\mathcal{H}$ as in Theorem 2. Let $\mathbf{X}$ follow a multivariate lognormal distribution with parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ positive definite. Define $\sigma_\circ^2 = \max_{i\in\widetilde{\mathcal{H}}} \left(\sigma_i^2 - (\boldsymbol{\Sigma}_{\widetilde{\mathcal{H}}\widetilde{\mathcal{H}}^c}\boldsymbol{\Sigma}^{-1}_{\widetilde{\mathcal{H}}^c\widetilde{\mathcal{H}}^c}\boldsymbol{\Sigma}_{\widetilde{\mathcal{H}}^c\widetilde{\mathcal{H}}})_{ii}\right)$ and $\sigma_*^2 = \max_{i\in\widetilde{\mathcal{H}}}\sigma_i^2$ where $\widetilde{\mathcal{H}} = \mathcal{H}\cup\mathcal{T}$. Assume that $\mathbf{X}$ is constrained to be in $[\mathbf{0},\mathbf{c}]$ almost surely. Then, for any $\mathbf{A}\in\mathcal{A}$,*

$$-\frac{1}{2\sigma_\circ^2} \le \liminf_{m\to\infty} \frac{1}{(\log m)^2}\log \mathsf{P}\left(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) \le \limsup_{m\to\infty} \frac{1}{(\log m)^2}\log \mathsf{P}\Big(\mathbf{X}^m \in \mathsf{D}_{\mathcal{A}}\Big) \le -\frac{1}{2\sigma_*^2}.$$

Proof of Theorem EC.1. We assume $\eta = 0$ without loss of generality. Let $\widetilde{\mathbf{X}}$ be an unconstrained version of $\mathbf{X}$. Then, as in the proof of Theorem 2, it is enough show

$$-\frac{1}{2\sigma_\circ^2} \le \liminf_{m\to\infty} \frac{1}{(\log m)^2}\log \mathsf{P}\left(\widetilde{\mathbf{X}}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) \le \limsup_{m\to\infty} \frac{1}{(\log m)^2}\log \mathsf{P}\Big(\widetilde{\mathbf{X}}^m \in \mathsf{D}_{\mathcal{A}}\Big) \le -\frac{1}{2\sigma_*^2}.$$

For simplicity, abusing notation, we write $\mathbf{X}$ as the unconstrained version in the rest of the proof. We note that if X_i follows a lognormal distribution with parameters μ_i and σ_i, then

$$\mathsf{P}(X_i > x) \sim \frac{\sigma_i/\sqrt{2\pi}}{\log x - \mu_i}\exp\left(-\frac{\left(\log x - \mu_i\right)^2}{2\sigma_i^2}\right). \tag{EC.16}$$

From the proof of Theorem 2, for large m, we observe the following relationship:

$$\begin{aligned}
\mathsf{P}\Big(\mathbf{X}^m \in \mathsf{D}_{\mathcal{A}}\Big) &\le \mathsf{P}\Big(\mathbf{1}^\top \mathbf{X}^m_{\widetilde{\mathcal{H}}} > \min_{i\in\mathcal{T}} w_i\Big) \\
&\le \mathsf{P}\left(\bigcup_{i\in\widetilde{\mathcal{H}}}\left\{X_i > m|\widetilde{\mathcal{H}}|^{-1}\left(\min_{i\in\mathcal{T}} w_i\right)\right\}\right) \\
&\le \sum_{i\in\widetilde{\mathcal{H}}} \mathsf{P}\Big(X_i > m|\widetilde{\mathcal{H}}|^{-1}\Big(\min_{i\in\mathcal{T}} w_i\Big)\Big) \\
&\sim \sum_{i\in\widetilde{\mathcal{H}}} \frac{\kappa_1^i}{\log m - \kappa_2^i}\exp\left(-\frac{\left(\log m - \kappa_2^i\right)^2}{2\sigma_i^2}\right) \\
&\le |\widetilde{\mathcal{H}}|\frac{\kappa_1}{\log m - \kappa_2}\exp\left(-\frac{\left(\log m - \kappa_2\right)^2}{2\sigma_*^2}\right)
\end{aligned}$$

where κ_1^i and κ_2^i are constants for $i \in \widetilde{\mathcal{H}}$, and $\kappa_j = \max_{i\in\widetilde{\mathcal{H}}} \kappa_j^i$ for $j = 1, 2$. The asymptotic equivalence is based on (EC.16). Thus,

$$\limsup_{m\to\infty} \frac{1}{(\log m)^2} \log \mathsf{P}\Big(\mathbf{X}^m \in \mathsf{D}_{\mathcal{A}}\Big) \le -\frac{1}{2\sigma_*^2}.$$

Next, we define $\tilde{\mathbf{e}}$ and $\tilde{\mathbf{c}}$ as in the proof of Theorem 2. Then, for large m, we observe

$$\begin{aligned}
\mathsf{P}\big(\tilde{\mathbf{e}}^\top \mathbf{X}^m > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m \in [\mathbf{0}, \tilde{\mathbf{c}}]\big) &= \mathsf{P}\big(\mathbf{1}^\top \mathbf{X}^m_{\widetilde{\mathcal{H}}} > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m_{\widetilde{\mathcal{H}}} \in [\mathbf{0}, \tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}}], \mathbf{X}^m_{\widetilde{\mathcal{H}}^{\mathrm{c}}} \in [\mathbf{0}, \tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}^{\mathrm{c}}}]\big) \\
&= \mathsf{P}\big(\mathbf{1}^\top \mathbf{X}^m_{\widetilde{\mathcal{H}}} > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m_{\widetilde{\mathcal{H}}} \in [\mathbf{0}, \tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}}], \mathbf{X}_{\widetilde{\mathcal{H}}^{\mathrm{c}}} \in [\mathbf{0}, m\tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}^{\mathrm{c}}}]\big) \\
&\ge \mathsf{P}\big(\mathbf{1}^\top \mathbf{X}^m_{\widetilde{\mathcal{H}}} > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m_{\widetilde{\mathcal{H}}} \in [\mathbf{0}, \tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}}], \mathbf{X}_{\widetilde{\mathcal{H}}^{\mathrm{c}}} \in [\mathbf{1}, \exp(1)\cdot\mathbf{1}]\big) \\
&= \int_{[\mathbf{1}, 2\cdot\mathbf{1}]} \mathsf{P}\big(\mathbf{1}^\top \mathbf{X}^m_{\widetilde{\mathcal{H}}} > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m_{\widetilde{\mathcal{H}}} \in [\mathbf{0}, \tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}}] | \mathbf{X}_{\widetilde{\mathcal{H}}^{\mathrm{c}}} = \mathbf{x}\big) f_{\widetilde{\mathcal{H}}^{\mathrm{c}}}(\mathbf{x}) d\mathbf{x}
\end{aligned}$$

where $f_{\widetilde{\mathcal{H}}^{\mathrm{c}}}(\cdot)$ is a probability density function of $\mathbf{X}_{\widetilde{\mathcal{H}}^{\mathrm{c}}}$.

Since $f_{\widetilde{\mathcal{H}}^{\mathrm{c}}}(\cdot)$ is positive and continuous, there exists $\delta > 0$ such that $f_{\mathbf{X}_{\widetilde{\mathcal{H}}^{\mathrm{c}}}}(\mathbf{x}) \ge \delta$ for all $\mathbf{x} \in [\mathbf{1}, 2\cdot\mathbf{1}]$. Also, since $X_1, \ldots, X_n$ are lognormally distributed, the probability in the integrand is positive and continuous in $\mathbf{x}$, and hence there exists $\mathbf{x}'$ such that $\mathsf{P}\big(\mathbf{1}^\top \mathbf{X}^m_{\widetilde{\mathcal{H}}} > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m_{\widetilde{\mathcal{H}}} \in [\mathbf{0}, \tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}}] | \mathbf{X}_{\widetilde{\mathcal{H}}^{\mathrm{c}}} = \mathbf{x}\big) \ge \mathsf{P}\big(\mathbf{1}^\top \mathbf{X}^m_{\widetilde{\mathcal{H}}} > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m_{\widetilde{\mathcal{H}}} \in [\mathbf{0}, \tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}}] | \mathbf{X}_{\widetilde{\mathcal{H}}^{\mathrm{c}}} = \mathbf{x}'\big)$. for all $\mathbf{x} \in [\mathbf{1}, 2\cdot\mathbf{1}]$. Thus, for large m, we have

$$\begin{aligned}
\mathsf{P}\big(\tilde{\mathbf{e}}^\top \mathbf{X}^m > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m \in [\mathbf{0}, \tilde{\mathbf{c}}]\big) &\ge \delta \mathsf{P}\big(\mathbf{1}^\top \mathbf{X}_{\widetilde{\mathcal{H}}} > m\tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}_{\widetilde{\mathcal{H}}} \in [\mathbf{0}, m\tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}}] | \mathbf{X}_{\widetilde{\mathcal{H}}^{\mathrm{c}}} = \mathbf{x}'\big) \\
&= \delta \mathsf{P}\big(\mathbf{1}^\top \mathbf{X}'_{\widetilde{\mathcal{H}}} > m\tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}'_{\widetilde{\mathcal{H}}} \in [\mathbf{0}, m\tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}}]\big) \\
&= \delta\Big(\mathsf{P}\big(\mathbf{1}^\top \mathbf{X}'_{\widetilde{\mathcal{H}}} > m\tilde{\mathbf{e}}^\top \mathbf{w}\big) - \mathsf{P}\big(\mathbf{X}'_{\widetilde{\mathcal{H}}} \in [\mathbf{0}, m\tilde{\mathbf{c}}_{\widetilde{\mathcal{H}}}]^{\mathrm{c}}\big)\Big) \\
&\ge \delta\Big(\mathsf{P}\big(\mathbf{1}^\top \mathbf{X}'_{\widetilde{\mathcal{H}}} > m\tilde{\mathbf{e}}^\top \mathbf{w}\big) - \mathsf{P}\big(\mathbf{1}^\top \mathbf{X}'_{\widetilde{\mathcal{H}}} > m(1 + \tilde{\mathbf{e}}^\top \mathbf{w})\big)\Big) \\
&\sim \delta \mathsf{P}\big(\mathbf{1}^\top \mathbf{X}'_{\widetilde{\mathcal{H}}} > m\tilde{\mathbf{e}}^\top \mathbf{w}\big) \\
&\sim \frac{\kappa_3 \sigma_\circ / \sqrt{2\pi}}{\log m + \log \tilde{\mathbf{e}}^\top \mathbf{w} - \mu_\circ} \exp\left(-\frac{\big(\log m + \log \tilde{\mathbf{e}}^\top \mathbf{w} - \mu_\circ\big)^2}{2\sigma_\circ^2}\right).
\end{aligned}$$

where $\mathbf{X}'_{\widetilde{\mathcal{H}}}$ is lognormally distributed with parameters $\widetilde{\boldsymbol{\mu}} = \boldsymbol{\mu}_{\widetilde{\mathcal{H}}} + \boldsymbol{\Sigma}_{\widetilde{\mathcal{H}}\widetilde{\mathcal{H}}^{\mathrm{c}}} \boldsymbol{\Sigma}^{-1}_{\widetilde{\mathcal{H}}^{\mathrm{c}}\widetilde{\mathcal{H}}^{\mathrm{c}}} (\log(\mathbf{x}') - \boldsymbol{\mu}_{\widetilde{\mathcal{H}}^{\mathrm{c}}})$ and $\widetilde{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}_{\widetilde{\mathcal{H}}\widetilde{\mathcal{H}}} - \boldsymbol{\Sigma}_{\widetilde{\mathcal{H}}\widetilde{\mathcal{H}}^{\mathrm{c}}} \boldsymbol{\Sigma}^{-1}_{\widetilde{\mathcal{H}}^{\mathrm{c}}\widetilde{\mathcal{H}}^{\mathrm{c}}} \boldsymbol{\Sigma}_{\widetilde{\mathcal{H}}^{\mathrm{c}}\widetilde{\mathcal{H}}}$, $\sigma_\circ^2 = \max_{i\in\widetilde{\mathcal{H}}} \tilde{\sigma}_{ii}$, $\mu_\circ = \max_{i\in\widetilde{\mathcal{H}}:\tilde{\sigma}_{ii}=\sigma_\circ^2} \tilde{\mu}_i$, and $\kappa_3 = \delta \sum_{i\in\widetilde{\mathcal{H}}} \mathbf{1}_{\{\tilde{\sigma}_{ii}=\sigma_\circ^2, \tilde{\mu}_i = \mu_\circ\}}$. Since $\boldsymbol{\Sigma}$ is positive definite, so is $\widetilde{\boldsymbol{\Sigma}}$. Then, $\tilde{\sigma}_{ij} < \sqrt{\tilde{\sigma}_{ii}\tilde{\sigma}_{jj}}$ for all $i, j \in \widetilde{\mathcal{H}}$

with $i \neq j$. Thus, the two asymptotic equivalences above hold by Theorem 1 of Asmussen and Rojas-Nandayapa (2008), and it follows that

$$\liminf_{m\to\infty} \frac{\log \mathsf{P}\big(\tilde{\mathbf{e}}^\top \mathbf{X}^m > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m \in [\mathbf{0}, \tilde{\mathbf{c}}]\big)}{(\log m)^2} \geq -\frac{1}{2\sigma_\circ^2}.$$

We now fix $\varepsilon > 0$ small enough. Then, for large m and for any $i \in \mathcal{T}$,

$$\begin{aligned}
&\log \mathsf{P}\left(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) \qquad \text{(EC.17)}\\
&\geq \log \mathsf{P}\big(\mathbf{X}^m \in \mathsf{D}_i\big)\\
&\geq \log \mathsf{P}\big(\mathbf{X}^m \in \overline{\mathsf{D}}_i \cap [\varepsilon\mathbf{1}, \mathbf{c}]\big)\\
&= \log \int \mathbf{1}_{\left\{\mathbf{x}/m \in \overline{\mathsf{D}}_i \cap [\varepsilon\mathbf{1}, \mathbf{c}]\right\}} f(\mathbf{x}) d\mathbf{x}\\
&= \log \int \mathbf{1}_{\left\{\mathbf{z} \in \overline{\mathsf{D}}_i \cap [\varepsilon\mathbf{1}, \mathbf{c}]\right\}} \kappa_4 \left(\prod_{i=1}^n z_i^{-1}\right) \exp\left(-\frac{1}{2}\big(\log(m\mathbf{z}) - \boldsymbol{\mu}\big)^\top \boldsymbol{\Sigma}^{-1}\big(\log(m\mathbf{z}) - \boldsymbol{\mu}\big)\right) d\mathbf{z}\\
&\geq -\frac{1}{2}(\log m)^2 \mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1} - (\log m)\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}(\log \mathbf{z}^* - \boldsymbol{\mu}) + \log \int \mathbf{1}_{\left\{\mathbf{z} \in \overline{\mathsf{D}}_i \cap [\varepsilon\mathbf{1}, \mathbf{c}]\right\}} f(\mathbf{z}) d\mathbf{z}\\
&\sim -\frac{1}{2}(\log m)^2 \mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1} - (\log m)\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}(\log \mathbf{z}_* - \boldsymbol{\mu}) + \log \int \mathbf{1}_{\left\{\tilde{\mathbf{e}}^\top \mathbf{z} \geq \tilde{\mathbf{e}}^\top \mathbf{w},\, \mathbf{z} \in [\varepsilon\mathbf{1}, \tilde{\mathbf{c}}]\right\}} f(\mathbf{z}) d\mathbf{z}\\
&\geq \log \int \mathbf{1}_{\left\{\tilde{\mathbf{e}}^\top \mathbf{z} \geq \tilde{\mathbf{e}}^\top \mathbf{w},\, \mathbf{z} \in [\varepsilon\mathbf{1}, \tilde{\mathbf{c}}]\right\}} \kappa_4 \left(\prod_{i=1}^n z_i^{-1}\right) \exp\left(-\frac{1}{2}\big(\log(m\mathbf{z}) - \boldsymbol{\mu}\big)^\top \boldsymbol{\Sigma}^{-1}\big(\log(m\mathbf{z}) - \boldsymbol{\mu}\big)\right) d\mathbf{z}\\
&= \log \int \mathbf{1}_{\left\{\tilde{\mathbf{e}}^\top \mathbf{x}/m \geq \tilde{\mathbf{e}}^\top \mathbf{w},\, \mathbf{x}/m \in [\varepsilon\mathbf{1}, \tilde{\mathbf{c}}]\right\}} f(\mathbf{x}) d\mathbf{x}\\
&= \log \mathsf{P}\big(\tilde{\mathbf{e}}^\top \mathbf{X}^m > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m \in [\varepsilon\mathbf{1}, \tilde{\mathbf{c}}]\big)
\end{aligned}$$

where $\overline{\mathsf{D}}_i$ is the closure of D_i, $\kappa_4 = (2\pi)^{-n/2}|\boldsymbol{\Sigma}|^{-1/2}$, and $f(\mathbf{x})$ is a probability density function of $\mathbf{X}$ given by $f(\mathbf{x}) = \kappa_4 \left(\prod_{i=1}^n x_i^{-1}\right) \exp\left(-\big(\log \mathbf{x} - \boldsymbol{\mu}\big)\boldsymbol{\Sigma}^{-1}\big(\log \mathbf{x} - \boldsymbol{\mu}\big)/2\right)$. Also, $\mathbf{z}^* = \arg\max_{\mathbf{z} \in \overline{\mathsf{D}}_i \cap [\varepsilon\mathbf{1}, \mathbf{c}]} \mathbf{1}^\top \boldsymbol{\Sigma}^{-1}(\log \mathbf{z} - \boldsymbol{\mu})$, and $\mathbf{z}_* = \arg\min_{\mathbf{z} \in [\varepsilon\mathbf{1}, \tilde{\mathbf{c}}] : \tilde{\mathbf{e}}^\top \mathbf{z} \geq \tilde{\mathbf{e}}^\top \mathbf{w}} \mathbf{1}^\top \boldsymbol{\Sigma}^{-1}(\log \mathbf{z} - \boldsymbol{\mu})$. Since ε is arbitrary, the above relationship implies that

$$\limsup_{m\to\infty} \frac{\log \mathsf{P}\big(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\big)}{\log \mathsf{P}\big(\tilde{\mathbf{e}}^\top \mathbf{X}^m > \tilde{\mathbf{e}}^\top \mathbf{w}, \mathbf{X}^m \in [\mathbf{0}, \tilde{\mathbf{c}}]\big)} \leq 1.$$

Therefore, we finally arrive at

$$\liminf_{m\to\infty} \frac{1}{(\log m)^2} \log \mathsf{P}\left(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) \geq -\frac{1}{2\sigma_\circ^2}.$$

Hence, the result follows. $\square$

EC.1.6. Proof of Theorem 4

Let us consider a feasible solution $\boldsymbol{\nu}$ of the problem (18). Since $\overline{\Phi}_i(\mathbf{X}) < M$ almost surely for all $i \in \mathcal{T}$, by adding redundant constraints $\nu_i \le M$, $i \in \mathcal{T}$, we can rewrite the problem (18) as:

$$\begin{aligned} \min \quad & \sum_{i\in\mathcal{T}} \nu_i \\ \text{s.t.} \quad & \mathsf{P}\left(\max_{i\in\mathcal{T}} \left\{\overline{\Phi}_i(\mathbf{X}) - w_i - \nu_i\right\} > 0\right) \le 1-\alpha, \\ & 0 \le \nu_i \le M, \quad i \in \mathcal{T}. \end{aligned} \tag{EC.18}$$

The left-hand side of the probabilistic constraint is strictly decreasing with respect to ν_i, $i \in \mathcal{T}$. Thus, there is an optimal solution $\boldsymbol{\nu}^\alpha$ of the problem (EC.18) such that $\nu_i^\alpha < M$ for each $i \in \mathcal{T}$, and for any $\varepsilon > 0$, $\mathsf{P}\left(\max_{i\in\mathcal{T}} \left\{\overline{\Phi}_i(\mathbf{X}) - w_i - \nu_i'\right\} > 0\right) < 1-\alpha$ where $\nu_i' = \nu_i^\alpha + \min\{\varepsilon, M - \max_{i\in\mathcal{T}} \nu_i^\alpha\}/n$ for $i \in \mathcal{T}$. Note that $\|\boldsymbol{\nu}' - \boldsymbol{\nu}^\alpha\| < \varepsilon$.

The sample average approximation of the problem (EC.18) is formulated as follows:

$$\begin{aligned} \min \quad & \sum_{i\in\mathcal{T}} \nu_i \\ \text{s.t.} \quad & \sum_{j=1}^{N} \mathbf{1}_{\{\max_{i\in\mathcal{T}}\left\{\overline{\Phi}_i(\mathbf{x}^j) - w_i - \nu_i\right\} > 0\}} \le N(1-\alpha), \\ & 0 \le \nu_i \le M, \quad i \in \mathcal{T}. \end{aligned} \tag{EC.19}$$

Define $G(\boldsymbol{\nu}, \mathbf{X}) \coloneqq \max_{i\in\mathcal{T}}(\overline{\Phi}_i(\mathbf{X}) - w_i - \nu_i)$. Then, it is easy to see that $G(\boldsymbol{\nu}, \cdot)$ is measurable for every $\boldsymbol{\nu}$ and $G(\cdot, \mathbf{X})$ is continuous for a.e. $\mathbf{X}$. Also, the set $[0, M]^{|\mathcal{T}|}$ is compact, and the function $\boldsymbol{\nu} \mapsto \sum_{i\in\mathcal{T}} \nu_i$ is continuous. Therefore, by Proposition 2.2 of Pagnoncelli et al. (2009), for any optimal solution $\boldsymbol{\nu}^N$ of (EC.19), we have $\sum_{i\in\mathcal{T}} \nu_i^N \to \nu_\circ^\alpha$ and $\inf_{\boldsymbol{\nu}^\alpha \in \mathsf{V}} \|\boldsymbol{\nu}^N - \boldsymbol{\nu}^\alpha\| \to 0$ with probability 1 as $N \to \infty$.

It remains to show that (EC.19) and (19) are equivalent. Let $\{\nu_i, z_j\}_{i\in\mathcal{T}, j=1,\ldots,N}$ be a feasible solution of (19). Then, from the second constraint of (19), $\mathbf{1}_{\{\max_{i\in\mathcal{T}}\left\{\overline{\Phi}_i(\mathbf{x}^j) - w_i - \nu_i\right\} > 0\}} \le z_j$ for all j, and hence from its first constraint, we have

$$\sum_{j=1}^{N} \mathbf{1}_{\{\max_{i\in\mathcal{T}}\left\{\overline{\Phi}_i(\mathbf{x}^j) - w_i - \nu_i\right\} > 0\}} \le \sum_{j=1}^{N} z_j \le N(1-\alpha).$$

Therefore, $\{\nu_i\}_{i\in\mathcal{T}}$ is feasible for (EC.19) with the same objective value. Conversely, let $\{\nu_i\}_{i\in\mathcal{T}}$ be a feasible solution of (EC.19), and we define $z_j \coloneqq \mathbf{1}_{\{\max_{i\in\mathcal{T}}\{\overline{\Phi}_i(\mathbf{x}^j)-w_i-\nu_i\}>0\}}$ for all j. Then, $\{\nu_i, z_j\}_{i\in\mathcal{T}, j=1,\dots,N}$ is feasible for (19) with the same objective value, which establishes the claim of the theorem. □

Appendix EC.2: Further Discussions on Asymptotic Default Probabilities

EC.2.1. Remarks on Theorems 2 and 3

In this subsection, we record detailed observations and remarks related to Theorems 2 and 3. Firstly, as mentioned earlier, both of the probabilities (3) and (4) turn out to be highly affected by the shocks' heavy-tailedness which is represented by $(-\rho_i)$ in the regularly varying case and σ_i in the lognormal case. Those parameters form ρ_* and σ_* based on the set $\mathcal{H}$ and thus mainly determine the decay rate of the probabilities when the shock size gets smaller, whereas the sizes of interbank liabilities and net worths do not matter at least asymptotically.

Secondly, let $\mathbf{\Pi} = (\pi_{ij}) \in \mathbb{R}_+^{n\times n}$ such that $\pi_{ij} = \epsilon \mathbf{1}_{\{(i,j)\notin\mathcal{K}\}}$ given $\mathcal{K}$ and ϵ in (8). Given the target set $\mathcal{T}$, the set $\mathcal{H}$ in Theorems 2 and 3 can be formally defined as $\mathcal{H} = \bigcup_{i\in\mathcal{T}}\{j : ((\mathbf{I}-\mathbf{\Pi}_{\mathcal{T}^c})^{-1}\boldsymbol{\pi}^i_{\mathcal{T}^c})_j > 0\}$. To see this, consider a Markov chain with $2n$ states and transition probabilities $\{q_{ij}\}_{i,j=1,\dots,2n}$ defined by

$$q_{ij} = \begin{cases} \pi_{ij}, & \text{if } i,j = 1,\dots,n; \\ 1-\sum_{j=1}^n \pi_{ij}, & \text{if } i = 1,\dots,n \text{ and } j = n+i; \\ \mathbf{1}_{\{i=j\}}, & \text{otherwise.} \end{cases}$$

States $n+1$ to $2n$ are absorbing states. Then, for each $i\in\mathcal{T}$ and $j\in\mathcal{T}^c$, it can be easily shown that $((\mathbf{I}-\mathbf{\Pi}_{\mathcal{T}^c})^{-1}\boldsymbol{\pi}^i_{\mathcal{T}^c})_j$ is the probability that the chain, starting from state j, eventually reaches state i without visiting the other states in $\mathcal{T}$. Thus, the set $\mathcal{H}$ corresponds to the set of banks having a directed path to a bank in $\mathcal{T}$.

Thirdly, in Theorems 2 and 3, we discuss asymptotic behaviors of (3) and (4) for a sequence of diminishing shock vectors. Since the default of banks is arguably a rare event, decreasing shock sizes can be considered as a mild assumption. The impact of small shocks to financial networks has been

considered in the literature (Acemoglu et al. 2015, Amini and Minca 2016, Amini et al. 2016). Under small shock regime, the first work addresses stable network structures, and the other two papers consider the fraction of defaults in a different default cascade model other than the Eisenberg-Noe framework. Similar approaches can also be found in the literature of portfolio credit risk, which focus on the asymptotic behaviors of portfolio loss probabilities (Glasserman et al. 2000a,b, 2002). Those papers discuss large loss threshold regimes that could be considered equivalent to the small shock regime.

Last but not least, one might be interested in the asymptotic behaviors of (3) and (4) in the case of correlated lognormal shocks, which are described in Theorem EC.1. We observe that when they are correlated, the probabilities (3) and (4) are not asymptotically equivalent in general, but the theorem gives us an asymptotic bound on the relative difference between (3) and (4).

EC.2.2. Asymptotic Default Probabilities under Other Distributions

In this subsection, we add the cases under two other distributions: multivariate normal distribution and heavy-tailed elliptical distribution.

THEOREM EC.2. *Let* $\mathbf{X}$ *has a truncated multivariate normal distribution with mean vector* $\boldsymbol{\mu}$ *and nonsingular covariance matrix* $\boldsymbol{\Sigma}$*. Assume that it is truncated to* $[\mathbf{0}, \mathbf{c}]$*. Then,*

$$\lim_{m\to\infty} \frac{1}{m^2} \log \mathsf{P}\left(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i \right) = -\frac{1}{2} \min_{\mathbf{x}\in\bigcup_{i\in\mathcal{T}} \overline{\mathsf{D}}_i} \mathbf{x}^\top \boldsymbol{\Sigma}^{-1} \mathbf{x} \tag{EC.20}$$

where $\overline{\mathsf{D}}_i$ *is the closure of the set* D_i*.*

Proof. Let $\widetilde{\mathbf{X}} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ without any truncation. According to the proof of Theorem 2, we need to show that

$$\lim_{m\to\infty} \frac{1}{m^2} \log \mathsf{P}\left(\widetilde{\mathbf{X}}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i \right) = -\frac{1}{2} \min_{\mathbf{x}\in\bigcup_{i\in\mathcal{T}} \overline{\mathsf{D}}_i} \mathbf{x}^\top \boldsymbol{\Sigma}^{-1} \mathbf{x}.$$

We first note that

$$\Xi(\boldsymbol{\lambda}) \coloneqq \lim_{m\to\infty} \frac{1}{m^2} \log \mathsf{E}\left[e^{m^2 \boldsymbol{\lambda}^\top \widetilde{\mathbf{X}}^m} \right] = \lim_{m\to\infty} \frac{1}{m^2} \log \mathsf{E}\left[e^{m \boldsymbol{\lambda}^\top \widetilde{\mathbf{X}}} \right] = \frac{1}{2} \boldsymbol{\lambda}^\top \boldsymbol{\Sigma} \boldsymbol{\lambda}.$$

Then, it is easy to see that its Fenchel-Legendre transform $\Xi^*(\mathbf{x}) \coloneqq \sup_{\boldsymbol{\lambda}}\{\boldsymbol{\lambda}^\top \mathbf{x} - \Xi(\boldsymbol{\lambda})\} = (1/2)\mathbf{x}^\top \boldsymbol{\Sigma}^{-1}\mathbf{x}$. Since the origin belongs to the interior of $\{\boldsymbol{\lambda} \in \mathbb{R}^n | \Xi(\boldsymbol{\lambda}) < \infty\}$, by the Gärtner-Ellis Theorem in Dembo and Zeitouni (2009), it holds that

$$\begin{aligned} -\inf_{\mathbf{x}\in\bigcup_{i\in\mathcal{T}} \mathsf{D}_i^\circ} \Xi^*(\mathbf{x}) &\le \liminf_{m\to\infty} \frac{1}{m^2} \log \mathsf{P}\left(\widetilde{\mathbf{X}}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) \\ &\le \limsup_{m\to\infty} \frac{1}{m^2} \log \mathsf{P}\left(\widetilde{\mathbf{X}}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) \\ &\le -\inf_{\mathbf{x}\in\bigcup_{i\in\mathcal{T}} \overline{\mathsf{D}}_i} \Xi^*(\mathbf{x}). \end{aligned}$$

where D_i° and $\overline{\mathsf{D}}_i$ are the interior and the closure of the set D_i, respectively. Hence, the result follows. $\square$

Clearly, the default probability depends both on distribution parameters and on the default region $\bigcup_{i\in\mathcal{T}} \mathsf{D}_i$. For this reason, it is worth noting that the mean vector $\boldsymbol{\mu}$ does not play any role in the limiting behavior and that the effects of the covariance matrix $\boldsymbol{\Sigma}$ and the default region $\bigcup_{i\in\mathcal{T}} \mathsf{D}_i$ are decoupled in Theorem EC.2. The former is in the objective function and the latter is in the constraint of the optimization in (EC.20). This can be useful when we want to see the difference of the behavior according to the change of $\boldsymbol{\Sigma}$ or $\bigcup_{i\in\mathcal{T}} \mathsf{D}_i$. Also, the above theorem tells us that, in the case of normal random shocks, the default probability can be approximated by

$$\mathsf{P}\left(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) \approx \exp\left(-\frac{m^2}{2} \min_{\mathbf{x}\in\bigcup_{i\in\mathcal{T}} \overline{\mathsf{D}}_i} \mathbf{x}^\top \boldsymbol{\Sigma}^{-1}\mathbf{x}\right) \tag{EC.21}$$

when m is large enough.

The approximation (EC.21) can be utilized to show an interplay between the covariance structure of the shock distribution and the network structure of the financial system. For a better understanding, we revisit the example of a 3-bank network in Figure EC.8 with $w_3 = 1$ and $\mathcal{T} = \{3\}$. In this example, it is easy to see that $\bar{p}_{12} = \bar{p}_{13} = \bar{p}_{21} = \bar{p}_{23} = 3/2$. Let X_i follow a normal distribution with mean 0 and variance σ_i^2 for $i = 1, 2$ and $r \coloneqq \mathrm{corr}(X_1, X_2)$. For each given shock distribution, let us assume that bank 3 can redistribute the amount of money it lends to bank 1 and bank 2, i.e., $\bar{p}_{13}$ and $\bar{p}_{23}$, while keeping the aggregate amount equal, i.e, $\bar{p}_{13} + \bar{p}_{23} = 3$. In this situation, we

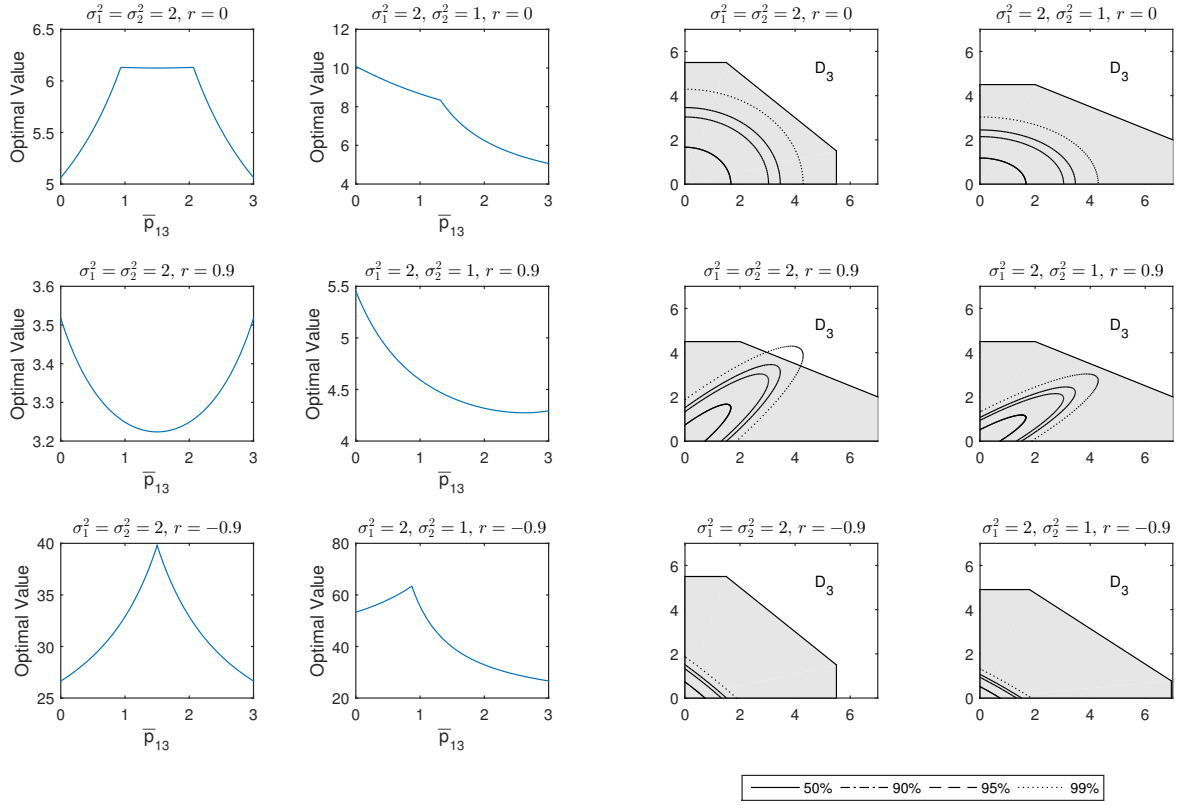


Figure EC.1 The optimal values of the optimization problem in (EC.21) with respect to $\bar{p}_{13}$ for different variances and correlations (left) and bank 3's default regions (unshaded) at the points of $(\bar{p}_{13}, \bar{p}_{23})$ where the optimal values are maximized (right)

consider six different cases where $(\sigma_1^2, \sigma_2^2, r) = (2,2,0)$, $(2,2,0.9)$, $(2,2,-0.9)$ $(2,1,0)$, $(2,1,0.9)$, and $(2,1,-0.9)$. For each case, we calculate the optimal values of the optimization problem in (EC.21) with respect to $\bar{p}_{13}$ between 0 and 3, which the left panels of Figure EC.1 illustrate. In other words, the figure shows how the approximate default probabilities change according to the amount of money bank 3 lends to bank 1.

According to (EC.21), the approximate default probability of bank 3 decreases as the optimal value gets bigger. Thus, for each case, we find $\bar{p}_{13}$ and $\bar{p}_{23}$ maximizing the optimal value, which provides the following interesting observations. First, in the case of equal variances, the plots in the first column of Figure EC.1 show the perfect symmetry. In particular, bank 3 should select

either bank 1 or 2 and lend all the interbank assets to the selected bank when shocks are positively correlated, whereas it is better for bank 3 to allocate the assets equally to bank 1 and 2 when shocks are negatively correlated. Second, when the shock X_1 is more volatile than X_2, the figures in the second column are skewed to the left. This implies that it is optimal for bank 3 to reduce the amount of money lending to bank 1, which coincides with our intuition. Especially, when shocks are uncorrelated or positively correlated, bank 3 needs to loan its total interbank assets to bank 2 only. Finally, with those $\bar{p}_{13}$ and $\bar{p}_{23}$ maximizing the optimal value for each case, we describe new default regions and compare them with 50%, 90%, 95%, and 99% confidence regions of the corresponding shock distributions in the right panels of Figure EC.1.

Next, we provide the case of dependent heavy-tailed distributions. In the following theorem, the random shock vector is assumed to have a multivariate elliptical distribution in which the radial component or the distance is regularly varying. The conclusion of Theorem 2 does not change, and we observe that the heavy-tailedness ρ is the only factor determining the asymptotic default probability.

THEOREM EC.3. *Let* $\mathbf{X}$ *have an elliptical distribution, given by* $\mathbf{X} = \boldsymbol{\mu} + R\boldsymbol{\Lambda}\boldsymbol{\Theta}$*, where* $\boldsymbol{\mu}$ *is an* n*-dimensional vector,* R *is a nonnegative continuous random variable,* $\boldsymbol{\Lambda}$ *is a* $n \times d$ *matrix such that* $\boldsymbol{\Sigma} := \boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top$ *is positive definite, and* $\boldsymbol{\Theta}$ *is uniformly distributed on the unit sphere in* $\mathbb{R}^d$ *independent of* R*. Assume that* $\mathbf{X}$ *is constrained to be in* $[\mathbf{0}, \mathbf{c}]$ *and* R *has a regularly varying distribution with index* $\rho > 1$*. Then,*

$$\lim_{m\to\infty} \frac{1}{\log m} \log \mathsf{P}\left(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) = -\rho + 1. \tag{EC.22}$$

Proof. Let $\widetilde{\mathbf{X}}$ be an unconstrained version of $\mathbf{X}$. According to the proof of Theorem 2, we need to show that

$$\lim_{m\to\infty} \frac{1}{\log m} \log \mathsf{P}\big(\widetilde{\mathbf{X}}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\big) = -\rho + 1.$$

For simplicity, abusing notation, we write $\mathbf{X}$ as the unconstrained version in the rest of the proof. We note that the complementary cumulative distribution function $\mathsf{P}(R > x)$ is also regularly varying with index $\rho - 1$ by Karamata's theorem.

Let $\mathsf{R}(\boldsymbol{\theta}) \coloneqq \{r \geq 0 | r\boldsymbol{\Lambda}\boldsymbol{\theta} \in \bigcup_{i\in\mathcal{T}} \overline{\mathsf{D}}_i\}$ for each $\boldsymbol{\theta}$ where $\overline{\mathsf{D}}_i$ is the closure of D_i. One can prove that $\mathsf{R}(\boldsymbol{\theta})$ is an interval if it is nonempty. It is known that the density of $\mathbf{X}$ is given by $f_{\mathbf{X}}(\mathbf{x}) = |\boldsymbol{\Sigma}|^{-1/2} g\big((\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\big)$ for some function g. Also, by Fang et al. (1990), the relationship between the function $g(\cdot)$ and the density function $f_R(\cdot)$ of R is

$$f_R(r) = \frac{2\pi^{n/2}}{\Gamma(n/2)} r^{n-1} g(r^2).$$

Then, we observe the following relationship:

$$\begin{aligned}
\mathsf{P}\left(\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) &= \mathsf{E}\big[\mathbf{1}_{\{\mathbf{X}^m \in \bigcup_{i\in\mathcal{T}} \overline{\mathsf{D}}_i\}}\big] \\
&= \mathsf{E}^0\left[\frac{f_{\mathbf{X}}(\mathbf{X})}{f^0_{\mathbf{X}}(\mathbf{X})} \mathbf{1}_{\{m^{-1}\mathbf{X} \in \overline{\mathsf{D}}_n\}}\right] \\
&= \mathsf{E}\left[\left[\frac{f_{\mathbf{X}}(R\boldsymbol{\Lambda}\boldsymbol{\Theta})}{f^0_{\mathbf{X}}(R\boldsymbol{\Lambda}\boldsymbol{\Theta})} \mathbf{1}_{\{R \in m\mathsf{R}(\boldsymbol{\Theta})\}} \middle| \boldsymbol{\Theta}\right]\right] \\
&= \int f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta} \int_{m\mathsf{R}(\boldsymbol{\theta})} \frac{f_{\mathbf{X}}(r\boldsymbol{\Lambda}\boldsymbol{\theta})}{f^0_{\mathbf{X}}(r\boldsymbol{\Lambda}\boldsymbol{\theta})} f_R(r) dr \\
&= m \int f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta} \int_{\mathsf{R}(\boldsymbol{\theta})} \frac{f_{\mathbf{X}}(m\tau\boldsymbol{\Lambda}\boldsymbol{\theta})}{f^0_{\mathbf{X}}(m\tau\boldsymbol{\Lambda}\boldsymbol{\theta})} f_R(m\tau) d\tau \\
&\sim m \int f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta} \int_{\mathsf{R}(\boldsymbol{\theta})} f_R(m\tau) d\tau \qquad \text{(EC.23)} \\
&= \int f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta} \int_{m\mathsf{R}(\boldsymbol{\theta})} f_R(r) dr \\
&= \mathsf{P}\left(m^{-1} R\boldsymbol{\Lambda}\boldsymbol{\Theta} \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right)
\end{aligned}$$

where the superscript of E^0 denotes the probability measure P^0 under which $\mathbf{X}$ is distributed as $R\boldsymbol{\Lambda}\boldsymbol{\Theta}$ with density $f^0_{\mathbf{X}}$. Since $\varphi_m(\tau\boldsymbol{\Lambda}\boldsymbol{\theta}) \coloneqq \sqrt{(\tau\boldsymbol{\Lambda}\boldsymbol{\theta} - \boldsymbol{\mu}/m)^\top \boldsymbol{\Sigma}^{-1}(\tau\boldsymbol{\Lambda}\boldsymbol{\theta} - \boldsymbol{\mu}/m)} \to \tau$ as m increases, $f_R\left(m\varphi_m(\tau\boldsymbol{\Lambda}\boldsymbol{\theta})\right) \sim f_R(m)\tau^{-\rho}$ as m increases. Thus, the asymptotic equivalence (EC.23) holds since

$$\frac{f_{\mathbf{X}}(m\tau\boldsymbol{\Lambda}\boldsymbol{\theta})}{f^0_{\mathbf{X}}(m\tau\boldsymbol{\Lambda}\boldsymbol{\theta})} = \left(\frac{\tau}{\varphi_m(\tau\boldsymbol{\Lambda}\boldsymbol{\theta})}\right)^{n-1} \frac{f_R\left(m\varphi_m(\tau\boldsymbol{\Lambda}\boldsymbol{\theta})\right)}{f_R(m\tau)} \to 1$$

uniformly on the compact set $\{(\tau, \boldsymbol{\theta}) \in \mathbb{R}_+ \times \mathbb{R}^n | \tau\boldsymbol{\Lambda}\boldsymbol{\theta} \in \bigcup_{i\in\mathcal{T}} \overline{\mathsf{D}}_i, \|\boldsymbol{\theta}\| = 1\}$ as m increases. Therefore, it remains to show that

$$\lim_{m\to\infty} \frac{1}{\log m} \log \mathsf{P}\left(m^{-1} R\boldsymbol{\Lambda}\boldsymbol{\Theta} \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) = -\rho + 1.$$

Define $\mathsf{Q}^{\mathcal{T}}$ as in the proof of Theorem 2. Note that $\boldsymbol{\xi}^\top \boldsymbol{\Lambda} \neq \mathbf{0}$ for all $\boldsymbol{\xi} \in \mathsf{Q}^{\mathcal{T}}$ since $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top$ is positive definite and $\boldsymbol{\xi}$ is a nonzero vector for all $\boldsymbol{\xi} \in \mathsf{Q}^{\mathcal{T}}$. Since $R\boldsymbol{\Theta}$ follows a spherical distribution, for all $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{u}^\top(R\boldsymbol{\Theta}) = \|\mathbf{u}\|(R\Theta_1)$ in distribution. Then, for large m, we observe the following relationship:

$$
\begin{aligned}
\mathsf{P}\left(m^{-1}R\boldsymbol{\Lambda}\boldsymbol{\Theta} \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) &= \mathsf{P}\left(\bigcup_{\boldsymbol{\xi}\in\mathsf{Q}^{\mathcal{T}}} \left\{\boldsymbol{\xi}^\top(R\boldsymbol{\Lambda}\boldsymbol{\Theta}) > m\boldsymbol{\xi}^\top\mathbf{w}\right\}, m^{-1}R\boldsymbol{\Lambda}\boldsymbol{\Theta} \leq \mathbf{c}\right) \\
&\leq \mathsf{P}\left(\bigcup_{\boldsymbol{\xi}\in\mathsf{Q}^{\mathcal{T}}} \left\{\boldsymbol{\xi}^\top(R\boldsymbol{\Lambda}\boldsymbol{\Theta}) > m\boldsymbol{\xi}^\top\mathbf{w}\right\}\right) \\
&\leq \sum_{\boldsymbol{\xi}\in\mathsf{Q}^{\mathcal{T}}} \mathsf{P}\left(\boldsymbol{\xi}^\top\boldsymbol{\Lambda}(R\boldsymbol{\Theta}) > m\boldsymbol{\xi}^\top\mathbf{w}\right) \\
&= \sum_{\boldsymbol{\xi}\in\mathsf{Q}^{\mathcal{T}}} \mathsf{P}\left(\|\boldsymbol{\xi}^\top\boldsymbol{\Lambda}\|(R\Theta_1) > m\boldsymbol{\xi}^\top\mathbf{w}\right) \\
&= \sum_{\boldsymbol{\xi}\in\mathsf{Q}^{\mathcal{T}}} \mathsf{P}\left(R\Theta_1 > m\frac{\boldsymbol{\xi}^\top\mathbf{w}}{\|\boldsymbol{\xi}^\top\boldsymbol{\Lambda}\|}\right) \\
&\leq \sum_{\boldsymbol{\xi}\in\mathsf{Q}^{\mathcal{T}}} \mathsf{P}\left(R > m\frac{\boldsymbol{\xi}^\top\mathbf{w}}{\|\boldsymbol{\xi}^\top\boldsymbol{\Lambda}\|}\right) \\
&\sim \sum_{\boldsymbol{\xi}\in\mathsf{Q}^{\mathcal{T}}} \mathsf{P}\left(R > m\right)\left(\frac{\boldsymbol{\xi}^\top\mathbf{w}}{\|\boldsymbol{\xi}^\top\boldsymbol{\Lambda}\|}\right)^{-\rho+1} \\
&= \mathsf{P}\left(R > m\right) \sum_{\boldsymbol{\xi}\in\mathsf{Q}^{\mathcal{T}}} \left(\frac{\boldsymbol{\xi}^\top\mathbf{w}}{\|\boldsymbol{\xi}^\top\boldsymbol{\Lambda}\|}\right)^{-\rho+1}.
\end{aligned}
$$

The second inequality holds since $|\Theta_1| \leq 1$ and $R \geq 0$ almost surely. Hence, we have

$$
\lim_{m\to\infty} \frac{\log \mathsf{P}\left(m^{-1}R\boldsymbol{\Lambda}\boldsymbol{\Theta} \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right)}{\log \mathsf{P}\left(R > m\right)} \geq 1.
$$

On the other hand, it is easy to see that there exist $r_1, r_2 > 0$ and a subset $\mathcal{B}$ of the unit sphere such that $r_1 < r_2$, $\mathsf{P}(\boldsymbol{\Theta} \in \mathcal{B}) > 0$, and $r\boldsymbol{\Lambda}\boldsymbol{\theta} \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i$ for all $r \in (r_1, r_2]$ and $\boldsymbol{\theta} \in \mathcal{B}$. For example, $\mathbf{x}^\circ := (c_1/2, \ldots, c_{n-1}/2, (w_n + c_n)/2) \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i^\circ$ where D_i° is the interior of D_i. Then, we can find $r_2 > 0$ and $\boldsymbol{\theta}'$ such that $\|\boldsymbol{\theta}'\| = 1$ and $\mathbf{x}^\circ = r_2\boldsymbol{\Lambda}\boldsymbol{\theta}'$. Also, there exists $\varepsilon > 0$ and $r_1 < r_2$ such that $B_{\varepsilon/2}(r_1\boldsymbol{\Lambda}\boldsymbol{\theta}') \subset B_\varepsilon(r_2\boldsymbol{\Lambda}\boldsymbol{\theta}') \subset \bigcup_{i\in\mathcal{T}} \mathsf{D}_i^\circ$ where $B_a(\mathbf{u})$ is an open ball with center $\mathbf{u}$ and radius a. Let $\mathcal{B} := \{\boldsymbol{\theta} | \|\boldsymbol{\theta}\| = 1, r_1\boldsymbol{\Lambda}\boldsymbol{\theta} \in B_{\varepsilon/2}(r_1\boldsymbol{\Lambda}\boldsymbol{\theta}'), r_2\boldsymbol{\Lambda}\boldsymbol{\theta} \in B_\varepsilon(r_2\boldsymbol{\Lambda}\boldsymbol{\theta}')\}$. Since $\boldsymbol{\theta}' \in \mathcal{B}$ and the mappings $\boldsymbol{\theta} \mapsto r_1\boldsymbol{\Lambda}\boldsymbol{\theta}$

and $\boldsymbol{\theta} \mapsto r_2 \boldsymbol{\Lambda}\boldsymbol{\theta}$ are continuous, $\mathcal{B}$ is nonempty and open, which implies that $\mathsf{P}(\boldsymbol{\Theta} \in \mathcal{B}) > 0$. Therefore, for large m,

$$\begin{aligned}\mathsf{P}\left(m^{-1}R\boldsymbol{\Lambda}\boldsymbol{\Theta} \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\right) &\geq \mathsf{P}\big(r_1 < m^{-1}R \leq r_2\big) \cdot \mathsf{P}\big(\boldsymbol{\Theta} \in \mathcal{B}\big) \\ &= \big(\mathsf{P}\big(R > mr_1\big) - \mathsf{P}\big(R > mr_2\big)\big)\, \mathsf{P}\big(\boldsymbol{\Theta} \in \mathcal{B}\big) \\ &\sim \mathsf{P}\big(R > m\big)\, \big(r_1^{-\rho+1} - r_2^{-\rho+1}\big)\, \mathsf{P}\big(\boldsymbol{\Theta} \in \mathcal{B}\big).\end{aligned}$$

Accordingly, we have

$$\lim_{m\to\infty} \frac{\log \mathsf{P}\big(m^{-1}R\boldsymbol{\Lambda}\boldsymbol{\Theta} \in \bigcup_{i\in\mathcal{T}} \mathsf{D}_i\big)}{\log \mathsf{P}\big(R > m\big)} \leq 1.$$

Hence, the result follows since $\lim_{m\to\infty} \log \mathsf{P}\big(R > m\big)/\log m = -\rho + 1$. $\square$

Appendix EC.3: Characterizing Solvency Regions

Define the solvency region of bank i as $\mathsf{S}_i \coloneqq [\mathbf{0}, \mathbf{c}] \setminus \mathsf{D}_i$ for $i = 1, 2, \ldots, n$. Then, the next corollary is a trivial consequence of Lemma 1.

COROLLARY EC.1. *Suppose that $x_n < w_n$. Then $\mathbf{x} \in \mathsf{S}_n$ if and only if $\mathbf{x} \in [0, \mathbf{c}]$ and*

$$\frac{\mathbf{x}_{-n} - \mathbf{w}_{-n}}{w_n - x_n} \in \mathsf{Q}_n^*$$

where Q_n^* is the polar set of Q_n defined by $\mathsf{Q}_n^* = \{\mathbf{z} | \boldsymbol{\zeta}^\top \mathbf{z} \leq 1, \forall \boldsymbol{\zeta} \in \mathsf{Q}_n\}$.

Based on the above corollary, we describe our result for bank n, still retaining the full generality. We further impose the next assumption for convenience sake. It can be done for a general network by renumbering banks.

ASSUMPTION EC.1. *There exists an integer m with $1 \leq m \leq n-1$ such that $a_{in} \neq 0$ for $i = 1, \ldots, m$ but $a_{in} = 0$ for $i = m+1, \ldots, n-1$.*

In the next theorem, we now see how the network information contained in the matrix

$$(\mathbf{I} - \mathbf{A}_{-n})^\top \begin{pmatrix} \operatorname{diag}(\mathbf{a}_{\mathcal{I}})^{-1} & 0 \\ 0 & \mathbf{I} \end{pmatrix}$$

affects a target bank's survivability.

THEOREM EC.4. *Suppose that Assumption EC.1 holds. We further assume* $x_n < w_n$ *and* $\eta = 0$. *Then,* $\mathbf{x} \in \mathsf{S}_n$ *if and only if* $\mathbf{x} \in [0, c]$

$$\mathbf{y} := \frac{\mathbf{x}_{-n} - \mathbf{w}_{-n}}{w_n - x_n} \in (\mathbf{I} - \mathbf{A}_{-n})^\top \begin{bmatrix} \operatorname{diag}\left(\mathbf{a}^n_{\{1,\dots,m\}}\right)^{-1} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \triangle_m \times \mathbb{R}_+^{n-m-1} - \mathbb{R}_+^{n-1}.$$

where $\triangle_m = \{\mathbf{z} \in \mathbb{R}_+^m | \mathbf{1}^\top \mathbf{z} \le 1\}$. *Here, the set* $U - V$ *for* $U, V \subset \mathbb{R}^{n-1}$ *is the set of vectors of form* $\mathbf{u} - \mathbf{v}$ *with* $\mathbf{u} \in U$ *and* $\mathbf{v} \in V$.

Proof. We prove this theorem in three steps.

Step 1. Let us denote $\{1, \dots, m\}$ by $\mathcal{I}$ and $\{1, \dots, n-1\} \backslash \mathcal{I}$ by $\mathcal{J}$ for brevity. We set

$$\mathbf{I} - \mathbf{A}_{-n} = \begin{pmatrix} \mathbf{M}^{\mathcal{I}} \\ \mathbf{M}^{\mathcal{J}} \end{pmatrix}$$

where $\mathbf{M}^{\mathcal{I}}$ is the matrix consisting of the first m rows of $\mathbf{I} - \mathbf{A}_{-n}$ and $\mathbf{M}^{\mathcal{J}}$ is the rest of $\mathbf{I} - \mathbf{A}_{-n}$. Define

$$\mathbf{M}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}} = \begin{pmatrix} \operatorname{diag}(\mathbf{a}_{\mathcal{I}})^{-1} \mathbf{M}^{\mathcal{I}} \\ \operatorname{diag}(\boldsymbol{\varepsilon})^{-1} \mathbf{M}^{\mathcal{J}} \\ -\operatorname{diag}(\boldsymbol{\tau})^{-1} \end{pmatrix},$$

where $\boldsymbol{\varepsilon}$ and $\boldsymbol{\tau}$ are strictly positive real vectors decreasing to zero. Here the operation diag transforms a given vector into a diagonal matrix with the same diagonal entries. Then, by letting $\mathsf{Q}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}} := \{\mathbf{z} \,|\, \mathbf{M}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}} \mathbf{z} \le \mathbf{1}\}$, we observe that

$$\mathsf{Q}_n = \lim_{\boldsymbol{\varepsilon},\boldsymbol{\tau} \downarrow 0} \mathsf{Q}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}}$$

if we understand the limit as a decreasing limit of sets.

Step 2. The following result based on the polar duality is well known (see for example Lemma 3.2 of Faigle et al. (2002)). For a nonzero $k \times d$ matrix M, it holds that

$$\left\{\mathbf{x} \in \mathbb{R}^d \,\middle|\, \mathbf{M}\mathbf{x} \le \mathbf{1}\right\}^* = \left\{\mathbf{x} \in \mathbb{R}^d \,\middle|\, \mathbf{x} = \mathbf{M}^\top \mathbf{w} \text{ for some } \mathbf{y} \text{ such that } \mathbf{1}^\top \mathbf{y} = 1, \mathbf{y} \ge \mathbf{0}\right\}.$$

This implies that the polar dual of $\mathsf{Q}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}}$ consists of vectors of form

$$(\mathbf{M}^{\mathcal{I}})^\top \operatorname{diag}(\mathbf{a}_{\mathcal{I}})^{-1} \mathbf{u} + (\mathbf{M}^{\mathcal{J}})^\top \operatorname{diag}(\boldsymbol{\varepsilon})^{-1} \mathbf{v} - \operatorname{diag}(\boldsymbol{\tau})^{-1} \mathbf{w} \tag{EC.24}$$

with $(\mathbf{u},\mathbf{v},\mathbf{w}) \in \mathbb{R}_+^{|\mathcal{I}|} \times \mathbb{R}_+^{|\mathcal{J}|} \times \mathbb{R}_+^{n-1}$ and $\mathbf{1}^\top(\mathbf{u},\mathbf{v},\mathbf{w}) = 1$.

We next claim that the polar dual of Q_n is given by

$$\mathsf{Q}_n^* = \overline{\bigcup_{\boldsymbol{\varepsilon},\boldsymbol{\tau}} \mathsf{Q}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}}^*}.$$

Since $\mathsf{Q}_n \subset \mathsf{Q}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}}$, we have $\mathsf{Q}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}}^* \subset \mathsf{Q}_n^*$ for all $\boldsymbol{\varepsilon}$ and $\boldsymbol{\tau}$. As a result,

$$\bigcup_{\boldsymbol{\varepsilon},\boldsymbol{\tau}} \mathsf{Q}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}}^* \subset \mathsf{Q}_n^*.$$

Since the polar set of any given set is convex and closed, Q_n^* is convex and closed. Hence, the closure of the left side is still a subset of Q_n^*.

To show the converse, let us consider the following set

$$\left(\mathsf{Q}_n^*\right)^\circ = \left\{ \mathbf{x} \;\middle|\; \max_{\boldsymbol{\zeta} \in \mathsf{Q}_n} \boldsymbol{\zeta}^\top \mathbf{x} < 1 \right\} \subset \mathsf{Q}_n^*.$$

This is nonempty ($\mathbf{0} \in (\mathsf{Q}_n^*)^\circ$) and open. To see this, for any $\mathbf{x}$ in the set, consider a small ball B with radius r. Then, we observe that

$$\max_{\boldsymbol{\zeta} \in \mathsf{Q}_n} \boldsymbol{\zeta}^\top (\mathbf{x} + \mathbf{v}) \le \max_{\boldsymbol{\zeta} \in \mathsf{Q}_n} \boldsymbol{\zeta}^\top \mathbf{x} + r \times \max_{\boldsymbol{\zeta} \in \mathsf{Q}_n} |\boldsymbol{\zeta}|$$

for any $\mathbf{v} \in \mathsf{B}$. Since Q_n is bounded, it is enough to choose a sufficiently small r that makes the right-hand side less than 1. On the other hand, it is clear that $\overline{(\mathsf{Q}_n^*)^\circ} = \mathsf{Q}_n^*$; we simply consider the sequence $\mathbf{x}_m = (1 - m^{-1})x$ for $\mathbf{x} \in \mathsf{Q}_n^*$ and check $\mathbf{x}_m \in (\mathsf{Q}_n^*)^\circ$.

Next, consider the following function

$$g(\boldsymbol{\varepsilon},\boldsymbol{\tau}) \coloneqq \max_{\boldsymbol{\zeta} \in \mathsf{Q}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}}} \boldsymbol{\zeta}^\top \mathbf{x}.$$

Let us fix $\mathbf{x} \in (\mathsf{Q}_n^*)^\circ$ with $\max_{\boldsymbol{\zeta} \in \mathsf{Q}_n} \boldsymbol{\zeta}^\top \mathbf{x} = 1 - \delta$ for some $\delta > 0$. Since the objective function is linear and the domain $\mathsf{Q}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}}$ continuously decreases to Q_n, there exists $(\boldsymbol{\varepsilon},\boldsymbol{\tau})$ such that $g(\boldsymbol{\varepsilon},\boldsymbol{\tau}) \le 1 - \delta/2 < 1$, which implies $\mathbf{x} \in \mathsf{Q}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}}^*$. Consequently,

$$(\mathsf{Q}_n^*)^\circ \subset \bigcup_{\boldsymbol{\varepsilon},\boldsymbol{\tau}} \mathsf{Q}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}}^* \Rightarrow \overline{(\mathsf{Q}_n^*)^\circ} \subset \overline{\bigcup_{\boldsymbol{\varepsilon},\boldsymbol{\tau}} \mathsf{Q}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}}^*}.$$

The claim is now proved.

Step 3. From the previous steps, we only need to find the closure of the set $\bigcup_{\boldsymbol{\varepsilon},\boldsymbol{\tau}} \mathsf{Q}^*_{\boldsymbol{\varepsilon},\boldsymbol{\tau}}$. For any $\mathbf{u} \in \triangle_{|\mathcal{I}|}$, we have the constraint $\mathbf{1}^\top(\mathbf{v},\mathbf{w}) = 1 - \mathbf{1}^\top\mathbf{u}$ and $\mathbf{v},\mathbf{w} \geq \mathbf{0}$. However, by scaling up arbitrarily by $\boldsymbol{\varepsilon},\boldsymbol{\tau}$, we include the nonnegative rays of $(\mathbf{v},\mathbf{w})$. This leads to

$$\mathbf{y} \in \mathsf{Q}^*_n \Leftrightarrow \mathbf{y} = (\mathbf{M}^{\mathcal{I}})^\top \mathrm{diag}(\mathbf{a}_{\mathcal{I}})^{-1}\mathbf{u} + (\mathbf{M}^{\mathcal{J}})^\top \mathbf{v} - \mathbf{w}$$

for some $\mathbf{u} \in \Delta_{|\mathcal{I}|}$, $\mathbf{v} \in \mathbb{R}_+^{|\mathcal{J}|}$, and $\mathbf{w} \in \mathbb{R}_+^{n-1}$. But note that if $\mathbf{1}^\top\mathbf{u} = 1$, then $\mathbf{v} = \mathbf{w} = \mathbf{0}$. Hence, more precisely,

$$(\mathbf{u},\mathbf{v},\mathbf{w}) \in \left\{\mathbf{x} \in \mathbb{R}_+^{|\mathcal{I}|} \;\middle|\; \mathbf{1}^\top\mathbf{x} < 1\right\} \times \left\{\mathbb{R}_+^{|\mathcal{J}|} \times \mathbb{R}_+^{n-1} \backslash (\mathbf{0},\mathbf{0})\right\} \cup \left\{\mathbf{x} \in \mathbb{R}_+^{|\mathcal{I}|} \;\middle|\; \mathbf{1}^\top\mathbf{x} = 1\right\} \times \{(\mathbf{0},\mathbf{0})\}.$$

The set of vectors generated by this set via (EC.24) is $\bigcup_{\boldsymbol{\varepsilon},\boldsymbol{\tau}} \mathsf{Q}^*_{\boldsymbol{\varepsilon},\boldsymbol{\tau}}$. However, it is clear that we can take a sequence of $(\mathbf{u},\mathbf{v},\mathbf{w})$ converging to an arbitrarily chosen element of $\Delta_{|\mathcal{I}|} \times \mathbb{R}_+^{|\mathcal{J}|} \times \mathbb{R}_+^{n-1}$. Hence,

$$\mathsf{Q}^*_n = \left\{\mathbf{y} \;\middle|\; \mathbf{y} = (\mathbf{I} - \mathbf{A}_{-n})^\top \begin{pmatrix} \mathrm{diag}(\mathbf{a}_{\mathcal{I}})^{-1} & 0 \\ 0 & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \mathbf{w}, (\mathbf{u},\mathbf{v},\mathbf{w}) \in \Delta_{|\mathcal{I}|} \times \mathbb{R}_+^{|\mathcal{J}|} \times \mathbb{R}_+^{n-1}\right\}.$$

The statement is then immediate using the previous corollary. □

Example EC.1. In general, the default region D_i or equivalently the solvency region S_i is not easy to picture, nor computationally simple due to exponentially many half-spaces. Rather, the results in Section EC.3 based on duality are helpful in visualizing the solvency region. For example, let us consider a network with $n = 3$ with $a_{13}, a_{23} > 0$. According to Theorem EC.4 in Section EC.3, the solvency region for bank 3, S_3, is equivalent to the region of $\mathbf{x} \in [\mathbf{0},\mathbf{c}]$ satisfying

$$x_3 < w_3 \quad \text{and} \quad \mathbf{y} \coloneqq \frac{\mathbf{x}_{-3} - \mathbf{w}_{-3}}{w_3 - x_3} \in \begin{bmatrix} \frac{1}{a_{13}} & -\frac{a_{21}}{a_{23}} \\ -\frac{a_{12}}{a_{13}} & \frac{1}{a_{23}} \end{bmatrix} \triangle_2 - \mathbb{R}_+^2 =: \mathsf{Q}^*_3.$$

Using this Q^*_3, it is easy to see how S_3 is affected by the relative liabilities $\mathbf{A}$. See Figure EC.2 for an illustration. For this, we use the setting in Example EC.3 and utilize $\gamma\mathbf{A}$ instead of $\mathbf{A}$ where γ is a positive multiplying factor. As γ gets bigger, a larger proportion of liabilities are concentrated

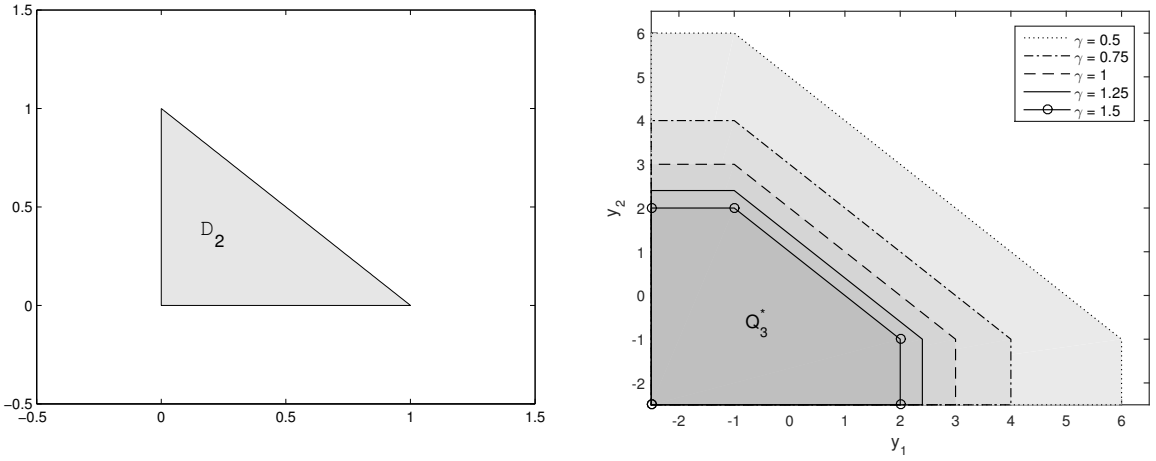


Figure EC.2 The 2-dimensional simplex (left) is transformed to the polar dual Q_3^* (right) for the normalized shock $\mathbf{y}$: $a_{ij} = 1/3$ for all $i \neq j$ and $\gamma = 0.5, 0.75, 1, 1.25, 1.5$.

on the inside of the network. The left side of Figure EC.2 represents the 2-dimensional simplex Δ_2, and the right side shows the polar dual Q_3^* for different γ values: 0.5, 0.75, 1, 1.25, and 1.5. We observe that Q_3^* gets smaller as γ increases. This is because more shocks are propagated as the interconnectedness between banks increases.

Appendix EC.4: Supplementary Numerical Results

EC.4.1. Results using the Full EBA Dataset

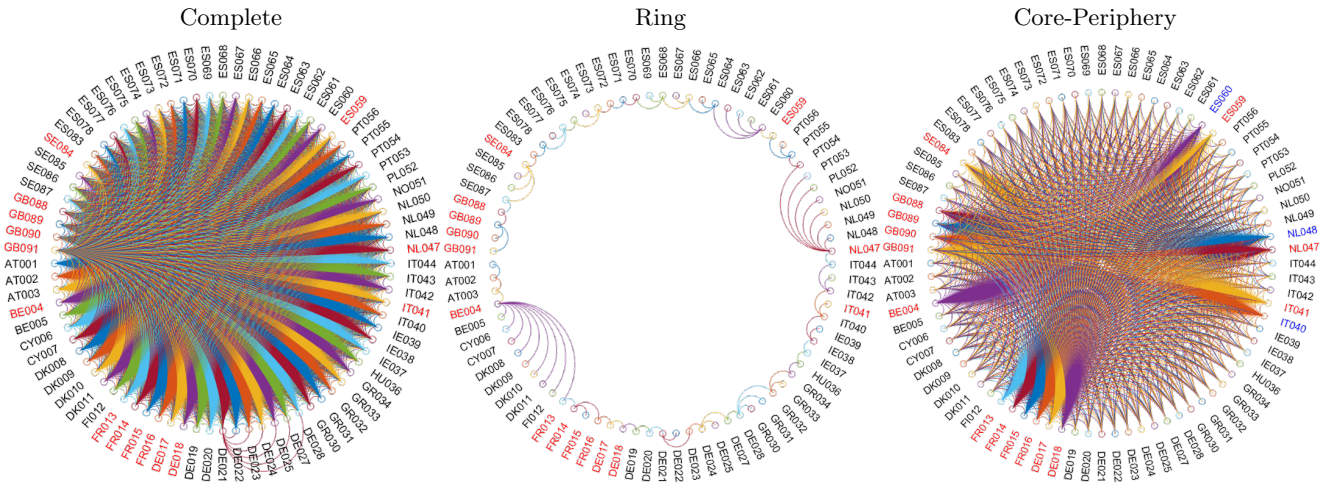
In this subsection, we conduct numerical experiments similar to those in Section 6 using the full dataset of the 2011 EBA stress test. We find that some small banks have problematic data, and hence, we omit the ten smallest banks as a simple rule. The dataset of the remaining 80 banks is given in Table EC.1. Note that for EAD, Section 6 uses the domestic exposure at default as in Chen et al. (2016) since the network is composed of only German banks, whereas this section uses the total exposure at default as in Glasserman and Young (2015). Our target banks include BE004, FR013, FR014, FR015, FR016, DE017, DE018, IT041, NL047, ES059, SE084, GB088, GB089, GB090, and GB091 since they were identified as G-SIBs by the Financial Stability Board in 2011. Using the same network reconstruction method as in Section 6, we form complete, ring-like, and

Table EC.1 Data of the largest 80 banks based on the 2011 EBA EU-wide stress test.

Code	Bank Name	Total Assets	EAD	Equity ($\mathbf{w}$)	External Assets ($\mathbf{c}$)
AT001	ERSTE BANK GROUP (EBG)	205.9	25.0	10.5	180.9
AT002	RAIFFEISEN BANK INTERNATIONAL (RBI)	131.2	30.4	7.6	100.8
AT003	OESTERREICHISCHE VOLKSBANK AG	44.7	10.8	1.8	34.0
BE004	DEXIA	548.1	228.2	17.0	319.9
BE005	KBC BANK	276.7	23.9	11.7	252.9
CY006	MARFIN POPULAR BANK PUBLIC CO LTD	42.6	7.9	2.0	34.7
CY007	BANK OF CYPRUS PUBLIC CO LTD	42.0	7.3	2.1	34.7
DK008	DANSKE BANK	402.6	75.9	14.6	326.7
DK009	JYSKE BANK	32.8	4.7	1.7	28.1
DK010	SYDBANK	20.2	3.7	1.2	16.6
DK011	NYKREDIT	175.9	8.6	6.6	167.3
FI012	OP-POHJOLA GROUP	74.7	8.2	5.2	66.5
FR013	BNP PARIBAS	1,998.2	90.3	55.4	1,907.8
FR014	CREDIT AGRICOLE	1,503.6	83.7	46.3	1,419.9
FR015	BPCE	1,000.7	35.0	31.9	965.7
FR016	SOCIETE GENERALE	1,051.3	100.0	27.8	951.3
DE017	DEUTSCHE BANK AG	1,905.6	194.4	30.4	1,711.2
DE018	COMMERZBANK AG	771.2	138.2	26.7	633.0
DE019	LANDESBANK B-W	374.4	133.9	9.8	240.5
DE020	DZ BANK AG DT.Z-G	323.6	135.9	7.3	187.7
DE021	BAYERISCHE LANDESBANK	316.4	97.3	11.5	219.0
DE022	NORDDEUTSCHE LANDESBANK-GZ	228.6	91.2	4.0	137.4
DE023	HYPO REAL ESTATE HOLDING AG	328.1	29.1	5.5	299.0
DE024	WESTLB AG, DÜSSELDORF	191.5	58.1	4.2	133.4
DE025	HSH NORDBANK AG	150.9	9.5	4.4	141.4
DE027	LANDESBANK BERLIN AG	133.9	49.3	5.2	84.6
DE028	DEKABANK DEUTSCHE GIROZENTRALE	130.3	41.3	3.4	89.0
GR030	EFG EUROBANK ERGASIAS S.A.	85.9	3.8	4.3	82.0
GR031	NATIONAL BANK OF GREECE	118.8	8.6	8.2	110.2
GR032	ALPHA BANK	66.8	3.5	5.3	63.3
GR033	PIRAEUS BANK GROUP	57.7	1.6	3.0	56.1
GR034	AGRICULTURAL BANK OF GREECE S.A.	31.2	1.7	0.8	29.6
HU036	OTP BANK NYRT.	35.2	2.5	3.3	32.7
IE037	ALLIED IRISH BANKS PLC	131.3	11.3	3.7	120.0
IE038	BANK OF IRELAND	156.7	17.3	7.0	139.5
IE039	IRISH LIFE AND PERMANENT	46.7	6.1	1.7	40.6
IT041	UNICREDIT S.P.A	929.5	106.7	35.7	822.8
IT040	INTESA SANPAOLO S.P.A	577.0	109.9	26.2	467.1
IT042	BANCA MONTE DEI PASCHI DI SIENA S.P.A	244.3	12.1	6.3	232.2
IT043	BANCO POPOLARE - S.C.	140.0	7.6	5.5	132.4
IT044	UNIONE DI BANCHE ITALIANE SCPA	130.6	19.8	6.6	110.8
NL047	ING BANK NV	933.1	111.8	30.9	821.3
NL048	RABOBANK NEDERLAND	607.5	37.5	27.7	569.9
NL049	ABN AMRO BANK NV	379.6	29.2	11.6	350.4
NL050	SNS BANK NV	78.9	0.4	1.8	78.5
NO051	DNB NOR BANK ASA	210.0	7.0	9.7	202.9
PL052	PKO BANK POLSKI	35.5	2.2	4.2	33.4
PT053	CAIXA GERAL DE DEPÓSITOS, SA	119.3	14.2	6.5	105.1
PT054	BANCO COMERCIAL PORTUGUÊS	100.0	7.7	3.5	92.3
PT055	ESPÍRITO SANTO FINANCIAL GROUP	85.6	8.7	4.5	77.0
PT056	BANCO BPI	43.8	5.5	2.1	38.4
ES059	BANCO SANTANDER S.A.	1,223.3	51.4	42.0	1,171.9
ES060	BANCO BILBAO VIZCAYA ARGENTARIA S.A.	540.9	110.5	24.9	430.5
ES061	BFA-BANKIA	327.9	39.5	13.9	288.4
ES062	CAJA DE AHORROS Y PENSIONES DE BARCELONA	275.9	5.5	11.1	270.3
ES063	EFFIBANK	54.5	4.1	2.7	50.4
ES064	BANCO POPULAR ESPAÑOL, S.A.	129.2	14.8	6.7	114.4
ES065	BANCO DE SABADELL, S.A.	96.7	3.7	3.5	93.0
ES066	CAIXA D'ESTALVIS DE CATALUNYA	76.0	8.2	3.1	67.8
ES067	CAIXA DE AFORROS DE GALICIA	73.3	2.9	2.8	70.4
ES068	GRUPO BMN	69.8	7.7	3.3	62.1
ES069	BANKINTER, S.A.	53.5	2.1	1.9	51.3
ES070	CAJA ESPAÑA DE INVERSIONES	45.7	7.2	2.1	38.4
ES071	GRUPO BANCA CIVICA	71.1	7.4	3.7	63.6
ES072	CAJA DE AHORROS Y M.P. DE ZARAGOZA	42.7	1.8	2.3	40.9
ES073	MONTE DE PIEDAD Y CAJA DE AHORROS DE RONDA	34.3	2.6	2.5	31.7
ES074	BANCO PASTOR, S.A.	31.1	1.7	1.4	29.5
ES075	GRUPO BBK	44.6	1.9	3.0	42.7
ES076	CAIXA D'ESTALVIS UNIO DE CAIXES DE MANLLEU	28.3	1.8	1.1	26.5
ES077	CAJA DE AHORROS Y M.P. DE GIPUZKOA Y SAN SEBASTIAN	20.8	0.3	1.9	20.5
ES078	GRUPO CAJA3	20.1	1.9	1.2	18.2
ES083	CAJA DE AHORROS DEL MEDITERRANEO	72.0	5.0	1.8	67.1
SE084	NORDEA BANK AB	542.9	61.4	19.1	481.4
SE085	SKANDINAVISKA ENSKILDA BANKEN AB	212.2	26.0	9.6	186.3
SE086	SVENSKA HANDELSBANKEN AB	240.2	20.9	8.2	219.3
SE087	SWEDBANK AB	191.4	17.4	7.4	174.0
GB088	ROYAL BANK OF SCOTLAND GROUP PLC	607.4	105.5	59.0	501.8
GB089	HSBC HOLDINGS PLC	1,783.2	212.1	86.9	1,571.1
GB090	BARCLAYS PLC	1,725.7	53.9	46.2	1,671.8
GB091	LLOYDS BANKING GROUP PLC	1,006.1	29.2	48.0	976.8

All quantities are exhibited in billions of euros.

Figure EC.3 The recovered interbank network structures from the EBA stress test data.

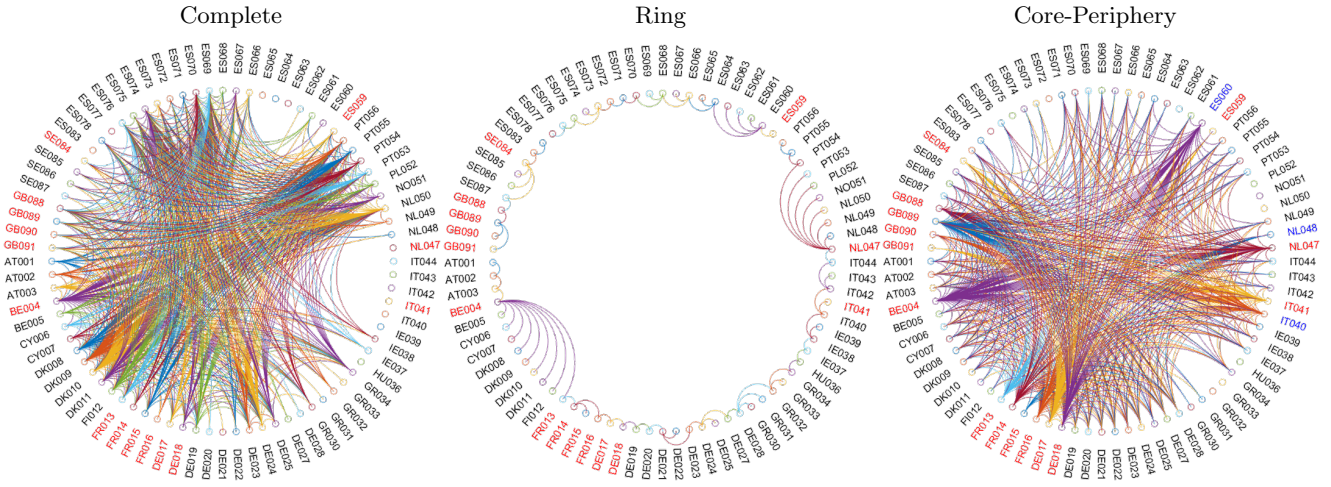


Note. The three structures represent complete, ring-like, and core-periphery networks, respectively. For each structure, nodes represent individual banks, and edges stand for their interbank exposures. Red nodes represent target banks, and blue nodes denote core banks except target banks.

core-periphery networks, which are illustrated in Figure EC.3. In the case of the core-periphery network, banks with total assets larger than 500 billion euros are selected as core banks. Further, we consider six different network information: full network information, Top 5 banks' information, SIFIs' information, large exposures' information, link information, and aggregate information when computing the worst-case quantities. For the top 5 banks' information, we select FR013, FR014, DE017, GB089, and GB090 based on the size of total assets. As noted in Section 4, we identify large exposures whose size is greater than or equal to 10% of the corresponding bank's equity; Figure EC.4 exhibit large exposures in the three networks in Figure EC.3.

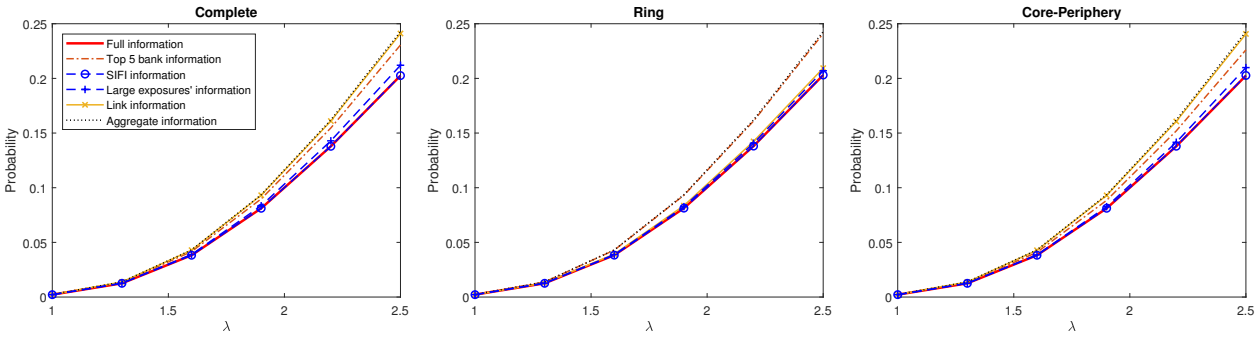
In Figures EC.5 and EC.6, under different network information and network structures, we illustrate the worst-case default probabilities and risk capitals with $\alpha = 0.95$. We assume that the random shocks $X_1, \ldots, X_{11}$ follow Pareto distributions: for each i, the probability density function of X_i is given by $f_i(x) = \theta_i^{-1}(1 + \lambda x/\theta_i)^{-(1/\lambda+1)}$, where $\theta_i = c_i/c_1$ and $\lambda \in [1, 2.5]$. It is also assumed that the shocks are independent and constrained to be in $[\mathbf{0}, \mathbf{c}]$ as in Section 6. We arbitrarily set $\epsilon = 0.0005$ for the link information and $\eta = 0$ for all cases. Unlike the results in Section 6, there exists a slight gap between the true quantities and the worst-case quantities based on the large

Figure EC.4 Large exposures in the recovered interbank network structures from the EBA stress test data.



Note. The lines in the three figures represent large exposures in the complete, ring-like, and core-periphery networks in Figure EC.3, respectively. Red nodes represent target banks, and blue nodes denote core banks except target banks.

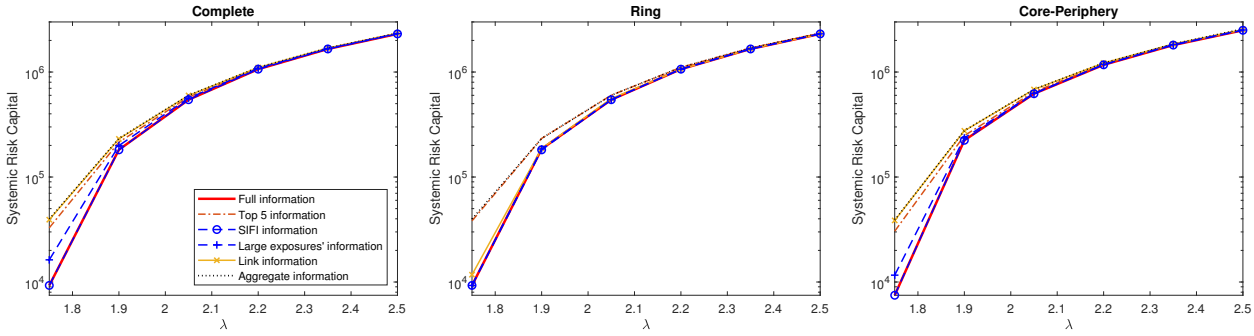
Figure EC.5 The worst-case default probabilities under different network information and different network structures.



Note. The Monte Carlo method with 10^5 replications is used for the estimation of the probabilities. The legend in the left subfigure applies to all the subfigures.

exposures' information because each exposure gets smaller in relative terms as the network size grows. We can also observe from Figures EC.3 and EC.4 that not many exposures are classified as large exposures. Nonetheless, since the gap is limited, this information is still good enough to approximate the true quantities, and thus, it demonstrates the effectiveness of collecting large exposures for regulatory purposes. Since the other results have the same patterns as the associated results in Section 6, we omit the explanation.

Figure EC.6 The worst-case risk capitals under different network information and different network structures. when $\alpha = 0.95$



Note. The number of samples (N) is 2,000. All subfigures share the same legend in the left subfigure.

EC.4.2. Examples of Shock Contagion

We introduce two examples to help the reader better understand the shock propagation and the default region, respectively. In Example EC.2, using two simple networks, we exemplify the procedure of contagion of shocks in the Eisenberg-Noe network model. In Example EC.3, we illustrate how the default region of a specific bank changes with respect to its net worth.

EXAMPLE EC.2. Let us first consider a network of n banks with the relative liability matrix $\mathbf{A}$ with $a_{ij} > 0$ if and only if $j = i+1$ for $i = 1, 2, \ldots, n-1$ as described in the left side of Figure EC.7. Assume that $\eta = 0$. Careful counting would tell us that $\mathsf{Q}_n = \{\boldsymbol{\zeta} | \zeta_1 \le \zeta_2 \le \cdots \le \zeta_{n-1},\ \zeta_i = \prod_{j=i}^{n-1} a_{j(j+1)}$ or $0,\ i = 1, \ldots, n-1\}$. It is then not difficult to see that the total shock to bank n is

$$\Phi_n(\mathbf{x}) = x_n + \Big(\big((x_1 - w_1)^+ a_{12} + x_2 - w_2 \big)^+ a_{23} + \cdots + x_{n-1} - w_{n-1} \Big)^+ a_{(n-1)n}.$$

It shows that banks 1 to $n-1$ sequentially add shocks if there is an overflow of shocks for each bank. For the case of a star network in the right side of Figure EC.7, let $n = 2m+1$ for some m. This network consists of one hub 1, debtors $\{2, 4, \ldots, 2m\}$, and creditors $\{3, 5, \ldots, 2m+1\}$. Then, it is easy to see that $\mathsf{Q}_n = \{\mathbf{0}\} \bigcup \{\boldsymbol{\zeta} | \zeta_1 = a_{1n},\ \zeta_3 = \zeta_5 = \cdots = \zeta_{2m-1} = 0,\ \zeta_{2k} = a_{(2k)1} a_{1n}$ or $0,\ k = 1, \ldots, m\}$. Thus, the total shock to bank n is equal to

$$\Phi_n(\mathbf{x}) = x_n + a_{1n} \left(x_1 - w_1 + \sum_{k=1}^{m} a_{(2k)1} (x_{2k} - w_{2k})^+ \right)^+ .$$

Figure EC.7 A ring network with n banks (left) and a star network with an odd number (n) of banks (right).

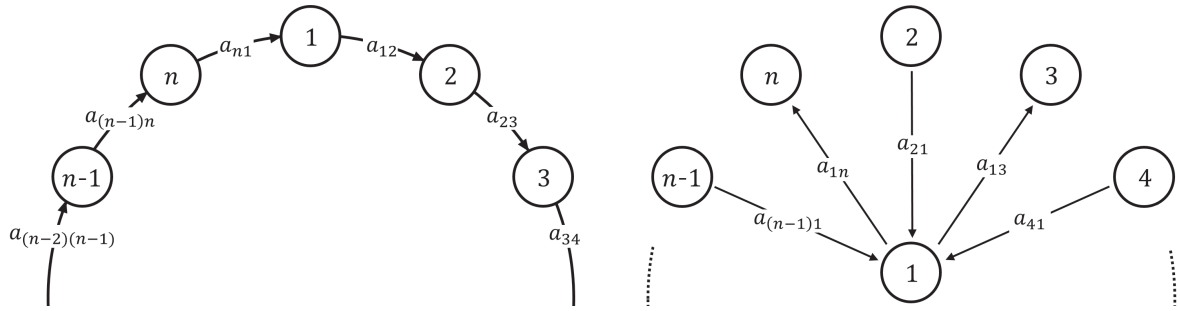
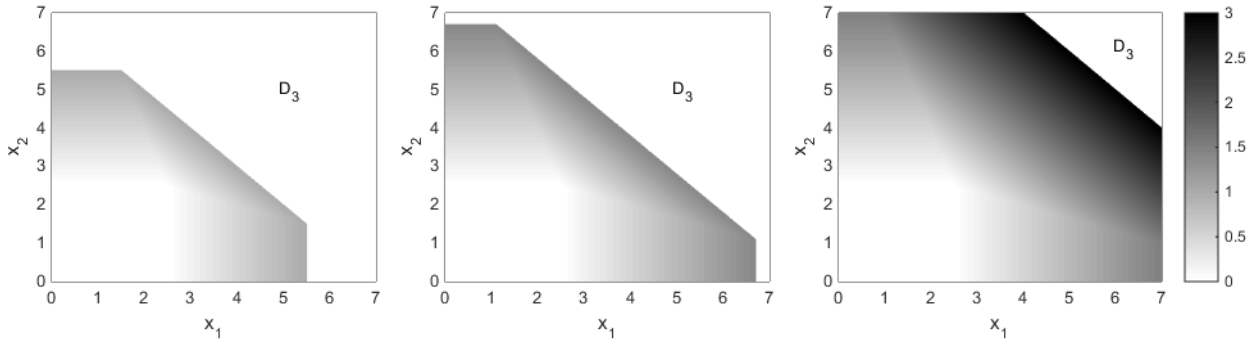


Figure EC.8 A graphical illustration of $\Phi_3(\mathbf{x})$ in a 3-bank network.



Note that the summation inside the second term is the indirect shock to bank 1 from other banks. This shows that the fraction a_{1n} of the excess shock at bank 1 affects bank n if bank 1 cannot afford to cover shocks, whereas it takes two steps for a shock at bank $2k$ to affect bank n.

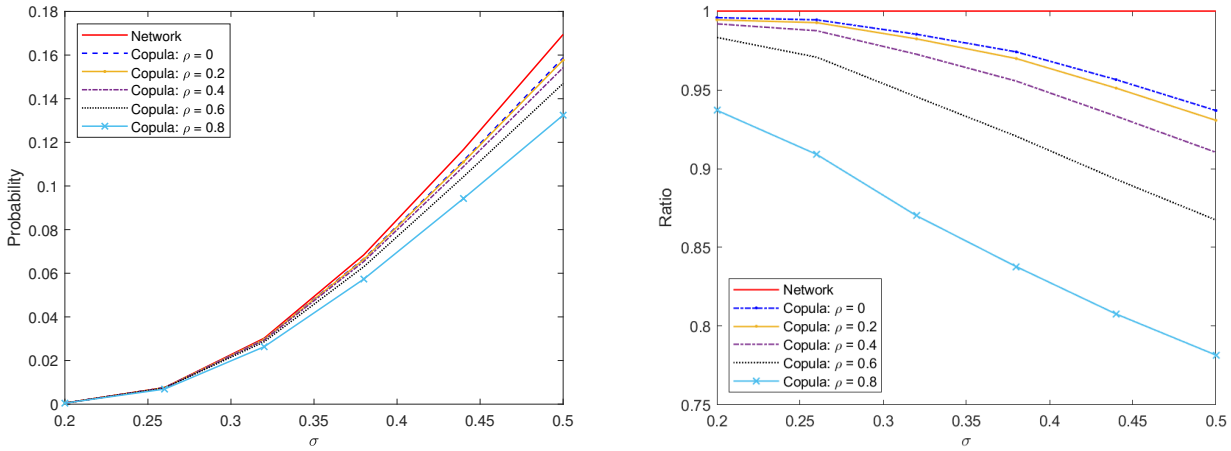
EXAMPLE EC.3. Figure EC.8 presents the amount of the total shock $\Phi_3(\mathbf{x})$ for $w_3 = 1, 1.4$, and 3 with respect to x_1 and x_2 in the case of a three-bank network with $c_1 = c_2 = 7$, $w_1 = w_2 = 2.5$, and

$$\mathbf{A} = \begin{bmatrix} 0 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 \\ 1/3 & 1/3 & 0 \end{bmatrix}.$$

Since our interest lies in the indirect shock, we assume $x_3 = 0$ and $\eta = 0$ in this example. The default region D_3 gets smaller for increasing w_3, and $\Phi_3(\mathbf{x})$ is zero when $x_1 < w_1$ and $x_2 < w_2$. The further the point (x_1, x_2) is away from the origin, the bigger the total shock is. Also, the shaded region can be divided into 3 parts: a trapezoid adjacent to x_1-axis, a trapezoid adjacent to x_2-axis, and the other part. Those three parts correspond to (6) when $\mathcal{I}$ is $\{1\}, \{2\}$, and $\{1, 2\}$, respectively.

EC.4.3. Auxiliary Numerical Results

Figure EC.9 The default probability (3) for $\mathcal{T}=\{1,2\}$ in Example 1 with Gaussian copula models.



Note. We assume the 5-bank homogeneous financial network in Example 1 and apply the same shock distribution used in that example to estimate the marginal default probability as well as the probability (3) under our network model. The left panel illustrates the banks' default probabilities under our network model and the Gaussian copula models with different correlation coefficients ρ, whereas the right panel describes the ratios of the default probabilities associated with the copula models to the probability under our network model.

Table EC.2 The matrices of relative liabilities for the complete network.

	DE017	DE018	DE019	DE020	DE021	DE022	DE023	DE024	DE025	DE027	DE028
DE017	0	0.0026	0.0053	0.0060	0.0036	0.0029	0.0004	0.0012	0.0002	0.0014	0.0015
DE018	0.0065	0	0.0143	0.0162	0.0097	0.0077	0.0010	0.0032	0.0006	0.0037	0.0041
DE019	0.0274	0.0291	0	0.0678	0.0405	0.0325	0.0043	0.0133	0.0025	0.0154	0.0174
DE020	0.0357	0.0381	0.0782	0	0.0530	0.0425	0.0056	0.0174	0.0032	0.0202	0.0227
DE021	0.0222	0.0236	0.0485	0.0549	0	0.0263	0.0035	0.0108	0.0020	0.0125	0.0141
DE022	0.0241	0.0257	0.0527	0.0598	0.0357	0	0.0038	0.0117	0.0022	0.0136	0.0153
DE023	0.0022	0.0024	0.0048	0.0055	0.0033	0.0026	0	0.0011	0.0002	0.0012	0.0014
DE024	0.0118	0.0126	0.0258	0.0293	0.0175	0.0140	0.0018	0	0.0011	0.0067	0.0075
DE025	0.0028	0.0030	0.0062	0.0070	0.0042	0.0033	0.0004	0.0014	0	0.0016	0.0018
DE027	0.0200	0.0213	0.0437	0.0496	0.0296	0.0237	0.0031	0.0097	0.0018	0	0.0127
DE028	0.0228	0.0243	0.0498	0.0565	0.0337	0.0270	0.0036	0.0111	0.0021	0.0129	0

The shaded components represent the amount that can be inferred from the large exposures' information.

Table EC.3 The matrices of relative liabilities for the ring-like network.

	DE017	DE018	DE019	DE020	DE021	DE022	DE023	DE024	DE025	DE027	DE028
DE017	0	0.0170	0	0	0	0	0	0	0	0	0.0081
DE018	0.0428	0	0.0242	0	0	0	0	0	0	0	0
DE019	0	0.0494	0	0.2007	0	0	0	0	0	0	0
DE020	0	0	0.2314	0	0.0851	0	0	0	0	0	0
DE021	0	0	0	0.0883	0	0.1300	0	0	0	0	0
DE022	0	0	0	0	0.1764	0	0.0354	0.0327	0	0	0
DE023	0	0	0	0	0	0.0247	0	0	0	0	0
DE024	0	0	0	0	0	0.0392	0	0	0.0248	0.0642	0
DE025	0	0	0	0	0	0	0	0.0317	0	0	0
DE027	0	0	0	0	0	0	0	0.0934	0	0	0.1219
DE028	0.1201	0	0	0	0	0	0	0	0	0.1236	0

The shaded components represent the amount that can be inferred from the large exposures' information.

Table EC.4 The matrices of relative liabilities for the core-periphery network.

	DE017	DE018	DE019	DE020	DE021	DE022	DE023	DE024	DE025	DE027	DE028
DE017	0	0.0014	0.0027	0.0030	0.0056	0.0046	0.0007	0.0020	0.0004	0.0023	0.0026
DE018	0.0035	0	0.0072	0.0080	0.0149	0.0123	0.0018	0.0054	0.0010	0.0062	0.0069
DE019	0.0138	0.0146	0	0.0313	0.0585	0.0483	0.0070	0.0211	0.0041	0.0244	0.0272
DE020	0.0176	0.0187	0.0361	0	0.0750	0.0619	0.0090	0.0270	0.0052	0.0312	0.0349
DE021	0.0342	0.0363	0.0700	0.0778	0	0	0	0	0	0	0
DE022	0.0383	0.0407	0.0784	0.0871	0	0	0	0	0	0	0
DE023	0.0039	0.0041	0.0079	0.0088	0	0	0	0	0	0	0
DE024	0.0201	0.0213	0.0411	0.0457	0	0	0	0	0	0	0
DE025	0.0050	0.0053	0.0102	0.0113	0	0	0	0	0	0	0
DE027	0.0337	0.0358	0.0690	0.0767	0	0	0	0	0	0	0
DE028	0.0381	0.0405	0.0781	0.0868	0	0	0	0	0	0	0

The shaded components represent the amount that can be inferred from the large exposures' information.

References

Ahn D (2020) Shock amplification in financial networks with applications to the CCP feasibility. *Quantitative Finance* 20(7):1045–1056.

Asmussen S, Rojas-Nandayapa L (2008) Asymptotics of sums of lognormal random variables with Gaussian copula. *Statistics and Probability Letters* 78(16):2709–2714.

Dembo A, Zeitouni O (2009) *Large Deviations Techniques and Applications* (Springer).

Eisenberg L, Noe TH (2001) Systemic risk in financial systems. *Management Science* 47(2):236–249.

Faigle U, Kern W, Still G (2002) *Algorithmic Principles of Mathematical Programming* (Dordrecht, Netherlands: Springer).

Fang KT, Kotz S, Ng KW (1990) *Symmetric Multivariate and Related Distributions* (London, UK: Chapman and Hall).

Pagnoncelli BK, Ahmed S, Shapiro A (2009) Sample average approximation method for chance constrained programming: Theory and applications. *Journal of Optimization Theory and Applications* 142(2):399–416.

Resnick SI (2007) *Heavy-tail Phenomena: Probabilistic and Statistical Modeling* (Springer).