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 $\tilde{\mathsf{Q}}_n := \{\mathbf{u} \in \mathbb{R}_+^{n-1} | (\mathbf{I} - \mathbf{A}_{-n})\mathbf{u} \le \mathbf{a}_{-n}^n\}$. 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

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

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

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

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:

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

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

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

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

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

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

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:

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

Fix $\widetilde{\mathbf{A}}\in\mathcal{A}$ and assume $\mathcal{T}^c=\{1,\ldots,d\}$ 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 $\mathbf{u}\le((1+\eta)^{-1}\mathbf{I}-\widetilde{\mathbf{A}}_{\mathcal{T}^c})^{-1}\tilde{\mathbf{a}}_{\mathcal{T}^c}^i$. In addition, consider a Markov chain with $2n$ states and transition probabilities $\{q_{ij}\}_{i,j=1,\ldots,2n}$ defined by

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

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 $1-(((1+\eta)^{-1}\mathbf{I}-\widetilde{\mathbf{A}}_{\mathcal{T}^c})^{-1}\tilde{\mathbf{a}}_{\mathcal{T}^c}^i)_j$, 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

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

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

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

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 $\mathsf{C}=\widetilde{\mathsf{C}}:=\bigcap_{j=1}^n\{\boldsymbol{\xi}:(1+\eta)(\sum_{\{k:(j,k)\in\mathcal{K}\cup\widetilde{\mathcal{K}}\}}a_{jk}\xi_k+\tilde{\beta}_j\xi_l)\leq\xi_j$ for all l 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 $\bar{F}_i(x) := \mathsf{P}(X_i > x)$ 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 $\mathsf{Q}^{\mathcal{T}} := \bigcup_{i\in\mathcal{T}} \{\boldsymbol{\xi} \in \mathbb{R}^n_+ \mid \xi_i = 1,\ \xi_j = 0 \text{ for } j \in \mathcal{T}\setminus\{i\},\ \xi_k = u_k \text{ for } k \in \mathcal{T}^{\mathrm{c}} \text{ and } \mathbf{u} \in \mathsf{Q}_i^{\mathcal{T}}\}$, 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 $\mathbf{A} \in \mathcal{A}$, $\{i \mid \xi_i > 0,\ \boldsymbol{\xi} \in \mathsf{Q}^{\mathcal{T}}\} \subset \widetilde{\mathcal{H}} := \mathcal{H} \cup \mathcal{T}$. 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 $i \in \mathcal{T}$,

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

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 $\sigma_\circ^2 = \max_{i\in\widetilde{\mathcal{H}}}\ \big(\sigma_i^2 - (\boldsymbol{\Sigma}_{\widetilde{\mathcal{H}}\widetilde{\mathcal{H}}^c}\boldsymbol{\Sigma}^{-1}_{\widetilde{\mathcal{H}}^c\widetilde{\mathcal{H}}^c}\boldsymbol{\Sigma}_{\widetilde{\mathcal{H}}^c\widetilde{\mathcal{H}}})_{ii}\big)$ 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

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

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

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

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

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

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

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

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

where $\mathbf{X}'_{\widetilde{\mathcal{H}}}$ is lognormally distributed with parameters $\widetilde{\boldsymbol{\mu}} = \boldsymbol{\mu}_{\widetilde{\mathcal{H}}} + \boldsymbol{\Sigma}_{\widetilde{\mathcal{H}}\widetilde{\mathcal{H}}^c} \boldsymbol{\Sigma}^{-1}_{\widetilde{\mathcal{H}}^c\widetilde{\mathcal{H}}^c} (\log(\mathbf{x}') - \boldsymbol{\mu}_{\widetilde{\mathcal{H}}^c})$ and $\widetilde{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}_{\widetilde{\mathcal{H}}\widetilde{\mathcal{H}}} - \boldsymbol{\Sigma}_{\widetilde{\mathcal{H}}\widetilde{\mathcal{H}}^c} \boldsymbol{\Sigma}^{-1}_{\widetilde{\mathcal{H}}^c\widetilde{\mathcal{H}}^c} \boldsymbol{\Sigma}_{\widetilde{\mathcal{H}}^c\widetilde{\mathcal{H}}}$, $\sigma_\circ^2 = \max_{i\in\widetilde{\mathcal{H}}} \tilde{\sigma}_{ii}$, $\mu_\circ = \max_{i\in\widetilde{\mathcal{H}}:\tilde{\sigma}_{ii}=\sigma_\circ^2} \tilde{\mu}_i$, 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\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:

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

Define $G(\boldsymbol{\nu}, \mathbf{X}) := \max_{i\in\mathcal{T}}(\overline{\Phi}_i(\mathbf{X}) - w_i - \nu_i)$. 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), $\mathbf{1}_{\{\max_{i\in\mathcal{T}}\{\overline{\Phi}_i(\mathbf{x}^j) - w_i - \nu_i\} > 0\}} \le z_j$ for all j, and hence from its first constraint, we have

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

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 $z_j := \mathbf{1}_{\{\max_{i\in\mathcal{T}}\{\overline{\Phi}_i(\mathbf{x}^j)-w_i-\nu_i\}>0\}}$ for all j. Then, $\{\nu_i, z_j\}_{i\in\mathcal{T}, j=1,\ldots,N}$ 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 $\{q_{ij}\}_{i,j=1,\ldots,2n}$ defined by

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

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 on financial networks has been considered in the literature; see, e.g., Acemoglu et al. (2015), Amini and Minca (2016), Amini et al. (2016). Under a 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 on portfolio credit risk, which focus on the asymptotic behaviors of portfolio loss probabilities; see 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

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

Then, it is easy to see that its Fenchel-Legendre transform $\Xi^*(\mathbf{x}) := \sup_{\boldsymbol{\lambda}}\{\boldsymbol{\lambda}^\top\mathbf{x} - \Xi(\boldsymbol{\lambda})\} = (1/2)\mathbf{x}^\top\boldsymbol{\Sigma}^{-1}\mathbf{x}$. 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 $r:=\operatorname{corr}(X_1,X_2)$. 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

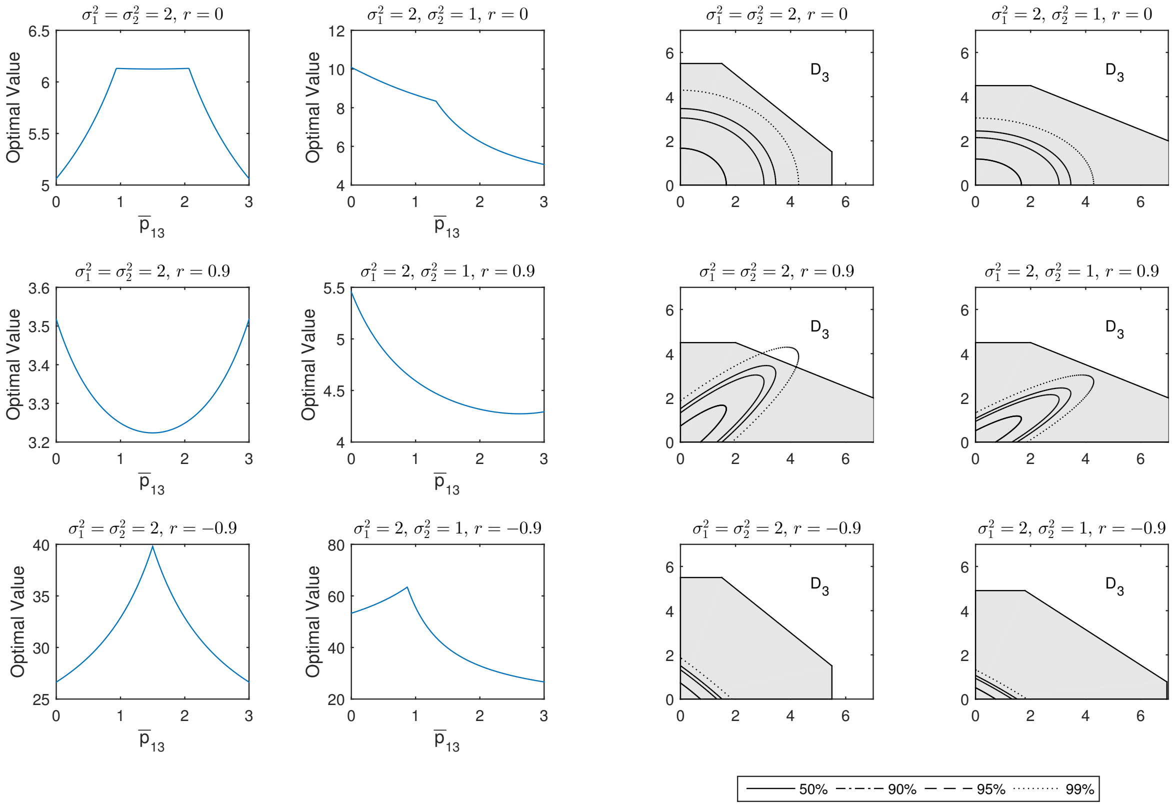


Figure EC.1 The optimal values of the optimization problem in (EC.21) with respect to $\bar{p}_{13}$ for different variances and correlations (left) and bank 3's default regions (unshaded) at the points of $(\bar{p}_{13}, \bar{p}_{23})$ where the optimal values are maximized (right)

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 $\mathsf{R}(\boldsymbol{\theta}) := \{r \geq 0 | r\boldsymbol{\Lambda}\boldsymbol{\theta} \in \bigcup_{i\in\mathcal{T}} \overline{\mathsf{D}}_i\}$ 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 $\varphi_m(\tau\boldsymbol{\Lambda}\boldsymbol{\theta}) := \sqrt{(\tau\boldsymbol{\Lambda}\boldsymbol{\theta} - \boldsymbol{\mu}/m)^\top \boldsymbol{\Sigma}^{-1} (\tau\boldsymbol{\Lambda}\boldsymbol{\theta} - \boldsymbol{\mu}/m)} \to \tau$ as m increases, $f_R(m\varphi_m(\tau\boldsymbol{\Lambda}\boldsymbol{\theta})) \sim f_R(m)\tau^{-\rho}$ as m increases. Thus, the asymptotic equivalence (EC.23) holds since

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

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:

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

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

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

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,

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

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 $\mathsf{S}_i := [\mathbf{0}, \mathbf{c}] \setminus \mathsf{D}_i$ 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 $\mathsf{Q}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}} := \{\mathbf{z} \mid \mathbf{M}_{\boldsymbol{\varepsilon},\boldsymbol{\tau}}\mathbf{z} \le \mathbf{1}\}$, 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

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

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 $g(\boldsymbol{\varepsilon},\boldsymbol{\tau})\le1-\delta/2<1$, 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

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

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

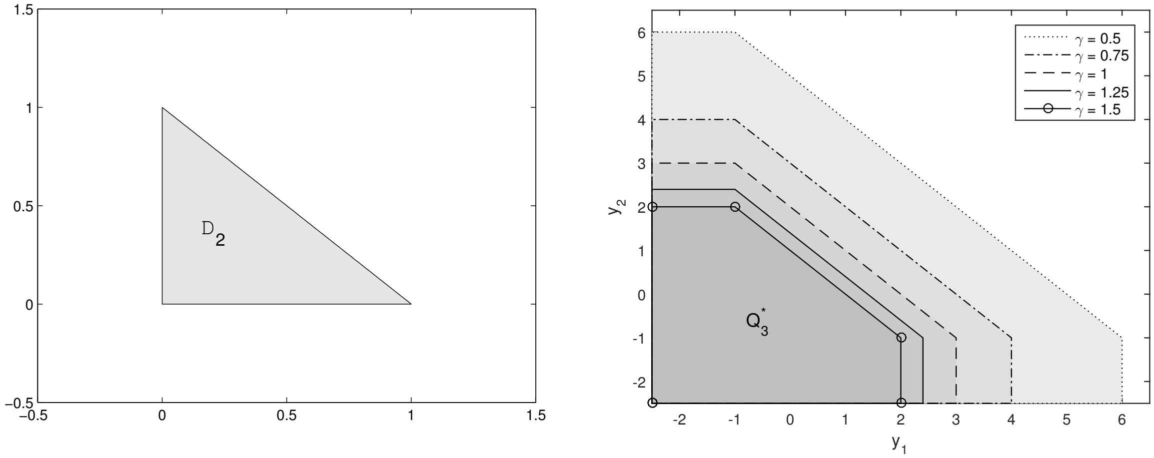


Figure EC.2 The 2-dimensional simplex (left) is transformed to the polar dual Q_3^* (right) for the normalized shock $\mathbf{y}$: $a_{ij} = 1/3$ for all $i \neq j$ and $\gamma = 0.5, 0.75, 1, 1.25, 1.5$.

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

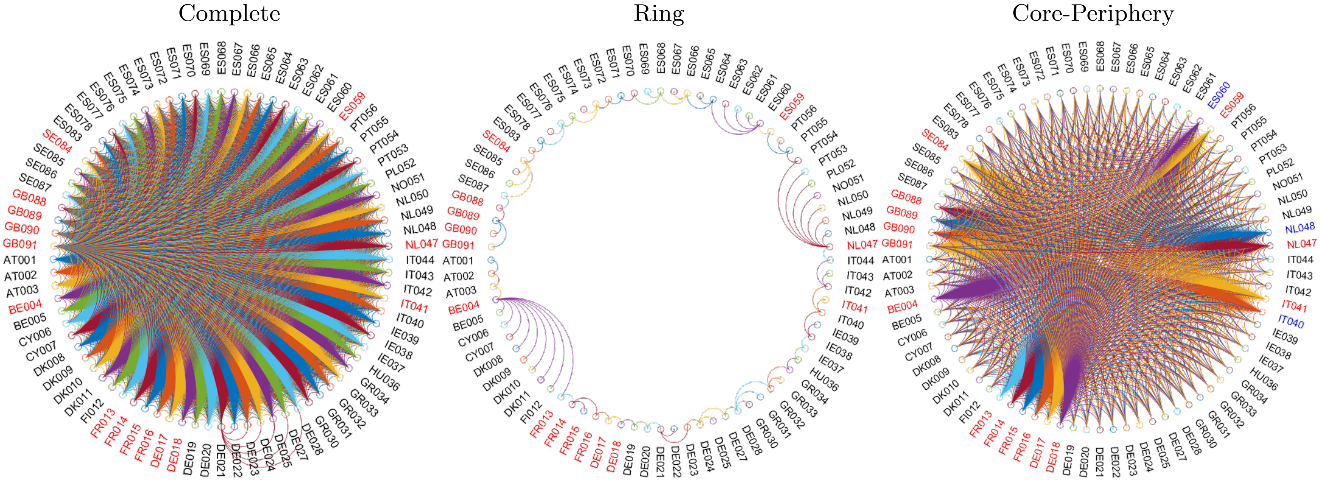
In this subsection, we conduct numerical experiments similar to those in Section 6 using the full dataset of the 2011 EBA stress test. We find that some small banks have problematic data, and hence, we omit the ten smallest banks as a simple rule. The dataset of the remaining 80 banks is given in Table 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.

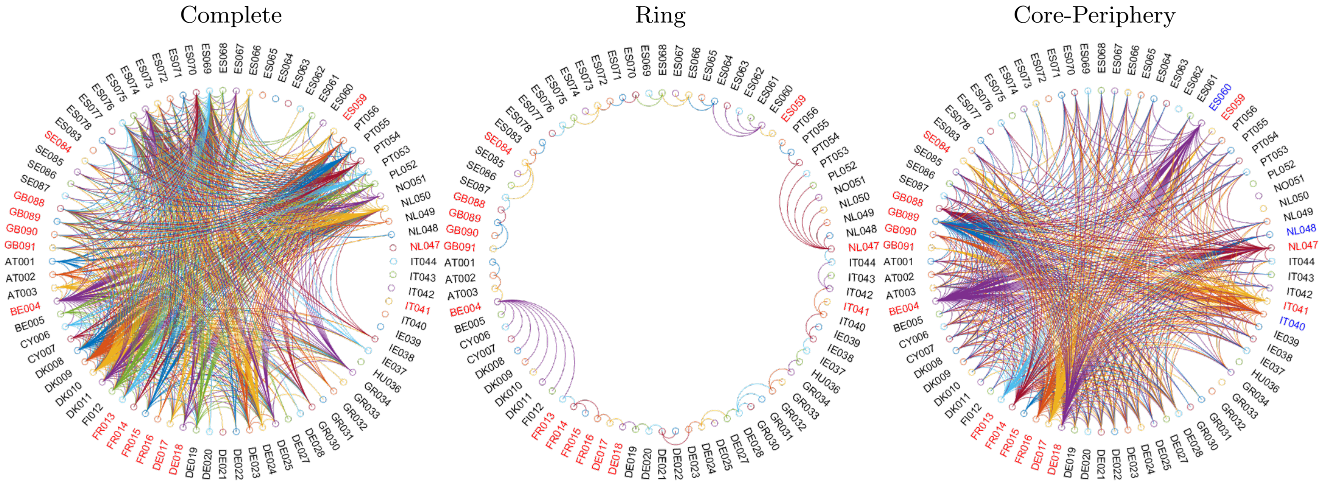


Note. The three structures represent complete, ring-like, and core-periphery networks, respectively. For each structure, nodes represent individual banks, and edges stand for their interbank exposures. Red nodes represent target banks, and blue nodes denote core banks except target banks.

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.

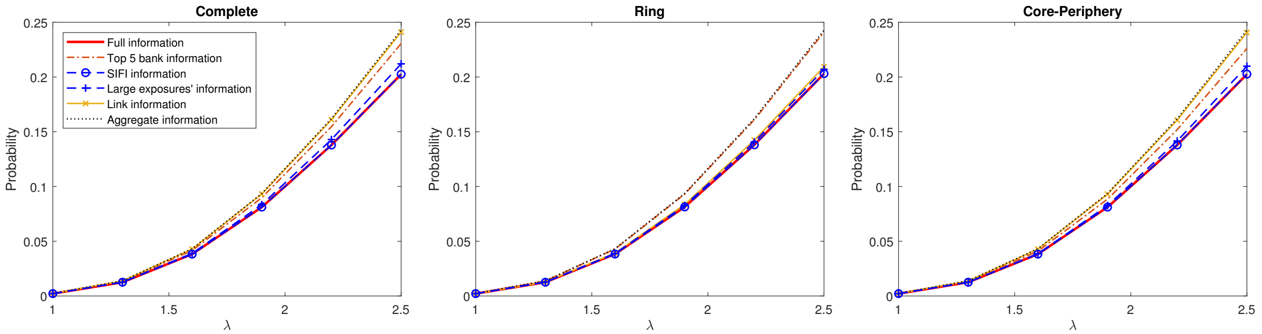
In Figures EC.5 and EC.6, under different network information and network structures, we illustrate the worst-case default probabilities and risk capitals with $\alpha = 0.95$. We assume that the random shocks $X_1, \ldots, X_{11}$ follow Pareto distributions: for each i, the probability density function 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.

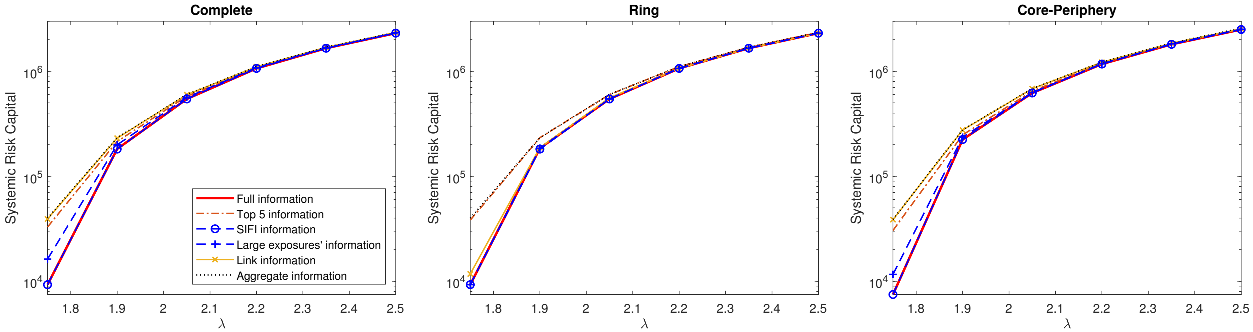
Figure EC.5 The worst-case default probabilities under different network information and different network structures.



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).

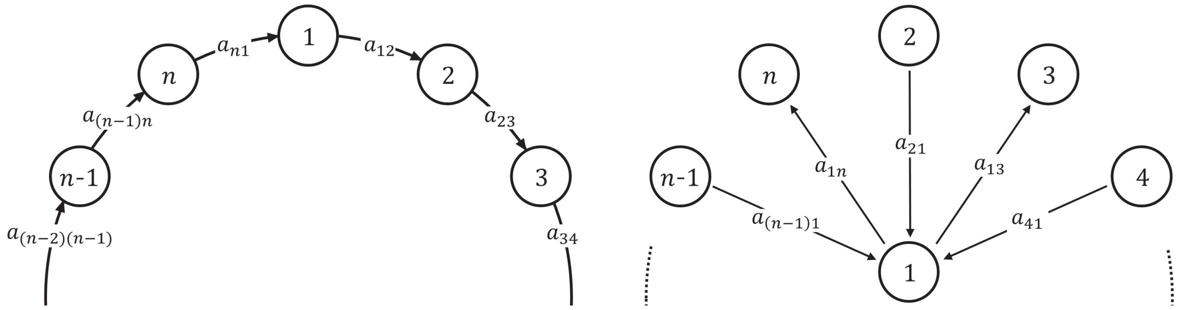
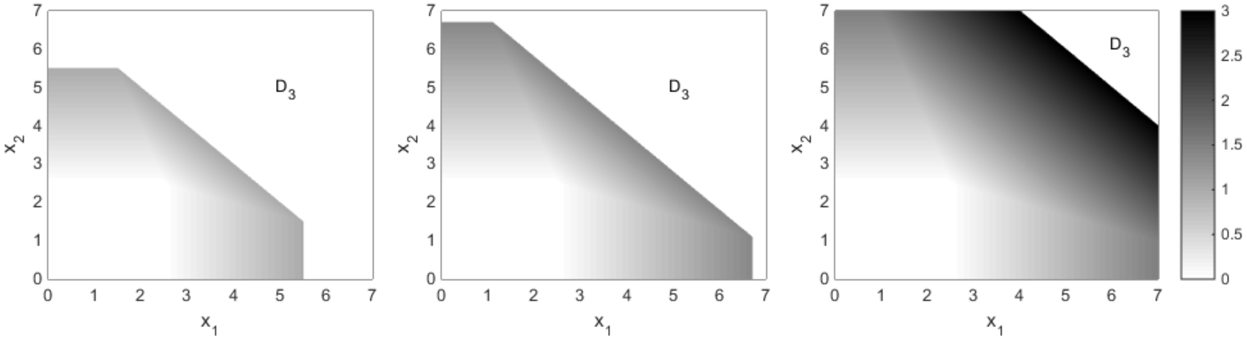


Figure EC.8 A graphical illustration of $\Phi_3(\mathbf{x})$ in a 3-bank network.



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.

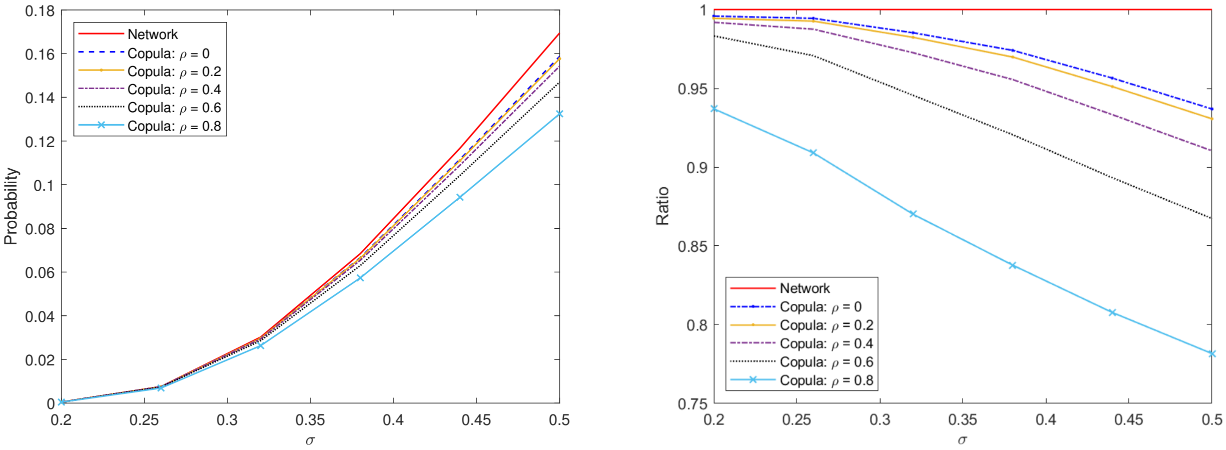
Example EC.3. Figure EC.8 presents the amount of the total shock $\Phi_3(\mathbf{x})$ for $w_3 = 1, 1.4$, and 3 with respect to x_1 and x_2 in the case of a three-bank network with $c_1 = c_2 = 7$, $w_1 = w_2 = 2.5$, and

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

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 matrix 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 matrix 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 matrix 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

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.

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

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

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

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

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.

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).

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.

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).