Asset Pricing in Large Information Networks*

Han N. Ozsoylev† Johan Walden ‡

October 19, 2010

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 network topology. We focus on networks that are sparse and have power law degree distributions, in line with empirical studies of large scale social 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 profit. Moreover, the profit distribution among investors and their trading volume are intimately linked to the topological properties of the information network. We also study agent welfare and show that uniform networks always dominate non-uniform networks with the same degree of connectedness, and that the network that optimizes total welfare is typically one with an intermediate degree of connectedness.

Keywords: *Information networks, noisy rational expectations equilibrium, power law.*

*Helpful discussions with Jonathan Berk, Sanjiv Das, Tina Eliassi-Rad, Xavier Gabaix, Nicolae Garleanu, Peter Jones, Pete Kyle, Dmitry Livdan, Santiago Oliveros, Christine Parlour, Matt Spiegel, Richard Stanton, and Marko Tervio are gratefully acknowledged. We benefited from comments by seminar audiences at UC Berkeley, National University of Singapore, IDEI - University of Toulouse, the 2008 Oxford Financial Research Summer Symposium, the 2008 NBER Summer Institute Asset Pricing Workshop, the 2008 NBER Behavioral Finance Workshop, the 2009 CRETA Workshop at University of Warwick, and the 2009 CARESS-Cowles Conference at Yale University. We also thank the Associate Editor and referee for valuable comments and suggestions.

†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 abundant casual evidence about this subject. The following recent example vividly displays the influence of information networks: Hedge fund manager John Paulson profited US$ 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 US$ 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.

The implications of information networks for the aggregate behavior of asset prices are potentially large. For example, an important stylized fact about stock markets is that price movements are not easily explained by public news: Both Cutler, Poterba, and Summers (1989) and Fair (2002) document that most large stock market movements are not associated with the arrival of public information. It also seems difficult to reconcile market-wide movements with the arrival of private information at the individual investor level. Instead, such movements are consistent with an economy in which information is gradually diffused into asset prices through an information network of small traders. This new channel of gradual information diffusion is also consistent with other puzzling stylized properties of stock markets, e.g., highly time-varying return volatility and trading volumes.

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

[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.

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 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 different approach may be possible, however. 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.

[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. Also, Xia (2007) develops an asset pricing model in which boundedly-rational agents communicate information in social networks.

[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.

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, so 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. We show the existence of—and completely characterize—equilibrium under general conditions. Our existence theorem provides a contribution in itself, since it provides a significant extension of Hellwig (1980) by allowing for information commonality across agents in a large economy noisy rational expectations equilibrium, i.e., unlike Hellwig (1980), our model allows for agents to have information with correlated error terms and with severely different signal precisions.

We find closed form expressions for price, expected profits, price volatility, trading volume and value of connectedness. We analyze how connectedness influences asset pricing and the expected profits of 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 and, in particular, what types of networks lead to more dispersed wealth distributions. We also study welfare across different networks, in terms of agents' certainty equivalents. Interestingly, several aggregate properties of the market are typically non-monotonic functions of network connectedness, e.g., price volatility, expected trading profits and agent welfare.

The rest of the paper is organized as follows. In section 2, we present the model and derive equilibrium in closed form for large economies. We also elaborate on the types of information networks that are socially plausible and the role such networks play in our analysis. Section 3 examines the implications of information networks for asset prices and agent welfare. Section 4 discusses various potential extensions and alternative assumptions, whereas Section 5 maps out how the asset pricing implications of our model can be empirically tested. Finally, we make some concluding remarks in section 6. Proofs are delegated to the Appendix.

2 Model

We follow the large economy analysis in Hellwig (1980) closely,[6] but extend the analysis to allow for network relationships: Agents communicate information to each other about asset payoffs, and this communication takes place according to an information network. In

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

particular, each agent has some information about her network neighbors' payoff-related information. The information network is exogenous, and can be considered to represent information sharing across social connections, such as friendships and acquaintanceships. Alternatively, being in the same network neighborhood can be interpreted as using similar information sources, such as newsletters or advisory services. Our approach to modeling information networks is similar to the approaches taken in Ozsoylev (2005) and Colla and Mele (2008).

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

2.1 Networks

There are n agents in the economy. The set of agents is $N = \{1, 2, \ldots, n\}$.[7] Agents are connected in a network: The relation, $\mathcal{E} \subset N \times N$, describes whether agent i and j are connected in the network. Specifically, the edge $(i, j) \in \mathcal{E}$, if and only if there is a connection between agent i and j. We use the convention that each agent is connected with herself, that is, $(i, i) \in \mathcal{E}$ for all $i \in N$. We also assume that connections are undirected. Thus, $\mathcal{E}$ is reflexive and symmetric. Formally, the n-agent network is described by the duple $\mathcal{G}^n = (N, \mathcal{E})$. We alternatively represent the network relation, $\mathcal{E}$, by the matrix $\mathbf{E} \in \mathbb{R}^{N\times N}$, with $(\mathbf{E})_{ij} = 1$ if $(i, j) \in \mathcal{E}$ and $(\mathbf{E})_{ij} = 0$ otherwise.[8]

We define the distance function $D(i, j)$ as the number of edges in the shortest path between i and j. We use the conventions that $D(i, i) = 0$, and that $D(i, j) = \infty$ whenever there is no path between node i and j. The set of nodes adjacent to node i 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 $R_i^m \stackrel{\text{def}}{=} \cup_{j=0}^m Q_i^j$. The number of nodes at a distance not further away than m from node i is $W_i^m \stackrel{\text{def}}{=} |R_i^m|$. For $m = 1$, we simply write R_i and W_i. R_i is the set of agent i's neighbors, and this set includes agent i himself. W_i is the *degree* of node i, which we also refer to as *agent i's*

[7]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. Calligraphed letters represent structures, e.g. graphs, and relations. The set of natural numbers is denoted by $\mathbb{N}$, the set of real numbers is denoted by $\mathbb{R}$, the set of positive real numbers is denoted $\mathbb{R}_+$, and the set of strictly positive real numbers is denoted by $\mathbb{R}_{++}$. 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 $\{i \in A : i \notin B\}$.

[8]We use the following vector and matrix notations: The ith element of the vector $\mathbf{v}$ is $(\mathbf{v})_i$, and the n elements v_i, $i = 1, \ldots, n$, form the vector $[v_i]_i$. 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 ith row and jth column of the matrix $\mathbf{A}$, or, if there can be no confusion, we write it as $\mathbf{A}_{ij}$. 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).

connectedness. The degree distribution is the function, $d \in S^n$,[9] defined as $d(i) = \frac{\left|\{j:W_j=i\}\right|}{n}$.

The common neighbors of agents i and j constitute the set $R_{ij} \stackrel{\text{def}}{=} R_i \cap R_j$. The number of such common neighbors is given by $W_{ij} = |R_{ij}|$. We define the symmetric neighborhood matrix $\mathbf{W}$ as being equal to $[W_{ij}]_{ij}$. The element on row i and column j of $\mathbf{W}$ thus represents the number of agents that are common neighbors of i and j.[10] The relation $\mathbf{W} = \mathbf{E}^2$ follows from standard graph theory. Clearly, we have

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

2.2 Agents, assets, and information structure

The economy operates in two periods. Trade takes place at $t = 0$ and asset payoffs realize at $t = 1$. Agents derive utility only from their final wealth at $t = 1$. Agents are price-takers. Also, they are expected utility maximizers and have CARA preferences. For simplicity, we assume that the constant absolute risk aversion coefficient of each agent is 1. Therefore, the expected utility derived by any agent from a risky gamble, $\tilde{\xi}$, is

$$E[U(\tilde{\xi})] = -E[e^{-\tilde{\xi}}].^{11}$$

We note that for agents with the above specifications, the certainty equivalent, CE, of the gamble $\tilde{\xi}$ is

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

There are two assets in the economy: one risk-free and one risky. Prior to trading, agents are not endowed with either asset. The price and payoff of the risk-free asset, which is in elastic supply, are normalized to 1.[12] The risky asset pays off a random liquidating dividend $\tilde{X}$ at $t = 1$, which is normally distributed with mean $\bar{X} \geq 0$ and variance σ^2. There is a random supply of the risky asset during the trading period, i.e., at $t = 0$: in the current n-agent setup, this supply is given by $\tilde{Z}_n = n\tilde{Z}$, where $\tilde{Z}$ is normally distributed with mean $\bar{Z} \geq 0$ and variance Δ^2. There are n distinct *primary* pieces of information, $\{\tilde{y}_k\}_{k=1}^n$, about the risky asset payoff $\tilde{X}$: $\tilde{y}_k$ communicates $\tilde{X}$ with some additive noise $\tilde{\epsilon}_k$. In particular,

[9] Here, $S^n \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n, x(i) \geq 0, \sum_{i=1}^n x(i) = 1\}$ is the unit simplex in $\mathbb{R}^n$. The unit simplex over the natural numbers is S^∞, with the natural interpretation that $S^1 \subset \cdots \subset S^n \subset S^{n+1} \subset \cdots \subset S^\infty$.

[10] This number includes nodes i and j if i and j are linked themselves.

[11] We use the following standard notation: The expectation and variance of a random variable, $\tilde{\xi}$, are $E[\tilde{\xi}]$ and $var(\tilde{\xi})$, respectively. The correlation and covariance between two random variables are $cov(\tilde{\xi}_1, \tilde{\xi}_2)$ and $corr(\tilde{\xi}_1, \tilde{\xi}_2)$, respectively.

[12] The terms *price* and *demand* will be exclusively used for the risky asset price and the risky asset demand, respectively, unless otherwise stated.

$\tilde{y}_k = \tilde{X} + \tilde{\epsilon}_k$, where $\tilde{\epsilon}_k$ is normally distributed with mean 0 and variance s^2. The random variables $\tilde{X}$, $\tilde{Z}$ and $\{\tilde{\epsilon}_k\}_{k=1}^n$ are jointly independent.

Prior to trading at $t = 0$, each agent observes a signal about the risky payoff. Formally, agent i receives the signal

$$\tilde{x}_i = F_i\left(\tilde{y}_1, \dots, \tilde{y}_n \mid \mathcal{G}^n\right),$$

for some function F_i. Hence, each agent's signal conveys and combines the primary information pieces, $\{\tilde{y}_k\}_{k=1}^n$, according to the information network $\mathcal{G}^n$. Here, $\tilde{y}_i$ can be viewed as agent i's information prior to any communication among neighbors, whereas $\tilde{x}_i$ can be interpreted as agent i's information after such communication takes place. In general, we want the topological properties of the network to carry over to agents' signals so that the following properties hold:

(i) Agents with more neighbors receive more precise signals about the risky payoff:

$$W_i > W_j \;\Rightarrow\; var(\tilde{X}|\tilde{x}_i) < var(\tilde{X}|\tilde{x}_j).$$

(ii) If two agents have no common neighbors, then their signals' error terms are uncorrelated:

$$R_i \cap R_j = \emptyset \;\Rightarrow\; cov(\tilde{x}_i, \tilde{x}_j) = var(\tilde{X}).$$

(iii) Two agents, who have the same neighbors,[13] receive the same signal:

$$R_i = R_j \;\Rightarrow\; \tilde{x}_i = \tilde{x}_j.$$

(iv) All else equal, the correlation between agent i's and j's signals is higher if they are connected than if they are not connected. That is, given two networks $\mathcal{G} = (N, \mathcal{E})$ and $\mathcal{G}' = (N, \mathcal{E}')$, which are identical except for that $(i, j) \in \mathcal{E}$ but $(i, j) \notin \mathcal{E}'$, then the correlation between $\tilde{x}_i$ and $\tilde{x}_j$ in network $\mathcal{G}$ is higher than that in network $\mathcal{G}'$.

A signal structure that satisfies the above 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 $\tilde{x}_i = \tilde{X} + \tilde{\eta}_i$, with $\tilde{\eta}_i = \frac{\sum_{k \in R_i} \tilde{\epsilon}_k}{W_i}$. The error terms, $\{\tilde{\eta}\}_i$, are multivariate normally distributed random variables with mean zero and covariance matrix

$$\mathbf{S} \stackrel{\text{def}}{=} [cov(\tilde{\eta}_i, \tilde{\eta}_j)]_{ij} = s^2 \mathbf{D}^{-1} \mathbf{W} \mathbf{D}^{-1}, \tag{6}$$

[13] According to our definition, the set of agent i's neighbors, namely R_i, includes agent i himself. Therefore, if two agents have the same neighbors, then they are also each other's neighbor.

where $\mathbf{D} = diag((\mathbf{W})_{11}, \ldots, (\mathbf{W})_{nn})$.[14] Clearly, all $\tilde{\eta}_i$'s, being linear combinations of $\tilde{\epsilon}_i$'s, are independent of $\tilde{Z}$ and $\tilde{X}$.

Agents have rational expectations concerning asset prices, therefore, they learn from the risky asset price, which aggregates all agents' signals. Agent i's information set, when he trades at $t = 1$, is thus[15]

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

where $\tilde{p}$ stands for the risky asset price. Agent i's risky asset demand schedule depends on his information as well as price, and is represented by $\psi_i(\tilde{x}_i, \tilde{p})$.

The key feature of our model is that the network topology maps to the information structure in the economy.[16] This modeling approach provides a useful framework in which information networks are employed to explore a wide range of information structures in a tractable manner.

2.3 Interpretation of network relations

As we elaborate above, information networks determine who shares information with whom. Arguably, the most natural interpretation of information networks is that they represent information sharing via direct social connections, such as friendships and acquaintanceships. According to this interpretation, $(i, j) \in \mathcal{E}$ represents information sharing between friends i and j. However, our forthcoming analysis is perfectly general and holds for other interpretations of the network relation $\mathcal{E}$, and thereby of the neighborhood matrix $\mathbf{W}$. In particular, network relations can represent information sharing via not only direct but also indirect social connections.

For instance, take the network relation $\mathcal{E} = \{(i, j) : D(i, j) \leq 1\}$ to represent information sharing via direct social connections, say friendships. We can define a new relation $\hat{\mathcal{E}} = \{(i, j) : D(i, j) \leq 2\}$, which represents relationships in which agents share information with not only their friends but also with their friends' friends. The relation $\hat{\mathcal{E}}$ leads to a new neighborhood matrix, $\hat{\mathbf{W}}$, and degrees, $\hat{W}_i = (\hat{\mathbf{W}})_{ii}$. It also allows for a new network metric, namely *centrality*, to play a role in the information structure. As a specific example, consider

[14]For a general vector $\mathbf{d}$, $diag(\mathbf{d})$ is the diagonal matrix with diagonal elements $(diag(\mathbf{d}))_{ii} = (\mathbf{d})_i$.

[15]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.

[16]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.

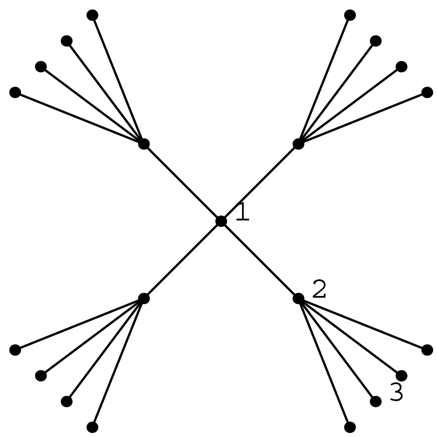

Figure 1: *According to relation $\mathcal{E}$, $W_1 = 5$, $W_2 = 6$, $W_3 = 2$, hence agent 2 has the most precise information. On the other hand, according to relation $\hat{\mathcal{E}}$, $\hat{W}_1 = 21$, $\hat{W}_2 = 9$, $\hat{W}_3 = 6$, and therefore agent 1 has the most precise information due to his centrality.*

the network shown in Figure 1. According to relation $\mathcal{E}$, agent 2 has more precise information about the asset payoff compared to agent 1, since $W_1 = 5$ and $W_2 = 6$. 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 connected, which should work to his advantage. This is captured in the definition of relation $\hat{\mathcal{E}}$, which also takes into account friends' friends. Observe that agent 1's degree is $\hat{W}_1 = 21$ according to relation $\hat{\mathcal{E}}$, whereas agent 2's degree is only 9. According to relation $\hat{\mathcal{E}}$, agent 1 is the one who is most connected, and this follows from his centrality in the network. Therefore, in general, an agent's connectedness, as defined in section 2.1, can be interpreted as that agent's centrality.

Other definitions of centrality also exist, as has been extensively discussed in the network literature. For example, in Das and Sisk (2005), the *centrality score*, which measures the centrality of a node taking into account even more distant indirect connections, is used to apply network methods to the analysis of asset prices. Their interpretation of what constitutes a network is somewhat different from ours, however, since they use nodes to represent stocks and connections to represent overlapping posters in Internet stock message boards.

Our forthcoming analysis is valid for any given network relation $\mathcal{E}$ and neighborhood matrix $\mathbf{W}$ as long as there is a set of nodes, R_i, associated with each node, i, such that $i \in R_i$, $(\mathbf{W})_{ij} \in N$, $(\mathbf{W})_{ij} \leq \min\{(\mathbf{W})_{ii}, (\mathbf{W})_{jj}\}$ and $(\mathbf{W})_{ii} \geq 1$, where $[\mathbf{W}]_{ij} \stackrel{\text{def}}{=} |R_i \cap R_j|$. We use $\mathcal{E}$, as defined in section 2.1, to represent the network relation going forward, keeping in mind that, depending on how connections are defined, this relation can take centrality into account.

2.4 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

- market always clears, i.e., $\tilde{Z}_n = \sum_{i=1}^{n} \psi_i(\tilde{x}_i, \tilde{p})$ for all realizations of $\{\tilde{x}_i\}_i$, $\tilde{X}$, $\tilde{Z}_n$, and
- each agent optimizes expected utility of his final wealth at $t = 1$, conditional on his information, under rational expectations.

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, in contrast to the analysis in Hellwig (1980), the existence of a linear NREE for a finite number of agents is not guaranteed here, because in our setup agents, who are each other's neighbors or who have common neighbors, receive signals with correlated error terms. However, as we show below, under some additional assumptions regarding the information structure, a linear NREE always exists when the number of agents is sufficiently high.

Formally, we study a sequence of markets, $\mathcal{M}^1, \ldots \mathcal{M}^n, \ldots$, with increasing number of agents, n. We use the following notation: For vectors, $\mathbf{y}$, we define the vector norms $\|\mathbf{y}\|_p = (\sum_i (\mathbf{y})_i^p)^{1/p}$ and $\|\mathbf{y}\|_\infty = \max_i |(\mathbf{y})_i|$. Similarly, we define the matrix norms, $\|\mathbf{A}\|_p = \sup_{\{\mathbf{y}:\|\mathbf{y}\|_p=1\}} \|\mathbf{A}\mathbf{y}\|_p$, $p \in [1, \infty]$. Moreover, we say that $f(n) = o(g(n))$ if $\lim_{n\to\infty} f(n)/g(n) = 0$, and 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, we say that $f \sim g$ at x if $\lim_{\epsilon \searrow 0} f(x+\epsilon)/g(x+\epsilon) = C$ for some $C > 0$.

Our main result is:

Theorem 1 *Assume a sequence of n-agent markets, $\mathcal{M}^n$, $n = 1, 2, \ldots$, 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), the neighborhood matrix $\mathbf{W}^n$ of market $\mathcal{M}^n$ satisfies equations (1)-(3), and*

$$\|\mathbf{W}^n\|_\infty = o_p(n), \tag{10}$$

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} (\mathbf{W}^n)_{ii}}{s^2 n} = \beta + o_p(1) > 0. \tag{11}$$

Then, with probability one, the equilibrium price converges to

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

where

$$\pi^* \quad = \quad \gamma^* \beta, \tag{13}$$

$$\gamma^* \quad = \quad \frac{\sigma^2 \Delta^2 + \sigma^2 \beta}{\beta \sigma^2 \Delta^2 + \Delta^2 + \beta^2 \sigma^2}, \tag{14}$$

$$\pi_0^* \quad = \quad \gamma^* \frac{\bar{X} \Delta^2 + \bar{Z} \beta \sigma^2}{\sigma^2 \Delta^2 + \sigma^2 \beta}. \tag{15}$$

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

Theorem 1 will be our workhorse when we analyze asset pricing and welfare implications of large information networks. In this theorem, β appears as a crucial parameter and affects the large-economy equilibrium price. From (11), it follows that β stands for the asymptotic average number of connections which agents possess as n tends to infinity, scaled by the precision of information noise $\{\varepsilon\}_i$, namely $\frac{1}{s^2}$. It is, therefore, natural to think of β as a measure of *network connectedness*. It is intuitively not surprising that β affects the equilibrium price. A high β suggests that agents on average have more precise information about the asset payoff due to their many connections, and this leads them to trade aggressively on their information, which in turn increases the prominence of the asset payoff $\tilde{X}$ relative to the asset supply $\tilde{Z}$ in the equilibrium price. A low β affects the relative prominence between the asset payoff and asset supply in the opposite direction. We note that, depending on how the network relation, $\mathcal{E}$, is defined, network connectedness can encompass different network metrics, such as network centrality, in line with our discussion in section 2.3.

Conditions (10) and (11), stated in Theorem 1, are sufficient for the existence of the linear NREE. Condition (11) ensures that the average number of connections of agents in the network is well defined as the economy grows. On the other hand, condition (10) imposes a restriction on the asymptotic behavior of individual agents' number of connections. Condition (11) implies that the information network of the large market is *sparse*, meaning that the number of connections between agents are of the same order as the number of nodes. Recall from our discussions in the introduction that sparseness is one of the most common empirical features of social networks, so the condition is well motivated from a social perspective.

Condition (10) also possesses a social meaning as well as an economic one. In particular, it ensures that no agent is *informationally superior* in the large market, i.e., that no agent possesses too much information. For example, suppose that condition (10) was dropped so that there were informationally superior agents. Consider the case where some of the agents were connected to everyone else in the network. Then those agents would know the exact value of the risky payoff $\tilde{X}$ in the large market and therefore trade such that price fully revealed payoff. With fully revealing prices, agents could afford to disregard their private information and connections, but if agents did not act on their own information and connections at all, it is unclear why the price should reveal the payoff in the first place. Hellwig (1980) refers to equilibria under such circumstances as *schizophrenic*.

Condition (10) is sufficient to avoid the conceptual difficulty described by Hellwig (1980), however it does not rule out existence of relatively well-informed agents. For instance, condition (10) allows for situations of the following nature. Consider an n-agent economy where some agents are connected to $\sqrt{n}$-many other agents. In the large economy, where n tends to infinity, these agents possess the information of many other agents. However, the total mass of these agents is too small to effectively make the price fully revealing. In fact, the information they possess is negligible compared to the residual uncertainty they face in the sense that these agents, individually, have no effect on the large-economy equilibrium price (12). If they had any effect, the equilibrium price would have reflected the error terms in their information. Since the equilibrium price is not fully revealing, agents use their private information, and condition (10) therefore ensures the internal consistency of the model under significantly weaker conditions than in Hellwig (1980).

Theorem 1 also generalizes the results in Hellwig (1980) by allowing agents to have information with correlated error terms via network connections. To the best of our knowledge, ours is the first NREE model to solve for equilibrium in closed form while allowing for correlation across agents' signal error terms.[17] Hence, earlier NREE models cannot investigate the implications of commonality of information across economic agents. Several extensions of the Kyle (1985) model have been introduced, which allow for dispersedly informed agents, who possess signals carrying correlated error terms (see, e.g., Foster and Viswanathan (1996) and Back, Cao, and Willard (2000)). However, in these market microstructure models, agents submit market orders and thereby do not learn from contemporaneous prices, whereas in the NREE models, agents do learn from contemporaneous prices. This makes the introduction of information commonality across agents in NREE models challenging from a technical standpoint. Theorem 1 shows that, even with correlated signal error terms, a large-economy NREE exists and can be solved for, provided that agents' signals do not become too correlated due to network connections. This is ensured by conditions (10,11).

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:

[17] Ozsoylev (2005) allows for correlation across agents' signal error terms in a finite-agent NREE model. However, a closed-form solution for equilibrium cannot be obtained in Ozsoylev (2005), which significantly restricts the equilibrium analysis.

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.

Assumption 1 is thus a purely technical assumption, needed since we are technically studying *sequences* of economies. We call d the degree distribution of the large network. To avoid confusion, we will specifically refer to highlighted assumptions, such as Assumption 1, in the statements of our results whenever they are needed.

In our subsequent analysis of individual agents, we will focus on agents for which the asymptotic degree exists, i.e., for which $\lim_{n\to\infty} \mathbf{W}^n_{ii}$ exists and is finite (with probability one). Similarly, when we compare pairs of agents in section 3.3, an additional underlying assumption is that $\lim_{n\to\infty} \mathbf{W}^n_{ij}$ exists and is finite. We could, alternatively, have focused on networks for which $\lim_{n\to\infty} \mathbf{W}^n_{ii}$ 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 $\lim_{n\to\infty} \mathbf{W}^n_{i_n i_n}$ exists and is finite, but we avoid this approach since it leads to a cumbersome notation.

2.5 Socially plausible networks

Given the enormous number of degrees of freedom in 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 $d(i) \sim i^{-\alpha}$, $\alpha \le 2$, 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* α, or simply to be power-law distributed.[18] Power-law distributed networks with low α's are said to be heavy-tailed.

[18]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.) of the degree distribution, whereas our definition is based on the p.d.f. The correspondence between $\hat{\alpha}$ and α is then $\hat{\alpha} = \alpha - 1$.

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. 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 correspondences and the World Wide Web do not fit into the random network framework.[19] 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.[20] 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.[21] Proposition 1 implies that a large-economy equilibrium characterized by Theorem 1 exists with power-law distributed information networks as long as their tail exponent is larger than 2.

We next analyze the relationship between tail exponent α and network connectedness β for a specific network type. In order to keep the number of parameters down, we assume:

Assumption 2 $s^2 = 1$.

Assumption 2 is purely for normalization, leading to simpler formulas; it does not restrict the model in any way. We will employ this assumption throughout most of the paper.

A convenient class of networks are the so-called Zipf-Mandelbrot distributed networks, with degree distribution, $d^n \sim ZM(\alpha, n)$. Here, the Zipf-Mandelbrot distribution, $d^n \sim ZM(\alpha, n)$, is a particular form of power-law distribution. For 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 ζ 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:

[19] 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.

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

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

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 $\beta(\alpha) = \zeta(\alpha-1)/\zeta(\alpha)$, where β is as defined in (11). If the tail-exponent, $\alpha \leq 2$, then $\beta = \infty$.*[22]

This result immediately leads to :

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

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

Propositions 1 and 2 make it quite clear when to expect the existence of the large-economy equilibrium characterized by Theorem 1. 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 a linear NREE may not exist in the asymptotic economy. 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 (i.e., non-price-taking) may be needed instead. Similar breakpoints occur in economic models with power-laws at $\alpha = 2$ in other contexts, see e.g., Ibragimov, Jaffee, and Walden (2009).

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 α 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)).

3 Asset pricing and welfare implications of networks

We examine asset pricing and welfare implications of information networks in the large-economy equilibrium characterized by Theorem 1. We identify novel relationships between asset prices and network connectedness. We also study how network connectedness affects agent welfare.

3.1 Price volatility and market efficiency

The unconditional variance of price is often used as a measure of price volatility in the rational expectations equilibrium literature since it lends itself to empirical testing – see, e.g. Vives (1995) and Wang (1993). Following this convention, we use the unconditional variance of price in our analysis of price volatility. From Theorem 1, we see that the price volatility is

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

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

Following the terminology in Ozsoylev (2005), the price volatility can be decomposed into an *information driven* volatility component, $(\pi^*)^2\sigma^2$, and a *liquidity (supply) driven* volatility component, $(\gamma^*)^2\Delta^2$. We would expect that when the network's connectedness becomes large, the price converges to the payoff since the aggregate information in the economy fully reveals the 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 $\beta \to \infty$. 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, β. 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:*

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

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

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

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

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

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

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. The intuition behind this result is as follows. Suppose, to begin with, that agents have no connections. As networks connectedness increases, agents learn more from each other, and equilibrium price becomes more informative. Therefore agents rely more on prices as an information source while forming their demands, which also makes their demands more dependent on liquidity and, in turn, renders a larger liquidity driven volatility component. On the other hand, above a certain level of network connectedness, agents become so informed due to their network connections that they hardly learn additional information from the price. As a result, agents rely less on price as an information source, which makes their demands less dependent on liquidity, and hence 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 respect to connectedness depends on which of the two components, information driven or liquidity driven, is the dominant one.

Proposition 3 shows that the underlying network topology is intimately connected to volatility in a nontrivial way, and that it may therefore be important in understanding real world volatility dynamics in capital markets. The result complements the analysis in Ozsoylev (2005), who focuses on economies in which the liquidity variance, Δ^2, is high, and who thereby provides a partial characterization of price volatility.

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.

Proposition 4 *In the large-economy equilibrium characterized by Theorem 1, market efficiency increases as the network's connectedness increases. That is,*

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

3.2 Trading profits

We now turn our attention to individual agents' trading profits. 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 assume:

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

Similar to Assumption 1, Assumption 3 is thus a purely technical assumption, needed since we are technically studying sequences of economies.

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

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

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

Proposition 5 *Consider the large-economy equilibrium characterized by Theorem 1. Assume that Assumption 3 holds for agent i. Then, agent i's ex-ante trading profit,* Π_i*, is linear in the agent's connectedness,* W_i*. In particular,*

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

Here, Π^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* Π_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 5 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 5, 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, Π^F is the information-free average trading profit and Π_i^I is the information-related average trading profit.

For simplicity, we make the following assumption:

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

Assumption 4 is effectively a normalization of the expected payoff and expected liquidity (supply), which is common in the literature, see, e.g., Brunnermeier (2005) and Spiegel (1998). It leads to simplified formulas, without restricting the intuition. We then have:

Proposition 6 *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, β, is held constant, then agent i's ex-ante trading profit increases as her own connectedness increases. That is,

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

(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,

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

The intuitions behind the proposition are straightforward. The higher the number of connections an agent has in an information network, the higher her profits, 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 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 7 *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,

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

(b) Π^F *is positive, decreasing in* β*, and approaches* 0 *as* β *tends to* ∞*.*

(c) Π^I *is positive, non-monotonic in* β*, and approaches* 0 *as* β *tends to* ∞*.*

(d) As β *tends to* ∞*,* Π *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 ∞. 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, Π^F, of average trading profit is decreasing in β. As we have mentioned before, the information-free component is the compensation agents need to take on risk. When β, 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., for Π^I being non-monotonic in β, has already been discussed following Proposition 6, and (d) is also natural, since all informational rents disappear in the limit, when connectedness grows.

3.3 Portfolio holdings and trading volume

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 8 *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 8 finds a positive relationship between informational proximity and correlated trading. Geographical proximity is expected to encourage communication, therefore, the empirical studies cited above lend support to this result.

The impact of information networks on trading volume is also straightforward to analyze. The trading volume of individual i, with connectedness W_i, is defined as his expected unsigned asset demand, $\psi_i^{unsigned}(W_i) = E[|\psi_i|]$, or equivalently as $\psi_i^{unsigned} = \sqrt{\frac{\pi}{2}E[\psi_i^2]}$. Here, π is the mathematical constant: $\pi = 3.1415....$ The aggregate trading volume is defined as $\psi^{market} = \lim_{n\to\infty}\sqrt{\frac{1}{n}\frac{\pi}{2}E[\sum_i \psi_i^2]}$.[23] It turns out that to characterize the aggregate trading volume, in addition to network connectedness, we also need to take into consideration the *variance* of network connectedness, defined as $\sigma_\beta^2 = \lim_{n\to\infty}\frac{1}{n}\sum_i\left(\frac{W_i^n}{s^2}-\beta\right)^2$. The variable σ_β is thus a measure of the spread of individual connectedness in the network. If some agents are much more connected than others, σ_β will be large, whereas if all agents have very similar connectedness σ_β will be small. We have

Proposition 9 *Consider the large-economy equilibrium characterized by Theorem 1, satisfying Assumption 4.*

(a) *The individual trading volume, $\psi_i^{unsigned}(W_i)$, is an increasing, concave function of connectedness with asymptote,*

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

for large W_i.

(b) *The aggregate trading volume, ψ^{market}, is increasing in network connectedness (β), in markets with low variance of network connectedness (σ_β^2), and is decreasing in network connectedness in markets with high variance of network connectedness.*

(c) *The aggregate trading volume, ψ^{market}, is increasing in the variance of network connectedness, σ_β^2.*

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 5 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. The aggregate trading volume, on the other hand, can be either

[23] Alternatively, we could have defined aggregate trading volume as $\lim_{n\to\infty}\frac{1}{n}E[\sum_i|\psi_i|]$. Such a definition is qualitatively similar to ours — although, contrary to an individual's trading volume, not identical — but complicates the analysis considerably.

increasing or decreasing in network connectedness. However, aggregate trading volume is always increasing in the *variance* of network connectedness. This is in line with the idea that information asymmetries across the trading population drive trading volume.

3.4 Welfare

In this section, we analyze the welfare implications of information networks. We base the analysis on the certainty equivalent of utility that an agent derives from trading in the market.[24] The *ex ante* certainty equivalent for an agent is $CE(W)$, where W is the agent's connectedness. This is the certainty equivalent, before the agent receives any information about the risky asset payoff. We distinguish this from the *ex interim* certainty equivalent, which is the certainty equivalent after an agent has received his information and traded but before the risky payoff is realized.

A closed form expression for the average *ex ante* certainty equivalent is given by:

Proposition 10 *Consider the large-economy equilibrium characterized by Theorem 1. Assume that assumptions 1, 2 and 4 hold.*

(a) For agent i, satisfying assumption 3, the ex ante certainty equivalent is

$$CE(W_i) = \frac{1}{2}\log\left(\frac{(\Delta^2+(\beta+\Delta^2)^2\sigma^2)\,(\beta^2 s^2\sigma^2+\Delta^2 s^2+W_i\Delta^2\sigma^2)}{s^2(\beta^2\sigma^2+\Delta^2+\Delta^2\beta\sigma^2)^2}\right). \quad (18)$$

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

$$\overline{CE} = \sum_j CE(j)d(j). \quad (19)$$

We will use Proposition 10 to analyze which networks are welfare optimal in the sense that they maximize the average certainty equivalent, taking into consideration possible costs involved in forming network connections. This is the first-best optimal solution that would occur in a centralized economy, in which a central planner, who has the power to redistribute wealth, chooses the network on behalf of the agents.

We first analyze which networks optimize the average certainty equivalent for the special case when there are no costs associated with link formation. As we shall see, the analysis is then easily extended to more general cost functions. From (19) it follows that maximizing the average certainty equivalent, $\overline{CE}$, over networks is equivalent to maximizing it over network

[24] A natural interpretation of the random asset supply, $\tilde{Z}$, is that it is due to noise trading. Therefore, it can be argued that the welfare of noise traders is not taken into account in our analysis. We are particularly interested in the welfare of rational agents, because we would like to understand how rational agents, who can coordinate their actions and assign a central planner to choose the information network for them, would behave.

degree distributions. We solve the maximization problem in two steps: We first maximize the average certainty equivalent over network degree distributions with the same network connectedness. To that end, we define $\overline{CE}^*(\beta)$ as the maximum value of $\sum_j CE(j)d(j)$ attained over all networks which have connectedness β (as defined in (11)), and satisfy Assumption 1 and the conditions in Theorem 1. We then maximize $\overline{CE}^*(\beta)$ over all feasible β's (that is, over $\beta \geq \frac{1}{s^2}$), to get the globally optimal solution.

Define the *support* of a degree distribution, d, as $supp[d] = \{j : d(j) > 0\}$. We have:

Proposition 11 *Assume that assumptions 2 and 4 hold.*

(a) For $\beta \in \mathbb{N}$, $\overline{CE}^*(\beta)$ is attained by a network with degree distribution d, if and only if $supp[d] = \{\beta\}$.

(b) For $\beta \in \mathbb{R}_+ \backslash \mathbb{N}$, $\overline{CE}^*(\beta)$ is attained by a network with degree distribution d, if and only if $supp[d] = \{\lfloor\beta\rfloor, \lceil\beta\rceil\}$, $d(\lceil\beta\rceil) = \beta - \lfloor\beta\rfloor$, and $d(\lfloor\beta\rfloor) = 1 - \beta + \lfloor\beta\rfloor$.

(c) $\overline{CE}^*(\beta)$ is either everywhere decreasing in β, or initially increasing in β and eventually decreasing, and attains a unique maximum.

We obtain an unambiguous ranking of information networks through Proposition 11. Uniform networks achieve a higher average certainty equivalent compared to non-uniform networks with the same degree of connectedness. Moreover, among uniform networks, the average certainty equivalent is either decreasing in β or hump-shaped, with a unique maximum. Therefore, the maximum average certainty equivalent is either achieved by a network in which no agent shares information, or by one where all agents basically have the same finite number of connections.

The results in Proposition 11 are quite intuitive. The intuition for part (c) is similar to that for our result on trading profits, discussed in Section 3.2. If agents, on average, receive too much information, the informational rents are competed away. If agents receive too little information, on the other hand, the uncertainty about the final payoff is high, which adversely affects risk averse agents' expected utilities. The optimum typically lies somewhere in-between. Parts (a) and (b) of Proposition 11 follow from the certainty equivalent of an agent being a concave function of his connectedness. Given the total information rents in the economy – which depend on average connectedness, β – any asymmetry in how these rents are divided between agents will lead to a lower average certainty equivalent, since $CE(W_i)$ is concave.

In practice, we would expect link formation to be costly. For example, expanding one's 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 if links are interpreted as those linked using or accessing the same information source, 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, namely the subscribers. We therefore extend our earlier analysis to an environment where link formation is costly.

We let $f(W)$ denote an agent's cost of having W connections, and assume that $f(1) = 0$ — an agent is always connected to himself, which carries no cost. It is natural to assume that more links will be more costly, so we have $f_W > 0$. We also assume that the cost is (weakly) convex in the number of links, i.e., that $f_{WW} \geq 0$.[25] We argue that the marginal cost for an additional link should be increasing, at least eventually, since agents have finite resources and capabilities for link formation (e.g., limited time, social barriers).

In the case with non-zero costs for link formation, the social planner's generalized welfare optimization problem is to maximize

$$\sum_j \Big(CE(j) - f(j)\Big)d(j) \tag{20}$$

over all networks that satisfy assumption 1 and the conditions in Theorem 1. We have:

Proposition 12 *Assume that assumptions 2 and 4 hold. Then, there is a network that maximizes (20). Moreover, any network that maximizes (20) has a degree of connectedness* $\beta < \infty$ *and a degree distribution with supp*$[d] \subset \{\lfloor\beta\rfloor, \lceil\beta\rceil\}$.

Proposition 12 thus shows that Proposition 11-(c) can be extended to general weakly convex cost functions, f. The intuition for the result is identical to that of Proposition 11, once it is noted that $CE - f$ is concave when CE is concave and f is convex.

4 Potential extensions and alternative assumptions

Our model of information networks is, of course, very stylized. In this section we discuss potential extensions and variations of the model, and further justify the assumptions we have made. For some of the extensions, the resulting effects are quite clear, whereas other parts of our discussion is speculative, opening up for future research.

In the model, we assume that network connections are bi-directional, that is, if agent i learns from j, then agent j also learns from i. In practice, there are many situations where one agent learns from the other but not vice versa. Technically, an information network with unidirectional connections will have a neighborhood matrix that is asymmetric, say $\mathbf{W}_{as}$. In principle, it is straightforward to generalize the current analysis to networks with unidirectional connections. Given a sequence of symmetric neighborhood matrices, $\mathbf{W}^n$,

[25] For simplicity we require that $f(W)$ is a twice continuously differentiable function in $\mathbb{R}_+$, even though W belongs to the set of natural numbers.

that satisfies (1)-(3) and the conditions of Theorem 1, it follows immediately that the same conditions also hold for the neighborhood matrices, $\mathbf{W}^n_{as}$, that are constructed by deleting some links of $\mathbf{W}^n$ unilaterally. This implies that the existence of a large-economy equilibrium with a unidirectional network can be easily established following the same proof method applied for Theorem 1. Much of the subsequent analysis is then straightforward.

Another assumption we make is that every agent is a price taker, an assumption that is rationally motivated in a "large" economy in which each agent is "small." In a large economy, agents have no incentive to not share information with their neighbors, nor do they have an incentive to lie about their private information, since they know that their neighbors have no price impact. In a "smaller" economy, where individual agents do impact prices, however, an incentive to hoard information exists. A full analysis of what happens in such an economy is outside the scope of this paper, but we can make some qualitative arguments.

Consider a small economy in which agents are non-price-takers and can credibly communicate information (i.e., they can commit not to lie). Two agents may agree to bilaterally share information as long as the marginal benefit of receiving information outweighs the marginal cost of higher price competition that follows from information sharing. Since price impact increases the cost of price competition, the marginal cost of link formation in the small economy is higher than in the one studied in Section 3.4. We would therefore expect the degree of connectedness to be lower in the small economy than in the large economy.

If agents can not credibly communicate information, they have an incentive to lie and the previous argument breaks down. In a dynamic setting, it may still be possible to have information sharing. For example, in an infinitely repeated game version of our model, an agent may be able to punish a neighbor who provides incorrect information. In the simplest case, where information is *ex post* verifiable after each time period, a grim trigger strategy where an observed lie is punished by perpetually cutting the link, agents may find it optimal to truthfully share information. Even if information is not *ex post* verifiable, a punishment strategy based on statistical inference would also severely limit the opportunities to deviations from truth-telling in the long run.

In summary, the analysis becomes much more complex when agents have price impact and there are interesting open questions in this setting, although the basic intuition for the effects that will come into play is quite clear.

One can also consider an economic environment where agents sell (rather than share) information to those with whom they are connected. Because of the externalities associated with information diffusion, and the global character of agents' optimization problems, the analysis of such an environment — although potentially very interesting — would be extremely complex. At the margin, however, the costs and benefits of selling information is clear. The benefit to the seller is the fee that the buyer is willing to pay for information, which is bounded above by the benefits the buyer receives from it. The cost to the seller of each sale has two parts: first, the cost of increased price competition from the buyer generated by the information he receives, and, second, the lower fee the seller's other connections are willing to pay because the information gets shared more broadly.

It may also be possible to extend our model to multiple risky assets, along the lines of Admati (1985), but with information networks. Here we can only speculate about the interesting effects such an extension might lead to. Admati (1985) allows for general correlation patterns across asset payoffs, but assumes that signal error terms are independent across agents. As we elaborated on in Section 2.2, information networks introduce correlation across agents' information error terms in a tractable fashion. Therefore, they may lead to interesting correlation patterns across equilibrium prices of assets. For instance, two assets with uncorrelated payoffs may have highly correlated prices due to information networks. Such an outcome would be consistent with the observed "excess co-movement" phenomenon. Moreover, the pattern of price co-movement would be affected by the network topology, since the signal error terms of highly connected agents would affect prices, and hence their correlations, more than those of agents with few connections.

Another possible extension is to make the current model dynamic so that there are multiple trading opportunities and gradual information diffusion over time. For example, consider a T-period version of the model, where information diffuses according to the rule that after t periods each agent gets to see the information of agents at a distance not further away than t in the network. Such an extension, which may be carried out along the lines of Vives (1995), would potentially explain some of the stylized facts about asset price dynamics discussed in the introduction. Specifically, gradual diffusion of an information shock may lead to large price movements over time, unrelated to public news. Further, interesting time variations in trading volume and price volatility would most certainly arise, as an information shock propagated through a network.

Finally, we note that we have made some strong assumptions about the efficiency with which agents share information, namely that each agent shares his private information completely only with his immediate neighbors but not at all with more distant agents. In practice, such an assumption may be too strong. First, some private information may be difficult to share (even if an agent wants to share it), and second, once information is shared with neighbors, it may be inferred by others who are not in that neighborhood. One could, for example, assume that agents receive noisy signals from agents in their neighborhood and that signals become more noisy with distance. Technically, this extension would lead to a redefinition of the neighborhood matrix, $\mathbf{W}$, that would take into account increasing noise with distance. The, already high, number of degrees of freedom in the model would increase significantly, but it is clear that, qualitatively, the model would not change. In fact, our freedom in choosing the network relation already gives us some freedom. With the network relation $\hat{\mathcal{E}}$, as defined in section 2.3, information spreads perfectly up to (but not beyond) a distance of two. Other variations are also possible within our framework. For example, one could assume that an agent gets perfect signals from one half of the agents at a distance of two, one quarter of the agents at a distance of three, and so on. This would be one possible way of modeling an economy in which signals travel long distances, albeit imperfectly.

5 Empirical implications

Our model suggests that the structure of information networks is important, both at the individual level, where individuals' positions in the network affect their behavior and profits, and in the aggregate, where aggregate network properties affect asset pricing features.

Empirically, it has been extensively verified that social networks are important in explaining individuals' economic behavior, e.g., in the studies by Myers and Shultz (1951) and Reiss and Shultz (1970) (labor markets), Uzzi (1996) (industrial relationships), Fafchamps and Lund (2003) (risk sharing), and in the more recent finance-focused studies by Hong, Kubik, and Stein (2004) and Ivković and Weisbenner (2007). In what follows, we outline how the predictions of our model can be tested.

Testing the model's predictions requires a way to identify information networks in financial markets. We propose two distinct approaches for the identification of information networks. The first approach makes use of information about individual households to build a proxy for network connections. The second approach identifies networks indirectly, based on the similarities of agents' trades or portfolio holdings. We discuss each of these approaches below in turn.

The first proposed approach relies on extensive datasets which provide detailed information at the individual household level. One potential dataset that can be employed is the Swedish dataset used by Calvet, Campbell, and Sodini (2007). The dataset comes from the Swedish government database for tax records and it covers the entire population over an extended time period. The dataset contains household information such as residential address, education level, employer, and demographic data, which can be used to create a proxy for social network connections between households. The dataset also contains household information on holdings of financial securities and bank accounts which are useful for testing the model's predictions about individual investors' trading behavior.[26]

The second proposed approach identifies information networks from agents' portfolios, using the property that the more similar the portfolio strategies of two individuals, the closer these two individuals are in the information network (i.e., the higher the number of neighbors they have in common; see Proposition 8). For example, agents who tend to trade in the same stock at similar points in time can be inferred to be close. Using datasets which contain

[26]Other similar datasets are the LINDA dataset used in Massa and Simonov (2006), which contains detailed personal and financial information for about 3% of the Swedish population, the Finnish Central Securities Depositary dataset used in Grinblatt and Keloharju (2000) and Grinblatt and Keloharju (2001), which is a comprehensive panel on Finnish stockholdings, and the dataset used in Ivković and Weisbenner (2007), which contains common-stock investments of 35,673 U.S. households made through a large discount brokerage in the period from 1991 to 1996.

complete trade-by-trade information at the individual investor level over an extensive time period, the network structure could therefore be inferred. Such datasets were, e.g., studied in Barber, Lee, Liu, and Odean (2009) and in Aragon, Bildik, and Yavuz (2007). A similar approach would be to use the correlation between portfolio holdings to provide a network distance proxy, in line with what is done in the recent empirical study by Pareek (2009), who uses correlations between mutual fund managers' portfolio holdings to identify information networks.

Given an empirical information network, identified by one of the methods proposed above, we can estimate individual and aggregate connectedness measures, namely W_i and β. These measures can be used to test whether a power-law distribution, verified in numerous empirical studies of networks in different social contexts, provides a valid approximation for the degree distribution of information networks among traders. The estimated individual connectedness measure W_i can then be used to directly test the model's predictions on individual investors' trading behavior. The empirical predictions listed below follow from the results derived in Section 3.

Prediction 1 *(a) Investors with higher connectedness, W_i, earn higher profits.*

(b) Investors with higher connectedness, W_i, trade more aggressively.

(c) Investors who are closer in the information network have more correlated trades.

(d) The distribution of investors' trading profits is an affine function of the degree distribution of the information network.

(e) The trading volume distribution of investors is an increasing concave function of the degree distribution of the information network.

Several empirical studies lend indirect support to predictions listed above. Hau (2001) investigates the implications of informational asymmetries across trader population for trading profits: the paper proxies informational advantage of traders by their geographical proximity to corporate headquarters of equities they trade in and then shows that informationally advantaged traders enjoy higher proprietary trading profits. Hau's finding is consistent with Prediction 1-(a). Dorn, Huberman, and Sengmueller (2008) show that correlated trading is greater in heavily traded stocks. This finding is in line with (b) and (c) of Prediction 1: stocks with dense and *tight-knit* information networks should exhibit both higher correlated trading and higher trading volume. Empirical validation of Prediction 1 would shed new light on findings of Hau (2001) and Dorn, Huberman, and Sengmueller (2008).

Also, the estimated network connectedness measure β can be used to test the asset pricing implications of the model, following the results derived in Section 3.

Prediction 2 *(a) Price volatility is high in markets with an intermediate level of network connectedness, β. It is lower in markets with high or low levels of network connectedness.*

(b) Trading profits are high in markets with an intermediate level of network connectedness, β. They are lower in markets with high or low levels of network connectedness.

(c) Aggregate trading volume is high in markets with high variance of network connectedness.

Empirical validation of Prediction 2 would indicate that information networks provide an important determinant of the aggregate behavior of financial markets. The prediction could be tested by comparing aggregate implications across markets. Different markets may interpreted as different asset classes (e.g., stocks versus commodities), different exchanges (e.g., NYSE versus NASDAQ), different stock types (e.g., value versus glamor stocks), or even different individual stocks. Comparisons may also be made across geographical regions (within a country or across countries). Gomez, Priestley, and Zapatero (2009) find that U.S. firms in regions with low population density have higher (risk adjusted) returns than firms in regions with high population density. They interpret this as an effect of relative wealth concerns of investors. However, this finding is also consistent with our model: it is reasonable to expect that population density is related to information network connectedness since densely populated regions tend to exhibit higher connectedness in social networks. Interestingly, the effect found in Gomez, Priestley, and Zapatero (2009) is non-monotone: Excess returns are lowest in the highest population density regions (New England and Middle Atlantic; see their Table 2 and Table 3 – panel A), but are not highest in the most sparsely populated regions (Mountain and West North Central). Instead, the highest excess returns are found in regions with intermediate population density (West South Central, Pacific and East South Central). This is in line with our Prediction 2-(b).

6 Concluding remarks

The properties of information networks have profound impact on asset prices. We have introduced a simple, parsimonious rational expectations equilibrium model with large information networks, in which the relationship between network properties and asset pricing can be conveniently analyzed. Our model suggests that various network metrics, such as connectedness and centrality, come into play in the analysis of information flow in financial markets. On the aggregate level, these network metrics affect asset prices, volatilities, trading volume and welfare in non-trivial ways. Our model therefore takes a first step towards an information network based explanation of observed price behavior, and of time varying volatility and trading volume in financial markets.

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

$$\begin{aligned} E[\tilde{X}|\mathcal{I}_i] &= \alpha_{0i} + \alpha_{1i}\tilde{x}_i + \alpha_{2i}\tilde{p}, && (21)\\ var(\tilde{X}|\mathcal{I}_i) &= \beta_i, && (22)\end{aligned}$$

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

$$\psi_i(\tilde{x}_i, p) = \frac{1}{\beta_i}\left(\alpha_{0i} + \alpha_{1i}\tilde{x}_i + (\alpha_{2i} - 1)\tilde{p}\right). \quad (23)$$

The market clearing condition now gives.

$$\begin{aligned} \pi_0 &= \gamma\sum_{i=1}^n \frac{\alpha_{0i}}{\beta_i}, && (24)\\ \pi_i &= \gamma\frac{\alpha_{1i}}{\beta_i}, && (25)\end{aligned}$$

where

$$\gamma = \left(\sum_{i=1}^n \frac{1-\alpha_{2i}}{\beta_i}\right)^{-1}. \quad (26)$$

When we wish to stress the dependence on n, we write π_0^n, π_i^n and γ^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

$$\begin{aligned} \alpha_{0i} &= \frac{\bar{X}}{b_i}\left(s_i^2(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi} + \gamma^2 n^2\Delta^2) - (\boldsymbol{\pi}^T\mathbf{s}_i)^2\right) - \alpha_{2i}(\pi_0 - \gamma n\bar{Z}), && (27)\\ \alpha_{1i} &= \frac{\sigma^2}{b_i}\left(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi} + \gamma^2 n^2\Delta^2 - (\mathbf{1}^T\boldsymbol{\pi})(\boldsymbol{\pi}^T\mathbf{s}_i)\right), && (28)\\ \alpha_{2i} &= \frac{\sigma^2}{b_i}\left((\mathbf{1}^T\boldsymbol{\pi})s_i^2 - (\boldsymbol{\pi}^T\mathbf{s}_i)\right), && (29)\\ \beta_i &= \frac{\sigma^2}{b_i}\left(s_i^2(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi} + \gamma^2 n^2\Delta^2) - (\boldsymbol{\pi}^T\mathbf{s}_i)^2\right), && (30)\end{aligned}$$

and where we have defined

$$b_i = (\sigma^2 + s_i^2)\left(\boldsymbol{\pi}^T\mathbf{S}\boldsymbol{\pi} + n^2\Delta^2\gamma^2 + (\mathbf{1}^T\boldsymbol{\pi})^2\sigma^2\right) - \left((\mathbf{1}^T\boldsymbol{\pi})\sigma^2 + (\boldsymbol{\pi}^T\mathbf{s}_i)\right)^2. \quad (31)$$

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 (27-31) — satisfy equations (25) and (26), this generates a NREE, where π_0 can be defined via (24).

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

$$\pi_i = \gamma \frac{\boldsymbol{\pi}^T \mathbf{S} \boldsymbol{\pi} + \gamma^2 n^2 \Delta^2 - (\mathbf{1}^T \boldsymbol{\pi})(\boldsymbol{\pi}^T \mathbf{s}_i)}{s_i^2(\boldsymbol{\pi}^T \mathbf{S} \boldsymbol{\pi} + \gamma^2 n^2 \Delta^2) - (\boldsymbol{\pi}^T \mathbf{s}_i)^2}, \tag{32}$$

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 γ:

$$(\mathbf{q})_i = \frac{1}{s_i^2} \times \frac{\mathbf{q}^T \mathbf{S} \mathbf{q} + n^2 \Delta^2 - (\mathbf{1}^T \mathbf{q})(\mathbf{q}^T \mathbf{s}_i)}{\mathbf{q}^T \mathbf{S} \mathbf{q} + n^2 \Delta^2 - (\mathbf{q}^T \mathbf{s}_i)^2 / s_i^2}. \tag{33}$$

Given $\mathbf{q}$, we get

$$\frac{1}{\gamma} = \sum_{i=1}^{n} \frac{\sigma^2 + s_i^2}{\sigma^2 s_i^2} + \sum_{i=1}^{n} \frac{(\mathbf{1}^T \mathbf{q} - \mathbf{s}_i^T \mathbf{q})^2 - \frac{1}{\gamma}(\mathbf{1}^T \mathbf{q} - \frac{\mathbf{s}_i^T \mathbf{q}}{s_i^2})}{\mathbf{q}^T \mathbf{S} \mathbf{q} + n^2 \Delta^2 - \frac{(\mathbf{s}_i^T \mathbf{q})^2}{s_i^2}}, \tag{34}$$

which leads to

$$\gamma = \frac{1 + \sum_{i=1}^{n} \frac{(\mathbf{1}^T \mathbf{q} - \frac{\mathbf{s}_i^T \mathbf{q}}{s_i^2})}{\mathbf{q}^T \mathbf{S} \mathbf{q} + n^2 \Delta^2 - \frac{(\mathbf{s}_i^T \mathbf{q})^2}{s_i^2}}}{\sum_{i=1}^{n} \frac{\sigma^2 + s_i^2}{\sigma^2 s_i^2} + \sum_{i=1}^{n} \frac{(\mathbf{1}^T \mathbf{q} - \mathbf{s}_i^T \mathbf{q})^2}{\mathbf{q}^T \mathbf{S} \mathbf{q} + n^2 \Delta^2 - \frac{(\mathbf{s}_i^T \mathbf{q})^2}{s_i^2}}}, \tag{35}$$

which is bounded, since $\mathbf{S}$ is strictly positive definite. From (24) and the definition of $\mathbf{q}$, we also have

$$\frac{\pi_0}{\gamma} = \frac{\bar{X} n}{\sigma^2} - \left(\frac{\pi_0}{\gamma} - n \bar{Z} \right) \gamma \times \sum_i \frac{(\mathbf{1}^T \boldsymbol{\pi}) s_i^2 - (\boldsymbol{\pi}^T \mathbf{s}_i)}{s_i^2(\boldsymbol{\pi}^T \mathbf{S} \boldsymbol{\pi} + \gamma^2 n^2 \Delta^2) - (\boldsymbol{\pi}^T \mathbf{s}_i)^2} \tag{36}$$

leading to

$$\pi_0 = \gamma n \left(\frac{\frac{\bar{X}}{\sigma^2} + \bar{Z} A}{1 + A} \right), \tag{37}$$

where

$$\begin{aligned} A &= \gamma \sum_i \frac{(\mathbf{1}^T \boldsymbol{\pi}) s_i^2 - (\boldsymbol{\pi}^T \mathbf{s}_i)}{s_i^2(\boldsymbol{\pi}^T \mathbf{S} \boldsymbol{\pi} + \gamma^2 n^2 \Delta^2) - (\boldsymbol{\pi}^T \mathbf{s}_i)^2} \\ &= \sum_{i=1}^{n} \frac{(\mathbf{1}^T \mathbf{q}) s_i^2 - (\mathbf{q}^T \mathbf{s}_i)}{s_i^2(\mathbf{q}^T \mathbf{S} \mathbf{q} + n^2 \Delta^2) - (\mathbf{q}^T \mathbf{s}_i)^2}. \end{aligned} \tag{38}$$

Thus, if the system of equations defined in (33) has a solution, it will generate a NREE. To show that a solution indeed exists for large enough n, we define

$$\mathbf{y} \stackrel{\text{def}}{=} s^2 \mathbf{D}^{-1} \mathbf{q}, \tag{39}$$

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 (33) is equivalent to $\mathbf{y}$ satisfying

$$(\mathbf{y})_i = \frac{\mathbf{y}^T\mathbf{W}^n\mathbf{y} + n^2\Delta^2 s^2 - (\mathbf{d}^T\mathbf{y})(\mathbf{d})_i^{-1}(\mathbf{W}^n\mathbf{y})_i}{\mathbf{y}^T\mathbf{W}^n\mathbf{y} + n^2\Delta^2 s^2 - (\mathbf{W}^n\mathbf{y})_i^2}. \tag{40}$$

We define the mapping $F_n : \mathbb{R}^n \to \mathbb{R}^n$ by the r.h.s. of (40), so a NREE can be derived from a solution to $\mathbf{y} = F_n(\mathbf{y})$. Now, F_n can be rewritten as:

$$(F(\mathbf{y}))_i = 1 + \frac{(\mathbf{W}^n\mathbf{y})_i^2/n^2 - (\mathbf{d}^T\mathbf{y})(\mathbf{d})_i^{-1}(\mathbf{W}^n\mathbf{y})_i/n^2}{(\mathbf{y}^T\mathbf{W}^n\mathbf{y})/n^2 + \Delta^2 s^2 - (\mathbf{W}^n\mathbf{y})_i^2/n^2}. \tag{41}$$

Clearly, F_n is a continuous mapping, as long as the denominator in (41) 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 (41). For any $\epsilon > 0$, for n large enough,

$$\mathbf{y} \in \mathbb{R}^n,\ \|\mathbf{y}\|_\infty \leq 2 \qquad \Rightarrow \qquad |(F_n(\mathbf{y}))_i - 1| \leq \frac{\epsilon\Delta s^2 + \epsilon\Delta s^2}{-\epsilon\Delta s^2 + \Delta s^2 - \epsilon\Delta s^2}, \tag{42}$$

implying that $F_n : [0,2]^n \to [1-4\epsilon, 1+4\epsilon]^n$. Because the denominator of (40) 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 (40) 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

$$\lim_{n\to\infty} \|\mathbf{y}_n - \mathbf{1}_n\|_\infty = 0. \tag{43}$$

We now use this result to derive expressions for π_0, π and γ, using equations (39), (35) and (36).

We have from (39), (43) and (11)

$$\lim_{n\to\infty} \frac{\mathbf{1}_n^T\mathbf{q}_n}{n} = \lim_{n\to\infty} \frac{(\mathbf{W}^n)_{ii}(\mathbf{y}_n)_i}{s^2 n} = \beta. \tag{44}$$

Moreover, using (39) (43) and (10), a similar argument shows that

$$\lim_{n\to\infty} \frac{\mathbf{s}_{i_n}^T\mathbf{q}_n}{n} = 0, \tag{45}$$

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

$$\lim_{n\to\infty} \frac{\mathbf{q}_n^T\mathbf{S}\mathbf{q}_n}{n^2} = 0. \tag{46}$$

We therefore have from (35)

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

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

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

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

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

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

Finally, from (38), it is clear that A approaches

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

so through (37), it is clear that π_0 converges to

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

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_ℓ^n clusters of tightly connected sets of agents, with ℓ 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 β. We have:

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

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

$$(\pi^*)^2\sigma^2 = \frac{\beta^2 \left(\beta + \Delta^2\right)^2 \sigma^6}{\left(\Delta^2 + \beta\left(\beta + \Delta^2\right)\sigma^2\right)^2}, \tag{47}$$

$$(\gamma^*)^2\Delta^2 = \frac{\Delta^2 \left(\beta + \Delta^2\right)^2 \sigma^4}{\left(\Delta^2 + \beta\left(\beta + \Delta^2\right)\sigma^2\right)^2}, \tag{48}$$

$$var(\tilde{p}) = \frac{\left(\beta + \Delta^2\right)^2 \sigma^4 \left(\Delta^2 + \beta^2\sigma^2\right)}{\left(\Delta^2 + \beta\left(\beta + \Delta^2\right)\sigma^2\right)^2}. \tag{49}$$

(47) implies that

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

and this proves part (a).

(48) implies that

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

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

Finally, (49) implies that

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

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

Proof of Proposition 4: It is straightforward from Theorem 1 and the projection theorem that

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

Hence the result follows. ∎

Proof of Proposition 5: From (23), 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 $\frac{\beta}{\Delta^2}$ for large n. Therefore

$$
\begin{aligned}
\frac{\alpha_{0i}}{\beta_i} &\xrightarrow{n\to\infty} \frac{\bar{X}\Delta^2 + \bar{Z}\beta\sigma^2}{\sigma^2\Delta^2 + \sigma^2\beta}, \\
\frac{\alpha_{1i}}{\beta_i} &= \frac{\boldsymbol{\pi}^T \mathbf{S} \boldsymbol{\pi} + \gamma^2 n^2 \Delta^2 - (\mathbf{1}^T \boldsymbol{\pi})(\boldsymbol{\pi}^T \mathbf{s}_i)}{s_i^2(\boldsymbol{\pi}^T \mathbf{S} \boldsymbol{\pi} + \gamma^2 n^2 \Delta^2) - (\boldsymbol{\pi}^T \mathbf{s}_i)^2} \\
&\xrightarrow{n\to\infty} \frac{1}{s_i^2} = \frac{W_i}{s^2}, \\
\frac{\alpha_{2i}}{\beta_i} &= \frac{(\mathbf{1}^T \boldsymbol{\pi}) s_i^2 - (\boldsymbol{\pi}^T \mathbf{s}_i)}{s_i^2(\boldsymbol{\pi}^T \mathbf{S} \boldsymbol{\pi} + \gamma^2 n^2 \Delta^2) - (\boldsymbol{\pi}^T \mathbf{s}_i)^2} \\
&\xrightarrow{n\to\infty} \frac{\beta}{\Delta^2\gamma^*}.
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
\frac{1}{\beta_i} &= \frac{(\sigma^2 + s_i^2)\left(\boldsymbol{\pi}^T \mathbf{S} \boldsymbol{\pi} + n^2\Delta^2\gamma^2 + (\mathbf{1}^T \boldsymbol{\pi})^2\sigma^2\right) - \left((\mathbf{1}^T \boldsymbol{\pi})\sigma^2 + (\boldsymbol{\pi}^T \mathbf{s}_i)\right)^2}{\sigma^2\left(s_i^2(\boldsymbol{\pi}^T \mathbf{S} \boldsymbol{\pi} + \gamma^2 n^2 \Delta^2) - (\boldsymbol{\pi}^T \mathbf{s}_i)^2\right)} \\
&= \frac{(\sigma^2 + s_i^2)\left(\boldsymbol{q}^T \mathbf{S} \boldsymbol{q}/n^2 + \Delta^2 + (\mathbf{1}^T \boldsymbol{q})^2\sigma^2/n^2\right) - \left((\mathbf{1}^T \boldsymbol{q})\sigma^2 + (\boldsymbol{q}^T \mathbf{s}_i)\right)^2/n^2}{\sigma^2\left(s_i^2(\boldsymbol{q}^T \mathbf{S} \boldsymbol{q}/n^2 + \Delta^2) - (\boldsymbol{q}^T \mathbf{s}_i)^2/n^2\right)} \\
&\xrightarrow{n\to\infty} \frac{(\sigma^2 + s_i^2)\left(\Delta^2 + \beta^2\sigma^2\right) - (\beta\sigma^2)^2}{\sigma^2 s_i^2 \Delta^2} = \frac{1}{s_i^2} + \frac{1}{\sigma^2} + \frac{\beta^2}{\Delta^2}.
\end{aligned}
\tag{50}
$$

Thus,

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

Since

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

the expression for the demand function reduces to

$$
\psi_i(\tilde{x}_i, \tilde{p}) = \frac{\bar{X}\Delta^2 + \bar{Z}\beta\sigma^2}{\sigma^2\Delta^2 + \sigma^2\beta} - \frac{\Delta^2}{\sigma^2(\Delta^2 + \beta)}\tilde{p} + \frac{W_i}{s^2}(\tilde{x}_i - \tilde{p}),
\tag{51}
$$

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 6: 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

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

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

$$\Pi=\gamma^*\Delta^2=\frac{\Delta^2\left(\beta+\Delta^2\right)\sigma^2}{\Delta^2+\beta\left(\beta+\Delta^2\right)\sigma^2}. \tag{52}$$

We also have

$$\begin{aligned}
\Pi_i &= \frac{\Delta^2}{\sigma^2(\Delta^2+\beta)}\left((\gamma^*)^2\Delta^2-\pi^*(1-\pi^*)\sigma^2\right)+\frac{W_i}{s^2}\left((1-\pi^*)^2\sigma^2+(\gamma^*)^2\Delta^2\right) \\
&= \frac{\Delta^4\left(W_i+s^2\Delta^2\right)\sigma^2+W_i\Delta^2\left(\beta+\Delta^2\right)^2\sigma^4}{s^2\left(\Delta^2+\beta\left(\beta+\Delta^2\right)\sigma^2\right)^2},
\end{aligned} \tag{53}$$

It then follows from (53) that

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

Hence the proposition follows. ∎

Proof of Proposition 7: (a) It follows from (52) that

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

Observe that the expression above is strictly negative if and only if $\sigma < \frac{1}{\Delta}$ and $\beta < \frac{\Delta}{\sigma}-\Delta^2$. This proves part (a).
(b) The decomposition into Π^F and Π^I follows immediately from (52,53). We have

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

which is positive, decreasing in β and approaches 0 as β tends to ∞.
(c) That Π^F is positive and approaches zero as $\beta \to \infty$ is immediate since

$$\pi^F = \frac{\beta\Delta^2\sigma^2(\beta^2\sigma^2 + \Delta^4\sigma^2 + \Delta^2(1 + 2\beta\sigma^2))}{(\beta^2\sigma^2 + \Delta^2(1 + \beta\sigma^2))^2}. \tag{54}$$

Non-monotonicity of Π^F in β can be easily observed from (54).
(d) This follows immediately from (b) and (c) . ∎

Proof of Proposition 8: Following Theorem 1 and (51), we can rewrite agent i's demand function as follows:

$$\psi_i(\tilde{x}_i, \tilde{p}) \quad = \quad c_i + \tfrac{\Delta^2(-\beta s^2 + W_i)}{s^2(\Delta^2 + \beta(\beta + \Delta^2)\sigma^2)}\tilde{X} + \tfrac{s^2\Delta^2 + (\beta + \Delta^2)\sigma^2 W_i}{s^2(\Delta^2 + \beta(\beta + \Delta^2)\sigma^2)}\tilde{Z} + \tfrac{\sum_{k \in W(i)} \tilde{\epsilon}_k}{s^2}, \tag{55}$$

where c_i is a constant scalar. Thus,

$$\begin{aligned} cov\left(\psi_i(\tilde{x}_i, \tilde{p}), \psi_j(\tilde{x}_j, \tilde{p})\right) \quad = \quad & \left(\tfrac{\Delta^2(-\beta s^2 + W_i)}{s^2(\Delta^2 + \beta(\beta + \Delta^2)\sigma^2)}\right)\left(\tfrac{\Delta^2(-\beta s^2 + W_j)}{s^2(\Delta^2 + \beta(\beta + \Delta^2)\sigma^2)}\right)\sigma^2 \\ & + \left(\tfrac{s^2\Delta^2 + (\beta + \Delta^2)\sigma^2 W_i}{s^2(\Delta^2 + \beta(\beta + \Delta^2)\sigma^2)}\right)\left(\tfrac{s^2\Delta^2 + (\beta + \Delta^2)\sigma^2 W_j}{s^2(\Delta^2 + \beta(\beta + \Delta^2)\sigma^2)}\right)\Delta^2 + W_{ij}. \end{aligned} \tag{56}$$

On the other hand, observe from (55) that variance of agent i's demand, $var\left(\psi_i(\tilde{x}_i, \tilde{p})\right)$, does not depend on W_{ij}. Therefore, following (56) we have

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

Hence we have the desired result. ∎

Proof of Proposition 9:
(a): From (8) and (51) it follows that $\psi_i \sim N(0, a_1 + a_2 W_i + a_3 W_i^2)$, where $a_1 = \frac{\Delta^6 + \beta^2\Delta^4\sigma^2}{a_4^2}$, $a_2 = \frac{1}{s^2}\left(1 + \frac{2\Delta^6\sigma^2}{a_4^2}\right)$, $a_3 = \frac{\Delta^2\sigma^2(\beta^2\sigma^2 + \Delta^4\sigma^2 + \Delta^2 + 2\Delta^2\beta\sigma^2)}{s^4 a_4^2}$, and $a_4 = \beta^2\sigma^2 + \Delta^2 + \Delta^2\beta\sigma^2$. Since, $E[|\tilde{z}|] = \sqrt{\frac{2A}{\pi}}$ for a general normally distributed random variable, $z \sim N(0, A)$, it follows that

$$\psi_i^{unsigned} = \sqrt{\frac{2(a_1 + a_2 W_i + a_3 W_i^2)}{\pi}}, \tag{57}$$

It immediately follows that this function is increasing and concave, with the given asymptotics. It is also clear that $E[\psi_i^2] = var(\psi_i) + E[\psi_i]^2 = var(\psi_i) = a_1 + a_2 W_i + a_3 W_i^2 = \frac{\pi}{2}(\psi_i^{unsigned})^2$, so it is indeed the case that $\psi_i^{unsigned} = \sqrt{\frac{2}{\pi}E[\psi_i^2]}$.

(b,c):

$$\begin{aligned} E\left[\sum_i \psi_i^2(W_i)d(i)\right] = \sum_i E[\psi_i^2(W_i)]d(i) = \sum_i (a_1 + a_2 W_i + a_3 W_i^2)d(i) = a_1 + a_2 s^2\beta + a_3 s^4(\beta^2 + \sigma_\beta^2) \\ = \beta^2 + \Delta^2 + a_3\sigma_\beta^2. \end{aligned}$$

Therefore, $\psi^{market} = \sqrt{\frac{\pi}{2}(\beta^2 + \Delta^2 + a_3\sigma_\beta^2)}$, and ψ^{market} is increasing in σ_β. Moreover, for small σ_β, ψ^{market} is increasing in β. Also, it is easy to show that $\frac{\partial a_3}{\partial \beta} < 0$, so for large σ_β, ψ^{market} is decreasing in β. ∎

Proof of Proposition 10: 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 $\sum_{i=1}^n d^n(i)f(i) \in [c_0, c_0 + c_1\beta + \epsilon]$, 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

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

where the last equality follows by using (12-15,51). 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 11: (a) This follows immediately from Jensen's inequality, since $CE(W)$ is a strictly convex function of $W \ge 1$.

(b) We first note that the "two-point distribution," for which a fraction $\beta - \lfloor\beta\rfloor$ of the agents has $\lfloor\beta\rfloor + 1$ connections and the rest has $\lfloor\beta\rfloor$ connections, has connectedness $(\beta - \lfloor\beta\rfloor)(\lfloor\beta\rfloor + 1) + (1-\beta+\lfloor\beta\rfloor)\lfloor\beta\rfloor = \beta$, 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 β, and for $\beta \notin \mathbb{N}$, there is no one-point distribution with connectedness β. We define $n = \lfloor\beta\rfloor$, $q_n = 1 - \beta + \lfloor\beta\rfloor$, $q_{n+1} = \beta - \lfloor\beta\rfloor$.

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 $\sum_i r_i x_i < \infty$.[27] 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, θ 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, $\mathcal{D}(d_1, d_2) = \inf\{\epsilon > 0 : F_{d_1}(x - \epsilon) - \epsilon \leq F_{d_2}(x) \leq F_{d_1}(x + \epsilon) \text{ for all } x \in \mathbb{R}_+\}$. 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 $\beta > 0$, is $D_\beta = \{d \in D : E[d] = \beta\}$.

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 $d \in D^1 \cap D_{\mathbb{N}} \cap D_\beta$, with $\beta \notin \mathbb{N}$, 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 $d \in D^1 \cap D_{\mathbb{N}} \cap D_\beta$. Clearly, since $\hat{d}$ is the only two-point distribution in $D^1 \cap D_{\mathbb{N}} \cap D_\beta$, and there is no one-point distribution in $D^1 \cap D_{\mathbb{N}} \cap D_\beta$, 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 $E[d] = \beta$. 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 $E[d] = \beta$. 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 $E[d] = \beta$ 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 $d_3 = (r_{n+1} - q_{n+1})\delta_{n+1} + \sum_{i>n+1} r_i \delta_i$.

These decompositions imply that

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

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

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, $d_m \in D^1 \cap D_\beta$.

Now, if $r_{n+1} \geq q_{n+1}$, since $d \in D^1$, it must be that $S(d_1) + S(d_3) = q_n - r_n$, and since $E[d_L + d_2 + d_R] = \beta = E[q_n\delta_n + q_{n+1}\delta_{n+1}]$ 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 $E[d_a + d_2] = \beta$, 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, $E[V_f(d)] \leq E[V_f(d_m)] < E[V_f(\hat{d})]$. A similar argument can be applied if $r_n \geq q_n$.

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 (a,b,18,19) it follows that $\overline{CE}^*(\beta)$ is of the form $\overline{CE}^*(\beta) = \frac{1}{2}\log(v(\beta))$, where $v(\beta) = \frac{(\Delta^2+(\beta+\Delta^2)^2\sigma^2)(\beta^2 s^2\sigma^2+\Delta^2 s^2+\beta\Delta^2\sigma^2)}{s^2(\beta^2\sigma^2+\Delta^2+\Delta^2\beta\sigma^2)^2}$. It immediately follows that $v'(\beta)$ is of the form $-v_2(\beta)(c_4\beta^4 + c_3\beta^3 + c_2\beta^2 + c_1\beta + c_0)$, where $v_2(\beta) > 0$ for all $\beta > 0$, $c_4 > 0$, $c_3 > 0$ and $c_2 > 0$, and where $c_1 = \Delta^2 + 4s^2 - 3$ and $c_0 = 2\Delta^2 s^2\sigma^2 - \Delta^2\sigma^2 - 1$. Moreover, since $c_4 > 0$, it follows that $v'(\beta) < 0$ for large β.

From Descartes' rule of signs, it follows that the maximum number of roots to $v'(\beta) = 0$ is two, and there can only be two roots if $c_1 < 0$ and $c_0 > 0$. The condition $c_0 > 0$ is equivalent to $2s^2 - 1 > \frac{1}{\Delta^2\sigma^2}$, which in particular implies that $s^2 > \frac{1}{2}$. Similarly, $c_1 < 0$ iff $3 - 4s^2 > \Delta^2\sigma^2$, which in particular implies that $s^2 < \frac{3}{4}$. Multiplying these two conditions, we get that a necessary condition for two roots to be possible is that $(3-4s^2)(2s^2-1) > 1$, for $s \in (1/2, 3/4)$, but it is easy to check that $(3-4s^2)(2s^2-1)$ is in fact less than one in this region. Therefore, it can not be the case that $c_1 < 0$ and $c_0 > 0$ at the same time, and there can be at most one root to the equation $v'(\beta) = 0$. Since $v'(\beta) < 0$ for large β, it must therefore be the case that $v(\beta)$ is either decreasing for all β, or initially increasing and then decreasing, with a unique maximum. It is easy to check numerically that both cases are in fact possible. We are done.

Proof of Proposition 12:

Since f is weakly concave and CE is concave, $CE - f$ is a concave function of W, and an identical argument as in the proofs of Proposition 11 (a),(b) can be made to show that a degenerate (i.e., uniform) network is optimal. Now, since $\overline{CE}^*$, as defined in Proposition 11, is decreasing for large β, it follows that $\overline{CE}^* - f$ is decreasing for large β, so the optimal β must be interior. ∎

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.

ARAGON, G., R. BILDIK, AND D. YAVUZ (2007): "Do institutional investors have an information advantage," Discussion paper, Working Paper, Arizona State University.

BACK, K., H. CAO, AND G. WILLARD (2000): "Imperfect competition among informed traders," *Journal of Finance*, 55, 21172155.

BARABASI, A. L., AND R. ALBERT (1999): "Emergence of scaling in random networks," *Science*, 286, 509–512.

BARBER, B., Y. T. LEE, Y. J. LIU, AND T. ODEAN (2009): "Just how much do individual investors lose by trading?," *Review of Financial Studies*, 22, 609–632.

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.

CALVET, L., J. CAMPBELL, AND P. SODINI (2007): "Down or Out: Assessing the welfare costs of household investment mistakes," *Journal of Political Economy*, 115, 707–747.

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.

CUTLER, D. M., J. M. POTERBA, AND L. H. SUMMERS (1989): "What moves stock prices," *Journal of Portfolio Management*, 15, 4–12.

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.

DORN, D., G. HUBERMAN, AND P. SENGMUELLER (2008): "Correlated trading and returns," *Journal of Finance*, 63, 885–920.

FAFCHAMPS, M., AND S. LUND (2003): "Risk-Sharing Networks in Rural Phillipines," *Journal of Development Economics*, 71, 261–287.

FAIR, R. (2002): "Events that shook the market," *Journal of Business*, 75, 713–731.

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

Foster, F., and S. Viswanathan (1996): "Strategic trading when agents forecast the forecasts of others," *Journal of Finance*, 51, 14371478.

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

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

Gomez, J., R. Priestley, and F. Zapatero (2009): "The effect of relative wealth concerns on the cross-section of stock returns," AFA2009 conference paper, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1102231.

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

Grinblatt, M., and M. Keloharju (2000): "The investment behavior and performance of various investor types: a study of Finland's unique data set," *Journal of Financial Economics*, 55, 43–57.

——— (2001): "How distance, language, and culture influence stockholdings and trades," *Journal of Finance*, 56, 1053–1073.

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

Hau, H. (2001): "Location matters: an examination of trading profits," *Journal of Finance*, 56, 1959–1983.

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

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 (2009): "Nondiversification traps in markets for catastophe insurance," *The Review of Financial Studies*, 22, 559–594.

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.

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): "Crawling the web for emerging cybercommunities," in *Proceedings of 8th WWW Conference*, pp. 403–416.

Kyle, A. S. (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315–1336.

Massa, M., and A. Simonov (2006): "Hedging, familiarity and portfolio choice," *Review of Financial Studies*, 19, 633–685.

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

Myers, C. A., and G. P. Shultz (1951): *The Dynamics of a Labor Market*. Prentice-Hall.

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.

Pareek, A. (2009): "Information networks: Implications for mutual fund trading behavior and stock returns," Working Paper, Yale University.

Reiss, A., and G. P. Shultz (1970): *Workers in an Urban Labor Market*. University of Chicago Press.

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.

Uzzi, B. (1996): "The Sources and Consequences of Embededdeness for the Economic Performance of Organizations: The Network Effect," *American Sociological Review*, 61, 674–698.

Vives, X. (1995): "Short-term investment and the information efficiency of the market," *Review of Financial Studies*, 8, 125–160.

Wang, J. (1993): "A model of intertemporal asset prices under asymmetric information," *Review of Economic Studies*, 60, 249–282.

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.

Xia, C. (2007): "Communication and confidence in financial markets," Working Paper, University of Minnesota.