# Internet Appendix for
# "Inter-firm Relationships and Asset Prices"

## CARLOS RAMIREZ[*]

This appendix contains supporting results, tables, and figures to supplement the analysis in the main paper. Section A provides a derivation of the conditions under which consumption growth follows a log-normal distribution as the economy grows large. Section B provides a detailed description of the algorithm used to simulate the model. Section C presents a table and figure that help to illustrate the results in section A.

## A. Asymptotic Normality of Consumption Growth

To fix notation, let $\mathcal{G}_{n+1}$ denote the network $\mathcal{G}_n$, to which I add one new firm and all the relationships the entrant firm may create with incumbent firms in $\mathcal{G}_n$. The following proposition imposes sufficient conditions on: (a) the sequence of production networks $\{\mathcal{G}_n\}_{n\geq 1}$ and (b) the values within the propensity matrix, $\widetilde{p}_{t+1}$, so that $\Delta\widetilde{c}_{t+1}$ is normally distributed as $n$ grows large.

PROPOSITION 1 (Asymptotic Normality of $\Delta\widetilde{c}_{t+1}$): *Given $0 < q < 1$ and a sequence of production networks, $\{\mathcal{G}_n\}_{n\geq 1}$, define the threshold probability $p_c^q$ and the set $\mathcal{C}_n$ as*

$$p_c^q \equiv \sup_{p\in(0,1)} \left\{p : \textit{If every relationship in } \mathcal{G}_n \textit{ has propensity } p, \textit{ then } \lim_{n\to\infty} P_q^\alpha(\mathcal{G}_n) = 0\right\}$$

$$\mathcal{C}_n \equiv \{G \textit{ is a connected component of } \mathcal{G}_n : \textit{Number of nodes in } G = O(n)\}$$

*where $P_q^\alpha(\mathcal{G}_n)$ denotes the probability that a shock to any given firm in $\mathcal{G}_n$ affects at least $\alpha n$ firms via shock propagation, with $\alpha > 0$. The graph $G$ is said to be a connected component of $\mathcal{G}_n$ if $G$ is a subset of $\mathcal{G}_n$ in which any two firms are connected to each other by sequences of relationships and which is connected to no additional firms in $\mathcal{G}_n$. Notation $x = O(n)$*

---

[*]Board of Governors of the Federal Reserve System. The information in this manuscript represents the view of the author, and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or other members of its staff. E-mail: carlos.ramirez@frb.gov.

*indicates that $x$ grows, at most linearly, with $n$. If*

$$\lim_{n \to \infty} \left\{ \max_{(i,j) \, \in \, \mathcal{C}_n} \widetilde{p}_{ij,t+1} \right\} \;\; < \;\; p_c^q,$$

*then $\sqrt{n}\widetilde{W}_{n,t+1}$ and $\Delta\widetilde{c}_{t+1}$ are normally distributed as $n$ grows large.*

Under the conditions of Proposition 1, the distribution of $\Delta\widetilde{c}_{t+1}$ can be characterized in terms of its mean and variance. Because the network topology is fixed, the dynamics of the mean and variance of consumption growth are fully determined by the dynamics of the propensity matrix $\widetilde{p}_{t+1}$. Provided that the dynamic of $\widetilde{p}_{t+1}$ is determined by $\zeta_{t+1}$, the economy follows a Markov process with a continuum of values for aggregate consumption and its growth rate, $\Delta\widetilde{c}_{t+1}$, but only four values for the first two moments of the distribution of consumption growth.

The following corollaries provide a more detailed characterization of those large network economies in which consumption growth is normally distributed. Corollary 1 focuses on large networks in which all firms have the same number of direct relationships.

COROLLARY 1 (Symmetric Production Networks): *Suppose $\widetilde{p}_{ij,t+1} = p_{t+1} > 0$, $\forall(i,j) \in \mathcal{G}_n$. Given a sequence of production networks, $\{\mathcal{G}_n\}_{n \geq 1}$, with limiting topology $\mathcal{G}_\infty$,[1]*

- $p_c^q = 1 - 2\sin\left(\frac{\pi}{18}\right) \approx 0.65$ *if $\mathcal{G}_\infty$ is the two-dimensional honeycomb lattice.*
- $p_c^q = \frac{1}{2}$ *if $\mathcal{G}_\infty$ is the two-dimensional square lattice.*
- $p_c^q = 2\sin\left(\frac{\pi}{18}\right) \approx 0.34$ *if $\mathcal{G}_\infty$ is the two-dimensional triangular lattice.*
- $p_c^q = \frac{1}{z-1}$ *if $\mathcal{G}_\infty$ is the Bethe lattice with $z$ neighbors per each firm.*

Figure IA.1 illustrates each of the network economies considered in corollary 1. Corollary 2 focuses on large networks in which the number of direct relationships differs across firms.

COROLLARY 2 (Asymmetric Production Networks): *Suppose $\widetilde{p}_{ij,t+1} = p_{t+1} > 0$, $\forall(i,j) \in \mathcal{G}_n$. Given a sequence of production networks, $\{\mathcal{G}_n\}_{n \geq 1}$,*

- $p_c^q = \frac{1}{branching\ number\ of\ \mathcal{G}_\infty}$ *if $\mathcal{G}_\infty$ is a tree. The branching number of a tree is the average number of relationships per firm in a tree.[2]*
- $p_c^q = \frac{1}{e_M}$ *if $\mathcal{G}_n$ is sparse and locally treelike. $\mathcal{G}_n$ is said to be sparse if the number of relationships in $\mathcal{G}_n$ increases linearly with $n$, as $n$ increases. $\mathcal{G}_n$ is said to be locally*

---

[1] A lattice is a graph whose drawing can be embedded in $\mathbb{R}^n$. The two-dimensional honeycomb lattice is a graph in 2D that resembles a honeycomb. The two-dimensional square lattice is a graph that resembles the $\mathbb{Z}^2$ grid. The two-dimensional triangular lattice is a graph in 2D in which each node has six neighbors.

[2] A tree is a network in which any two nodes are connected by exactly one sequence of edges. A forest is a network whose components are trees.

*treelike if an arbitrarily large neighborhood around any given firm takes the form of a tree. For finite $n$, parameter $e_M$ is the leading eigenvalue of the matrix*

$$M_n = \begin{pmatrix} A_n & \mathbb{I}_n - D_n \\ \mathbb{I}_n & 0 \end{pmatrix} \tag{IA.1}$$

*where $A_n$ is the adjacency matrix of $\mathcal{G}_n$, i.e. the $n \times n$ matrix in which $A_{ij} = 1$ if there is a direct relationship between firms $i$ and $j$ and zero otherwise. $\mathbb{I}_n$ is the $n \times n$ identity matrix, and $D_n$ is the diagonal matrix that contains the number of relationships per firm along the diagonal.*

*Sketch of proof of Proposition 1 and Corollaries 1 and 2.* Given a sequence of network topologies $\{\mathcal{G}_n\}_{n \geq 1}$ and the realization of the matrix $\widetilde{p}_t$, the goal is to find the conditions under which $\sqrt{n}\widetilde{W}_{n,t}$ is normally distributed as $n$ grows large. Without loss of generality, fix period $t$ so that subscript $t$ on the sequence $\{\widetilde{\varepsilon}_{i,t}\}_{i=1}^n$ and on the matrix $\widetilde{p}_t$ can be eliminated. If the sequence of Bernoulli random variables $\{\widetilde{\varepsilon}_i\}_{i=1}^n$ is independent, the Lindeberg-Lévy central limit theorem implies that $\sqrt{n}\widetilde{W}_n$ is normally distributed as $n$ grows large. Consequently, if $\widetilde{p}$ is a matrix of zeros then $\{\widetilde{\varepsilon}_i\}_{i=1}^n$ is a sequence of independent random variables and $\sqrt{n}\widetilde{W}_n$ is asymptotically normally distributed. In the presence of inter-firm relationships, however, some elements in the matrix $\widetilde{p}$ are different from zero. In particular, $\widetilde{\varepsilon}_i$ and $\widetilde{\varepsilon}_j$ are correlated if there exists a sequence of relationships between firms $i$ and $j$ in $\mathcal{G}_n$. In this case, $\{\widetilde{\varepsilon}_i\}_{i=1}^n$ is a sequence of dependent random variables and the conditions under which the Lindeberg-Lévy central limit theorem hold are not satisfied.

Despite the fact that the sequence $\{\widetilde{\varepsilon}_i\}_{i=1}^n$ may be dependent, $\sqrt{n}\widetilde{W}_n$ may still be asymptotically normally distributed if the dependence among variables in $\{\widetilde{\varepsilon}_i\}_{i=1}^n$ is sufficiently weak. In particular, if the sequence $\{\widetilde{\varepsilon}_i\}_{i=1}^n$ is $\alpha$-mixing and stationary, $\sqrt{n}\widetilde{W}_n$ follows a normal distribution as $n$ grows large—see Billingsley (1995, Theorem 27.4). Generally speaking, the sequence $\{\widetilde{\varepsilon}_i\}_{i=1}^n$ is said to be $\alpha$-mixing if $\widetilde{\varepsilon}_k$ and $\widetilde{\varepsilon}_{k+n}$ are approximately independent for large $n$ and $k \geq 1$.[3] The sequence is said to be stationary if the distribution of the subsequence $\{\widetilde{\varepsilon}_l, \widetilde{\varepsilon}_{l+1}, \cdots, \widetilde{\varepsilon}_{l+j}\}$ does not depend on the subscript $l$. A special case of the above result occurs if there exists an ordering of the sequence $\{\widetilde{\varepsilon}_i\}_{i=1}^n$ such that the dependence between variables $\widetilde{\varepsilon}_k$ and $\widetilde{\varepsilon}_j$ decreases as the distance between them in such an ordering

---

[3]To be more specific, let $\alpha_n$ be a non-negative number such that

$$\left| \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B) \right| \leq \alpha_n \tag{IA.2}$$

with $A \in \sigma(\widetilde{\varepsilon}_1, \cdots, \widetilde{\varepsilon}_k)$, $B \in \sigma(\widetilde{\varepsilon}_{k+n}, \widetilde{\varepsilon}_{k+n+1}, \cdots)$, $k \geq 1$ and $n \geq 1$; where $\sigma(\cdot)$ denotes the $\sigma$-algebra defined on the power set of $\{0,1\}^n \equiv \{0,1\} \times \cdots \times \{0,1\}$. The sequence $\{\widetilde{\varepsilon}_i\}_{i=1}^n$ is said to be $\alpha$-mixing if $\alpha_n \to 0$ as $n$ grows large.

increases. In particular, if there exists such an ordering and a positive constant $m$ such that the subsequences $\{\widetilde{\varepsilon}_1, \cdots, \widetilde{\varepsilon}_k\}$ and $\{\widetilde{\varepsilon}_{1,k+s}, \cdots, \widetilde{\varepsilon}_{k+s+l}\}$ are independent whenever $s > m$, the sequence $\{\widetilde{\varepsilon}_i\}_{i=1}^n$ is said to be $m$-dependent in which case $\sqrt{n}\widetilde{W}_n$ follows a normal distribution as $n$ grows large.[4]

In what follows, I apply an idea similar to that behind the notion of $m$-dependent random variables to sketch the proof of the asymptotic normality of $\sqrt{n}W_n$. In particular, I impose sufficient conditions on the sequence of networks $\{\mathcal{G}_n\}_{n \geq 1}$ and on the matrix $\widetilde{p}$ so that negative shocks to individual firms tend to remain locally confined as $n$ grows large. In case the shocks remain locally confined, shocks would almost surely spread only over sets of firms of finite size—with all sets independent among each other—whose size would become negligible compared to the size of the economy as $n$ grows large. As a consequence, one would almost surely be able to find an index ordering $\mathcal{I}$ and a positive constant $\widehat{m}$ under which the sequence $\{\widetilde{\varepsilon}_i\}_{i \in \mathcal{I}}$ would be $\widehat{m}$-dependent and, thus, $\sqrt{n}W_n$ would follow a normal distribution as $n$ grows large.

For the sake of illustration, suppose that $\forall\ (i,j) \in \mathcal{G}_n$, $p_{ij} = p_0 > 0$, with $i \neq j$. Given $0 < q < 1$ and a sequence of networks, $\{\mathcal{G}_n\}_{n \geq 1}$, define the threshold probability $p_c^q$ as

$$p_c^q \ \equiv\ \sup_{p \in (0,1)} \left\{ p : \lim_{n \to \infty} P_q^\alpha(\mathcal{G}_n) = 0 \right\}, \tag{IA.3}$$

where $P_q^\alpha(\mathcal{G}_n)$ denotes the probability that a shock to any individual firm in $\mathcal{G}_n$ at least affects other $\alpha n$ firms via shock propagation, with $\alpha > 0$. Determining the conditions under which a CLT applies to $\sqrt{n}W_n$ is related to determining the threshold probability $p_c^q$. If every relationship in $\mathcal{G}_n$ has a propensity $p_0$ and $p_0 < p_c^q$, then shocks would remain locally confined because the number of firms affected by the shock would become negligible compared to the size of the economy as $n$ grows large. Then, one would almost surely be able to find an index ordering $\mathcal{I}$ and a positive constant $\widehat{m}$ under which the sequence $\{\widetilde{\varepsilon}_i\}_{i \in \mathcal{I}}$ would be $\widehat{m}$-dependent.

The condition $p_0 < p_c^q$ may be stronger than necessary to prove the asymptotic normality of $\sqrt{n}W_n$. Imposing such a condition, however, greatly facilitates the proof, as the determination of the threshold $p_c^q$ has been extensively studied in percolation theory, e.g. Grimmett (1989) and Stauffer and Aharony (1994). In percolation, $p_c^q$ is sometimes called the critical probability or critical phenomenon because it indicates the arrival of an infinite, connected component as $n$ grows large. A connected component of a graph is a subgraph in which any two nodes are connected to each other by sequences of edges, and which is connected to no additional nodes in the original graph.

---

[4]An independent sequence is 0-dependent using this terminology.

To illustrate how $p_c^q$ can be determined, consider the following two examples:

- Imagine $n$ firms are arranged in a straight line and each relationship may transmit shocks with probability $p_0$. The probability that every relationship in the line transmits shocks is $p_0^{n-1}$. Given how shocks spread from one firm to another, the probability that at least one negative shock spreads over $n-1$ different firms equals

$$(1 - (1-q)^n) \, \mathbb{P} \, [\text{every relationship in the line transmits shocks}] \qquad \text{(IA.4)}$$
$$\approx \; (1 - e^{-nq})p_0^{n-1} \quad \text{(for large } n)$$

  which tends toward zero as $n \to \infty$. Thus, $p_c^q = 1$.
- Suppose $n$ firms are arranged in a circle. The probability that every relationships in the circle transmits shocks is $p_0^n$, which tends toward zero as $n \to \infty$. Following the previous argument, it is easy to see that $p_c^q = 1$.

Taking results from bond percolation, Table IA.1 reports critical probabilities for several symmetric network topologies assuming that every relationship in the graph transmits shocks with the same probability. As Table IA.1 shows, $p_c^q$ varies across network topologies. For instance, if the limiting topology of the sequence $\{\mathcal{G}_n\}_{n \geq 1}$, $\mathcal{G}_\infty$, is the two-dimensional honeycomb lattice, then $p_c^q = 1 - 2\sin\left(\frac{\pi}{18}\right) \approx 0.65$, whereas if $\mathcal{G}_\infty$ is the two-dimensional square lattice then $p_c^q = \frac{1}{2}$.

The previous analysis determines conditions under which $\sqrt{n}\widetilde{W}_n$ is normally distributed for some symmetric network topologies. But what happens in other networks? In particular, under what conditions is $\sqrt{n}\widetilde{W}_n$ asymptotically normally distributed in large asymmetric networks? Using random walks on trees, Lyons (1990) shows that if $\mathcal{G}_\infty$ is a tree then

$$p_c = \frac{1}{\text{branching number of } \mathcal{G}_\infty}, \qquad \text{(IA.5)}$$

where the branching number of a tree is the average number of branches per node in a tree.[5] A tree is a connected graph in which two given nodes are connected by exactly one sequence of edges. A tree is said to be $z$-regular if each node has degree $z$. If $\mathcal{G}_\infty$ is a $z$-regular tree, the average number of branches per node is $z-1$ so $p_c = \frac{1}{z-1}$; which is consistent with Table IA.1.[6]

---

[5]For a concrete definition of the branching number see Lyons (1990, page 935).

[6]To motivate the previous result, it is informative to compute the percolation threshold in the Bethe lattice with $z$ neighbors per every node. Start at the root and check whether there is a chance of finding an infinite open path from the root. Starting from the root, one has $(z-1)$ new edges emanating from each new node in each layer of the lattice. Each of these $(z-1)$ new edges leads to one new node, which is affected with probability $p$. On average, $(z-1)p$ nodes are affected at each layer of the lattice. If $(z-1)p < 1$ then

One can generalize the previous result for topologies where $\mathcal{G}_\infty$ is sparse and locally treelike. $G_n$ is said to be sparse if $G_n$ has $m$ edges and $m = O(n)$. Notation $m = O(n)$ indicates that $m$ grows, at most, linearly with $n$ so there exists a positive number $c$ such that $\left|\frac{m}{n}\right| < c$ for all $n$. Namely, $G_n$ is sparse if only a small fraction of the possible $\frac{n(n-1)}{2}$ edges are present. $\mathcal{G}_\infty$ is said to be locally treelike if in the limit an arbitrarily large neighborhood around any node takes the form of a tree. Reformulating percolation in trees as a message passing process, Karrer et al. (2014) show that if $\mathcal{G}_\infty$ is sparse and locally treelike then

$$p_c = \frac{1}{\epsilon_H} \tag{IA.6}$$

where $\epsilon_H$ is the leading eigenvalue of the $2n \times 2n$ matrix

$$M = \begin{pmatrix} A & \mathbb{I} - D \\ \mathbb{I} & 0 \end{pmatrix} \tag{IA.7}$$

where $A$ is the adjacency matrix that represents $\mathcal{G}_n$, $\mathbb{I}$ is the $n \times n$ identity matrix, and $D$ is the diagonal matrix with the number of relationships per firm along the diagonal.

Now to allow for propensities to vary across relationships, one can extend the previous argument in the following way. Let $\mathcal{C}_n$ be the set of graphs composed by those connected components of $\mathcal{G}_n$ whose number of nodes grows, at most, linearly with $n$. Mathematically,

$$\mathcal{C}_n \equiv \{G \text{ is a connected component of } \mathcal{G}_n : \text{the number of nodes in } G = O(n)\} \tag{IA.8}$$

Because in the limit one only needs to pay attention to components of $G_n$ that potentially grow linearly with $n$, the following inequality

$$\lim_{n \to \infty} \left\{ \max_{(i,j) \in \mathcal{C}_n} \widetilde{p}_{ij} \right\} < p_c^q, \tag{IA.9}$$

ensures that shocks to individual firms would remain locally confined provided that sets of firms of finite size become negligible as the economy grows large. As a consequence, if inequality (IA.9) holds then $\sqrt{n}\widetilde{W}_n$ and $\Delta\widetilde{c}$ are normally distributed as $n$ grows large. $\qquad\square$

---

the average number of affected nodes decreases in each layer by a factor of $(z-1)p$. As a consequence, if $(z-1)p < 1$ the probability of finding an infinite open path goes to zero exponentially in the path length. Thus, $p_c = \frac{1}{z-1}$ for the Bethe lattice with $z$ neighbors for every node.

## B. Simulation of the Model

This section describes the algorithm I use to compute firms' probabilities of facing negative cash-flow shocks in each state of nature so one can compute asset prices and returns at the firm level using proposition 4. Let $s_t$ denote the state of the parameter vector $\zeta_t$ at period $t$. To simplify the computation of probabilities $\{\pi_i(s_t)\}_{i=1}^n$, I restrict the topology of $\mathcal{G}_n$. In general topologies, computing $\{\pi_i(s_t)\}_{i=1}^n$ is difficult, because the number of states that need to be considered increases exponentially with $n$. In economies with no cycles, however, computing $\{\pi_i(s_t)\}_{i=1}^n$ is easier. In those economies, computing $\{\pi_i(s_t)\}_{i=1}^n$ can be framed as a recursive problem, as the following algorithm describes.

**Algorithm** *Firms' Probabilities ($\mathcal{G}_n$, $s_t$, q)*
(∗ Description: Algorithm that computes firms' probabilities of facing shocks if $G_n$ is a forest ∗)

**Input:** $G_n$ (a forest), $s_t$, $q$.
**Output:** The set of probabilities of firms facing a negative cash-flow shock at time $t$,
$\quad \{\pi_i(s_t)\}_{i=1}^n$
1.  **for** each firm $i \in G_n$
2.  $\qquad$ Determine the subgraph of $G_n$ wherein firm $i$ participates. Denote such a graph
    $\qquad$ as $T_i$ and label firm $i$ as its root.[7]
3.  $\qquad$ **if** firm $i$ has a no connections
4.  $\qquad\qquad$ **return** $\pi_i(s_t) = q$
5.  $\qquad\quad$ **else** **return** $\text{Prob}(i, T_i, s_t, q)$

where $\text{Prob}(i, T_i, s_t, q)$ corresponds to the following recursive program:

**Algorithm** *Prob(i,$T_i$,$s_t$,q)*
(∗ Description: Recursive algorithm that computes firm $i$'s probability of facing a shock ∗)
**Input:** A node $i$ in $\mathcal{G}_n$, the tree $T_i$ wherein node $i$ is the root, $s_t$, and $q$.
**Output:** $\pi_i(s_t)$
1.  Determine the set of children of node $i$ in $T_i$, say $\mathcal{C}_i$.[8]
2.  **if** $\mathcal{C}_i = \emptyset$
3.  $\qquad$ **return** $\pi_i(s_t) = q$
4.  $\quad$ **else** **if** every node in $\mathcal{C}_i$ has no children
5.  $\qquad\qquad$ **return** $\pi_i(s_t) = q + (1-q)\left(1 - \mathbb{E}\left[\prod_{k \in \mathcal{C}_i}(1 - q\widetilde{p}_{ikt}) \big| s_t\right]\right)$

---

[7]Note that such a graph is a tree provided that $G_n$ is a forest.
[8]In a rooted tree, the parent of a node is the node connected to it on the path to the root. Every node except the root has a unique parent. A child of a node $v$ is a node of which $v$ is the parent.

6.  $\quad\quad\quad\quad\quad$ **else return** $\pi_i(s_t) = q + (1-q)\left(1 - \mathbb{E}\left[\prod_{k\in\mathcal{C}_i}\left(1 - \widetilde{p}_{ikt}\text{Prob}(k, T_{i,k}, s_t, q)\right)\right)\bigg|s_t\right]$

where tree $T_{i,k}$ denotes the branch of tree $T_i$ that starts at node $k$. In economies with no cycles, it is also simple to compute the first two moments of the distribution of $\sqrt{n}\widetilde{W}_{n,t+1}$ at $t+1$. Let $\mu_s$, $\sigma_s^2$ denote the mean and variance of $\sqrt{n}\widetilde{W}_{n,t+1}$ if $s_{t+1} = s$, respectively. In other words,

$$\mu_s = \lim_{n\to\infty}\mathbb{E}\left(\sum_{i=1}^n\frac{\widetilde{\varepsilon}_{i,t+1}}{\sqrt{n}}\bigg|s\right) = \lim_{n\to\infty}\sum_{i=1}^n\frac{\pi_i(s)}{\sqrt{n}}\quad\quad s = LL, LH, HL, HH. \text{ (IA.10)}$$

and

$$\sigma_s^2 = \lim_{n\to\infty}\text{Var}\left(\sum_{i=1}^n\frac{\widetilde{\varepsilon}_{i,t+1}}{\sqrt{n}}\bigg|s\right) \quad\quad\quad\quad\quad\quad\quad\quad\quad \text{(IA.11)}$$

$$= \lim_{n\to\infty}\left\{\frac{1}{n}\sum_{i=1}^n\pi_i(s)\left(1 - \pi_i(s)\right) + \frac{1}{n}\sum_{(i,j)\in\mathcal{G}_n}\text{Cov}\left(\widetilde{\varepsilon}_{i,t+1}, \widetilde{\varepsilon}_{j,t+1}\big|s_t = s\right)\right\}\quad s = LL, LH, HL, HH.$$

The second term in the equation above can be simplified further. If there exists a path between firms $i$ and $j$ after edges are removed at period $t+1$, then $\widetilde{\varepsilon}_{i,t+1} = \widetilde{\varepsilon}_{j,t+1}$. If there is no path between firms $i$ and $j$ in $\mathcal{G}_n$, variables $\widetilde{\varepsilon}_{i,t+1}$ and $\widetilde{\varepsilon}_{j,t+1}$ are independent. It then follows that

$$\mathbb{E}_t\left[\widetilde{\varepsilon}_{i,t}\widetilde{\varepsilon}_{j,t}\bigg|\text{ there is a path between }i\text{ and }j\right] = \text{Var}_t\left[\widetilde{\varepsilon}_{i,t}\right] + \mathbb{E}_t^2\left[\widetilde{\varepsilon}_{i,t}\right] = \pi_i(s)(1 - \pi_i(s)) + \pi_i^2(s)$$

$$\mathbb{E}_t\left[\widetilde{\varepsilon}_{i,t}\widetilde{\varepsilon}_{j,t}\bigg|\text{ there is no path between }i\text{ and }j\right] = \mathbb{E}_t\left[\widetilde{\varepsilon}_{i,t}\right]\mathbb{E}_t\left[\widetilde{\varepsilon}_{j,t}\right] = \pi_i(s)\pi_j(s).$$

Hence,

$$\mathbb{E}_t\left[\widetilde{\varepsilon}_{i,t}\widetilde{\varepsilon}_{j,t}\right] = \mathbb{E}_t\left[\widetilde{\varepsilon}_{i,t}\widetilde{\varepsilon}_{j,t}\bigg|\text{ there is a path between }i\text{ and }j\right]\mathbb{P}\left[\text{there is a path between }i\text{ and }j\text{ at }t\right]$$

$$+ \mathbb{E}_t\left[\widetilde{\varepsilon}_{i,t}\widetilde{\varepsilon}_{j,t}\bigg|\text{ there is no path between }i\text{ and }j\right]\mathbb{P}\left[\text{there is no path between }i\text{ and }j\text{ at }t\right]$$

$$= \left(\pi_i(s)(1 - \pi_i(s)) + \pi_i^2(s)\right)\mathbf{P}_{ij}(s) + \pi_i(s)\pi_j(s)\left(1 - \mathbf{P}_{ij}(s)\right),$$

where $\mathbf{P}_{ij}(s) \equiv \mathbb{P}\left[\text{there is a path between }i\text{ and }j\text{ if }s_t = s\right]$. Thus,

$$\text{Cov}_t\left[\widetilde{\varepsilon}_{i,t}, \widetilde{\varepsilon}_{j,t}\right] = \mathbb{E}_t\left[\widetilde{\varepsilon}_{i,t}\widetilde{\varepsilon}_{j,t}\right] - \mathbb{E}_t\left[\widetilde{\varepsilon}_{i,t}\right]\mathbb{E}_t\left[\widetilde{\varepsilon}_{j,t}\right]$$

$$= \left(\pi_i(s)(1 - \pi_i(s)) + \pi_i^2(s)\right)\mathbf{P}_{ij}(s) + \pi_i(s)\pi_j(s)\left(1 - \mathbf{P}_{ij}(s)\right) - \pi_i(s)\pi_j(s)$$

$$= \pi_i(s)\left(1 - \pi_j(s)\right)\mathbf{P}_{ij}(s).$$

Therefore,

$$\sigma_s^2 = \lim_{n\to\infty} \left\{ \frac{1}{n} \sum_{i=1}^{n} \pi_i(s)\left(1 - \pi_i(s)\right) + \frac{1}{n} \sum_{(i,j)\in\mathcal{G}_n} \pi_i(s)\left(1 - \pi_j(s)\right) \mathbf{P}_{ij}(s) \right\} \qquad s = LL, LH, HL, HH.$$

To compute $\mathbf{P}_{ij}(s)$ I need to determine the set of paths that connect firms $i$ and $j$ on $\mathcal{G}_n$. If there is more than one path connecting firms $i$ and $j$, computing $\mathbf{P}_{ij}(s)$ is difficult because shocks can be transmitted by any of those paths. On the other hand, if there is only one path connecting any two given firms, say firms $i$ and $j$, computing $\mathbf{P}_{ij}(s)$ becomes easier because there is a unique path connecting firms $i$ and $j$. It then becomes handy to restrict the topology of $\mathcal{G}_n$ so that it does not have cycles. The following remark describes $\mathbf{P}_{ij}(s)$ when $\mathcal{G}_n$ is a forest.

REMARK 1: *Suppose $\mathcal{G}_n$ is a forest; namely, there are no cycles. Then, every component of $\mathcal{G}_n$ is a tree. Provided that any two given firms are jointed by a unique path (in case such a path exists),*

$$\boldsymbol{P}_{ij}(s) = \begin{cases} \mathbb{E}\left[ \prod_{(k,l)\in\mathcal{P}_{i,j}} \widetilde{p}_{kl} \Big| s \right] & \text{where } \mathcal{P}_{i,j} \text{ is the (unique) path between } i \text{ and } j \text{ in } \mathcal{G}_n \\ 0 & \text{there is no path between } i \text{ and } j. \end{cases} \qquad \text{(IA.12)}$$

## C. Tables and Figures

This section contains a table and figure mentioned in section A.
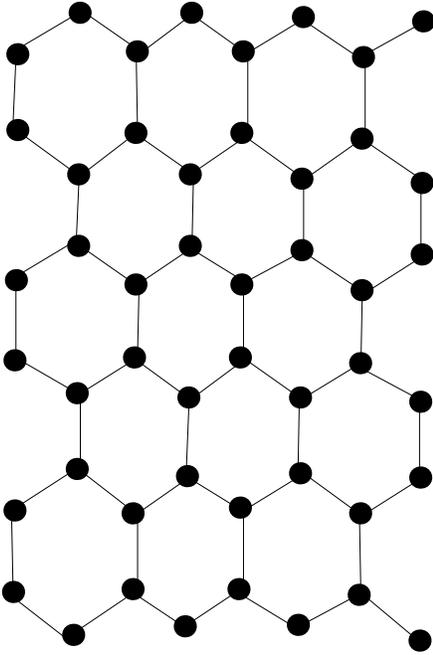
## Table IA.1
## Critical probability for different symmetric network topologies

The table reports critical probabilities for different symmetric network topologies. Besides reporting the two examples described in section A, the table reproduces a subset of the values reported in Stauffer and Aharony (1994, Table 1). The first column reports the topology of $\mathcal{G}_\infty$. The second column reports the number of neighbors of any given node in $\mathcal{G}_\infty$. The third column reports the critical probability, $p_c(\mathcal{G}_\infty)$. Despite the fact that $\mathcal{G}_\infty$ may be highly connected, if $\max \widetilde{p}_t < p_c(\mathcal{G}_\infty)$, then no infinite component emerges as $n \to \infty$, and thus $\sqrt{n}\widetilde{W}_n$ is asymptotically normally distributed. For illustrative purposes, figure 1(a) depicts a 2D honeycomb lattice, figure 1(b) depicts a 2D squared lattice, figure 1(c) depicts a 2D triangular lattice, and figure 1(d) depitcs a Bethe lattice with $z = 3$. The Bethe lattice of degree $z$ is defined as an infinite tree in which any node has degree $z$. For finite $n$ such topologies are called Cayley Trees.
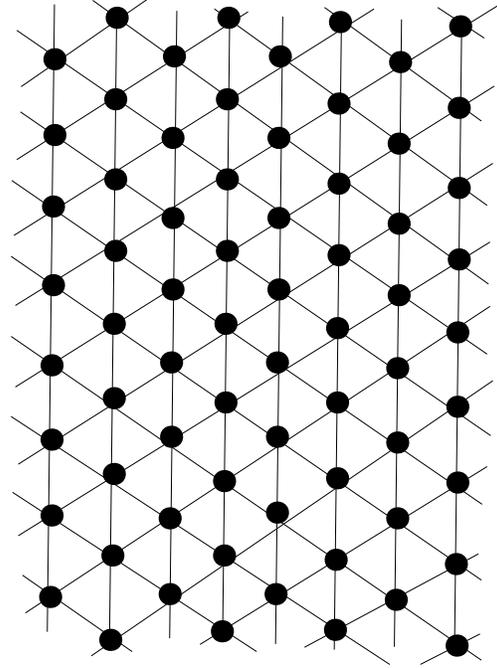
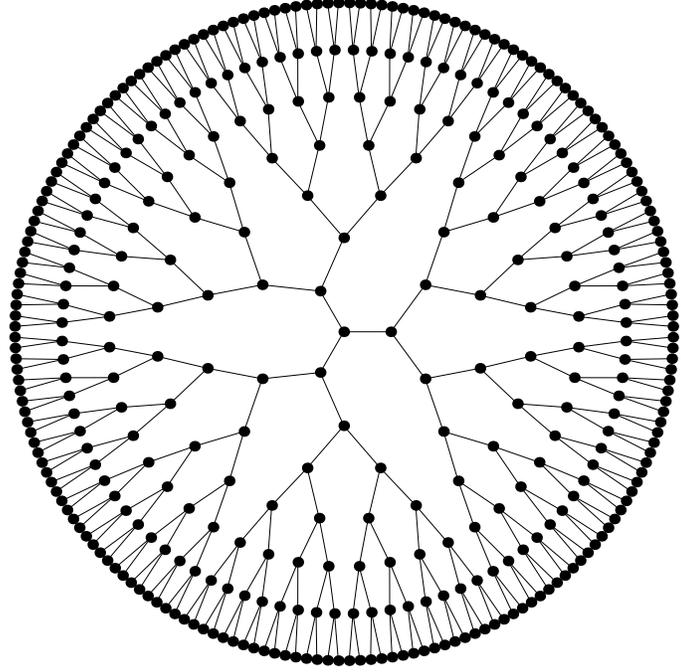| Topology of $\mathcal{G}_\infty$ | Number of neighbors | $p_c(\mathcal{G}_\infty)$ |
|---|---|---|
| Infinite line (1D lattice) | 2 | 1 |
| Infinite circle | 2 | 1 |
| 2D honeycomb lattice | 3 | $1 - 2\sin\left(\frac{\pi}{18}\right)$ |
| 2D squared lattice | 4 | $\frac{1}{2}$ |
| 2D triangular lattice | 6 | $2\sin\left(\frac{\pi}{18}\right)$ |
| Bethe lattice | $z$ | $\frac{1}{z-1}$ |

(a) 2D honeycomb lattice
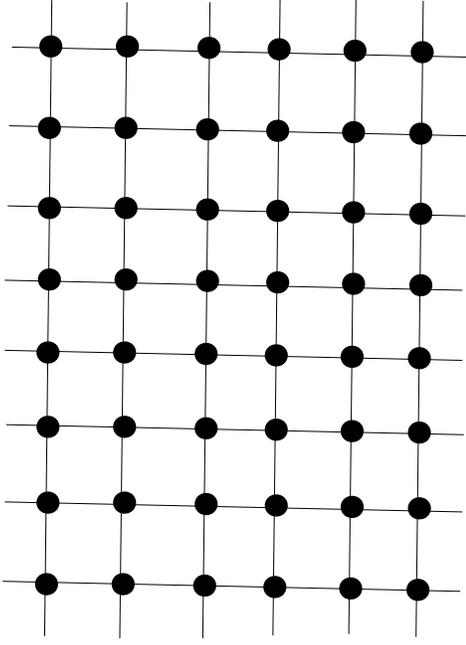
(b) 2D square lattice

(c) 2D triangular lattice

(d) Bethe lattice with degree 3

**Figure IA.1.** The figure shows the topologies of the symmetric networks considered in corollary 1 and Table IA.1.

# REFERENCES

Billingsley, Patrick, 1995, *Probability and Measure*, third edition (John Wiley and Sons, Inc.).

Grimmett, Geoffrey, 1989, *Percolation* (Springer-Verlag).

Karrer, Brian, M.E.J. Newman, and Lenka Zdeborová, 2014, Percolation on sparse networks, *Physical Review Letters* 113.

Lyons, Russell, 1990, Random walks and percolation on trees, *Annals of Probability* 18, 931–958.

Stauffer, Dietrich, and Amnon Aharony, 1994, *Introduction to Percolation Theory*, second edition (Taylor and Francis).