# Asset Pricing in Large Information Networks*

Han Ozsoylev† Johan Walden ‡

November 12, 2008

**Abstract**

We study asset pricing in economies with large information networks. We derive closed form expressions for price, volatility, profitability and several other key variables, as a function of the topological structure of the network. We focus on networks that are sparse and have power law degree distributions, in line with empirical studies of large scale human networks. Our analysis allows us to rank information networks along several dimensions and to derive several novel results. For example, price volatility is a non-monotone function of network connectedness, as is average expected profits. Moreover, the profit distribution among investors is intimately linked to the properties of the information network. We also examine which networks are stable, in the sense that no agent has an incentive to change the network structure. We show that if agents are ex ante identical, then strong conditions are needed to allow for non-degenerate network structures, including power-law distributed networks. If, on the other hand, agents face different costs of forming links, which we interpret broadly as differences in social skills, then power-law distributed networks arise quite naturally.

*Helpful discussions with Jonathan Berk, Sanjiv Das, Tina Eliassi-Rad, Nicolae Garleanu, Peter Jones, Pete Kyle, Dmitry Livdan, Santiago Oliveros, Christine Parlour, Richard Stanton and Marko Tervio are gratefully acknowledged. We benefited from comments by seminar audiences at UC Berkeley, the 2008 Oxford Financial Research Summer Symposium and the 2008 NBER Summer Institute Asset Pricing Workshop.

†Said Business School, University of Oxford, Park End Street, Oxford, OX1 1HP, United Kingdom. *E-mail*: `han.ozsoylev@sbs.ox.ac.uk`, *Phone:* +44-1865-288490. *Fax:* +44-1865-278826.

‡Haas School of Business, University of California at Berkeley, 545 Student Services Building #1900, CA 94720-1900. *E-mail*: `walden@haas.berkeley.edu`, *Phone:* +1-510-643-0547. *Fax:* +1-510-643-1420. Support from the Institute for Pure and Applied Mathematics (IPAM) at UCLA is gratefully acknowledged.

# 1 Introduction

Network theory provides a promising tool to help us understand how information is incorporated into asset prices. Empirically, social networks — or more generally *information networks*[1] — have been shown to be important in explaining investors' trading decisions and portfolio performance; see, for instance, Hong, Kubik, and Stein (2004), Ivković and Weisbenner (2007) and Cohen, Frazzini, and Malloy (2007).[2]

There is also casual evidence that information networks influence how investors manage their portfolios. Hedge fund manager John Paulson profited USD 15 billion in 2007, speculating against the subprime mortgage market by shorting risky collateralized debt obligations and buying credit default swaps. During the same time period, mogul Jeff Greene, a friend of Mr. Paulson, used similar mortgage-market trading strategies and made USD 500 million, after having been informed by Mr. Paulson about his ideas in the spring of 2006.[3] Clustering of investors in online financial communities on the Internet, as well as geographical clustering of investors in financial hubs, is also consistent with a world in which information networks play an important role in the functioning of financial markets.

Theoretically, the presence of information networks leads to several important questions, as, for instance, analyzed in recent papers by Ozsoylev (2005) and Colla and Mele (2008). Ozsoylev (2005) studies how informational efficiency depends on the structure — that is, the *topology* — of a social network, in which investors share information with their peers, and shows that for economies with large liquidity variance, price volatility decreases with the average number of information sources agents have. Colla and Mele (2008) study a cyclical network and show that agents who are close in the network have positively correlated trades, whereas agents who are distant may have negatively correlated trades.

One limitation of current theoretical models is the absence of closed form solutions, due to the complexity of the combination of networks, rational agents and endogenous price formation.[4] For example, the analysis in the static model of Ozsoylev (2005), although it allows for general networks, does not lead to closed form solutions for prices, which restricts the analysis to cases in which liquidity variance is high. The analysis in Colla and Mele (2008), on the other hand, leads to strong asset pricing implications in a dynamic model with strategic investors, but only for the very

[1] In this paper, we study general information networks. Social networks, i.e. personal and professional relationships between individuals, may make two individuals "close" in an information network, as may other factors, e.g., if two investors base their trading on the same information source. For our analysis, specific reasons for "informational proximity" between investors are not important since the proximity is modeled by a general metric.

[2] Hong, Kubik, and Stein (2004) provide evidence that fund managers' portfolio choices are influenced by word-of-mouth communication. Ivković and Weisbenner (2007) find similar evidence for households: they attribute more than a quarter of the correlation between households' stock purchases and stock purchases made by their neighbors to word-of-mouth communication. Cohen, Frazzini, and Malloy (2007) posit that there is communication via shared education networks between fund managers and corporate board members, manifested in the abnormal returns managers earn on firms they are connected to through their network.

[3] See The Wall Street Journal, January 15, 2008. Mr. Paulson and Mr. Greene are now former friends.

[4] If one is willing to drop the assumption of rationality, i.e. of having networks of expected utility optimizing agents with rational expectations, then the analysis is significantly simplified. For instance, DeMarzo, Vayanos, and Zwiebel (2003) propose a boundedly-rational model of opinion formation in social networks, and show that agents, who are "well-connected", may have more influence in the overall formation of opinions regardless of their information accuracies. DeMarzo, Vayanos, and Zwiebel (2004) apply the same model to financial markets.

special cyclical network topology. These limitations are not surprising, given the large number of degrees of freedom in a general large-scale network.[5]

A slightly different approach, however, may be possible. Several studies have shown remarkable similarities between different large-scale networks that arise when humans interact, like friendship networks, networks of co-authorship and networks of e-mail correspondence – see e.g., Milgram (1967), Barabasi and Albert (1999), Watts and Strogatz (1998), and also Chung and Li (2006) for a general survey of the literature. Specifically, these networks tend to be sparse (the number of connections between nodes are of the same order as number of nodes, where in our networks the nodes represent individuals), they have small effective diameter (the so-called small world property) and power laws govern their degree distributions (i.e. the distribution of the number of connections associated with a specific node is power law distributed).

It may therefore be fruitful to study a subclass of the general class of large-scale networks that satisfy these properties, and focus on asset pricing implications for this subclass of networks. Such an approach — in the spirit of statistical mechanics — rests on the assumption that for large-scale networks, the overwhelming majority of degrees of freedom average out, and only a few key statistical properties are important.

Indeed, the number of agents in the stock market's investor network is very large. For example, the number of investors participating in the stock market in the United States is in the tens of millions. A large economy approximation to the economy with a finite number of investors therefore seems to be in place. Theoretically, such an approximation may be helpful, since we know, e.g., from the study of noisy rational expectations equilibria, that tractable solutions often can be found in large economies, see Hellwig (1980) and Admati (1985).

In this paper, we carry out a large economy analysis for a general class of large-scale networks. It turns out that the analysis, indeed, simplifies significantly compared with the economy with a finite number of agents. We find closed form expressions for price, expected profits, price volatility, trading volume and value of connectedness. We compare networks with respect to connectedness, and see how connectedness influences, e.g., volatility and expected profits of different agents in the model. The distribution of expected profits among traders is a simple function of the topological properties of the network, which allows us to understand the wealth implications of information networks, i.e. to understand what type of networks lead to more disperse wealth distributions. The first contribution of the paper is thus the general existence theorem and the subsequent analysis of implications for asset pricing and welfare of agents.

The second contribution of the paper is to study welfare across different networks, in terms of agents' certainty equivalents, and relate this to the conditions under which a network will be *stable* in the sense that no agent has an incentive to change his position in the network, by either adding or dropping connections. In our model, if there is no dispersion in social skills among investors, strong

[5]The theoretical literature on networks and asset pricing is quite limited. There are, however, several other papers that apply network theory to other financial market settings. For example, Khandani and Lo (2007) argue that networks of hedge funds, linked through their portfolio holdings can explain liquidity driven systemic risks in capital markets. Brumen and Vanini (2008) show how firms, linked in buyer-supplier networks, will have similar credit risk. Recent empirical and theoretical work have done much to advance the more general proposition that social networks have important consequences for a number of other economic outcomes, including collaboration among firms, success in job search, educational attainment and participation in crime. Jackson (2008a,b) provide extensive surveys of the diverse literature on social networks in economics.

conditions are needed for any other network than a fully symmetric one to be stable. Specifically, entry costs need to be high and the cost function of forming connections needs to have a specific concave form. In contrast, when there is dispersion in investors' social skills, power-law distributed stable networks arise quite naturally. Our analysis is related to the endogenous network formation literature, but the concept of stable networks is weaker, since it does not take into account how a network was formed.[6]

The rest of the paper is organized as follows. In section 2 we describe the notational conventions employed in the paper. In section 3 we present the model and derive equilibrium prices in closed form for large economies. In section 4 we elaborate on the types of information networks that are socially plausible and the role such networks play in our analysis. Section 5 examines asset pricing implications of information networks while section 6 focuses on welfare and stability. Finally, we make some concluding remarks in section 7. Proofs are delegated to the Appendix.

# 2 Notation

We use the following conventions: lower case thin letters represent scalars, upper case thin letters represent sets and functions, lower case bold letters represent vectors and upper case bold letters represent matrices. For a general set, $W$, $|W|$ denotes the number of elements in the set. For two sets, $A$ and $B$, $A\backslash B$ represents the set $\{i \in A : i \notin B\}$. The $i$:th element of the vector $\mathbf{v}$ is $(\mathbf{v})_i$, and the $n$ elements $v_i$, $i = 1, \ldots, n$ form the vector $[v_i]_i$. We use $T$ to denote the transpose of vectors and matrices. One specific vector is $\mathbf{1}_n = (\underbrace{1, 1, \ldots, 1}_{n})^T$, (or just $\mathbf{1}$ when $n$ is obvious).

For vectors, $\mathbf{y}$, we define the vector norms $\|\mathbf{y}\|_p = \left(\sum_i (\mathbf{y})_i^p\right)^{1/p}$ and $\|\mathbf{y}\|_\infty = \max_i |(\mathbf{y})_i|$. Similarly, we define the matrix norms, $\|\mathbf{A}\|_p = \sup_{\{\mathbf{y}:\|\mathbf{y}\|_p=1\}} \|\mathbf{A}\mathbf{y}\|_p$, $p \in [1, \infty]$. For a vector, $\mathbf{d}$, we define the diagonal matrix $\mathbf{D} = diag(\mathbf{d})$, with $(\mathbf{D})_{ii} = (\mathbf{d})_i$. A matrix is defined by the $[\cdot]$ operator on scalars, e.g., $\mathbf{A} = [a_{ij}]_{ij}$. We write $(\mathbf{A})_{ij}$ for the scalar in the $i$th row and $j$th column of the matrix $\mathbf{A}$, or, if there can be no confusion, $\mathbf{A}_{ij}$.

Calligraphed letters represent structures, e.g. graphs, and relations. The set of natural numbers is $\mathbb{N} = \{1, 2, 3, \ldots\}$, the set of real numbers is $\mathbb{R}$, and the set of strictly positive real numbers is $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$ and $\mathbb{R}_{++} = \{x \in \mathbb{R} : x > 0\}$. For $x \in \mathbb{R}$, $\lfloor \mathrm{x} \rfloor$ denotes the largest integer not larger than $x$, and $\lceil \mathrm{x} \rceil$ denotes the smallest integer not smaller than $x$.

We say that $f(n) = o(g(n))$ if $\lim_{n\to\infty} f(n)/g(n) = 0$. Moreover, we say that $f(n) = O(g(n))$ if there is a $C > 0$ such that $f(n) \leq Cg(n)$ for all $n$. Similarly, if the conditions hold in probability, we say that $f(n) = o_p(g(n))$ and $f(n) = O_p(n)$ respectively. If there is a constant $C > 0$, such that $\lim_{n\to\infty} f(n)/g(n) = C$ then we say that $f(n) \sim g(n)$, and similarly we define $f(n) \sim_p g(n)$. Also, for a function $z : \mathbb{R} \to \mathbb{R}$, we define $f \sim g$ at $x$ if $\lim_{\epsilon \searrow 0} f(x + \epsilon)/g(x + \epsilon) = C$ for some $C > 0$.

The expectation and variance of a random variable, $\tilde{\xi}$, are denoted by $E[\tilde{\xi}]$ and $var(\tilde{\xi})$ respectively. The correlation and covariance between two random variables are denoted by $cov(\tilde{\xi}_1, \tilde{\xi}_2)$ and $corr(\tilde{\xi}_1, \tilde{\xi}_2)$ respectively.

[6]The dynamic question of how power-law distributed networks form, although of high interest, is outside of the scope of this paper. Many different network formation models that lead to power law degree distributions have been introduced since the original work by Simon (1955). For economic models, see, e.g., Jackson and Rogers (2007) and references therein.

We will use the unit simplex over the natural numbers, $S^\infty = \{x \in \mathbb{R}^{\mathbb{N}}, x(i) \geq 0, \sum_{i=1}^{\infty} x(i) = 1\}$. Similarly, we define $S^n$ to be the unit simplex in $\mathbb{R}^n_+$, with the natural interpretation that $S^1 \subset \cdots \subset S^n \subset S^{n+1} \subset \cdots \subset S^\infty$. The support of an element $d \in S^\infty$ is $supp[d] = \{i : d(i) > 0\}$. If $|supp[d]| \leq 2$, then we say that $d$ is degenerate. A specific degree distribution is $\delta_i \in S^\infty$, which has $\delta_i(i) = 1$.

# 3 Model

We follow the large economy analysis in Hellwig (1980) closely,[7] but extend to allowing for network relationships between agents in the model in the sense that agents can infer information about the signals given to their neighbors. This is similar to the approaches taken in Ozsoylev (2005) and Colla and Mele (2008).

We first study a network, $\mathcal{G}^n$, with a fixed number, $n$, of agents (also called nodes) and then use the results to study a growing sequence of networks $(\mathcal{G}^1, \ldots, \mathcal{G}^n, \ldots)$ to infer asymptotic properties, when $n$ approaches infinity.

## 3.1 Networks of agents

There are $n$ agents (investors) enumerated by the natural numbers, $N = \{1, 2, \ldots, n\}$, connected in a network. The relation, $\mathcal{E} \subset N \times N$, describes whether agent $i$ and $j$ are linked. Specifically, the edge $(i, j) \in \mathcal{E}$, if and only if there is a link between agent $i$ and $j$. We use the convention that agent $i$ is connected with herself, $(i, i) \in \mathcal{E}$ for all $i$, and that connections are undirected. Thus, $\mathcal{E}$ is reflexive and symmetric. Formally, the network is described by the duple $\mathcal{G} = (N, \mathcal{E})$. One way of representing $\mathcal{E}$ is by the matrix $\mathbf{E} \in \mathbb{R}^{N \times N}$, with $(\mathbf{E})_{i,j} = 1$ if $(i, j) \in \mathcal{E}$ and $(\mathbf{E})_{i,j} = 0$ otherwise.

We define the distance function $D(i, j)$ as the number of steps in the shortest path between $i$ and $j$, where we use the conventions that $D(i, i) = 0$, and $D(i, j) = \infty$ whenever there is no path between node $i$ and $j$. The set of nodes adjacent to node $i$ (i.e., node $i$'s neighbors) is $Q_i = \{j \neq i : (i, j) \in \mathcal{E}\} = \{j : D(i, j) = 1\}$. More generally, the set of nodes at distance $m$ from node $i$ is $Q_i^m = \{j : D(i, j) = m\}$, and the set of nodes at distance not further away than $m$ is then $R_i^m \stackrel{\text{def}}{=} \cup_{j=0}^m Q_i^j$.

The number of nodes not further away from node $i$ than $m$ is $W_i^m \stackrel{\text{def}}{=} |R_i^m|$. For $m = 1$, we simply write $R_i$ and $W_i$, so $W_i$ is the *degree* of node $i$, which we also call *node $i$'s connectedness*. The degree distribution is the function, $d : N \to [0, 1]$, such that

$$d(i) = \frac{\big|\{j : W_j = i\}\big|}{n}.$$

The common neighbors of nodes $i$ and $j$ are $R_{ij} = R_i \cap R_j$, and the number of such common neighbors is $W_{ij} = |R_{ij}|$, which can be used to define the symmetric neighborhood matrix $\mathbf{W} = [W_{ij}]_{ij}$. The element on row $i$ and column $j$ of $\mathbf{W}$ thus represents the number of nodes that are

[7] Our model is also related to the model of Diamond and Verrecchia (1981), however Diamond and Verrecchia (1981) only analyze a finite-agent economy.

common neighbors to nodes $i$ and $j$, (including nodes $i$ and $j$ if nodes $i$ and $j$ are linked). The relation $\mathbf{W} = \mathbf{E}^2$ follows from standard graph theory.

Clearly, we have

$$(\mathbf{W})_{ij} \in N, \tag{1}$$
$$(\mathbf{W})_{ij} \leq \min\{W_i, W_j\}, \tag{2}$$
$$(\mathbf{W})_{ii} = W_i \geq 1. \tag{3}$$

### 3.2 Information structure and agent characteristics

Following Hellwig (1980), we make the following assumptions: The economy operates in times $t = 0$ and $t = 1$. There are $n$ CARA agents, and for simplicity we assume that they all have risk-aversion of unity, $U = -E[e^{-\tilde{\xi}}]$. We note that for such agents, the certainty equivalent, $CE$, of a gamble, $\tilde{\xi}$, is

$$CE = -\log\left(E\left[e^{-\tilde{\xi}}\right]\right). \tag{4}$$

There is one asset, paying a liquidating dividend at $t = 1$ of $\tilde{X} \sim N(\bar{X}, \sigma^2)$ with $\bar{X} \geq 0$. The supply of the asset is stochastic, $\tilde{Z}_n = n \times \tilde{Z}$, where $\tilde{Z} \sim N(\bar{Z}, \Delta^2)$ and $\bar{Z} \geq 0$. Thus, $\sigma^2$ is the value variance and $\Delta^2$ is the liquidity variance. There are also $n$ information signals $\{\tilde{y}_k\}_{k=1}^n$ about the asset payoff $\tilde{X}$: signal $\tilde{y}_k$ communicates $\tilde{X}$ with some error $\tilde{\epsilon}_k$ so that $\tilde{y}_k = \tilde{X} + \tilde{\epsilon}_k$, where $\tilde{\epsilon}_k \sim N(0, s^2)$. The random variables $\tilde{X}$, $\tilde{Z}$ and $\{\tilde{\epsilon}_k\}_{k=1}^n$ are jointly independent.

Agents are price takers and they trade in period $t = 0$. Prior to trading, each agent receives a signal about the asset payoff. The relationship between agents' signals depends on the network's topology. Formally, agent $i$ has the signal

$$\tilde{x}_i = F_i(\tilde{y}_1, \ldots, \tilde{y}_n | \mathcal{G}_n),$$

for some function $F_i$, such that $E[\tilde{x}_i] = E[\tilde{X}]$. In general, we wish the topological properties of the network to carry over to the following *network signal properties*:

- Agents with more neighbors receive more precise signals, $W_i > W_j \Rightarrow var(\tilde{x}_i) < var(\tilde{x}_j)$.
- If two agents have no common neighbors, then their signals' error terms are uncorrelated,
$$R_i \cap R_j = \emptyset \Rightarrow cov(\tilde{x}_i, \tilde{x}_j) = var(\tilde{X}).$$
- Two agents, who have the same neighbors, receive the same signal, $R_i = R_j \Rightarrow \tilde{x}_i = \tilde{x}_j$.
- All else equal, the correlation between agent $i$'s and $j$'s signal is higher if they are connected than if they are not connected, i.e., given two economies, $\mathcal{G}$ and $\mathcal{G}'$ that are identical, except for that $(i, j) \in \mathcal{E}$, but $(i, j) \notin \mathcal{E}'$, then $corr(\tilde{x}_i, \tilde{x}_j) > corr(\tilde{x}'_i, \tilde{x}'_j)$.

A signal structure that satisfies these properties, which will be very convenient to work with, is given by

$$\tilde{x}_i \stackrel{\text{def}}{=} \frac{\sum_{k \in R_i} \tilde{y}_k}{W_i}, \tag{5}$$

which immediately implies that $\tilde{x}_i = \tilde{X} + \tilde{\eta}_i$, where the $\tilde{\eta}$'s are multivariate normally distributed random variables with mean zero and covariance matrix, $\mathbf{S} = [cov(\tilde{\eta}_i, \tilde{\eta}_j)]_{ij}$,

$$\mathbf{S} = s^2 \mathbf{D}^{-1} \mathbf{W} \mathbf{D}^{-1}, \tag{6}$$

where $\mathbf{D} = diag((\mathbf{W})_{11}, \ldots, (\mathbf{W})_{nn})$. Agent $i$'s information set is thus[8]

$$\mathcal{I}_i = \{\tilde{x}_i, \tilde{p}\}, \tag{7}$$

and his demand schedule takes the form $\psi_i(\tilde{x}_i, \tilde{p})$. Clearly, $\{\tilde{\eta}_i\}_i$, being linear combinations of $\{\tilde{\epsilon}_i\}_i$, are independent of $\tilde{Z}$, and $\tilde{X}$.

In our model the topology of the network maps to the information structure, which — as we shall see — in turn determines the asset pricing properties of the model.[9] Such a mapping from networks to information structures is also present in the models of Ozsoylev (2005) and Colla and Mele (2008). The approach provides a powerful way to use information networks to put restrictions on the information structure, in line with real world observations.

## 3.3 Interpretation of links

Although it is natural to think of $(i, j) \in \mathcal{E}$ as representing the relationship of agent $i$ being acquainted with agent $j$, our analysis is perfectly general and holds for other interpretations of $\mathcal{E}$, and thereby of $\mathbf{W}$. As long as there is a set of nodes, $R_i$, associated with each node, $i$, where the only requirement is that $i \in R_i$, we can define $[\mathbf{W}]_{ij} \stackrel{\text{def}}{=} |R_i \cap R_j|$, which leads to $(\mathbf{W})_{ij} \in N$, $(\mathbf{W})_{ij} \leq \min\{(\mathbf{W})_{ii}, (\mathbf{W})_{jj}\}$ and $(\mathbf{W})_{ii} \geq 1$. These are the conditions needed in the subsequent analysis.

For example, given a connection relation $\mathcal{E}$, we can define the connection relation $\hat{\mathcal{E}} = \{(i, j) : D(i, j) \leq 2\}$, representing a situation in which an agent's signal is also related to signals of neighbors to neighbors. This leads to a neighborhood matrix, $\hat{\mathbf{W}}$, and degrees, $\hat{W}_i = (\hat{\mathbf{W}})_{ii}$. The new relation, $\hat{\mathcal{E}}$, represents a situation in which *centrality* is important for an agent's signal, as opposed to $\mathcal{E}$, which only depends on direct links.

As a specific example, consider the network with 21 agents shown in Figure 1 below. Under the relation $\mathcal{E}$, agent 2 receives a more precise signal about the asset's value than agent 1, since $R_1 = 5$ and $R_2 = 6$. One might argue, however, that agent 1 is more central than agent 2 in the sense that although he has fewer connections than agent 2, his connections are themselves better

[8]Since $\tilde{x}_i$ is a sufficient statistic for $\tilde{X}$ conditioned on $\{\tilde{y}_k : k \in R_i\}$, agent $i$'s information set $\mathcal{I}_i$ is essentially equivalent to $\left\{E[\tilde{X}|\{\tilde{y}_k : k \in R_i\}], \tilde{p}\right\}$. A slightly different approach is taken in Ozsoylev (2005), who assumes that agent $i$'s information set is $\mathcal{I}_i = \left\{\tilde{y}_i, E[\tilde{X}|\{\tilde{y}_k : k \in R_i \backslash \{i\}\}], \tilde{p}\right\}$. We have also carried out the analysis with Ozsoylev's (2005) approach, with qualitatively similar — although somewhat more complex — results. The analysis is available upon request.

[9]The information structure in our model cannot be mapped to the information structures of Hellwig (1980) and Diamond and Verrecchia (1981). In Hellwig (1980) and Diamond and Verrecchia (1981) agents' private signals carry independent error terms whereas in our model signals have correlated error terms. It is in effect the correlated error terms that proxy the network connections. Also, as we shall see, in our model some agents are allowed to receive very precise signals. This is in contrast to Hellwig (1980), where there is a common upper bound on the precision of all signals.

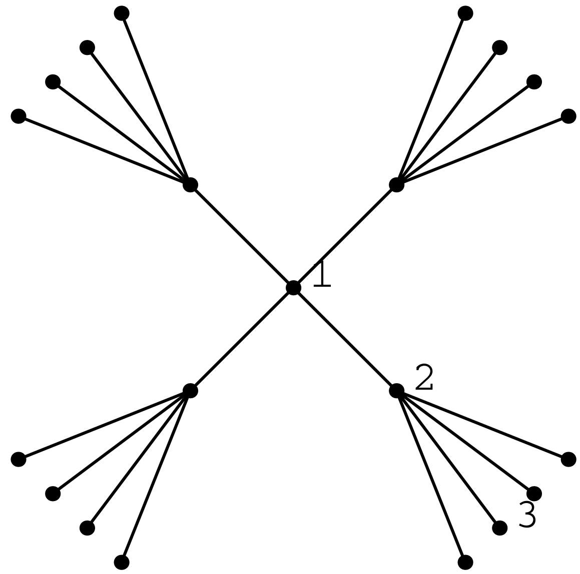


Figure 1: *Network: For direct neighbor relation, $\mathcal{E}$, $W_1 = 5$, $W_2 = 6$, $W_3 = 2$, so agent 2 receives most precise signal. For $\hat{\mathcal{E}}$, $\hat{W}_1 = 21$, $\hat{W}_2 = 9$, $\hat{W}_3 = 6$, so agent 1 receives most precise signal, due to centrality.*

connected, which should work to his advantage. This is captured in the $\hat{\mathcal{E}}$ definition, which also takes into account neighbors to neighbors. Since agent 1 can reach all agents in two steps, his degree is $\hat{W}_1 = 21$ in the $\hat{\mathcal{E}}$ metric, whereas agent 2's degree is only 9. Thus, in the $\hat{\mathcal{E}}$ metric, agent 1 is the one who is most connected.

The concept of centrality is not new to the asset pricing literature. In Das and Sisk (2005), the *centrality score*, which measures how central a node is, taking into account the connectedness of its neighbors, neighbors' neighbors, etc., is used to apply network methods to asset pricing. Their interpretation of what constitutes a network is somewhat different, however, since their nodes are interpreted as stocks, and links represent overlapping posters in Internet stock message boards.

Our analysis is valid for arbitrary connection relations, as long as some technical conditions are satisfied. We could assume that signals travel even further distances, perhaps with increased noisiness as distances increase. The degree would, in this case, be similar to the centrality score used in Das and Sisk (2005). Thus, general interpretations of $\mathbf{W}$ are allowed, although, for convenience we interpret $\mathcal{E}$ as representing direct links, going forward.

### 3.4 Equilibrium

A linear noisy rational expectations equilibrium (NREE) with $n$ agents is defined as a price function

$$\tilde{p} = \pi_0 + \sum_{i=1}^{n} \pi_i \tilde{x}_i - \gamma \tilde{Z}_n, \tag{8}$$

such that

- *Markets always clear*: $\tilde{Z}_n = \sum_{i=1}^{n} \psi_i(\tilde{x}_i, \tilde{p})$ for all realizations of $\{\tilde{x}_i\}_i$, $\tilde{X}$, and $\tilde{Z}_n$.
- *Agents optimize rationally*: each agent optimizes expected utility under rational expectations, given the agent's information.

It follows from our CARA-normal setup that agent $i$'s optimal demand schedule takes the form

$$\psi_i(\tilde{x}_i, \tilde{p}) = \frac{E[\tilde{X}|\mathcal{I}_i] - \tilde{p}}{Var[\tilde{X}|\mathcal{I}_i]}. \tag{9}$$

We are interested in the existence of a linear NREE in a "large" market. We note that, contrary to the analysis in Hellwig (1980), the existence of a linear NREE for a fixed number of agents, $n$, is not guaranteed, since the signals, $\{\tilde{x}_i\}_i$, have correlated noise terms for connected agents and agents with common neighbors in our set-up. However, as we shall show, under some additional assumptions, for large enough $n$, a linear NREE is guaranteed.

We study a sequence of markets, $\mathcal{G}^1, \dots \mathcal{G}^n, \dots$, with increasing number of agents, $n$. Our main result for a sequence of markets with covariance matrices defined by (6) is:

**Theorem 1** *Assume a sequence of $n$-agent markets, $\mathcal{M}^n$, $n = 1, 2, \dots$, in which agents' information sets are defined by (7), the covariance matrix $\mathbf{S}^n$ of market $\mathcal{M}^n$ is defined via equation (6), where each matrix $\mathbf{W}^n$ satisfies equations (1)-(3), and also*

$$\begin{aligned} \|\mathbf{W}^n\|_\infty &= o_p(n), && (10) \\ \lim_{n\to\infty} \frac{\sum_{i=1}^{n}(\mathbf{W}^n)_{ii}}{s^2 n} &= B + o_p(1) > 0. && (11) \end{aligned}$$

*Then, with probability one, the equilibrium price converges to*

$$\tilde{p} = \pi_0^* + \pi^* \tilde{X} - \gamma^* \tilde{Z}, \tag{12}$$

*where*

$$\begin{aligned} \pi^* &= \gamma^* B, && (13) \\ \gamma^* &= \frac{\sigma^2\Delta^2 + \sigma^2 B}{B\sigma^2\Delta^2 + \Delta^2 + B^2\sigma^2}, && (14) \\ \pi_0^* &= \gamma^* \frac{\bar{X}\Delta^2 + \bar{Z}B\sigma^2}{\sigma^2\Delta^2 + \sigma^2 B}. && (15) \end{aligned}$$

Theorem 1 is our main workhorse in analyzing economies with large information networks.

**Remark 1** *Since an agent is always connected to himself, $B \geq 1/s^2$.*

It is clear that the average number of links, $B$, is a crucial statistic for asset prices. It is natural to think of $B$ as a measure of the *network's connectedness*, since it — up to a scaling factor, $s^2$, — measures how many connections agents have on average (including their connection to themselves).

Even though Theorem 1 does not depend on the existence of an asymptotic degree distribution, $d$, as $n$ tends to infinity, we will throughout the rest of the paper restrict our attention to sequences of networks for which such a distribution exists, i.e., we assume that

**Assumption 1** *There is a degree distribution, $d \in S^\infty$, such that* $\lim_{n\to\infty} \sum_{i=1}^n |d^n(i) - d(i)| = 0$, *with probability one, where $d^n$ is the degree distribution for the economy with $n$ agents.*

We call $d$ the degree distribution of the large network.

In our subsequent analysis of individual agents, we will focus on agents for which the asymptotic degree exists, i.e., for which $\lim_{n\to\infty} \mathbf{W}_{ii}^n$ exists and is finite (with probability one). Similarly, when we compare pairs of agents in section 5.4, an additional underlying assumption is that $\lim_{n\to\infty} \mathbf{W}_{ij}^n$ exists and is finite. We could, alternatively, have focused on networks for which $\lim_{n\to\infty} \mathbf{W}_{ii}^n$ exist for all $i$, but this would be unnecessarily restrictive and would rule out many important random network models. The issue can be avoided completely by interpreting "agent $i$" with connectedness $W_{ii}$ as a sequence of different agents $i_1, \ldots, i_n, \ldots$, such that $\lim_{n\to\infty} \mathbf{W}_{i_n i_n}$ exists and is finite, but we avoid this approach since it leads to a cumbersome notation.

# 4 Socially plausible networks

Given the enormous number of degrees of freedom in constructing a general large network, it is not surprising that any degree distribution can be supported by a large economy. We have the following existence result.

**Proposition 1** *Given a degree distribution $d \in S^\infty$, there is a sequence of networks, $\mathcal{G}^n$, with degree distributions, $d^n \in S^n$, such that* $\lim_{n\to\infty} \sum_{i=1}^n |d^n(i) - d(i)| = 0$. *If $d(i) = O(i^{-\alpha})$, $\alpha > 2$, then the sequence of networks can be constructed to satisfy the conditions of Theorem 1. If $d(i) \sim i^{-\alpha}$, $\alpha \leq 2$, then condition (11) will fail.*

Networks that satisfy

$$d(i) \sim i^{-\alpha},$$

are said to have *power-law* distributed degree distributions, with *tail exponent* $\alpha$, or simply to be power-law distributed.[10] Power-law distributed networks with low $\alpha$'s are said to be heavy-tailed.

Theorem 1 derives a large-economy equilibrium by studying the limit of a sequence of economies with increasing number of agents. A large-economy scenario makes sense for US and European capital markets, where market participation is in the tens of millions. However, one may question the plausibility of network topologies that arise in our large-economy equilibrium. After all, certain conditions are needed, namely (10)-(11), which constrain the types of network topologies that can be analyzed.

[10] Alternatively, one can define the tail exponent to be $\hat{\alpha}$ when $\sum_{i=n}^\infty d(i) \sim n^{-\hat{\alpha}}$, as, e.g., done in Gabaix (1999). Such a definition is based on the c.d.f. (or, strictly speaking, on one minus the c.d.f.), whereas our definition is based on the p.d.f. The correspondence between $\hat{\alpha}$ and $\alpha$ is then $\hat{\alpha} = \alpha - 1$.

Condition (11) ensures that the average number of connections for agents in the network is well defined as the economy grows. Condition (10) imposes a restriction on the asymptotic behavior of agents' degrees. Below we argue that our results are applicable to *socially plausible* networks.

If we were to generate a social network in a random manner by creating links between people independently with some probability $p$, then the fraction of people with $k$ many links would decrease exponentially in $k$. This is a classical random network approach and, the tail exponent is $\alpha = \infty$, so our theory applies.

However, most large social networks, including collaboration networks, friendship networks, networks of e-mail correspondence and the World Wide Web do not fit into the random network framework.[11] Instead, in these social networks, the fraction of people with $k$ many links decreases only polynomially in $k$. In other words, the degree distributions of many large social networks satisfy power-laws.[12]

Our focus is on how information disseminates in social networks, i.e., we are interested in *information networks*. Recent studies show that information flow in social groups also exhibit a pattern which is consistent with an underlying network with a power-law degree distribution.[13] Therefore, we specifically study the implications of Theorem 1 in the context of power-law networks.

In order to keep the number of parameters down, we will throughout most of the paper assume that

**Assumption 2** $s^2 = 1$.

It is convenient to study networks that are Zipf-Mandelbrot distributed, $d^n \sim ZM(\alpha, n)$, which is a particular form of power-law distribution. With a Zipf-Mandelbrot distribution, $d^n(i) = c(\alpha, n) i^{-\alpha}$, where $c(\alpha, n) = (\sum_{i=1}^{n} i^{-\alpha})^{-1}$. For $\alpha > 2$, this implies that $c(\alpha, n) \to \zeta(\alpha)^{-1}$ as $n \to \infty$, where $\zeta$ is the Riemann Zeta function (see Abramowitz and Stegun (1970), page 807). For the large network degree distribution, we write $d \sim ZM(\alpha)$. We have

**Proposition 2** *For large networks, satisfying assumptions 1 and 2, with degrees that are Zipf-Mandelbrot distributed, $d \sim ZM(\alpha)$ with tail exponent $\alpha > 2$, the conditions for Theorem 1 are satisfied with $B(\alpha) = \zeta(\alpha - 1)/\zeta(\alpha)$, where $B$ is defined in (11). If the tail-exponent, $\alpha \leq 2$, then $B = \infty$.*[14]

This immediately leads to

**Corollary 1** *$B(\alpha)$ is a decreasing, strictly convex function of $\alpha$, such that $\lim_{\alpha \to \infty} B(\alpha) = 1$, $\lim_{\alpha \searrow 2} B(\alpha) = \infty$.*

[11] Newman (2001) shows that the data on scientific collaboration are well fitted by a power-law form with an exponential cutoff. Grabowskia (2007) study friendship networks, Adamic and Adar (2005) look at e-mail correspondences, and Kumar, Raghavan, Rajagopalan, and Tomkins (1999) at the World Wide Web.

[12] Simon (1955) wrote arguably the first paper which rigorously defined and analyzed a model for power-law distributions.

[13] See, e.g., Wu, Huberman, Adamic, and Tyler (2004).

[14] For general $s$, the expression becomes $B(\alpha) = \zeta(\alpha - 1)/(s^2 \zeta(\alpha))$.

We can therefore write $\alpha = F_{ZM}(B)$, where $F_{ZM} : (1, \infty) \to (2, \infty)$.

Propositions 1 and 2 make it quite clear when we expect Theorem 1 to fail. In the case when the degree distribution satisfies a power law with a heavy-tailed degree distribution, $\alpha \leq 2$, the information asymmetry between informed and uninformed investors is so large, that the informed investors may basically infer the true value of the asset, and an asymptotic large-scale NREE may not exist. If the connectedness of the most connected agents grows faster than implied by $\alpha > 2$, a model in which the most connected agents are strategic may instead be needed. Similar breakpoints occur in economic models with power-laws at $\alpha = 2$ in other contexts, see e.g., Ibragimov, Jaffee, and Walden (2008).

Although, power laws with heavier tails do occur in social sciences (e.g., distributions that satisfy Zipf's law, which in our notation corresponds to $\alpha = 2$, see Gabaix (1999)), it has been argued that $\alpha$ is typically larger than 2 but smaller than 3 in power-law networks (see, e.g., Grossman, Ion, and Castro (2007) and Barabasi and Albert (1999)).

# 5 Financial relevance of networks

In this section, we examine asset pricing implications of information networks. Our analysis assumes that the conditions in Theorem 1 hold, so that the equilibrium price converges to (12). In other words, we confine our analysis to the large-economy equilibrium characterized by Theorem 1. As we shall see, the closed-form expressions obtained in this large economy allows us to identify novel relationships between asset prices and network connectedness.

## 5.1 Network effects on price volatility and market efficiency

From Theorem 1, we see that the price volatility is

$$var(\tilde{p}) = (\pi^*)^2\sigma^2 + (\gamma^*)^2\Delta^2. \tag{16}$$

Thanks to the linearity of the equilibrium, the price volatility can be decomposed into the *information driven* volatility component, $(\pi^*)^2\sigma^2$, and the *liquidity (supply) driven* volatility component, $(\gamma^*)^2\Delta^2$.[15]

We would expect that when the network's connectedness becomes large, price converges to payoff since the aggregate information in the economy fully reveals payoff. Indeed, it is easy to check from equations (13)-(15) that such a convergence occurs, i.e. $\pi \to 1$, $\pi_0 \to 0$ and $\gamma^* \to 0$, as $B \to \infty$. As a direct corollary, volatility becomes solely driven by information rather than liquidity in the limit. However, the convergence need not be monotone in the level of network connectedness, $B$. The following proposition completely characterizes the behavior of volatility with regard to connectedness:

**Proposition 3** *The following hold for the large-economy equilibrium characterized by Theorem 1:*

[15] We use the terminology of Ozsoylev (2005) in the decomposition of price volatility.

(a) *The information driven volatility component increases as network connectedness increases. That is,*

$$\frac{\partial\ (\pi^*)^2\sigma^2}{\partial B} \quad > \quad 0.$$

(b) *The liquidity driven volatility component is a non-monotonic function of network connectedness. In particular,*

$$\begin{aligned} \frac{\partial\ (\gamma^*)^2\Delta^2}{\partial B} &< 0, \qquad \textit{if} \quad B > \frac{\Delta}{\sigma} - \Delta^2, \\ \frac{\partial\ (\gamma^*)^2\Delta^2}{\partial B} &\geq 0, \qquad \textit{otherwise}. \end{aligned}$$

(c) *The price volatility is a non-monotonic function of network connectedness. In particular,*

$$\begin{aligned} \frac{\partial\ var(\tilde{p})}{\partial B} &> 0, \qquad \textit{if} \quad \Delta^2 < \frac{1 - B\sigma^2}{2\sigma^2} + \frac{1}{2}\sqrt{\frac{1 - 2B\sigma^2 + 5B^2\sigma^4}{\sigma^4}}, \\ \frac{\partial\ var(\tilde{p})}{\partial B} &\leq 0, \qquad \textit{otherwise}. \end{aligned}$$

As network connectedness increases, agents become, on average, better informed about the payoff. Better informed agents' demands become more aggressive, rendering the information driven volatility component to increase. This is shown in part (a) of Proposition 3.

Part (b) shows that the liquidity driven volatility component behaves in a non-monotonic fashion with regard to network connectedness. In particular, when connectedness is initially small, this component decreases as connectedness increases. The intuition is as follows. Increasing connectedness allows agents to better disentangle noise, i.e. liquidity, from payoff while using price as a public signal. Hence agents rely more on price as an information source while forming their demands which, in turn, makes demands more dependent on liquidity. This renders a larger liquidity driven volatility component. On the other hand, when connectedness is initially large, the liquidity driven volatility component decreases as connectedness increases. When the premise is a highly connected network, agents rely less on price while forming their demands with increasing levels of connectedness, because information derived from a large number of agents in the network renders price almost useless as a signal. As a result, agents' demands become less dependent on liquidity and the liquidity driven volatility component diminishes.

Due to the non-monotonicity of liquidity driven volatility component price volatility also behaves in a non-monotonic fashion, as shown in part (c) of Proposition 3. The direction of its movement with regard to connectedness depends on which of the two components, information driven or liquidity driven, is the dominant one.

A well-established empirical regularity regarding volatility is the *excess volatility* phenomenon: stock price fluctuations appear to exceed what would be explained by rational fundamental value adjustment based on random news. LeRoy and Porter (1981) and Shiller (1981) were the first to draw attention to excess volatility in the US markets. In our model, excess volatility corresponds to price being more volatile than the payoff, i.e.

$$var(\tilde{p}) \; > var(\tilde{X}) \; = \; \sigma^2.$$

The following proposition reveals the relationship between excess volatility and network connectedness.

**Proposition 4** *In the large-economy equilibrium characterized by Theorem 1, there is excess volatility if and only if $B < \Delta^2$ and $\sigma > \sqrt{\frac{\Delta^2}{\Delta^4 - B^2}}$.*

Propositions 3 and 4 complement the results of Ozsoylev (2005), who focuses on economies in which the liquidity variance, $\Delta^2$, is high, and who thereby provides a partial characterization of price volatility. In particular, Proposition 4 shows that even for modest values of liquidity variance, $\Delta^2$, there can be excess volatility when the network connectedness is low in a large economy.[16] Actually we can establish a sharp bound for the volatility ratio $\frac{var(\tilde{p})}{var(\tilde{X})}$:

**Proposition 5** *Consider the large-economy equilibrium characterized by Theorem 1. When $\frac{\Delta^2}{\sigma^2}$ is held constant, a sharp upper bound for the volatility ratio $\frac{var(\tilde{p})}{var(\tilde{X})}$ is*

$$\sup_{\Delta^2 \; > \; 0} \frac{var(\tilde{p})}{var(\tilde{X})} \; = \; 1 + \frac{\frac{\Delta^2}{\sigma^2}}{B^2}.$$

Observe from (11) that when $s^2$ equals 1 excess volatility can never be higher than $\frac{\Delta^2}{\sigma^2}$ since $B$ is always greater than or equal to 1. If $s^2$ is large, however, and the average number of connections is low, the volatility ratio can be arbitrarily large. Thus, we may expect high excess volatility in markets in which private signals are noisy and there is limited information spread between agents through network connections.

As is common in the literature, we measure market efficiency by the precision of payoff conditional on price. Even though the relationship between price volatility and network connectedness is non-monotonic, an increase in connectedness unambiguously leads to higher market efficiency, i.e., to more information revelation via price.

[16] Excess volatility can also be interpreted as return being more volatile than the payoff, i.e.

$$var(\tilde{X} - \tilde{p}) \; > var(\tilde{X}).$$

In the large-economy equilibrium characterized by Theorem 1, $var(\tilde{X} - \tilde{p}) \; > var(\tilde{X})$ if and only if

$$B \; < \; \frac{1}{3}\left(-\Delta^2 + \frac{\Delta^2\left(-3+\Delta^2\sigma^2\right)}{\left(18\Delta^4\sigma^4 - \Delta^6\sigma^6 + 3\sqrt{3}\sqrt{\Delta^6\sigma^6\left(1+11\Delta^2\sigma^2 - \Delta^4\sigma^4\right)}\right)^{1/3}} + \frac{\left(18\Delta^4\sigma^4 - \Delta^6\sigma^6 + 3\sqrt{3}\sqrt{\Delta^6\sigma^6\left(1+11\Delta^2\sigma^2 - \Delta^4\sigma^4\right)}\right)^{1/3}}{\sigma^2}\right),$$

$$\Delta \; < \; \sqrt{\frac{1}{2}\left(11 + 5\sqrt{5}\right)}\frac{1}{\sigma}.$$

The inequalities above can be attained, therefore excess volatility on return is also feasible in our model.

**Proposition 6** *In the large-economy equilibrium characterized by Theorem 1, market efficiency increases as the network's connectedness increases. That is,*

$$\frac{\partial\, Var\left(\tilde{X}|\tilde{p}\right)}{\partial B} \quad < \quad 0.$$

This result is expected since higher connectedness implies that agents' demands are based on better information, rendering price to be a more precise signal of payoff.

## 5.2 Network effects on trading profits

We now turn our attention to individual agents' trading profits. Since we are actually studying a large economy, we need to be careful when carrying out agent-level analysis: it is easy to create a sequence of finite-agent economies that satisfies the conditions of Theorem 1, in which individual agents' connectedness do not converge. In other words, we may end up with a large economy where the numbers of individual agents' neighbors are not well-defined. A simple example is constructed by alternating the indices of connected and unconnected agents as the number of agents, $n$, grows. To avoid such situations, in line with our discussion in Section 3.3, we restrict our agent-level analysis to those agents in large economies, whose connectedness are well-defined and bounded. That is, when we analyze agent $i$'s trading profit, we will assume the following:

**Assumption 3** $W_i \stackrel{\text{def}}{=} \lim_{n\to\infty} \mathbf{W}^n_{i,i}$ *exists and is finite with probability one.*

Agent $i$'s ex-ante (expected) trading profit is given by

$$\Pi_i = E\left[\left(\tilde{X} - \tilde{p}\right)\psi_i(\tilde{x}_i, \tilde{p})\right],$$

where agent $i$'s demand function, $\psi_i(\tilde{x}_i, \tilde{p})$, is of the form

$$\psi_i(\tilde{x}_i, \tilde{p}) = \frac{\bar{X}\Delta^2 + \bar{Z}B\sigma^2}{\sigma^2\Delta^2 + \sigma^2 B} - \frac{\Delta^2}{\sigma^2(\Delta^2 + B)}\tilde{p} + \frac{W_i}{s^2}(\tilde{x}_i - \tilde{p}).$$

Under assumption 3, the following proposition derives individual agents' ex-ante trading profits in a large economy.

**Proposition 7** *Consider the large-economy equilibrium characterized by Theorem 1. Assume Assumption 3 holds for agent i. Then, agent i's ex-ante trading profit,* $\Pi_i$*, is linear in the agent's connectedness,* $W_i$*. In particular,*

$$\Pi_i \;=\; \underbrace{\frac{\bar{Z}\Delta^2\left(\bar{X}\Delta^2 + B\bar{Z}\sigma^2\right)}{\left(B + \Delta^2\right)\left(\Delta^2 + B\left(B + \Delta^2\right)\sigma^2\right)} - \frac{\Delta^2}{\sigma^2(\Delta^2 + B)}\; E\left[p(\tilde{X} - \tilde{p})\right]}_{\Pi^F} \;+\; \underbrace{\frac{W_i}{s^2}\; E\left[(\tilde{X} - \tilde{p})^2\right]}_{\Pi^I_i}. \quad (17)$$

*Here, $\Pi^F$ is the information-free ex-ante trading profit, common for all agents, which is driven by the compensation an agent needs to take on risk, and $\Pi_i^I$ is the information-related ex-ante trading profit, which varies by agent.*

This result immediately implies that there is a tight connection between the network degree distribution and the distribution of agents' ex-ante trading profits:

**Corollary 2** *In a large economy characterized by Theorem 1, which satisfies assumption 1, the distribution of agents' ex-ante trading profits is an affine transformation of the network's degree distribution.*

We use Proposition 7 to examine the relationship between information networks and ex-ante trading profits in a large economy. First we focus on the impact of an individual agent's network position on her ex-ante trading profit. Then we analyze the impact of network connectedness on the average ex-ante trading profit. The average ex-ante trading profit is given by

$$\Pi \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{\sum_{i=1}^{n} E\left[\left(\tilde{X} - \tilde{p}^n\right)\psi_i^n(\tilde{x}_i^n, \tilde{p}^n)\right]}{n},$$

where $\tilde{p}^n$ and $\{\psi_i^n(\tilde{x}_i^n, \tilde{p}^n)\}_{i=1}^n$ are equilibrium prices and demands, respectively, of $n$-agent economies. Similar to what we did for individual agents in Proposition 7, we decompose the average trading profit as follows:

$$\Pi = \Pi^F + \Pi^I,$$

where $\Pi^I \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{\sum_{i=1}^{n}\Pi_i^I}{n}$. Here, $\Pi^F$ is the information-free average trading profit and $\Pi_i^I$ is the information-related average trading profit.

For simplicity, we make the following assumption:

**Assumption 4** $\bar{X} = \bar{Z} = 0$.

This normalization of the expectations of payoff and liquidity is fairly common (see, e.g., Brunnermeier (2005) and Spiegel (1998)). We then have for individual agents

**Proposition 8** *Consider the large-economy equilibrium characterized by Theorem 1. Assume that Assumption 4 holds, and that Assumption 3 holds for agent $i$.*

*(a) If the network connectedness, $B$, is held constant, then agent $i$'s ex-ante trading profit increases as her own connectedness increases. That is,*

$$\frac{\partial \Pi_i}{\partial W_i} > 0.$$

*(b) If agent $i$'s connectedness, $W_i$, is held constant, then agent $i$'s ex-ante trading profit decreases as the network's connectedness increases. That is,*

$$\frac{\partial \Pi_i}{\partial B} \quad < \quad 0.$$

The intuitions behind the proposition are straightforward. The higher the number of connections one has in an information network, her trading profit is bound to increase due to her increasing informational advantage. On the other hand, when an agent's number of connections is held constant that agent's trading profit decreases as the network connectedness increases since more information is compounded into price, diminishing the agent's informational rent.

The two effects put together make the relationship between network connectedness and average trading profit non-trivial. On the one hand, higher network connectedness implies an increase in the average profit since everyone is, on average, better informed. On the other hand, it can also imply a decrease in the average profit, because more information is compounded into price and that diminishes everyone's informational rent. This is shown in

**Proposition 9** *Consider the large-economy equilibrium characterized by Theorem 1. Assume that Assumption 4 holds.*

*(a) The average ex-ante trading profit is a non-monotonic function of network connectedness. In particular,*

$$\begin{aligned} &\frac{\partial \Pi}{\partial B} > 0, \qquad \textit{if} \quad \sigma < \frac{1}{\Delta} \quad \textit{and} \quad B < \frac{\Delta}{\sigma} - \Delta^2, \\ &\frac{\partial \Pi}{\partial B} \leq 0, \qquad \textit{otherwise.} \end{aligned}$$

*(b) $\Pi^F$ is positive, decreasing in $B$, and approaches $0$ as $B$ tends to $\infty$.*

*(c) $\Pi^I$ is positive, non-monotonic in $B$, and approaches $0$ as $B$ tends to $\infty$.*

*(d) As $B$ tends to $\infty$, $\Pi$ approaches $0$.*

Part (a) of the proposition shows that there is an optimal level of network connectedness for average trading profit. Provided that $\sigma < \frac{1}{\Delta}$, the optimal level is neither 0 nor $\infty$. If network connectedness is very low, the average agent enjoys a higher trading profit as the number of connections increases since she is getting better informed.

Part (b) tells us that the information-free component, $\Pi^F$, of average trading profit is decreasing in $B$. As we have mentioned before, the information-free component is the compensation agents need to take on risk. When $B$, i.e. the network connectedness, increases, the risk perceived by agents decreases since they become better informed. As a result, the compensation required for the perceived risk decreases.

The intuition behind part (c) of the proposition, i.e. $\Pi^I$ being non-monotonic in $B$, has already been discussed following Proposition 8. When the network connectedness is very high, agents do, on average, receive a lot of information, but they then compete away the informational rents and the trading profits vanish, as shown by part (d) of of the proposition.

If we dispense with Assumption 4, the results on trading profit will not be clear cut as in Propositions 8 and 9. However, the main result of this section, namely the non-monotonic relationship between average trading profit and network connectedness, will remain unchanged.

## 5.3 Network effects on risk-return trade-off

Next we examine the effect of information networks on the risk-return trade-off. As is common in the literature, we make use of the Sharpe ratio as the metric for this trade-off. In the computation of the Sharpe ratio, we use ex-ante expected (dollar) return and ex-ante standard deviation of return so that the ratio is given by

$$S \stackrel{\text{def}}{=} \frac{E\left[\tilde{X}-\tilde{p}\right]}{\sqrt{Var\left(\tilde{X}-\tilde{p}\right)}}.$$

We have

**Proposition 10** *The following hold for the large-economy equilibrium characterized by Theorem 1:*

*(a) Ex-ante expected return decreases as the network's connectedness increases provided that $\bar{Z} \neq 0$. That is,*

$$\frac{\partial\, E\left[\tilde{X}-\tilde{p}\right]}{\partial B} < 0, \quad \textit{if} \;\; \bar{Z} \neq 0.$$

*(b) Ex-ante return volatility decreases as the network's connectedness increases. That is,*

$$\frac{\partial\, Var\left(\tilde{X}-\tilde{p}\right)}{\partial B} < 0.$$

*(c) The Sharpe ratio is a decreasing function of network connectedness, provided that $\bar{Z} \neq 0$. That is,*

$$\frac{\partial S}{\partial B} < 0, \quad \textit{if} \;\; \bar{Z} \neq 0.$$

When network connectedness is high, agents trade aggressively based on better information, rendering price approach payoff. Therefore, both expected return and return volatility are decreasing functions of $B$. It turns out that, as connectedness increases, expected return diminishes faster than volatility (measured in standard deviation). This in turn implies that the Sharpe ratio is a decreasing function of $B$.

A well-known shortcoming of the CAPM is that the empirically estimated security market line is flatter than that predicted by the CAPM, as, e.g., shown in Black, Jensen, and Scholes (1972).

Our one-asset model cannot offer a rigorous explanation of this shortcoming, however our results in this section are encouraging in the sense that introducing information networks into multi-asset models may diminish the discrepancy between theory and observation.

### 5.4 Network effects on portfolio holdings

Arguably, the most observable effect of information networks is on portfolio holdings. For instance, Hong, Kubik, and Stein (2004) show that the trades of any given fund manager respond more sensitively to the trades of other managers in the same city than to the trades of managers in other cities. The authors interpret this empirical regularity as managers spreading information to one another directly through word-of-mouth communication. Using account-level data from People's Republic of China, Feng and Seasholes (2004) find that trades are highly correlated when investors are divided geographically. In a similar spirit to the interpretation made by Hong, Kubik, and Stein (2004), the finding of Feng and Seasholes (2004) can be attributed to the positive relationship between geographical proximity and likelihood of communication among investors. Our model provides a theoretical justification of these empirical findings.

**Proposition 11** *Consider the large-economy equilibrium characterized by Theorem 1. Assume that, for agents $i, j$, Assumption 3 holds and also that $W_{ij} \stackrel{\text{def}}{=} \lim_{n\to\infty} \mathbf{W}^n_{i,j}$ exists and is bounded, with probability one. All else held constant, the demand correlation of agents $i$ and $j$ increases as the number of their common neighbors increases. That is,*

$$\frac{\partial\ corr\left(\psi_i(\tilde{x}_i, \tilde{p}), \psi_j(\tilde{x}_j, \tilde{p})\right)}{\partial W_{ij}} \quad > \quad 0.$$

Proposition 11 finds a positive relationship between informational proximity and correlated trading. Geographical proximity is expected to encourage communication, therefore, arguably, the empirical studies cited above lend support to this result.

The impact of information networks on demand is shown in the following proposition:

**Proposition 12** *Consider the large-economy equilibrium characterized by Theorem 1, satisfying Assumption 4. For an agent satisfying Assumption 3, the expected unsigned asset demand, $\psi_i^{unsigned} = E\left[|\psi_i|\right]$, is an increasing, concave function of connectedness with asymptote,*

$$\psi_i^{unsigned} \sim W_i \sqrt{\frac{2\Delta^2\sigma^2(B^2\sigma^2 + \Delta^4\sigma^2 + \Delta^2 + 2\Delta^2 B\sigma^2)}{\pi(B^2\sigma^2 + \Delta^2 + \Delta^2 B\sigma^2)^2}}$$

*for large $W_i$. Here, $\pi$ is the mathematical constant: $\pi = 3.1415...$.*

Trading volume of individual agents is thus increasing in connectedness, with a higher slope for low degrees of connectedness. Moreover, it directly follows from Proposition 7 that trading profits and trading volume move together, i.e., higher trading volume leads to higher profits. The relationship is stronger for agents with high trading volume, since trading volume is a concave function of connectedness, whereas expected profits is a linear function of connectedness.

**Corollary 3** *An agent's expected trading profit is a convex, increasing function of expected trading volume.*

## 6 Welfare and stability in networks

In this section, we analyze the welfare implications of information networks. We base the analysis on the certainty equivalents obtained by the agents.[17]

The *ex ante* certainty equivalent of participating in the market for an agent is $CE(W)$, where $W$ is the connectedness of the agent under consideration. This is the certainty equivalent, before the agent receives his signals, and it equals the price the agent is willing to pay to participate in the stock market, before he receives his signals. We distinguish this from the *ex interim* certainty equivalent, which is the certainty equivalent after an agent has received the signals and traded, but before the value of the asset is realized.

We derive the average ex ante certainty equivalent in an economy, which we can then use to analyze which structures are optimal in that they maximize average certainty equivalence. This is the first-best optimal solution that would occur in a centralized economy, in which a central planner chose the network structure om behalf of the agents.

We also discuss the costs involved in forming networks. Specifically, we discuss entry costs, and the costs of forming links, which could in principle vary by agent. Most of the analysis, however is performed in a simplified framework in which there are no entry costs and all agents have the same constant costs of forming links.

We then move to considering a decentralized economy in which agents take actions unilaterally. In a decentralized economy, we would expect connections to be formed between agents when in is mutually beneficial for them to link. It is outside of the scope of this paper to answer the general question of which specific network topology will form in a large economy. Instead, we ask ourselves a weaker question: If agents in a network can delete or add new connections, which network topologies are "stable" in the sense that no agent has an incentive to change her position in the network, once it exists? Not surprisingly, the stable networks are not the same as the first-best optimal ones. In fact, the competition for information rents will always lead to overinvestment in connectedness in stable networks.

Finally, we study some assumptions on network costs that are consistent with stable power-law distributed networks. When all agents' cost functions are the same, strong assumptions about the cost function are needed for there to be any nondegerate stable networks, including power-law distributed networks. On the contrary, when there is dispersion between agents' cost functions, perhaps due to differences in social skills, power-law degree distributed stable networks will arise

[17] Since a natural interpretation of the stochastic supply is that it is due to noise trading, it can be argued that the welfare of noise traders is not taken into account with this measure. The welfare of the agents who are not noise traders is still important though, since it has implications for endogenous network formation, i.e., for which types of networks these agents would prefer if they could coordinate their actions and assign a central planner to decide the network structure.

naturally, under assumptions about skills that are in the same spirit as previously introduced in the literature.

## 6.1 Welfare

For a given large network, it is straightforward to show the following

**Proposition 13** *If the conditions of Theorem 1 and assumptions 1, 2 and 4 hold, then:*

*(a) For agent $i$, satisfying assumption 3, the certainty equivalent is*

$$CE(W_i) = \frac{1}{2}\log\left(\frac{\left(\Delta^2+(B+\Delta^2)^2\sigma^2\right)\left(B^2s^2\sigma^2+\Delta^2s^2+W_i\Delta^2\sigma^2\right)}{(sB^2\sigma^2+\Delta^2+\Delta^2B\sigma^2)^2}\right). \tag{18}$$

*(b) The average ex ante certainty equivalent of trading across agents is*

$$\overline{CE} = \sum_i CE(W_i)d(i). \tag{19}$$

As long as connections are cost-free, any agent will always have an incentive to form new connections. In practice, however, we would expect the formation of new links to be costly for both the connector and the connectee. For example, expanding ones' social network is time consuming and may also carry monetary costs, e.g., the costs of joining a posh golf club to connect with other investors, or the costs of moving to and living in New York City to interact with investment bankers. Even with an interpretation of the links in the network as describing pure information links, a cost may be motivated. For example, companies like Forrester Research, Inc. charge for their research — an example of proprietary costly information that is shared between a subgroup of the population (the subscribers).

In general, we would expect the cost of agent $i$ of participating in a network, forming $W$ connections, to be of the form $f(t(i),W)$, where $t(i)\in\mathbb{R}_{++}$ represents the *skill* of agent $i$.[18] We assume that $f(t,1)=0$ for all $t$, i.e., there is no cost if no connections are formed. It is also natural to assume that $f_W>0$ and that $f_t\geq 0$.

There are, however, arguments both for $f$ to be concave and convex in $W$. On the one hand, one can argue that the cost should be concave in $W$, since social networking should have fixed costs, like developing social skills or moving to an area where networking is possible. On the other hand, one can argue that the marginal costs of maintaining a network should be increasing, at least eventually, since agents have finite resources (e.g., limited time). *A priori*, we therefore do not make any assumption about the sign of $f_{WW}$. Similarly, we do make no assumptions about the signs of $f_{tt}$ and $f_{tW}$.

[18] In light of our previous discussions, the skill could either be interpreted as a social skill, or more broadly as any skill that allows the agent to gather information, e.g., proficiency in data analysis. Also, even though $W$ belongs to the set of natural numbers, for simplicity we require that $f$ is a twice continuously differentiable function on the whole of $\mathbb{R}^2_+$.

In general, we may also impose entry costs. We assume that a fraction of the population, $0 \leq r_0 \leq 1$, already participates in the market (the participants), whereas $1 - r_0$ (the nonparticipants) do not. Nonparticipants can become participants, by paying an irreversible fixed cost, $C_r \geq 0$ and participants can choose to opt out and thereby become nonparticipants. We may think of this as the cost of learning about the stock market, finding a broker, etc. The certainty equivalent of a nonparticipant is then 0 and the fraction of participants after entry/exit decisions are made is $0 \leq r \leq 1$. Agents who are already participants (the $r_0$ fraction) will always get some *ex ante* surplus from trading in the stock (they can always choose to invest their wealth in the risk-free asset once they receive their signals, so they can never be forced to be *ex interim* worse off than a nonparticipant, since they do not pay entry cost). Therefore, $r \geq r_0$.

If connections are costly to form and wealth transfers between agents are possible, which type of network topologies are the most efficient from a welfare perspective? We analyze this question under the specific assumption of a constant cost per link, no entry costs, and no variation of agents' skills. We thus have $C_p = 0$, $f(W) = c \times (W - 1)$. The average net certainty equivalent of a network is then $\overline{CE} - f(B)$. Given the connectedness, $B$, of the network, optimizing the network boils down to deciding which structure has the highest possible average net certainty equivalent,

$$\max \overline{CE} - c(B - 1). \tag{20}$$

This is *the central planner's problem.* We have

**Proposition 14** *Given $\Delta > 0$, $\sigma > 0$ and $B \geq 1$, if assumptions 2 and 4 are satisfied, then among all large networks satisfying the conditions in Theorem 1 and assumption 1:*

*(a) Given $B \in \mathbb{N}$, the maximum $\overline{CE}$ is realized by a network if and only if the networks degree distribution is $supp[d] = \{B\}$.*

*(b) Given $B \in \mathbb{R}_+ \setminus \mathbb{N}$, $B \geq 1$, the maximum $\overline{CE}$ is realized by a network if and only if $supp[d] = \{\lfloor B \rfloor, \lceil B \rceil\}$, $d(\lceil B \rceil) = B - \lfloor B \rfloor$ and $d(\lfloor B \rfloor) = 1 - B + \lfloor B \rfloor$.*

*(c) There is a solution to the central planner's problem. Moreover, any solution to the central planner's problem has connectedness $B < \infty$, and $supp[d] \subset \{\lfloor B \rfloor, \lceil B \rceil\}$.*

**Remark 2** *Condition (b) reduces to (a) when $B \in \mathbb{N}$, but we write it out for clarity.*

Thus, under the assumption of constant cost of connections, the optimal network topology is a highly uniform one: every agent basically has the same number of connections. One could, of course, argue that it would be even better for the agents in the network if the information could be shared without the formation of costly connections. There may be reasons, however, why such an outcome is infeasible, e.g., if signals can not be credibly shared unless agents have invested in connections.

## 6.2 Stability

The previous discussion relied on the underlying assumption that, somehow, global coordination could occur when deciding the network structure — a cooperative approach. We now focus on the noncooperative setting, in which two agents form connections, only if it is mutually beneficial for them to do so.

An agent in a finite network is said to be *content*, if she has no incentive to change her number of connections. A finite network is *stable* if all of its agents are content.[19] A large network that satisfies assumption 1 is stable if all agents with connections $W \in supp[d]$ are content. This is a simple definition, in the spirit of our large network approach, that allows us to derive clear results, although it may be argued that it provides a somewhat weak requirement. A stronger requirement would be to assume that for all $\mathcal{G}^n$, $n \geq n_0$, for some $n_0 \geq 1$, all agents are content. This alternative definition would, however, make the analysis quite intractable, since we have no closed form solutions for agents' utility in the finite economy setting. Moreover, the "for all agents" part of the alternative definition leads to technical difficulties, since even if $d(i) = 0$, it might be the case that $d^n(i) = o(n)$.

We now study individual agents' optimization problems. In the the same setting as in the previous section, with $C_p = 0$ and $f(W) = c \times (W - 1)$, an agent's optimization problem is to find

$$W^* = \arg\max_{W \in \mathbb{N}} \ \overline{CE}(W) - c(W - 1). \tag{21}$$

It is interesting to compare the solution to the central planner's problem with the stable networks that arise in the noncooperative setting. Obviously the welfare will be higher in the network that solves the central planner's problem, but it is *a priori* unclear whether agents will overinvest or underinvest in network connections in the noncooperative economy. We have

**Theorem 2** *Given $\Delta > 0$, $\sigma > 0$ and assumptions 1, 2 and 4, with no entry costs, $C_p = 0$, and constant costs of connections, $f(W) = c \times (W - 1)$.*

*(a) There exists a stable network in which the degree distribution has $supp[d] \subset \{\lfloor B \rfloor, \lceil B \rceil\}$, for some $B > 1$. Moreover, any stable network will have a degree distribution with $supp[d] \subset \{i, i + 1\}$ for some $i \in \mathbb{N}$.*

*(b) If the central planner's problem has a solution with connectedness $B_S$, then for any stable network with connectedness $B$, the inequality $B \geq \lceil B_S \rceil - 1$ holds.*

From part (a) of theorem 2, for economies with cost functions of the form $f(W) = c(W - 1)$, stable networks will have the same structure as networks that solve the central planner's problem: They will have support on one or two points. From part (b), it follows that for any economy with first-best optimal connectedness $B_S$, the stable network with connectedness $B$ is overinvested in connectedness. Technically, the bound is only approximate, $B \geq \lceil B_S \rceil - 1$. This is because agents

[19]This type of stability concept was introduced in Jackson and Wolinksy (1996) and Jackson and Watts (2002), in a more general game theoretic setting.

are restricted to choose $W$ from the set of natural numbers: As shown in the proof of theorem 2, in an unrestricted setting, in which both the central planner and agents can choose $W \geq 1$ from the set of real numbers when solving (20) and (21), the inequality is strict, $B > B_S$. A mechanism that helps agents coordinate toward the first-best optimum in this case will therefore need to restrict the number of connections agents can form. Such restrictions could, e.g., be implemented through exclusive "clubs," like alumni networks of Ivy league institutions.

Our result on overinvestment in network connections is reminiscent of the propositions of Hirshleifer (1971) and Marshall (1974) that competitive markets need not reflect the social value of information. It is often the case that private incentives to acquire information are excessive in competitive markets rendering the social value of information to be smaller than its private value.

The stable networks guaranteed by theorem 2 are quite different from the power-law distributed networks discussed in section 4, but we have so far only analyzed one specific form of cost function — the linear one. We may ask ourselves, which cost-functions are consistent with power-law degree distributions. If there is no dispersion among agents in their cost functions, i.e., if $f$ does not depend on type, then it turns out that only very restrictive cost functions are consistent with power-law distributed networks, or with any non-degenerate networks for that matter.

**Theorem 3** *Assume a large network, satisfying the assumptions made in Theorem 1, and assumptions 1, 2 and 4, with a degree distribution, $d$, with $S = supp[d] \subset \mathbb{N}$, $B = \sum_i i \times d(i)$. If agents face a cost function on the form $f = f(W)$, and there are entry costs $C_p \geq 0$, then such a network is stable if and only if, the cost function is on the form:*

$$f(W) = g_r(W), \qquad \textit{for some } 0 < r \leq 1, \ \textit{for all } W \in S, \tag{22}$$

*where*

$$g_r(W) \stackrel{\text{def}}{=} \frac{1}{2} \log \left( \frac{\left(\Delta^2 + (Br + \Delta^2/r)^2\sigma^2\right)\left(B^2 s^2 \sigma^2 r^2 + \Delta^2 s^2 + W\Delta^2\sigma^2\right)}{(sB^2\sigma^2 r^2 + \Delta^2 + \Delta^2 B\sigma^2)^2} \right) - c, \tag{23}$$

*for some $c \geq c'$, where*

$$c' = \frac{1}{2} \log \left( \frac{\left(\Delta^2 + (Br + \Delta^2/r)^2\sigma^2\right)\left(B^2 s^2 \sigma^2 r^2 + \Delta^2 s^2 + \Delta^2\sigma^2\right)}{(sB^2\sigma^2 r^2 + \Delta^2 + \Delta^2 B\sigma^2)^2} \right),$$

*$f(W) \geq g_r(W)$ for all $W \notin S$, and, if $r < 1$, then $C_r \geq c$. Here, $0 < r \leq 1$ is the fraction of agents participating in the market. For such a network, if $1 \in S$, then $c = c'$.*

From theorem 3, the following corollary immediately follows

**Corollary 4** *For an economy with an agent-independent cost function, $f = f(W)$, there is a stable large network with a Zipf-Mandelbrot degree distribution, $d \sim ZM(F_{ZM}(B))$, $B > 1$, if and only if $f(W) = g_r(W)$ for* all *$W \in \mathbb{N}$, for some $0 < r \leq 1$, and either the entry cost satisfies $C_r \geq c$, or $r = 1$.*

From Corollary 4, it follows that with agent-independent cost functions, a stable network with a Zipf-Mandelbrot degree distribution may only occur if the cost function is concave in $W$, i.e., if the learning component of connecting outweighs the increasing marginal costs of maintaining a network – In fact, this is true for *any* stable network with a degree distribution with support on the whole of $\mathbb{N}$.

We next study the case in which $f$ is linear in the number of connections, but may vary with agent type. Given types, $t_i^n$, for agent $i$ in economy $n$, the type distribution is characterized by the c.d.f., $G$, if $G(x) = \lim_{n\to\infty} \sup_x \left|G(x) - n^{-1}|\{i : t_i^n \leq x\}|\right| = 0$. In this case, we say that the large economy has type distribution $G$. For example, if types are i.i.d. draws from a distribution with c.d.f. $G$, then, by the Cantelli-Glivenko theorem, the type distribution in the large network is almost surely $G$. We now have

**Theorem 4** *Assume a large network, satisfying the assumptions made in Theorem 1, and assumptions 1, 2 and 4, with a power-law degree distribution with tail exponent $\alpha > 2$, that the entry cost is $C_r = 0$, and that the cost function is of the form $f(t, W) = t(W - 1)$. Then,*

*(a) There is a type distribution, $G$, that satisfies $G \sim t^{\alpha-1}$ for small $t$, for which the network is stable and all agents participate, $r = 1$.*

*(b) Any type distribution, $G$ that is twice continuously differentiable and such that $|G''| = O(x^{\delta-1})$ close to 0 for some $\delta > 0$, for which the network is stable, must have $G \sim t^{\alpha-1}$.*

As we saw, in the case of agent-independent cost functions, quite strong conditions need to be imposed to get power-law distributed stable networks: The cost function needs to take on a specific form and the entry costs need to be high. On the contrary, as Theorem 4 shows, when the cost function is agent-dependent, power-law distributed networks occur quite easily. Specifically, the behavior of the type distribution close to zero decides the tail behavior of the type distribution. A limit case occurs when the type distribution is uniform close to 0, $G \sim t$. This leads to a Zipf distribution for the tail exponent of the network, i.e., to $\alpha = 2$, which is precisely the point at which the model breaks down. Another case is when the p.d.f. of the type distribution is linear close to 0, implying that $G \sim t^2$, which leads to a tail exponent of $\alpha = 3$. As mentioned previously, empirically, networks often have tail exponents between 2 and 3, which thus corresponds to type distributions between uniform and linear.

The functional form of the type/talent distribution is in line with Gabaix and Landier (2008), who use an assignment model to study executive payment, and who empirically find that a talent distribution with an upper bound (in our case a lower bound on the cost function, 0), where the type distribution in a neighborhood of this bound satisfies a power-law relation (in our case characterized by the exponent $\alpha - 1$), fits the data well.[20] The analogy can, of course, not be taken too far, since the talent in Gabaix and Landier (2008) is interpreted as a CEO's skill to increase the revenue of a firm, whereas our measure is the cost of connecting in a network.

Although theorem 4 is only proved for the case of costs functions that are linear in the number of connections, as discussed in the proof of the theorem, it is quite straightforward to generalize

[20] A similar assignment model has been studied in Tervio (2008), with similar implications.

the result to more general settings, in which the cost function is on the form $f(t, W) = tg(W)$, and $g(W) \sim W^\beta$, $\beta > 0$. The tail exponent will in this case also depend on $\beta$.

# 7 Concluding remarks

The properties of information networks have profound impact on the prices of assets. We have introduced a simple, parsimonious model of an economy with large information networks, in which the relationship between network properties and asset pricing can be conveniently analyzed. We used the model to derive novel predictions about excess volatility, expected profits, trading volume and Sharpe ratios. For example, price volatility can be a non-monotone function of network connectedness, as can trading profits. For networks with power-law distributed degree distributions — a common property of networks in practice — trading profits are also power-law distributed.

To understand which types of network topologies may occur as outcomes of endogenous network formation, we introduced the concept of network stability, with the interpretation that in stable networks, agents have no incentive to change their positions. When costs are agent-independent and linear in the number of connections, all agents will, roughly speaking, choose the same connectedness and the network will typically overinvest in connectedness compared with what is socially optimal.

With agent-independent cost functions, stable networks with large dispersion in agent connectedness can only occur for very specific forms of the cost function of forming connections. If, on the other hand, the cost function depends on agents' types, which may interpreted as agents having different social skills, then power-law distributed stable networks occur quite naturally.

The model could potentially be extended to multiple assets, which would allow for cross-asset comparisons of information networks. For example, one may consider economies in which the information diffusion is different for different assets, e.g. using $\mathcal{E}$ — which only takes direct connections to neighbors into account — for some assets, and $\hat{\mathcal{E}}$ — which also includes neighbors of neighbors — for other. It is an open question how connectedness in one asset will influence the pricing of other assets in such a setting.

## Appendix

*Proof of Theorem 1*: We prove the result for the case when (10-11) hold surely. The proof is identical for the case stated in the theorem, when the conditions only hold in probability.

For the economy with $n$ agents, we decompose the covariance matrix, $\mathbf{S}$, into column vectors, $\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_n]$, and also define the scalars $s_i^2 = [\mathbf{S}]_{ii} = s^2/[\mathbf{W}]_{ii}$. We are interested in the existence of a linear NREE for a fixed $n$. Following the analysis Hellwig (1980), it is clear that, given a pricing relationship (8) and demand functions of the form (9), and multivariate conditional expectations on the form

$$E[\tilde{X}|\mathcal{I}_i] = \alpha_{0i} + \alpha_{1i}\tilde{x}_i + \alpha_{2i}\tilde{p}, \tag{24}$$

$$var(\tilde{X}|\mathcal{I}_i) = \beta_i, \tag{25}$$

agent $i$'s demand function (under rational expectations) is on the form

$$\psi_i(\tilde{x}_i, p) = \frac{1}{\beta_i}\left(\alpha_{0i} + \alpha_{1i}\tilde{x}_i + (\alpha_{2i} - 1)\tilde{p}\right). \tag{26}$$

The market clearing condition now gives.

$$\pi_0 = \gamma \sum_{i=1}^{n} \frac{\alpha_{0i}}{\beta_i}, \tag{27}$$

$$\pi_i = \gamma \frac{\alpha_{1i}}{\beta_i}, \tag{28}$$

where

$$\gamma = \left(\sum_{i=1}^{n} \frac{1-\alpha_{2i}}{\beta_i}\right)^{-1}. \tag{29}$$

When we wish to stress the dependence on $n$, we write $\pi_0^n$, $\pi_i^n$ and $\gamma^n$, respectively. We define the vector $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_n)^T$. The projection theorem for multivariate normal distributions, given a linear pricing function, now guarantees multivariate conditional distributions, and the following relations

$$\alpha_{0i} = \frac{\bar{X}}{b_i}\left(s_i^2(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi} + \gamma^2 n^2 \Delta^2) - (\boldsymbol{\pi}^T\mathbf{s}_i)^2\right) - \alpha_{2i}(\pi_0 - \gamma n \bar{Z}), \tag{30}$$

$$\alpha_{1i} = \frac{\sigma^2}{b_i}\left(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi} + \gamma^2 n^2 \Delta^2 - (\mathbf{1}^T\boldsymbol{\pi})(\boldsymbol{\pi}^T\mathbf{s}_i)\right), \tag{31}$$

$$\alpha_{2i} = \frac{\sigma^2}{b_i}\left((\mathbf{1}^T\boldsymbol{\pi})s_i^2 - (\boldsymbol{\pi}^T\mathbf{s}_i)\right), \tag{32}$$

$$\beta_i = \frac{\sigma^2}{b_i}\left(s_i^2(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi} + \gamma^2 n^2 \Delta^2) - (\boldsymbol{\pi}^T\mathbf{s}_i)^2\right), \tag{33}$$

and where we have defined

$$b_i = (\sigma^2 + s_i^2)\left(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi} + n^2\Delta^2\gamma^2 + (\mathbf{1}^T\boldsymbol{\pi})^2\sigma^2\right) - \left((\mathbf{1}^T\boldsymbol{\pi})\sigma^2 + (\boldsymbol{\pi}^T\mathbf{s}_i)\right)^2. \tag{34}$$

Thus, given a $\boldsymbol{\pi}$ and a scalar, $\gamma \neq 0$, which — when $\{\alpha_{1i}\}$, $\{\alpha_{2i}\}$, $\{\beta_i\}$ and $\{b_i\}$ are defined via equations (30-34) — satisfy equations (28) and (29), this generates a NREE, where $\pi_0$ can be defined via (27).

Elimination of $\{\alpha_{1i}\}$, $\{\alpha_{2i}\}$, $\{\beta_i\}$ and $\{b_i\}$ now gives

$$\pi_i = \gamma \frac{\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi} + \gamma^2 n^2 \Delta^2 - (\mathbf{1}^T\boldsymbol{\pi})(\boldsymbol{\pi}^T\mathbf{s}_i)}{s_i^2(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi} + \gamma^2 n^2\Delta^2) - (\boldsymbol{\pi}^T\mathbf{s}_i)^2}, \tag{35}$$

and by defining $\mathbf{q} = \boldsymbol{\pi}/\gamma$ (also denoted by, $\mathbf{q}_n$, when we wish to stress the size of the vector) we get a system of equations that does not depend on $\gamma$:

$$(\mathbf{q})_i = \frac{1}{s_i^2} \times \frac{\mathbf{q}^T\mathbf{S}\mathbf{q} + n^2\Delta^2 - (\mathbf{1}^T\mathbf{q})(\mathbf{q}^T\mathbf{s}_i)}{\mathbf{q}^T\mathbf{S}\mathbf{q} + n^2\Delta^2 - (\mathbf{q}^T\mathbf{s}_i)^2/s_i^2}. \tag{36}$$

Given $\mathbf{q}$, we get

$$\frac{1}{\gamma} = \sum_{i=1}^{n} \frac{\sigma^2 + s_i^2}{\sigma^2 s_i^2} + \sum_{i=1}^{n} \frac{(\mathbf{1}^T\mathbf{q} - \mathbf{s}_i^T\mathbf{q})^2 - \frac{1}{\gamma}(\mathbf{1}^T\mathbf{q} - \frac{\mathbf{s}_i^T\mathbf{q}}{s_i^2})}{\mathbf{q}^T\mathbf{S}\mathbf{q} + n^2\Delta^2 - \frac{(\mathbf{s}_i^T\mathbf{q})^2}{s_i^2}}, \tag{37}$$

which leads to

$$\gamma = \frac{1 + \sum_{i=1}^{n} \frac{(\mathbf{1}^T\mathbf{q} - \frac{\mathbf{s}_i^T\mathbf{q}}{s_i^2})}{\mathbf{q}^T\mathbf{S}\mathbf{q} + n^2\Delta^2 - \frac{(\mathbf{s}_i^T\mathbf{q})^2}{s_i^2}}}{\sum_{i=1}^{n} \frac{\sigma^2 + s_i^2}{\sigma^2 s_i^2} + \sum_{i=1}^{n} \frac{(\mathbf{1}^T\mathbf{q} - \mathbf{s}_i^T\mathbf{q})^2}{\mathbf{q}^T\mathbf{S}\mathbf{q} + n^2\Delta^2 - \frac{(\mathbf{s}_i^T\mathbf{q})^2}{s_i^2}}}, \tag{38}$$

which is bounded, since $\mathbf{S}$ is strictly positive definite. From (27) and the definition of $\mathbf{q}$, we also have

$$\frac{\pi_0}{\gamma} = \frac{\bar{X}n}{\sigma^2} - \left(\frac{\pi_0}{\gamma} - n\bar{Z}\right)\gamma \times \sum_i \frac{(\mathbf{1}^T\boldsymbol{\pi})s_i^2 - (\boldsymbol{\pi}^T\mathbf{s}_i)}{s_i^2(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi} + \gamma^2 n^2\Delta^2) - (\boldsymbol{\pi}^T\mathbf{s}_i)^2} \tag{39}$$

leading to

$$\pi_0 = \gamma n \left(\frac{\frac{\bar{X}}{\sigma^2} + \bar{Z}A}{1 + A}\right), \tag{40}$$

where

$$\begin{aligned} A &= \gamma \sum_i \frac{(\mathbf{1}^T\boldsymbol{\pi})s_i^2 - (\boldsymbol{\pi}^T\mathbf{s}_i)}{s_i^2(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi} + \gamma^2 n^2\Delta^2) - (\boldsymbol{\pi}^T\mathbf{s}_i)^2} \\ &= \sum_{i=1}^{n} \frac{(\mathbf{1}^T\mathbf{q})s_i^2 - (\mathbf{q}^T\mathbf{s}_i)}{s_i^2(\mathbf{q}^T\mathbf{S}\mathbf{q} + n^2\Delta^2) - (\mathbf{q}^T\mathbf{s}_i)^2}. \end{aligned} \tag{41}$$

Thus, if the system of equations defined in (36) has a solution, it will generate a NREE. To show that a solution indeed exists for large enough $n$, we define

$$\mathbf{y} \stackrel{\text{def}}{=} s^2\mathbf{D}^{-1}\mathbf{q}, \tag{42}$$

and the vector $\mathbf{d}$, with $(\mathbf{d})_i = \mathbf{D}_{ii}$ (We also use the notation $\mathbf{y}_n$ when we wish to stress the size of the vector). Clearly, the condition that $\mathbf{q}$ satisfies (36) is equivalent to $\mathbf{y}$ satisfying

$$(\mathbf{y})_i = \frac{\mathbf{y}^T\mathbf{W}^n\mathbf{y} + n^2\Delta^2 s^2 - (\mathbf{d}^T\mathbf{y})(\mathbf{d})_i^{-1}(\mathbf{W}^n\mathbf{y})_i}{\mathbf{y}^T\mathbf{W}^n\mathbf{y} + n^2\Delta^2 s^2 - (\mathbf{W}^n\mathbf{y})_i^2}. \tag{43}$$

We define the mapping $F_n : \mathbb{R}^n \to \mathbb{R}^n$ by the r.h.s. of (43), so a NREE can be derived from a solution to $\mathbf{y} = F_n(\mathbf{y})$. Now, $F_n$ can be rewritten as:

$$(F(\mathbf{y}))_i = 1 + \frac{(\mathbf{W}^n\mathbf{y})_i^2/n^2 - (\mathbf{d}^T\mathbf{y})(\mathbf{d})_i^{-1}(\mathbf{W}^n\mathbf{y})_i/n^2}{(\mathbf{y}^T\mathbf{W}^n\mathbf{y})/n^2 + \Delta^2 s^2 - (\mathbf{W}^n\mathbf{y})_i^2/n^2}. \tag{44}$$

Clearly, $F_n$ is a continuous mapping, as long as the denominator in (44) is not zero. We are interested in the properties of $F_n$ for $\mathbf{y}$ that are uniformly bounded in infinity-norm, i.e., $\|\mathbf{y}\|_\infty \leq C$ for some $C > 0$, regardless of $n$.

For $\mathbf{y}$ uniformly bounded in infinity norm, we have from (10) and Hölder's inequality (see Golub and van Loan (1989)), $\mathbf{a}^T\mathbf{b} \leq \|\mathbf{a}\|_1\|\mathbf{b}\|_\infty$, that $\mathbf{y}^T\mathbf{W}^n\mathbf{y}/n^2 \leq \|\mathbf{y}\|_1\|\mathbf{W}^n\|_\infty\|\mathbf{y}\|_\infty/n^2 \leq n\|\mathbf{W}^n\|_\infty\|\mathbf{y}\|_\infty^2/n^2 = no(n)/n^2 = o(1)$.

A similar argument, based on (10), implies that $(\mathbf{W}^n\mathbf{y})_i = o(n)/n = o(1)$, and therefore that $(\mathbf{W}^n\mathbf{y})_i^2/n^2 = o(1)$.

Finally, $|(\mathbf{d})_i^{-1}| \leq 1$ and $\mathbf{d}^T\mathbf{y} \leq \|\mathbf{d}\|_1 \times \|\mathbf{y}\|_\infty = \sum_i \mathbf{W}_{ii}^n \times \|\mathbf{y}\|_\infty$, and since (11) implies that $\sum_i \mathbf{W}_{ii}^n = O(n)$, we altogether get that $(\mathbf{d}^T\mathbf{y})(\mathbf{d})_i^{-1}(\mathbf{W}^n\mathbf{y})_i/n^2 = o(1)$.

These asymptotic results, together, imply that we know the behavior of $F_n$ for large $n$, through (44). For any $\epsilon > 0$, for $n$ large enough,

$$\mathbf{y} \in \mathbb{R}^n,\ \|\mathbf{y}\|_\infty \leq 2 \qquad \Rightarrow \qquad |(F_n(\mathbf{y}))_i - 1| \leq \frac{\epsilon\Delta s^2 + \epsilon\Delta s^2}{-\epsilon\Delta s^2 + \Delta s^2 - \epsilon\Delta s^2}, \tag{45}$$

implying that $F_n : [0, 2]^n \to [1 - 4\epsilon, 1 + 4\epsilon]^n$. Because the denominator of (43) is not zero in this case, we therefore have a continuous mapping $F_n : [1 - 4\epsilon, 1 + 4\epsilon]^n \to [1 - 4\epsilon, 1 + 4\epsilon]^n$ which, by Brouwer's theorem implies that there there is a $\mathbf{y} \in [1 - 4\epsilon, 1 + 4\epsilon]^n$ that solves (43) and thereby provides a NREE.

We have thus shown that for all $n \geq n_0$ for some large $n_0$, there is a NREE, defined by $\mathbf{y}_n$, such that

$$\lim_{n\to\infty} \|\mathbf{y}_n - \mathbf{1}_n\|_\infty = 0. \tag{46}$$

We now use this result to derive expressions for $\pi_0$, $\pi$ and $\gamma$, using equations (42), (38) and (39).

We have from (42), (46) and (11)

$$\lim_{n\to\infty} \frac{\mathbf{1}_n^T\mathbf{q}_n}{n} = \lim_{n\to\infty} \frac{(\mathbf{W}^n)_{ii}(\mathbf{y}_n)_i}{s^2 n} = B. \tag{47}$$

Moreover, using (42) (46) and (10), a similar argument shows that

$$\lim_{n\to\infty} \frac{\mathbf{s}_{i_n}^T\mathbf{q}_n}{n} = 0, \tag{48}$$

for any sequence of $i_n$, where $0 \leq i_n \leq n$, and similarly, via (10),

$$\lim_{n\to\infty} \frac{\mathbf{q}_n^T\mathbf{S}\mathbf{q}_n}{n^2} = 0. \tag{49}$$

We therefore have from (38)

$$\begin{aligned}
\gamma^* &= \lim_{n\to\infty} n \times \frac{1 + \sum_{i=1}^n \frac{(\mathbf{1}^T\mathbf{q} - \frac{\mathbf{s}_i^T\mathbf{q}}{s_i^2})}{\mathbf{q}^T\mathbf{S}\mathbf{q} + n^2\Delta^2 - \frac{(\mathbf{s}_i^T\mathbf{q})^2}{s_i^2}}}{\sum_{i=1}^n \left(\frac{1}{s_i^2} + \frac{1}{\sigma^2}\right) + \sum_{i=1}^n \frac{(\mathbf{1}^T\mathbf{q} - \mathbf{s}_i^T\mathbf{q})^2}{\mathbf{q}^T\mathbf{S}\mathbf{q} + n^2\Delta^2 - \frac{(\mathbf{s}_i^T\mathbf{q})^2}{s_i^2}}} \\
&= \lim_{n\to\infty} n \times \frac{1 + \sum_{i=1}^n \frac{Bn - 0}{0 + n^2\Delta^2 - 0}}{nB + \frac{n}{\sigma^2} + \sum_{i=1}^n \frac{(Bn-0)^2}{0 + n^2\Delta^2 - 0}} \\
&= \lim_{n\to\infty} n \times \frac{1 + \frac{Bn^2}{n^2\Delta^2}}{n(B + \frac{1}{\sigma^2} + \frac{(Bn)^2}{n^2\Delta^2})} \\
&= \frac{1 + \frac{B}{\Delta^2}}{B + \frac{1}{\sigma^2} + \frac{B^2}{\Delta^2}} \\
&= \frac{\sigma^2\Delta^2 + B\sigma^2}{B\sigma^2\Delta^2 + \Delta^2 + B^2\sigma^2}.
\end{aligned}$$

Similarly, by defining $\pi^* \stackrel{\text{def}}{=} \lim_{n\to\infty} \sum_{i=1}^n \pi_i^n$, we get

$$\pi^* = \lim_{n\to\infty} \gamma^* \sum_{i=1}^n \frac{(\mathbf{W}^n)_{ii}(\mathbf{y}_n)_i}{s^2 n} = \gamma^* B.$$

We need to show that $\sum_{i=1}^n \pi_i^n \tilde{\eta}_i \to_p 0$. Clearly, via Hölder's inequality and (10), we have

$$\begin{aligned} Var\left(\sum_{i=1}^n \pi_i^n \tilde{\eta}_i\right) &= (\gamma^n \times n)^2 \frac{\mathbf{1}_n^T \mathbf{W}^n \mathbf{1}_n}{n^2} \\ &\leq (\gamma^n \times n)^2 \frac{\|\mathbf{1}_n\|_1 \|\mathbf{W}^n\|_\infty \|\mathbf{1}_n\|_\infty}{n^2} \\ &= ((\gamma^*)^2 + o(1)) \times \frac{no(n)}{n^2} \to 0, \end{aligned}$$

so by Chebyshev's inequality, it is clear that $\sum_{i=1}^n \pi_i^n \tilde{\eta}_i \to_p 0$.

Finally, from (41), it is clear that $A$ approaches

$$n \times \frac{n(B-0)}{n^2(0+\Delta^2-0)} = \frac{B}{\Delta^2},$$

so through (40), it is clear that $\pi_0$ converges to

$$\gamma^* \left( \frac{\frac{\bar{X}}{\sigma^2} + \bar{Z}\frac{B}{\Delta^2}}{1 + \frac{B}{\Delta^2}} \right),$$

which after multiplying the denominator and numerator with $\sigma^2\Delta^2$ leads to the form in (15). We are done.

We stress, again, that the derivation goes through step-by-step if conditions (10-11) are expressed in probability instead. ∎

*Proof of Proposition 1*: We construct a growing sequence of "caveman" networks that converge to a given degree distribution. A caveman network is one which partitions the set of agents in the sense that if agent $i$ is connected with $j$ and $j$ is connected with $k$, then $i$ is connected with $k$ (see Watts (1999)).

We proceed as follows: First we observe that for $d(1) = 1$, the result is trivial, so we assume that $d(1) \neq 1$. For a given $d \in S^\infty$, define $k = \min_i\{i \neq 1 : i \in supp[d]\}$. For $m > k$, we define $\hat{d}^m \in S^m$ by $\hat{d}^m(i) = d(i)/\sum_{j=1}^m d(j)$. Clearly, $\lim_{m\to\infty} \sum_{i=1}^m |\hat{d}^m(i) - d(i)| = 0$. For an arbitrary $n \geq k^3$, choose $m = \lfloor n^{1/3} \rfloor$. For $1 < \ell \leq m$, $\ell \neq k$, choose $z_\ell^n = \lfloor \hat{d}^m(\ell) \times n/\ell \rfloor$, and $z_k^n = \lfloor (n - \sum_{\ell \neq k} z_\ell^m \ell)/k \rfloor$.

Now, define $\mathcal{G}^n$ as a network in which there are $z_\ell^n$ clusters of tightly connected sets of agents, with $\ell$ members, $1 < \ell \leq m$ and $n - \sum_{\ell=2}^m \ell z_\ell^n$ singletons. With this construction, $|z_\ell^n \ell/n - \hat{d}^m(i)| \leq \ell/n$ for $\ell > 2$ and $\ell \neq k$. Moreover, $|z_1^n/n - \hat{d}^m(1)| \leq (k+1)/n$, and $|z_k^n k/n - \hat{d}^m(k)| \leq (k+1)/n + m^2/n$, so $\sum_{\ell=1}^m |z_\ell^n \ell - \hat{d}^m(\ell)| \leq 2(k+1)/n + 2m^2/n = O(n^{-1/3})$.

Thus, $\sum_{i=1}^{\lfloor n^{1/3} \rfloor} |d^n(i) - \hat{d}^{\lfloor n^{1/3} \rfloor}(i)| \to 0$, when $n \to \infty$ and since $\sum_{i=1}^{\lfloor n^{1/3} \rfloor} |\hat{d}^{\lfloor n^{1/3} \rfloor}(i) - d(i)| \to 0$, when $n \to \infty$, this sequence of caveman networks indeed provides a constructive example for which the degree distribution converges to $d$.

Moreover, it is straightforward to check that if $d(i) = O(i^{-\alpha})$, $\alpha > 1$, then (10) is satisfied in the previously constructed sequence of caveman networks, and that if $\alpha > 2$, then (11) is satisfied.

If $d(i) \sim i^{-\alpha}$, $\alpha \leq 2$, on the other hand, then clearly $\sum_i d(i)i = \infty$, so (11) will fail. ∎

*Proof of Proposition 2*: We first show the form for $B$. We have:

$$\lim_{n\to\infty} \frac{\sum_{k=1}^n (\mathbf{W}^n)_{ii}}{s^2 n} = \lim_{n\to\infty} \sum_k k \times c_\alpha^n k^{-\alpha}$$
$$= \zeta(\alpha)^{-1} \sum_{k=1}^{\infty} k^{-(\alpha-1)} = \zeta(\alpha)^{-1}\zeta(\alpha-1).$$

For (10), we notice that for a network with $n = m^\alpha$ nodes, the maximum degree, $(\mathbf{W}^n)_{ii}$ will not be larger than $m$. However, since each of the neighbors to that node has no more than $m$ neighbors, $\|\mathbf{W}^n\|_\infty = \sum_j (\mathbf{W}^n)_{ij} \leq m^2 = n^{2/\alpha} = o(n)$ when $\alpha > 2$. ∎

*Proof of Proposition 3:* It follows from Theorem 1 that

$$(\pi^*)^2\sigma^2 = \frac{B^2\left(B+\Delta^2\right)^2\sigma^6}{\left(\Delta^2 + B\left(B+\Delta^2\right)\sigma^2\right)^2}, \tag{50}$$

$$(\gamma^*)^2\Delta^2 = \frac{\Delta^2\left(B+\Delta^2\right)^2\sigma^4}{\left(\Delta^2 + B\left(B+\Delta^2\right)\sigma^2\right)^2}, \tag{51}$$

$$var(\tilde{p}) = \frac{\left(B+\Delta^2\right)^2\sigma^4\left(\Delta^2+B^2\sigma^2\right)}{\left(\Delta^2 + B\left(B+\Delta^2\right)\sigma^2\right)^2}. \tag{52}$$

(50) implies that

$$\frac{\partial\ (\pi^*)^2\sigma^2}{\partial B} = \frac{2B\Delta^2\left(B+\Delta^2\right)\left(2B+\Delta^2\right)\sigma^6}{\left(\Delta^2 + B\left(B+\Delta^2\right)\sigma^2\right)^3} > 0,$$

and this proves part (a).

(51) implies that

$$\frac{\partial\ (\gamma^*)^2\Delta^2}{\partial B} = \frac{2\Delta^4\left(B+\Delta^2\right)\sigma^4 - 2\Delta^2\left(B+\Delta^2\right)^3\sigma^6}{\left(\Delta^2 + B\left(B+\Delta^2\right)\sigma^2\right)^3}.$$

The expression above is strictly negative if and only if $B > \frac{\Delta}{\sigma} - \Delta^2$. This proves part (b).

Finally, (52) implies that

$$\frac{\partial\ var(\tilde{p})}{\partial B} = \frac{2\Delta^4\left(B+\Delta^2\right)\sigma^4 - 2\Delta^2\left(-B^3+2B\Delta^4+\Delta^6\right)\sigma^6}{\left(\Delta^2 + B\left(B+\Delta^2\right)\sigma^2\right)^3}.$$

The expression above is strictly positive if and only if $\Delta^2 < \frac{1-B\sigma^2}{2\sigma^2} + \frac{1}{2}\sqrt{\frac{1-2B\sigma^2+5B^2\sigma^4}{\sigma^4}}$. This proves part (c). ∎

*Proof of Proposition 4:* By definition, there is excess volatility if $var(\tilde{p}) > \sigma^2$. From (52), this reduces to

$$\frac{\left(B+\Delta^2\right)^2\sigma^2\left(\Delta^2+B^2\sigma^2\right)}{\left(\Delta^2 + B\left(B+\Delta^2\right)\sigma^2\right)^2} > 1.$$

The latter holds if and only if $B < \Delta^2$ and $\sigma > \sqrt{\frac{\Delta^2}{\Delta^4 - B^2}}$. ∎

*Proof of Proposition 5:* Let $c \equiv \frac{\Delta^2}{\sigma^2}$. From (13), (14) and (16), it immediately follows that the volatility ratio is

$$\frac{var(\tilde{p})}{\sigma^2} = \left(\frac{\Delta^2 + B}{B\Delta^2 + c + B^2}\right)^2 \left(B^2 + c\right).$$

Since this is a continuous function of $\Delta^2$, the supremum will either be realized when $\Delta \to \infty$, $\Delta \to 0$, or at an interior point, at which the slope w.r.t. $\Delta^2$ is zero. The first order condition is $\frac{\partial}{\partial \Delta^2}\left[\frac{var(\tilde{p})}{\sigma^2}\right] = \left(1 - \frac{B(\Delta^2+B)}{B\Delta^2+c+B^2}\right) v(\Delta^2, c, B^2) = 0$, where $v(\Delta^2, c, B^2)$ is strictly positive. The first order condition therefore implies that

$$\left(1 - \frac{B\Delta^2 + B^2}{B\Delta^2 + c + B^2}\right) = 0,$$

which will not be satisfied for a strictly positive $c$. Therefore, there is no interior maximum.

Moreover,

$$\lim_{\Delta^2 \to 0} \frac{var(\tilde{p})}{\sigma^2} = \left(\frac{B}{c + B^2}\right)^2 \left(B^2 + c\right) < 1,$$

whereas

$$\lim_{\Delta^2 \to \infty} \frac{var(\tilde{p})}{\sigma^2} = \left(\frac{1}{B}\right)^2 \left(B^2 + c\right) = 1 + \frac{c}{B^2} > 1,$$

so the supremum is realized at the limit $\Delta^2 \to \infty$ and is therefore $1 + \frac{c}{B^2}$. ∎

*Proof of Proposition 6:* It is straightforward from Theorem 1 and the projection theorem that

$$\begin{aligned} var\left(\tilde{X}\middle|\tilde{p}\right) &= \sigma^2 - \frac{\left(B\frac{\sigma^2\Delta^2+\sigma^2 B}{B\sigma^2\Delta^2+\Delta^2+B^2\sigma^2}\sigma^2\right)^2}{\left(B\frac{\sigma^2\Delta^2+\sigma^2 B}{B\sigma^2\Delta^2+\Delta^2+B^2\sigma^2}\right)^2\sigma^2 + \left(\frac{\sigma^2\Delta^2+\sigma^2 B}{B\sigma^2\Delta^2+\Delta^2+B^2\sigma^2}\right)^2\Delta^2} \\ &= \frac{\Delta^2\sigma^2}{\Delta^2 + B^2\sigma^2}. \end{aligned}$$

Hence the result follows. ∎

*Proof of Proposition 7:* From (26), we know that agent $i$'s demand will take the form

$$\psi_i(\tilde{x}_i, \tilde{p}) = \frac{\alpha_{0i}}{\beta_i} + \frac{\alpha_{1i}}{\beta_i}\tilde{x}_i + \left(\frac{\alpha_{2i}}{\beta_i} - \frac{1}{\beta_i}\right)\tilde{p}.$$

Similar arguments as in the proof of Theorem 1 shows that

$$\frac{\alpha_{0i}}{\beta_i} = \frac{\bar{X}}{\sigma^2} - \left(\frac{\pi_0}{\gamma n} - \bar{Z}\right) A_i,$$

where $A_i = \gamma n \frac{(\mathbf{1}^T\boldsymbol{\pi})s_i^2 - (\boldsymbol{\pi}^T\mathbf{s}_i)}{s_i^2(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi}+\gamma^2 n^2\Delta^2) - (\boldsymbol{\pi}^T\mathbf{s}_i)^2}$ converges to $\frac{B}{\Delta^2}$ for large $n$. Therefore

$$\frac{\alpha_{0i}}{\beta_i} \xrightarrow{n\to\infty} \frac{\bar{X}\Delta^2 + \bar{Z}B\sigma^2}{\sigma^2\Delta^2 + \sigma^2 B}.$$

Similarly, we have

$$
\begin{aligned}
\frac{\alpha_{1i}}{\beta_i} &= \frac{\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi}+\gamma^2n^2\Delta^2-(\mathbf{1}^T\boldsymbol{\pi})(\boldsymbol{\pi}^T\mathbf{s}_i)}{s_i^2(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi}+\gamma^2n^2\Delta^2)-(\boldsymbol{\pi}^T\mathbf{s}_i)^2} \\
&\xrightarrow{n\to\infty} \frac{1}{s_i^2} = \frac{W_i}{s^2}, \\
\frac{\alpha_{2i}}{\beta_i} &= \frac{(\mathbf{1}^T\boldsymbol{\pi})s_i^2-(\boldsymbol{\pi}^T\mathbf{s}_i)}{s_i^2(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi}+\gamma^2n^2\Delta^2)-(\boldsymbol{\pi}^T\mathbf{s}_i)^2} \\
&\xrightarrow{n\to\infty} \frac{B}{\Delta^2\gamma^*}, \\
\frac{1}{\beta_i} &= \frac{(\sigma^2+s_i^2)\left(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi}+n^2\Delta^2\gamma^2+(\mathbf{1}^T\boldsymbol{\pi})^2\sigma^2\right)-\left((\mathbf{1}^T\boldsymbol{\pi})\sigma^2+(\boldsymbol{\pi}^T\mathbf{s}_i)\right)^2}{\sigma^2\left(s_i^2(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi}+\gamma^2n^2\Delta^2)-(\boldsymbol{\pi}^T\mathbf{s}_i)^2\right)} \\
&= \frac{(\sigma^2+s_i^2)\left(\boldsymbol{q}^T\mathbf{S}\boldsymbol{q}/n^2+\Delta^2+(\mathbf{1}^T\boldsymbol{q})^2\sigma^2/n^2\right)-\left((\mathbf{1}^T\boldsymbol{q})\sigma^2+(\boldsymbol{q}^T\mathbf{s}_i)\right)^2/n^2}{\sigma^2\left(s_i^2(\boldsymbol{q}^T\mathbf{S}\boldsymbol{q}/n^2+\Delta^2)-(\boldsymbol{q}^T\mathbf{s}_i)^2/n^2\right)} \\
&\xrightarrow{n\to\infty} \frac{(\sigma^2+s_i^2)\left(\Delta^2+B^2\sigma^2\right)-(B\sigma^2)^2}{\sigma^2s_i^2\Delta^2} = \frac{1}{s_i^2}+\frac{1}{\sigma^2}+\frac{B^2}{\Delta^2}.
\end{aligned}
\tag{53}
$$

Thus,

$$
\psi_i(\tilde{x}_i,\tilde{p}) = \frac{\bar{X}\Delta^2+\bar{Z}B\sigma^2}{\sigma^2\Delta^2+\sigma^2B}+\frac{W_i}{s^2}(\tilde{x}_i-\tilde{p})+\left(\frac{B}{\Delta^2\gamma^*}-\frac{1}{\sigma^2}-\frac{B^2}{\Delta^2}\right)\tilde{p}.
$$

Since

$$
\begin{aligned}
\frac{B}{\Delta^2\gamma^*}-\frac{1}{\sigma^2}-\frac{B^2}{\Delta^2} &= \frac{B(B\sigma^2\Delta^2+\Delta^2+B^2\sigma^2)}{\Delta^2(\sigma^2\Delta^2+\sigma^2B)}-\frac{\Delta^4+B\Delta^2}{\Delta^2(\sigma^2\Delta^2+\sigma^2B)}-\frac{B^2\sigma^2(\Delta^2+B)}{\Delta^2(\sigma^2\Delta^2+\sigma^2B)} \\
&= -\frac{\Delta^2}{\sigma^2(\Delta^2+B)},
\end{aligned}
$$

the expression for the demand function reduces to

$$
\psi_i(\tilde{x}_i,\tilde{p}) = \frac{\bar{X}\Delta^2+\bar{Z}B\sigma^2}{\sigma^2\Delta^2+\sigma^2B}-\frac{\Delta^2}{\sigma^2(\Delta^2+B)}\tilde{p}+\frac{W_i}{s^2}(\tilde{x}_i-\tilde{p}),
\tag{54}
$$

Expected profits are of the form $E[\psi_i(\tilde{x}_i,\tilde{p})(\tilde{X}-\tilde{p})]$, and therefore (17) immediately follows. ∎

*Proof of Proposition 8:* We define the average expected profit in economy $n$,

$$
\Pi^n = \frac{\sum_{i=1}^n E\left[\left(\tilde{X}-\tilde{p}^n\right)\psi_i^n(\tilde{x}_i^n,\tilde{p}^n)\right]}{n}.
$$

From Theorem 1, we know that the market clearing condition $\sum_{i=1}^n \psi_i(\tilde{x}_i, \tilde{p})/n \equiv \tilde{Z}_n$. We therefore have

$$\begin{aligned}
\Pi^n &= E\left[\left(\tilde{X} - \tilde{p}^n\right)\tilde{Z}_n\right] \\
&= E\left[\left(\tilde{X} - \pi_0^n - \sum_{i=1}^n \pi_i^n(\tilde{X} + \tilde{\eta}_i^n) + \gamma^n \tilde{Z}_n\right)\tilde{Z}_n\right] \\
&= \left(1 - \sum_{i=1}^n \pi_i^n\right) E\left[\tilde{X}\tilde{Z}_n\right] - \pi_0^n E\left[\tilde{Z}_n\right] + \gamma^n E\left[\tilde{Z}_n \tilde{Z}_n\right] \\
&= \left(1 - \sum_{i=1}^n \pi_i^n\right) \bar{X}\bar{Z} - \pi_0^n \bar{Z} + \gamma^n(\Delta^2 + \bar{Z}^2) \\
&\xrightarrow{n\to\infty} (1-\pi^*)\bar{X}\bar{Z} - \pi_0^* \bar{Z} + \gamma^*(\Delta^2 + \bar{Z}^2).
\end{aligned}$$

Now, since $\bar{X} = \bar{Z} = 0$ it follows that

$$\Pi = \gamma^*\Delta^2 = \frac{\Delta^2\left(B + \Delta^2\right)\sigma^2}{\Delta^2 + B\left(B + \Delta^2\right)\sigma^2}. \tag{55}$$

We also have

$$\begin{aligned}
\Pi_i &= \frac{\Delta^2}{\sigma^2(\Delta^2 + B)}\left((\gamma^*)^2\Delta^2 - \pi^*(1-\pi^*)\sigma^2\right) + \frac{W_i}{s^2}\left((1-\pi^*)^2\sigma^2 + (\gamma^*)^2\Delta^2\right) \\
&= \frac{\Delta^4\left(W_i + s^2\Delta^2\right)\sigma^2 + W_i\Delta^2\left(B + \Delta^2\right)^2\sigma^4}{s^2\left(\Delta^2 + B\left(B + \Delta^2\right)\sigma^2\right)^2},
\end{aligned} \tag{56}$$

It then follows from (56) that

$$\begin{aligned}
\frac{\partial \Pi_i}{\partial W_i} &= \frac{\Delta^4\sigma^2 + \Delta^2\left(B + \Delta^2\right)^2\sigma^4}{s^2\left(\Delta^2 + B\left(B + \Delta^2\right)\sigma^2\right)^2} > 0, \\
\frac{\partial \Pi_i}{\partial B} &= -\frac{2\Delta^4\left(s^2\Delta^4 + B\left(W + 2s^2\Delta^2\right)\right)\sigma^4 + 2W_i\Delta^2\left(B + \Delta^2\right)^3\sigma^6}{s^2\left(\Delta^2 + B\left(B + \Delta^2\right)\sigma^2\right)^3} < 0.
\end{aligned}$$

Hence the proposition follows. ∎

*Proof of Proposition 9:* (a) It follows from (55) that

$$\frac{\partial \Pi}{\partial B} = \frac{\Delta^4\sigma^2 - \Delta^2\left(B + \Delta^2\right)^2\sigma^4}{\left(\Delta^2 + B\left(B + \Delta^2\right)\sigma^2\right)^2}.$$

Observe that the expression above is strictly negative if and only if $\sigma < \frac{1}{\Delta}$ and $B < \frac{\Delta}{\sigma} - \Delta^2$. This proves part (a).
(b) The decomposition into $\Pi^F$ and $\Pi^I$ follows immediately from (55,56). We have

$$\Pi^I = \frac{\Delta^6\sigma^2}{B^2\sigma^2 + \Delta^2(1 + B\sigma^2))^2},$$

which is positive, decreasing in $B$ and approaches 0 as $B$ tends to $\infty$.

(c) That $\Pi^F$ is positive and approaches zero as $B \to \infty$ is immediate since

$$\pi^F = \frac{B\Delta^2\sigma^2(B^2\sigma^2 + \Delta^4\sigma^2 + \Delta^2(1+2B\sigma^2))}{(B^2\sigma^2 + \Delta^2(1+B\sigma^2))^2}. \tag{57}$$

Non-monotonicity of $\Pi^F$ in $B$ can be easily observed from (57).
(d) This follows immediately from (b) and (c) . ∎

*Proof of Proposition 10:* From Theorem 1, we have

$$\begin{aligned} E\left[\tilde{X} - \tilde{p}\right] &= (1-\pi^*)\bar{X} \qquad (58)\\ &= \frac{\Delta^2\sigma^2}{\Delta^2 + B\left(B+\Delta^2\right)\sigma^2}\bar{Z}, \qquad (59)\\ var\left(\tilde{X} - \tilde{p}\right) &= (1-\pi^*)^2\sigma^2 + (\gamma^*)^2\Delta^2 \qquad (60)\\ &= \frac{\Delta^4\sigma^2 + \Delta^2\left(B+\Delta^2\right)^2\sigma^4}{\left(\Delta^2 + B\left(B+\Delta^2\right)\sigma^2\right)^2}, \qquad (61)\end{aligned}$$

thus

$$\begin{aligned} \frac{\partial\, E\left[\tilde{X} - \tilde{p}\right]}{\partial B} &= -\frac{\Delta^2\left(2B+\Delta^2\right)\sigma^4}{\left(\Delta^2 + B\left(B+\Delta^2\right)\sigma^2\right)^2}\bar{Z} \;<\; 0,\\ \frac{\partial\, var\left(\tilde{X} - \tilde{p}\right)}{\partial B} &= -\frac{2B\Delta^4\sigma^4 + 2\Delta^2\left(B+\Delta^2\right)^3\sigma^6}{\left(\Delta^2 + B\left(B+\Delta^2\right)\sigma^2\right)^3} \;<\; 0.\end{aligned}$$

Hence we have proved parts (a) and (b).
On the other hand, (59) and (61) imply that

$$S = \frac{\frac{\Delta^2\sigma^2}{\Delta^2+B(B+\Delta^2)\sigma^2}\bar{Z}}{\sqrt{\frac{\Delta^4\sigma^2+\Delta^2(B+\Delta^2)^2\sigma^4}{(\Delta^2+B(B+\Delta^2)\sigma^2)^2}}}, \tag{62}$$

which further implies

$$\frac{\partial S}{\partial B} = -\frac{\bar{Z}\Delta^2\left(B+\Delta^2\right)\sigma^4}{\left(\Delta^2 + B\left(B+\Delta^2\right)\sigma^2\right)\left(\Delta^2 + \left(B+\Delta^2\right)^2\sigma^2\right)\sqrt{\frac{\Delta^4\sigma^2+\Delta^2(B+\Delta^2)^2\sigma^4}{(\Delta^2+B(B+\Delta^2)\sigma^2)^2}}} \;<\; 0.$$

This proves part (c). ∎

*Proof of Proposition 11:* Following Theorem 1 and (54), we can rewrite agent $i$'s demand function as follows:

$$\psi_i(\tilde{x}_i, \tilde{p}) = c_i + \frac{\Delta^2\left(-Bs^2+W_i\right)}{s^2(\Delta^2+B(B+\Delta^2)\sigma^2)}\tilde{X} + \frac{s^2\Delta^2+\left(B+\Delta^2\right)\sigma^2 W_i}{s^2(\Delta^2+B(B+\Delta^2)\sigma^2)}\tilde{Z} + \frac{\sum_{k\in W(i)}\tilde{\epsilon}_k}{s^2}, \tag{63}$$

where $c_i$ is a constant scalar. Thus,

$$\begin{aligned} cov\left(\psi_i(\tilde{x}_i,\tilde{p}), \psi_j(\tilde{x}_j,\tilde{p})\right) &= \left(\frac{\Delta^2\left(-Bs^2+W_i\right)}{s^2(\Delta^2+B(B+\Delta^2)\sigma^2)}\right)\left(\frac{\Delta^2\left(-Bs^2+W_j\right)}{s^2(\Delta^2+B(B+\Delta^2)\sigma^2)}\right)\sigma^2\\ &\quad + \left(\frac{s^2\Delta^2+\left(B+\Delta^2\right)\sigma^2 W_i}{s^2(\Delta^2+B(B+\Delta^2)\sigma^2)}\right)\left(\frac{s^2\Delta^2+\left(B+\Delta^2\right)\sigma^2 W_j}{s^2(\Delta^2+B(B+\Delta^2)\sigma^2)}\right)\Delta^2 + W_{ij}. \qquad (64)\end{aligned}$$

On the other hand, observe from (63) that variance of agent $i$'s demand, $var\,(\psi_i(\tilde{x}_i, \tilde{p}))$, does not depend on $W_{ij}$. Therefore, following (64) we have

$$\frac{\partial\; corr\,(\psi_i(\tilde{x}_i, \tilde{p}), \psi_j(\tilde{x}_j, \tilde{p}))}{\partial W_{ij}} \quad = \quad \frac{1}{\sqrt{var\,(\psi_i(\tilde{x}_i, \tilde{p}))\;\; var\,(\psi_j(\tilde{x}_j, \tilde{p}))}} \quad > \quad 0.$$

Hence we have the desired result. ▮

*Proof of Proposition 12:* From (8) and (54) it follows that $\psi_i \sim N(0, a_1 + a_2 W_i + a_3 W_i^2)$, where $a_1 = \frac{\Delta^6 + B^2\Delta^4\sigma^2}{a_4^2}$, $a_2 = \frac{1}{s^2}\left(1 + \frac{2\Delta^6\sigma^2}{a_4^2}\right)$, $a_3 = \frac{\Delta^2\sigma^2(B^2\sigma^2 + \Delta^4\sigma^2 + \Delta^2 + 2\Delta^2 B\sigma^2)}{s^4 a_4^2}$, and $a_4 = B^2\sigma^2 + \Delta^2 + \Delta^2 B\sigma^2$. Since, $E[|\tilde{z}|] = \sqrt{\frac{2A}{\pi}}$ for a general normally distributed random variable, $z \sim N(0, A)$, it follows that

$$\psi_i^{unsigned} = \sqrt{\frac{2(a_1 + a_2 W_i + a_3 W_i^2)}{\pi}}, \tag{65}$$

It immediately follows that this function is increasing and concave, with the given asymptotics. ▮

*Proof of Proposition 13:* The following lemma ensures that the limit of average certainty equivalents is equal to the average certainty equivalent in the large economy.

**Lemma 1** *If Assumption 1 and the conditions of Theorem 1 are satisfied and the function $f : \mathbb{N} \to \mathbb{R}$ is concave and increasing, then $\lim_{n\to\infty} \sum_{i=1}^n d^n(i) f(i) = \sum_{i=1}^\infty d(i) f(i)$ with probability one.*

*Proof:* Since $f$ is concave, it is clear that $f \le g$, where $g(i) \stackrel{\text{def}}{=} f(1) + (f(2) - f(1))i \stackrel{\text{def}}{=} c_0 + c_1 i$. From (11), and since $f$ is increasing, it is therefore clear that $\sum_{i=1}^n d^n(i) f(i) \in [c_0, c_0 + c_1 B + \epsilon]$, for arbitrary small $\epsilon > 0$.

Now, for an arbitrary $m$ and $\epsilon > 0$, by Assumption 1, for large enough $n_0$, for all $n \ge n_0$, $|d^n(i) - d(i)| \le \frac{\epsilon}{m(c_0 + c_1)}$. Also, for large enough $m$ and $n_0'$, for all $n \ge n_0'$, $\sum_{i=m+1}^n d^n(i) f(i) \le \epsilon$, from (11). Finally, from Assumption 1, for large enough $m$, $\sum_{i=m+1}^\infty d(i) f(i) \le \epsilon$.

Thus, for an arbitrary $\epsilon > 0$, a large enough $m$ can be chosen and $n_0^* = \max(n_0, n_0', n_0'')$ such that for all $n \ge n_0^*$,

$$\begin{aligned}
\left| \sum_{i=1}^n d^n(i) f(i) - \sum_{i=1}^\infty d(i) f(i) \right| &= \left| \sum_{i=1}^m d^n(i) f(i) - \sum_{i=1}^m d(i) f(i) \right| \\
&+ \left| \sum_{i=m+1}^n d^n(i) f(i) - \sum_{i=m+1}^n d(i) f(i) \right| + \left| \sum_{i=n+1}^\infty d(i) f(i) \right| \\
&\le \epsilon + \epsilon + \epsilon,
\end{aligned}$$

and since $\epsilon > 0$ is arbitrary, convergence follows.

The expected utility in the large economy of an agent with $W$ connections is

$$\begin{aligned}
U(W) = E\left[-e^{-d(\tilde{x}_i, p)(\tilde{X} - p)}\right] &= \frac{1}{\sqrt{8\pi^3\sigma^2\Delta^2 s^2/W}} \int\int\int -e^{-d(X+\epsilon, p)(X + \tilde{\epsilon} - p) - X^2/(2\sigma^2) - Z^2/(2\Delta^2) - \tilde{\epsilon}^2/(2s^2/W)} dX\, dZ\, d\epsilon \\
&= -\frac{sB^2\sigma^2 + \Delta^2 + \Delta^2 B\sigma^2}{\sqrt{(\Delta^2 + (B + \Delta^2)\sigma^2)(B^2 s^2\sigma^2 + \Delta^2\sigma^2 + \Delta^2\sigma^2 W)}},
\end{aligned}$$

where the last equality follows by using (12-15,54). Since $U(W) = -e^{-CE(W)}$, condition (a) immediately follows.

Moreover, since the function $CE(W)$ is increasing and concave in $W$, from Lemma 1, it is clear that the average certainty equivalent is as defined in (b). ∎

*Proof of Proposition 14:* (a) This follows immediately from Jensen's inequality, since $CE(W)$ is a strictly convex function of $W \geq 1$.

(b) We first note that the "two-point distribution," for which a fraction $B - \lfloor \mathrm{B} \rfloor$ of the agents has $\lfloor \mathrm{B} \rfloor + 1$ connection and the rest, $1 - B + \lfloor \mathrm{B} \rfloor$, has $\lfloor \mathrm{B} \rfloor$ connections has connectedness $(B - \lfloor \mathrm{B} \rfloor)(\lfloor \mathrm{B} \rfloor + 1) + (1 - B + \lfloor \mathrm{B} \rfloor)\lfloor \mathrm{B} \rfloor = B$, so the two-point distribution is indeed a candidate for an optimal distribution. Clearly, this is the only two-point distribution with support on $\{n, n+1\}$ that has connectedness $B$, and for $B \notin \mathbb{N}$, there is no one-point distribution with connectedness $B$. We define $n = \lfloor \mathrm{B} \rfloor$, $q_n = 1 - B + \lfloor \mathrm{B} \rfloor$, $q_{n+1} = B - \lfloor \mathrm{B} \rfloor$.

We introduce some new notation. We wish to study a larger space of distributions than the ones with support on the natural numbers. Therefore, we introduce the space of discrete distributions with finite first moment, $D = \{\sum_{i=0}^{\infty} r_i \delta_{x_i}\}$, where $r_i \geq 0$, and $0 \leq x_i$ for all $i$, $0 < \sum_i r_i < \infty$ and $\sum_i r_i x_i < \infty$.[21] The subset, $D^1 \subset D$, in addition satisfies $\sum_i r_i = 1$.

The c.d.f. of a distribution in $D$ is a monotone function, $F_d : \mathbb{R}_+ \to \mathbb{R}_+$, defined as $F_d(x) = \sum_{i \geq 0} r_i \theta(x - x_i)$, where $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$. Here, $\theta$ is the Heaviside step function. Clearly, $F_d$ is bounded: $\sup_{x \geq 0} F_d(x) = \sum_i r_i < \infty$. We use the Lévy metric to separate distributions in $D$, $\mathcal{D}(d_1, d_2) = \inf\{\epsilon > 0 : F_{d_1}(x - \epsilon) - \epsilon \leq F_{d_2}(x) \leq F_{d_1}(x + \epsilon)$ for all $x \in \mathbb{R}_+\}$. We thus identify $d_1 = d_2$ iff $\mathcal{D}(d_1, d_2) = 0$.

For $d \in D$, we define the operation of addition and multiplication: $d_1 = \sum_i r_i^1 \delta_{x_i^1}$, $d_2 = \sum_i r_i^2 \delta_{x_i^2}$ leads to $d_1 + d_2 = \sum_i r_i^1 \delta_{x_i^1} + \sum_i r_i^2 \delta_{x_i^2}$ and $\alpha d_1 = \sum_i \alpha r_i^1 \delta_{x_i^1}$, for $\alpha > 0$. The two-point distribution can then be expressed as $\hat{d} = q_n \delta_n + q_{n+1} \delta_{n+1}$.

The support of a distribution $d = \sum r_x \delta_{x_i}$ in $D$ is now $supp[d] = \{x_i : r_i > 0\}$. A subset of $D$ is the set of distributions with support on the integers, $D_{\mathbb{N}} = \{d \in D : supp[d] \subset \mathbb{N}\}$. For this space, we can without loss of generality assume that the $x$'s are ordered, $x_i = i$. The expectation of a distribution is $E[d] = \sum_i r_i x_i$ and the total mass is $S(d) = \sum_i r_i$. Both the total mass and expectations operators are linear. Another subset of $D$, given $B > 0$, is $D_B = \{d \in D : E[d] = B\}$.

Given a strictly concave, function $f : \mathbb{R}_+ \to \mathbb{R}$, we define the operator $V_f : D \to D$, such that $V_f(d) = \sum_i r_i \delta_{f(x_i)}$. The function $f(x) = CE(x)$, is, of course, strictly concave $\mathbb{R}_+$. Clearly, $V_f$ is a linear operator, $V_f(d_1 + d_2) = V_f(d_1) + V_f(d_2)$.

The second part of the theorem, which we wish to prove, now states that for all $d \in D^1 \cap D_{\mathbb{N}} \cap D_B$, with $B \notin \mathbb{N}$, if $d \neq \hat{d}$, it is the case that $E[V_f(d)] > E[V_f(\hat{d})]$. It turns out that the inequality holds for any strictly concave function on $f : \mathbb{R}_+ \to \mathbb{R}$. To prove this, we use Jensen's inequality, which in our notation reads:

**Lemma 2 (Jensen)***: For any $d \in D$, with support on more than one point, and for a strictly concave function, $f : \mathbb{R}_+ \to \mathbb{R}$, the following inequality holds: $E[V_f(d)] < S(d)E[V_f(\delta_{E[d]/S(d)})] = E[V_f(S(d)\delta_{E[d]/S(d)})]$.*

Now, let's take a candidate function for an optimal solution, $d \neq \hat{d}$, such that $d \in D^1 \cap D_{\mathbb{N}} \cap D_B$. Clearly, since $\hat{d}$ is the only two-point distribution in $D^1 \cap D_{\mathbb{N}} \cap D_B$, and there is no one-point distribution in $D^1 \cap D_{\mathbb{N}} \cap D_B$, the support of $d$ is at least on three points.

Also, since $q_n + q_{n+1} = 1$, and $d \in D^1$, it must either be the case that $r_n < q_n$, or $r_{n+1} < q_{n+1}$, or both. We will now decompose $d$ into three parts, depending on which situation holds: First, let's assume that $r_{n+1} \geq q_{n+1}$. If, in addition, $r_{n+1} > q_{n+1}$, then it must be that $r_n < q_n$, and $r_i > 0$ for at least one $i < n$. Otherwise, it could not be that $E[d] = B$. In this case, we define $d_1 = \sum_{i<n} r_i \delta_i$, $d_2 = r_n \delta_n + q_{n+1} \delta_{n+1}$ and $d_3 = (r_{n+1} - q_{n+1})\delta_{n+1} + \sum_{i>n+1} r_i \delta_i$. If, on the other hand, $r_{n+1} = q_{n+1}$, then, there must be an $i < n$ such that $r_i > 0$ and also a $j > n+1$ such that $r_j > 0$, since otherwise, it would not be possible to have $E[d] = B$. In this case, we define, $d_1 = \sum_{i<n} r_i \delta_i$, $d_2 = r_n \delta_n + q_{n+1} \delta_{n+1}$ and $d_3 = \sum_{i>n+1} r_i \delta_i$. Exactly the same technique can be applied in the case of $r_n \geq q_n$ and $r_{n+1} < q_{n+1}$.

Finally, in the case of $r_n < q_n$ and $r_{n+1} < q_{n+1}$, there must, again, be an $i < n$ such that $r_i > 0$ and a $j > n+1$, such that $r_j > 0$, otherwise $E[d] = B$ would not be possible. In this case, we decompose $d_1 = \sum_{i<n} r_i \delta_i$, $d_2 = r_n \delta_n + q_{n+1} \delta_{n+1}$ and $d_3 = (r_{n+1} - q_{n+1})\delta_{n+1} + \sum_{i>n+1} r_i \delta_i$.

[21] Distribution here is in the sense of a functional on the space of infinitely continuous functions with compact support, $C_0^{\infty}$ (see Hörmander (1983)), and $\delta_x$ is the Dirac distribution, defined by $\delta_x(f) = f(x)$ for $f \in C_0^{\infty}$.

These decompositions imply that

$$\begin{aligned} E[V_f(d)] &= E[V_f(d_1)] + E[V_f(d_2)] + E[V_f(d_3)] \\ &\le S(d_1)E[V_f(\delta_{E[d_1]/S(d_1)})] + E[V_f(d_2)] + S(d_3)E[V_f(\delta_{E[d_3]/S(d_3)})] \\ &= E\left[V_f\left(S(d_1)\delta_{E[d_1]/S(d_1)} + d_2 + S(d_3)\delta_{E[d_3]/S(d_3)}\right)\right] \\ &= E[V_f(d_m)], \end{aligned}$$

where $d_m = d_L + d_2 + d_R$, $d_L = S(d_1)\delta_{E[d_1]/S(d_1)}$ and $d_R = S(d_3)\delta_{E[d_3]/S(d_3)}$. Clearly, $d_m \in D^1 \cap D_B$.

Now, if $r_{n+1} \ge q_{n+1}$, since $d \in D^1$, it must be that $S(d_1)+S(d_3) = q_n - r_n$, and since $E[d_L + d_2 + d_R] = B = E[q_n\delta_n + q_{n+1}\delta_{n+1}]$ it must be that $E[d_L + d_R] = (q_n - r_n)E[\delta_n] = E[(S(d_1)+S(d_2))\delta_n] = E[d_a]$, where $d_a = (S(d_1)+S(d_2))\delta_n$. Moreover, since $d_a + d_2$ has support on $\{n, n+1\}$ and $E[d_a + d_2] = B$, it is clear that $d_a + d_2 = \hat{d}$.

From Jensen's inequality, it is furthermore clear that $E[V_f(d_L + d_R)] < E[V_f(d_a)]$, and therefore $E[V_f(d_m)] = E[V_f(d_L + d_R + d_2)] < E[V_f(d_a + d_2)] = E[V_f(\hat{d})]$. Thus, all in all, $E[V_f(d)] \le E[V_f(d_m)] < E[V_f(\hat{d})]$. A similar argument can be applied if $r_n \ge q_n$.

Finally, in the case in which $r_n < q_n$ and $r_{n+1} < q_{n+1}$, we define $\alpha = E[d_1]/S(d_1)$ and $\beta = E[d_3]/S(d_3)$. Obviously, $\alpha < n < n+1 < \beta$. Now, we can define $g_1 = \frac{\beta-n}{\beta-\alpha}(q_n - r_n)\delta_\alpha + \frac{n-\alpha}{\beta-\alpha}(q_n - r_n)\delta_\beta$ and $g_2 = \frac{\beta-n-1}{\beta-\alpha}(q_{n+1} - r_{n+1})\delta_\alpha + \frac{n+1-\alpha}{\beta-\alpha}(q_n - r_n)\delta_\beta$. Clearly, $g_1 \in D$ and $g_2 \in D$ and, moreover, $g_1 + g_2 + d_2 = d_1 + d_2 + d_3 = d$. Also, Jensen's inequality implies that $E[V_f(g_1)] < E[V_f((q_n - r_n)\delta_n)]$ and $E[V_f(g_2)] < E[V_f((q_{n+1} - r_{n+1})\delta_{n+1})]$, so $E[V_f(d)] = E[V_f(g_1 + g_2 + d_2)] < E[V_f((q_n - r_n)\delta_n + (q_{n+1} - r_{n+1})\delta_{n+1} + d_2)] = E[V_f(\hat{d})]$. We are done.

(c) From (18) and $f(W) = c(W-1)$, the central planner's problem is to maximize

$$Q(B) \stackrel{\text{def}}{=} \frac{1}{2}\log\left(\frac{\left(\Delta^2 + (B+\Delta^2)^2\sigma^2\right)\left(B^2s^2\sigma^2 + \Delta^2s^2 + B\Delta^2\sigma^2\right)}{(sB^2\sigma^2 + \Delta^2 + \Delta^2B\sigma^2)^2}\right) - c(B-1).$$

We know from (a) and (b) that, given $B$, the network will have a distribution with support on $\{\lfloor \mathrm{B} \rfloor, \lceil \mathrm{B} \rceil\}$, so it is degenerate, and moreover, $\Theta(B) \stackrel{\text{def}}{=} \frac{1}{2}\left((1 - B + \lfloor B \rfloor)CE(\lfloor B \rfloor) + (B - \lfloor B \rfloor)CE(\lfloor B \rfloor + 1)\right)$, is a continuous function of $B \ge 1$. Clearly, $Q(B) \sim -cB < 0$, for large $B$, so there can be no solutions with arbitrary large $B$. In fact any solution must have $B \le q$, for some $q < \infty$. Since $\Theta(B) - c(B-1)$ is continuous, it must therefore be the case that $\max_{B\in[1,q]} Q(B)$ has a solution. The set of such solutions are the solutions to the central planner's problem. We are done. ∎

*Proof of Theorem 2:* First, we note that since $CE(W) > 0$ for all $W \ge 1$ and $B \ge 1$, and $f(1) = 0$, it is clear that $CE(1) - f(1) > 0$ regardless of $B$, so if $C_p = 0$, it will always be optimal for an agent to participate in the market.

(a) We begin by studying the solution to the unrestricted problem, i.e., the problems in which agents can choose $W \in \mathbb{R}$, such that $W \ge 1$. We show the strict inequality $B > B_S$ for this case, and then use the result to analyze the case when $W$ is restricted to take on integer values.

From the discussion, an agent's optimization problem is to maximize $g(W|B)$, where

$$\begin{aligned} g(W|B) &\stackrel{\text{def}}{=} \frac{1}{2}\log\left(\frac{\left(\Delta^2 + (B+\Delta^2)^2\sigma^2\right)\left(B^2s^2\sigma^2 + \Delta^2s^2 + W\Delta^2\sigma^2\right)}{(sB^2\sigma^2 + \Delta^2 + \Delta^2B\sigma^2)^2}\right) - c(W-1) && (66) \\ &= \frac{1}{2}\log\left(\frac{\left(\Delta^2 + (B+\Delta^2)^2\sigma^2\right)\left(B^2s^2\sigma^2 + \Delta^2s^2\right)}{(sB^2\sigma^2 + \Delta^2 + \Delta^2B\sigma^2)^2}\right) + 1 && (67) \\ &+ \frac{1}{2}\log\left(1 + \frac{\Delta^2\sigma^2}{B^2s^2\sigma^2 + \Delta^2s^2}\right) - cW \stackrel{\text{def}}{=} c_0 + \log(1 + c_1W) - cW. && (68) \end{aligned}$$

For any $B$, $g(W|B)$ is obviously strictly concave in $W$, so any optimum, given $B$ is unique. Moreover, due to the slow growth of the log function, $g(W|B)$ will obviously be negative for large $W$.

Let $W^*(B) : [1,\infty) \to [1,\infty)$ denote the optimal $W$, given $B$. It is easy to check that for large $B$,

$W^*(B) = 1$, which is intuitive, since in this case, the value of information, as well as of risk-reduction, is very low. A stable solution is then a $B$, $1 \le B < \infty$, such that $W^*(B) = B$, and we wish to prove that such a $B$ exists.

In fact, by taking the first order conditions on $g(W|B)$ with respect to $W$, it is easy to check that

$$W^* = \max\left\{\frac{1}{2c} - \frac{1}{c_1}, 1\right\},$$

from which it follows that the unique solution to $W^*(B) = B$ is

$$B = \max\left\{\frac{\Delta}{2s^2} + \sqrt{\frac{\Delta^4}{4s^4} + \frac{1}{2c} + \frac{s^2}{\sigma^2}}, 1\right\}.$$

We now apply these results to the restricted problem, i.e., the problem in which $W \in \mathbb{N}$. Let $B \ge 1$ be a solution to the unrestricted problem, which we have shown exists. If $B \in \mathbb{N}$, the result follows from (a). If $B \notin \mathbb{N}$ and $g(\lfloor \mathrm{B} \rfloor | B) = g(\lceil \mathrm{B} \rceil | B)$, where $g$ is defined in (66), then from the strict concavity of $g(W|B)$ in $W$, it is clear that a stable solution to the restricted problem is one in which the fraction $B - \lfloor \mathrm{B} \rfloor$ chooses $\lfloor \mathrm{B} \rfloor + 1$ connections and the rest of the agents choose $\lfloor \mathrm{B} \rfloor$ connections. For the case $B \notin \mathbb{N}$, $g(\lfloor \mathrm{B} \rfloor | B) \neq g(\lceil \mathrm{B} \rceil | B)$, we assume w.l.o.g. that $g(\lfloor \mathrm{B} \rfloor | B) > g(\lceil \mathrm{B} \rceil | B)$. From (66) it follows that $\frac{dW^*}{dB} < 0$.

Now, let $r(B') \stackrel{\text{def}}{=} g(\lfloor \mathrm{B} \rfloor | B') - g(\lfloor \mathrm{B} \rfloor + 1 | B')$ for $\lfloor \mathrm{B} \rfloor \le B' \le B$. There are two cases: The first case is $r(B') < 0$ for all $\lfloor \mathrm{B} \rfloor \le B' \le B$. In this case, $g(\lfloor \mathrm{B} \rfloor | \lfloor \mathrm{B} \rfloor) > g(\lfloor \mathrm{B} \rfloor + 1 | \lfloor \mathrm{B} \rfloor)$, and since $W^*(\lfloor \mathrm{B} \rfloor) > W^*(B)$ (because $\frac{dW^*}{dB} < 0$) it must be that $W^*(\lfloor \mathrm{B} \rfloor) < \lfloor \mathrm{B} \rfloor + 1$, since otherwise $W^*(B') = \lfloor \mathrm{B} \rfloor + 1$ for some $\lfloor \mathrm{B} \rfloor \le B' \le B$, in which case $g(\lfloor \mathrm{B} \rfloor + 1 | B') > g(\lfloor \mathrm{B} \rfloor | B')$, contrary to the assumption. In this case, a stable network is given by all agents having $\lfloor \mathrm{B} \rfloor$ connections, due to the concavity of agents' optimization problem.

The second case occurs if $r(B') = 0$ for some $\lfloor \mathrm{B} \rfloor \le B' \le B$. However, in this the network in which a fraction $B' - \lfloor B' \rfloor$ chooses $\lfloor B + 1 \rfloor$ connections and the rest chooses $\lfloor B \rfloor$ connections is obviously stable. A similar argument can be made if $B \notin \mathbb{N}$, $g(\lfloor \mathrm{B} \rfloor | B) < g(\lceil \mathrm{B} \rceil | B)$. Thus, there exists a stable solution with a degree distribution with support on $\{\lfloor \mathrm{B} \rfloor, \lceil \mathrm{B} \rceil\}$.

Now, assume that there is a stable network with degree distribution $d$, $i \in supp[d]$, $j \in supp[d]$ and $j \ge i + 2$. Then there is a $k \in \mathbb{N}$ such that $i < k < j$, and, due to the strictly concave optimization problem that agents face, it must be that $g(k|B) > \frac{k-i}{j-i} g(i|B) + \frac{j-k}{j-i} g(j|B) = g(i|B)$, which contradicts that $i \in supp[d]$, so no such solution can exist. We are done.

(b) We first note that the result follows trivially for $B_S = 1$, and therefore focus on the case for which $B_S > 1$. An optimal solution, $B^* > 1$ to the central planner's unrestricted problem is characterized by

$$\left.\frac{\partial \Theta}{\partial B}\right|_{B=B^*} = 0,$$

i.e.,

$$h(B^*) \stackrel{\text{def}}{=} -\Delta^2\sigma^2 \frac{B^2\sigma^2 + 3B^3\Delta^2\sigma^4 + \Delta^6\sigma^2(-1 + 2s^2 + B\sigma^2) + \Delta^4(-1 + B(-3 + 4s^2)\sigma^2 + 3B^2\sigma^4)}{2(B^2s^2\sigma^2 + \Delta^2(s^2 + B\sigma^2))(B^4\sigma^4 + B^2\Delta^2\sigma^2(2 + 3B\sigma^2) + \Delta^6(\sigma^2 + B\sigma^4) + \Delta^4(1 + 3B\sigma^2 + 3B^2\sigma^4))} = c.$$

The first order equation in the unrestricted noncooperative setting, from (68), is

$$f(B^{**}) \stackrel{\text{def}}{=} \frac{c_1}{2(1 + c_1 B)} = c.$$

Now,

$$\begin{aligned} f(B) - h(B) &= \Delta^2\sigma^4 \frac{\Delta^6\sigma^2 + B^4\sigma^2 + 3B^2\Delta^2\sigma^2 + B^2(\Delta^2 + 3\Delta^4\sigma^2) + B(2\Delta^4\sigma^2 + \Delta^6\sigma^2)}{2(B^2s^2\sigma^2 + \Delta^2(s^2 + B\sigma^2))(B^4\sigma^4 + B^2\Delta^2\sigma^2(2 + 3B\sigma^2) + \Delta^6(\sigma^2 + B\sigma^4) + \Delta^4(1 + 3B\sigma^2 + 3B^2\sigma^4))} \\ &> 0 \end{aligned}$$

Thus, if $h(B^*) = c$, then $f(B^*) > c$, and therefore $W^*(B^*) > B^*$. Moreover, for $B$ large enough, $W^*(B) < B$, and since $W^*$ is a continuous function of $B$, from the intermediate value theorem it follows that for some $B^{**} > B^*$, $W^*(B^{**}) = B^{**}$. This is the unique solution to the unrestricted noncooperative problem, so

$B^{**} > B^*$.

For the central planner's restricted problem, it is clear that for any solution with a degree distribution with support on $\{\lceil B_S\rceil - 1, \lceil B_S\rceil\}$ for some, $B_S > 1$ it must be that $h(B^*) - c = 0$ for some $B^* \in [\lceil B_S\rceil - 1, \lceil B_S\rceil]$. Otherwise, either choosing a network with a degree distribution with support on $\{\lceil B_S\rceil - 1\}$ or on $\{\lceil B_S\rceil\}$ would strictly improve the average certainty equivalent, contrary to the assumption. Similarly, for any solution with a degree distribution with support on $\{\lceil B_S\rceil\}$, $B_S > 1$, there must exist a $B^* \in [\lceil B_S\rceil - 1, \lceil B_S\rceil + 1]$ such that $h(B^*) = 0$.

In line with our previous discussion, at any such point, $g(B^*|B^*) > 0$, and thereby there is a stable solution to the unrestricted noncooperative problem, with connectedness $B^{**}$, $B^{**} > B^*$. However, this implies that there is a solution to the restricted noncooperative problem, with support on $\{\lfloor B^{**}\rfloor, \lceil B^{**}\rceil\}$, and since $B^{**} > B^* \geq \lceil B_S\rceil - 1$, the result follows. We are done.

∎

*Proof of Theorem 3:* Clearly, there can be no equilibrium in which $r = 0$, since markets will not clear in this case. Moreover, in any stable equilibrium, in which the degree distribution has support on $S$, agents must be indifferent between between any number of connections $i$, for $i \in S$. Now, an economy in which a fraction $0 < r < 1$ participates will be identical to the one in which $r = 1$, except for that the stochastic supply *per participant* will be higher. Technically this leads to the transformation $\Delta \mapsto \Delta/r$.

For the moment neglecting the cost function, $f$, from (18) and the fact that $\Delta \mapsto \Delta/r$ in the economy in which only the fraction $r$ of agents participates, the certainty equivalent for an agent with $j$ connections is

$$CE(j, r) = \frac{1}{2}\log\left(\frac{\left(\Delta^2 + (Br + \Delta^2/r)^2\sigma^2\right)\left(B^2 s^2\sigma^2 r^2 + \Delta^2 s^2 + j\Delta^2\sigma^2\right)}{(sB^2\sigma^2 r^2 + \Delta^2 + \Delta^2 B\sigma^2)^2}\right).$$

Define $U_j \stackrel{\text{def}}{=} CE(i, r) - f(j)$, $j \in \mathbb{N}$, to be the utility a participating agent gets by choosing to have $j$ links. Let $i \in S$, and define $c = U_i$. Then, since agents are indifferent between choosing $j$ connections, for $j \in S$, $U_j = U_i$ for all $j \in S$, which immediately implies that it is necessary for $f(j)$ to have the prescribed form, with $c = U_i$. Moreover, for $j \notin S$, it must be that $U_j \leq U_i$, which immediately implies that $f(j) \geq g(j, r)$.

Since, by assumption $f(1) = 0$, it must also be that $c = CE(1, r)$ if $1 \in S$ and $c \geq CE(1, r)$ if $1 \notin S$.

As discussed in the proof of Theorem 2, it is always the case that $CE(1, r) > 0$, so $c > 0$.

Finally, it is clear that if $r < 1$, it must be the case that $C_r \geq c$, since otherwise nonparticipants would be strictly better off by becoming participants. This, however, also provides a sufficient condition, since if it is satisfied nonparticipants are (weakly) better off staying outside of the market, and no participant has an incentive, either to become a nonparticipant, nor to change his number of links. We are done.

∎

Proof of Theorem 4: An agent of type $t$ will solve the optimization problem $\max_W U_t(W)$, where $U_t(W) = CE(W) - t(W-1) = \frac{1}{2}\log(q_0 + q_1 W) - t(W-1)$, where $q_0$ and $q_1$ follow from (18). The unrestricted problem (in that $W$ is not required to be a natural number, but can be any positive real number), is a concave optimization problem and, for $t$ close to 0, the $W$ that satisfies the first order condition gives the optimal solution (for large $t$, $W = 1$ is the optimal solution), i.e., $W = \frac{1}{2q_1 t} - \frac{q_0}{q_1}$.

Clearly, the optimal $W$ increases as $t$ decreases, and since the optimization problem is strictly concave, any agent of type $t \in (t_{W+1}, t_W)$, where $U_{t_W}(W-1) = U_{t_W}(W)$ and $U_{t_{W+1}}(W) = U_{t_{W+1}}(W+1)$ will therefore choose the connectedness to be $W$. The solution to the equation $U_{t_W}(W) = U_{t_W}(W+1)$ is

$$t_W = \frac{1}{2}\log\left(1 + \frac{q_1}{q_0 + q_1 W}\right), \tag{69}$$

so agents of type $t \in \left[\frac{1}{2}\log\left(1 + \frac{q_1}{q_0 + q_1(W-1)}\right), \frac{1}{2}\log\left(1 + \frac{q_1}{q_0 + q_1 W}\right)\right]$ will choose to have $W$ connections for $W > 1$. For $t > \frac{1}{2}\log\left(1 + \frac{q_1}{q_0}\right)$, they will choose to have 1 connection.

The network will be stable if the type distribution satisfies

$$1 - G(t_1) = d(1), \text{ and } G(t_i) - G(t_{i+1}) = d(i), \quad 1 \leq i, \tag{70}$$

where $d$ is the degree distribution of the network, as long as every agent with $W$ connections is "matched" with a type $t \in (t_{W+1}, t_W)$.

There are obviously many c.d.f.'s that satisfy these constraints, e.g., a piecewise linear function, which corresponds to a piecewise uniform type distribution. Moreover, any choice of type distribution that does not satisfy (70) will not match the degree distribution, so it will not lead to a stable network. Thus, (70) is a sufficient and necessary condition on the type distribution for it to be possible to create a network with degree distribution $d$. Specifically, we can choose $G$ to be twice continuously differentiable close to 0, such that $G'' \sim t^{\delta-1}$ for some $\delta > 0$, which immediately implies that $g \stackrel{\text{def}}{=} G' \sim x^{\delta}$ and $G \sim x^{\delta+1}$.

Since $t \in \mathbb{R}_{++}$, we have $G(0) = 0$. It also follows from a Taylor expansion of $1/W$ close to 0 in (69) that

$$t_W = \frac{1}{2}\log\left(1 + \frac{1}{W} \times \frac{1}{1 + \frac{q_0}{q_1 W}}\right) = \frac{1}{2}\log\left(1 + \frac{1}{W} \times \left(1 - \frac{q_0}{q_1 W} + O(W^{-2})\right)\right) = \frac{1}{2W} + O(W^{-2}),$$

and that

$$\begin{aligned}
t_W - t_{W+1} &= \frac{1}{2}\log\left(1 + \frac{q_1}{q_0 + q_1 W}\right) - \frac{1}{2}\log\left(1 + \frac{q_1}{q_0 + q_1(W+1)}\right) \\
&= \frac{1}{2}\left(\left(\frac{1}{W} - \frac{q_0}{q_1} \times \frac{1}{W^2} + O(W^{-3})\right) - \frac{1}{2}\left(\frac{1}{W} - \frac{q_0}{q_1} \times \frac{1}{W^2} + O(W^{-3})\right)^2\right) \\
&- \frac{1}{2}\left(\left(\frac{1}{W+1} - \frac{q_0}{q_1} \times \frac{1}{(W+1)^2} + O(W^{-3})\right) - \frac{1}{2}\left(\frac{1}{(W+1)} - \frac{q_0}{q_1} \times \frac{1}{(W+1)^2} + O(W^{-3})\right)^2\right) \\
&= \frac{1}{2}\left(\left(\frac{1}{W} - \frac{q_0}{q_1} \times \frac{1}{W^2} + O(W^{-3})\right) - \frac{1}{2}\left(\frac{1}{W^2} + O(W^{-3})\right)\right) \\
&- \frac{1}{2}\left(\left(\frac{1}{W} - \frac{1}{W^2} + O(W^{-3})\frac{q_0}{q_1} \times \frac{1}{W^2}(1 + O(W^{-1})) + O(W^{-3})\right) - \frac{1}{2}\left(\frac{1}{W^2} + O(W^{-3})\right)\right) \\
&= \frac{1}{2W^2} + O(W^{-3}).
\end{aligned}$$

For (70) to be satisfied, with $d(W) \sim W^{-\alpha}$, it must be that $G(t_W) - G(t_{W+1}) \sim W^{-\alpha}$ and since

$$\begin{aligned}
G(t_W) - G(t_{W+1}) &= g(t_W)(t_W - t_{W+1}) + \frac{g'(\xi)}{2}(t_W - t_{W+1})^2, \quad \text{where } \xi \in [t_{W+1}, t_W], \\
&= g(t_W)\left(\frac{1}{2W^2} + O(W^{-3})\right) + \frac{g'(\xi)}{2}\left(\frac{1}{4W^4} + O(W^{-5})\right) \\
&\sim W^{-\delta}\left(W^{-2} + O(W^{-3})\right) + W^{-\delta+1}\left(W^{-4} + O(W^{-5})\right) \\
&\sim W^{-\alpha},
\end{aligned}$$

where the last second to last inequality holds if and only if $g \sim W^{-\delta}$, and the last inequality holds if and only if $-\delta - 2 = -\alpha$, i.e., if and only if $\delta = \alpha - 2$, (70) will be satisfied if and only if $g(t) \sim W^{\alpha-2}$, and therefore $G \sim W^{\alpha-1}$. We are done.

We note that the proof could be extended to cost functions on the form $f = t \times h(w)$, where $h(w)$ is a nonlinear function, such that $h(w) \sim h^{\beta}$, for some $\beta > 0$. In this case the type distribution that corresponds to a specific degree distribution will also be a function of $\beta$. ∎

# References


Abramowitz, M., and I. A. Stegun (1970): *Handbook of Mathematical Functions*. Dover.

Adamic, L., and E. Adar (2005): "How to search a social network," *Social Networks*, 27, 187–203.

Admati, A. R. (1985): "A noisy rational expectations equilibrium for multi-asset securities markets," *Econometrica*, 53, 629–658.

Barabasi, A. L., and R. Albert (1999): "Emergence of scaling in random networks," *Science*, 286, 509–512.

Black, F., M. Jensen, and M. Scholes (1972): "The capital asset pricing model: Some empirical tests," in *Studies in the Theory of Capital Markets*, ed. by M. Jensen. Praeger, New York, NY.

Brumen, G., and P. Vanini (2008): "Pricing credit risk in buyer-supplier networks," SSRN Electronic paper collection, http://ssrn.com/abstract=1107700.

Brunnermeier, M. (2005): "Information leakage and market efficiency," *Review of Financial Studies*, 18, 418–457.

Chung, F., and L. Li (2006): *Complex Graphs and Networks*. American Mathematical Society.

Cohen, L., A. Frazzini, and C. Malloy (2007): "The small world of investing: Board connections and mutual fund returns," *Journal of Political Economy*, forthcoming.

Colla, P., and A. Mele (2008): "Information linkages and correlated trading," Working paper, presented at AFA 2008.

Das, S. J., and J. Sisk (2005): "Financial communities," *Journal of Portfolio Management*, 31, 112–123.

DeMarzo, P. M., D. Vayanos, and J. Zwiebel (2003): "Persuasion bias, social influence, and uni-dimensional opinions," *Quarterly Journal of Economics*, 118, 909–968.

——— (2004): "Financial markets and social networks," Working paper.

Diamond, D., and R. Verrecchia (1981): "Information aggregation in a noisy rational expectations equilibria," *Journal of Financial Economics*, 9, 219–317.

Feng, L., and M. Seasholes (2004): "Correlated trading and location," *Journal of Finance*, 59, 2117–2144.

Gabaix, X. (1999): "Zipf's law for cities: An explanation," *Quarterly Journal of Economics*, 114, 739–767.

Gabaix, X., and A. Landier (2008): "Why has CEO pay increased so much," *Quarterly Journal of Economics*, forthcoming.

Golub, G. H., and C. F. van Loan (1989): *Matrix Computations*. John Hopkins University Press, 2 edn.

Grabowskia, A. (2007): "Interpersonal interactions and human dynamics in a large social network," *Physica A: Statistical Mechanics and its Applications*, 385, 363–369.

Grossman, J., P. Ion, and R. D. Castro (2007): "The Erdos number project," http://www.oakland.edu/enp/.

Hellwig, M. F. (1980): "On the aggregation of information in competitive markets," *Journal of Ecnomic Theory*, 22, 477–498.

Hirshleifer, J. (1971): "The private and social value of information and the reward to inventive activity," *American Economic Review*, 61, 561–574.

HONG, H., J. D. KUBIK, AND J. C. STEIN (2004): "Social interaction and stock-market participation," *Journal of Finance*, 49, 137–163.

HÖRMANDER, L. (1983): *The analysis of linear partial differential operators, I-IV*. Springer Verlag.

IBRAGIMOV, R., D. JAFFEE, AND J. WALDEN (2008): "Nondiversification traps in markets for catastophe insurance," *The Review of Financial Studies*, Forthcoming.

IVKOVIĆ, Z., AND S. WEISBENNER (2007): "Information diffusion effects in individual investors' common stock purchases: Covet thy neighbors' investment choices," *Review of Financial Studies*, 20, 1327–1357.

JACKSON, M. O. (2008a): *Social and Economic Networks*. Princeton University Press.

——— (2008b): "Social networks in economics," in *Handbook of Social Economics*, ed. by J. Benhabib, A. Bisin, and M. O. Jackson. Elsevier, forthcoming.

JACKSON, M. O., AND B. W. ROGERS (2007): "Meeting strangers and friends of friends: How random are social networks," *American Economic Review*, 97, 890–915.

JACKSON, M. O., AND A. WATTS (2002): "The evolution of social and economic networks," *Journal of Economic Theory*, 106, 265–295.

JACKSON, M. O., AND A. WOLINKSY (1996): "A strategic model of social and economic networks," *Journal of Economic Theory*, 71, 44–74.

KHANDANI, A., AND A. W. LO (2007): "What happened to the quants in August 2007?," SSRN Electronic paper collection, http://ssrn.com/abstract=1015987.

KUMAR, S., P. RAGHAVAN, S. RAJAGOPALAN, AND A. TOMKINS (1999): "Trawling the web for emerging cybercommunities," p. 403416.

LEROY, S., AND R. PORTER (1981): "The present-value relation: Tests based on implied variance bounds," *Econometrica*, 49, 555–574.

MARSHALL, J. (1974): "Private incentives and public information," *American Economic Review*, 64, 373–390.

MILGRAM, S. (1967): "The small world problem," *Psychology Today*, 2, 60–67.

NEWMAN, M. (2001): "The structure of scientific collaboration networks," *Proeedings of the National Academy of Sciences USA*, 98, 404–409.

OZSOYLEV, H. (2005): "Asset pricing implications of social networks," Working paper, presented at AFA 2005.

SHILLER, R. (1981): "Do stock prices move too much to be justified by subsequent changes in dividends?," *American Economic Review*, 71, 421–436.

SIMON, H. A. (1955): "On a class of skew distribution functions," *Biometrika*, 42, 425–440.

SPIEGEL, M. (1998): "Stock Price volatility in a multiple security overlapping generations model," *Review of Financial Studies.*, 11, 419–447.

TERVIO, M. (2008): "The difference that CEOs make: An assignment model approach," *American Economic Review*, forthcoming.

WATTS, D. (1999): *Small worlds*. Princeton University Press.

WATTS, D. J., AND S. H. STROGATZ (1998): "Collective dynamis of 'small world' networks," *Nature*, 393, 440–442.

WU, F., B. HUBERMAN, L. ADAMIC, AND J. TYLER (2004): "Information flow in social groups," *Physica A*, 337, 327–335.