

Trading, Profits, and Volatility in a Dynamic Information Network Model*

Johan Walden[†]

November 14, 2013

Abstract

We introduce a dynamic noisy rational expectations model, in which information diffuses through a general network of agents. In equilibrium, agents' trading behavior and profits are determined by their position in the network. Agents who are more closely connected have more similar period-by-period trades, and an agent's profitability is determined by a centrality measure that is closely related to so-called Katz centrality. The model generates rich dynamics of aggregate trading volume and volatility, beyond what can be generated by heterogeneous preferences in a symmetric information setting. An initial empirical investigation suggests that price and volume dynamics of small stocks may be especially well explained by such asymmetric information diffusion. The model could potentially be used to study individual investor behavior and performance, and to analyze endogenous network formation in financial markets.

Keywords: *Information diffusion, information networks, heterogeneous investors, portfolio choice, asset pricing.*

*I thank seminar participants at the Duisenberg School of Finance, Maastricht University, the Oxford-Man Institute, Rochester (Simon School), Stockholm School of Economics, and USC, Bradyn Breon-Drish, Laurent Calvet, Nicolae Gârleanu, Brett Green, Pierre de la Harpe, Philipp Illeditsch, Matt Jackson, Svante Jansson, Ron Kaniel, Anders Karlsson, Martin Lettau, Dmitry Livdan, Hanno Lustig, Gustavo Manso, Christine Parlour, Han Ozsoylev, Elvira Sojli, Matas Šileikis, and Deniz Yavuz for valuable comments and suggestions.

[†]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.

1 Introduction

There is extensive evidence that heterogeneous and decentralized information diffusion influences investors' trading behavior. Shiller and Pound (1989) survey institutional investors in the NYSE, and find that a majority attribute their most recent trades to discussions with peers. Ivković and Weisbenner (2007) find similar evidence for households. Hong, Kubik, and Stein (2004) find that fund managers' portfolio choices are influenced by word-of-mouth communication. Heimer and Simon (2012) find similar influence from on-line communication between retail foreign exchange traders.

Such information diffusion may help explain several fundamental stylized facts of stock markets. First, investors are known to hold vastly different portfolios, in contrast to the prediction of classical models that everyone should hold the market portfolio. The standard explanations for such diverse portfolio holdings are hedging motives and heterogeneous preferences, but several studies indicate that there are limitations to how well such motives can explain observed heterogeneous investor behavior.¹ With decentralized information diffusion, however, it is unsurprising if significantly heterogeneous behavior of investors is observed in the market. Second, stock markets are known to experience large price movements that are unrelated to public news, as documented in Cutler, Poterba, and Summers (1989), and Fair (2002). These studies find that over two thirds of major stock market movements cannot be attributed to public news events, suggesting that there are other channels through which information is incorporated into asset prices. Third, the dynamics of trading volume and asset prices are known to be very rich. Returns and trading volume in many markets are heavy-tailed, time varying, show “long memory,” and are related to each other in a complex way (see Gabaix, Gopikrishnan, Plerou, and Stanley 2003; Karpoff 1987; Gallant, Rossi, and Tauchen 1992; Bollerslev and Jubinski (1999); Lobato and Velasco 2000). Lumpy information diffusion provides a potential explanation for such behavior. In periods when more information diffuses, volatility is higher, as is trading volume (see, Clark 1973; Epps and Epps 1976; Andersen 1996). To generate rich trading volume dynamics, however, such information diffusion must necessarily be heterogeneous.

In this paper, we follow a recent strand of literature that uses information networks to model information diffusion (see Colla and Mele 2010, Ozsoylev and Walden 2011, Han and Yang 2013, and Ozsoylev, Walden, Yavuz, and Bildik 2013). Agents who are directly linked in a network share information, which consequently diffuses among the population over time in a well-specified manner. This literature has made important observations about effects of information networks, but several key questions remain open. How does the network structure in a market determine the dynamic trading behavior of its agents, and their performance? How does the network structure influence aggregate properties of the market, e.g., price volatility and trading volume? What does heterogeneous information diffusion “add” compared with, e.g., what can be

¹For example, Massa and Simonov (2006) find that hedging motives for human capital risk—a fundamental source of individual investor risk—does not explain heterogeneous investment behavior among individual investors well. Similarly, Calvet, Campbell, and Sodini (2007) and Calvet, Campbell, and Sodini (2009) find that diversification and portfolio rebalancing motives do not explain investors' portfolio holdings well.

generated by heterogeneous preferences alone? These questions are obviously fundamental for our understanding of the impact of information networks on financial markets, and ultimately for how well such information networks can explain observed investor and market behavior. In this paper we analyze these questions.

We introduce a dynamic noisy rational expectations model in which agents in a network share information with their neighbors. Agents receive private noisy signals about the unknown value of an asset in stochastic supply, and trade in a market over multiple time periods. In each period, they share all the information they have received up until that point with their direct neighbors, leading to gradual diffusion of the private signals. The structure of the network is completely general.

As a first contribution, we prove the existence of a noisy rational expectations equilibrium, and present closed-form expressions for all variables of interest. Theorem 1 provides the main existence and characterization result for a Walrasian equilibrium in a large network economy. To define the large economy equilibrium, we use the concept of replica networks, assuming that there is a local network structure (e.g., at the level of a municipality) and that there are many similar such local network structures in the economy. This allows for a clean characterization of equilibrium, as well as justifies the assumption that agents act as price takers and are willing to share information. We know of no other network model of information diffusion in a centralized financial market (i.e., exchange) that allows for a complete characterization of equilibrium, that is completely general with respect to network structure, and that is based on first principles of financial economics.

The structure of the network is crucial in determining asset pricing dynamics. We show in a simple example that price informativeness and volatility at any given point in time does not only depend on the specific information agents in the model have obtained at that point, but also on how the information has diffused through the network. The equilibrium outcome thus depends on complex properties of the network, beyond the mere precision of agents' signals at any specific point in time.

We next study how the network structure determines the trading behavior and profitability of agents. A priori, one may expect that portfolio *holdings* of agents who are close in the network should be positively correlated, whereas their *trades* may be negatively correlated in some periods, because some agents trade earlier on information than others and may be ramping down their investments to realize profits when other agents are ramping up their investments. We show that, contrary to this intuition, the period-by-period trades of agents in the network are always positively related, and increasingly so in the degree of the overlap in their connectedness. This result justifies the use of data on trades to draw inferences about network structure, as has been done in previous studies, providing our second contribution.

As a third contribution, we study what determines who in the network makes profits. It is argued in Ozsoylev and Walden (2011) and elsewhere that some type of centrality measure should determine agent profitability.² Centrality—a fundamental concept in network theory—

²In an empirical study, Ozsoylev, Walden, Yavuz, and Bildik (2013) study the trades of all investors on the

captures the concept that it is not only who your direct neighbors are that matters, but also who your neighbors' neighbors are, who your neighbors' neighbors' neighbors are, etc. The argument is that agents who are centrally placed tend to receive information signals early, and therefore perform better in the market than peripheral agents, who tend to receive information later. It is not a priori clear, however, which centrality measure—among several—that is appropriate in this context.

We show that profitability is determined by a centrality measure related to eigenvector centrality or, more generally, to so-called Katz centrality. To the best of our knowledge, this is the first complete characterization of the relationship between agent centrality and performance in a general information network model of financial markets. For large random networks, Katz centrality dominates other common centrality measures in our model, i.e., it dominates degree centrality, average distance and closeness centrality. We show in simulations that the ranking of centrality measures also holds in medium-sized random networks. Thus, the use of centrality measures when studying individual investor behavior (e.g., in Ozsoylev, Walden, Yavuz, and Bildik 2013) is theoretically justified. The result is quite intuitive: The Katz centrality measure provides the best balance between direct connections and connections at farther distances, whereas degree centrality focuses exclusively on direct connections, and average distance and closeness centrality are tilted toward connections at far distances. Our main results that characterize trading behavior, profitability and welfare of agents, and the relationship between centrality and profitability, are Theorems 2-4.

Our fourth contribution is to derive and analyze several aggregate results regarding the dynamic behavior of price volatility and trading volume in the model. Specifically, the network structure in an economy is an important determinant of volatility and trading volume in the time series, as well as of the relationship between the two. Several stylized properties naturally arise in the model, for example, persistence of shocks to volatility and trading volume, as well as lead-lag relationships between the two.

We show that the rich dynamics of volatility and volume in our general information network model cannot be generated by heterogeneous preferences alone in a symmetric network. Again, the intuition is straightforward: With symmetry, there can be no bottlenecks in the information network. This in turn implies that the rate of information diffusion over time must be unimodal, i.e., it cannot slow down and then speed up again. Consequently, aggregate market dynamics are restricted under symmetry, whereas (as we show) there are no such unimodality restrictions in the general case.

Finally, we study volatility and volume of a sample of U.S. stocks. We compare Large Cap and Small Cap companies, and find that the behavior of Small Cap companies is better explained by asymmetric networks with a large degree of nonpublic information diffusion, suggesting that information networks may be especially important for such stocks. Our theoretical results on

Istanbul Stock Exchange in 2005, and find a positive relationship between investors' so-called eigenvector centrality and profitability, but this choice of centrality measure is not theoretically justified. Several other finance papers discuss and use various centrality measures without a complete theoretical justification, see e.g., Das and Sisk (2005), Adamic, Brunetti, Harris, and Kirilenko (2010), Li and Schurhoff (2012) and Buraschi and Porchia (2012).

aggregate volume and volatility are summarized in Theorems 5-7.

The rest of the paper is organized as follows. In the next section, we discuss related literature. In Section 3, we introduce the model and characterize equilibrium. In Section 4, we analyze trading behavior and profitability of individual agents. In Section 5, we study the implications of network structure for aggregate volatility and trading volume. Finally, Section 6 concludes. All proofs are delegated to the appendix.

2 Related Literature

Our paper is most closely related to the recent strand of literature that studies the effects of information diffusion on trading and asset prices. Colla and Mele (2010) show that the correlation of trades among agents in a network varies with distance, so that close agents naturally have positively correlated trades, whereas the correlation may be negative between agents who are far apart. Their model is dynamic, and assumes a very specific symmetric network structure, namely a circle, where each agent has exactly two neighbors. This restricts the type of dynamics that can arise in their model. Ozsoylev and Walden (2011) introduce a static rational expectations model that allows for general network structures and study, among other things, how price volatility varies with network structure. Their model is not appropriate for studying dynamic information diffusion, however, and is therefore not well-suited for several of the questions analyzed in this paper, e.g., the relationship between agent profitability and centrality, and the short-term correlation between agents' trade.

Han and Yang (2013) study the effects of information diffusion on information acquisition. They show that in equilibrium, information diffusion may reduce the amount of aggregate information acquisition, and therefore also the informational efficiency and liquidity in the market. Their model is also static, and does thereby not allow for dynamic effects. In an empirical study, Ozsoylev, Walden, Yavuz, and Bildik (2013) test the relationship between centrality—constructed from the realized trades of all investors in the market—and profitability. They find that more central agents, as measured by eigenvector centrality, are more profitable. However, they do not justify this choice of centrality measure theoretically. Pareek (2012) studies how information networks—proxied by the commonality in stock holdings—among mutual is related to return momentum.

A different strand of literature studies information diffusion through so-called information percolation (Duffie and Manso 2007, Duffie, Malamud, and Manso 2009). In the original setting, a large number of agents meet randomly in a bilateral decentralized (OTC) market and share information, and the distribution of beliefs over time can then be strongly characterized. Recently, the model has been adapted to centralized markets, with exchange traded assets and observable prices—a setting more closely related to ours. For example, Andrei (2012), shows that persistent price volatility can arise in such a model. In contrast to our model, in which some agents may be better positioned than others, these models are *ex ante* symmetric in that all agents have the same chance of meeting and sharing information.

Babus and Kondor (2013) also introduce a model of information diffusion in a bilateral OTC market. As in our paper, their network can be perfectly general. In contrast to our model, there is no centralized information aggregation mechanism in their setting, and therefore no interplay between diffusion through public and private channels. Moreover, agents have private values in their model, and their model is static.

This paper is also related to the literature on information diffusion and trading volume (Clark 1973). Lumpy information diffusion was suggested to explain heavy-tailed unconditional volatility of asset prices, as an alternative to the stable Paretian hypothesis. Under the Mixture of Distributions Hypothesis (MDH), lumpiness in the arrival of information lead to variation in return volatility and trading volume, as well as a positive relation between the two (see Epps and Epps 1976 and Andersen 1996). Foster and Viswanathan (1995) build upon this intuition to develop a model with endogenous information acquisition, leading to a positive autocorrelation of trading volume over time. Similar results arise in He and Wang (1995), in a model where an infinite number of ex ante identical agents receive noisy signals about an asset's fundamental value. Admati and Pfleiderer (1988) explain U-shaped intra-daily trading volume in a model with endogenous information acquisition.

Our paper further explores the richness of the dynamics of volatility and volume that arises when agents share their signals, allowing for completely general asymmetry in how some agents are better positioned than others. This extension may potentially shed further light on the very rich dynamics of volatility and volume, and the relationship between the two (see Karpoff 1987, Gallant, Rossi, and Tauchen 1992, Bollerslev and Jubinski 1999, Lobato and Velasco 2000, and references therein). A related strand of literature explores the role of trading volume in providing further information to investors about the market, see Blume, Easley, and O'Hara (1994), Schneider (2009), and Breon-Drish (2010). Our model does not explore this potential informational role of trading volume.

Our study is related to the large literatures on games on networks, see the survey of Jackson and Zenou (2012). The games in these models are typically not directly adaptable to a finance setting. Our existence result and the characterization of equilibrium in a model based on first principles of financial economics are therefore of interest. Since the welfare of agents in equilibrium can be simply characterized, our model could potentially also be used to study endogenous network formation, see Jackson (2005) for a survey of this literature.

Finally, our paper is related to the (vast) general literature on asset pricing with heterogeneous information (see, e.g., the seminal papers by Grossman 1976, Hellwig 1980, Kyle 1985, and Glosten and Milgrom 1985). Technically, we build upon the model in Vives (1995), who introduces a multi-period noisy rational expectations model in a similar spirit as the static model in Hellwig (1980). Like Vives, we assume the presence of a risk-neutral competitive market maker, to facilitate the analysis in a dynamic setting. This simplifies the characterization of equilibrium considerably. The cost of this assumption is that asset prices simply reflect the expected terminal value of the asset conditioned on public information at all points in time. Since our focus is on volatility and trading volume, this is a marginal cost for us. Unlike Vives,

we allow for information diffusion among agents, through general network structures.

3 Model

There are N agents, enumerated by $a \in \mathcal{N} = \{1, \dots, N\}$, in a $T + 1$ -period economy, $t = 0, \dots, T + 1$, where $T \geq 2$. We define $\mathcal{T} = \{1, \dots, T\}$. Each agent, a , maximizes expected utility of terminal wealth, and has constant absolute risk aversion (CARA) preferences with risk aversion coefficient γ_a , $a = 1, \dots, N$,

$$U_a = E[-e^{-\gamma_a W_{a,T+1}}].$$

We summarize agents' risk aversion coefficients in the N -vector $\Gamma = (\gamma_1, \dots, \gamma_N)$.

There is one asset with terminal value $v = \bar{v} + \eta$, where $\eta \sim N(0, \sigma_v^2)$, i.e., the value is normally distributed with mean \bar{v} and variance σ_v^2 . Here, \bar{v} is known by all agents, whereas η is unobservable.

Agents are connected in a network, represented by a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$. The relation $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ describes which agents (vertices, nodes) are connected in the network. Specifically, $(a, a') \in \mathcal{E}$, if and only if there is a connection (edge, link) between agent a and a' . We will subsequently assume that there are many identical "replica" copies of this network in the economy, each copy representing a "local" network structure. This will make the economy "large" and justify price taking behavior of agents, as well as simplify the characterization of equilibrium. For the time being, we focus on one representative copy of this large network.

We use the convention that each agent is connected to himself, $(a, a) \in \mathcal{E}$ for all $a \in \mathcal{N}$, i.e., \mathcal{E} is reflexive. We also assume that connections are bidirectional, i.e., that \mathcal{E} is symmetric. A convenient representation of the network is by the *adjacency matrix* $E \in \{0, 1\}^{N \times N}$, with $(E)_{aa'} = 1$ if $(a, a') \in \mathcal{E}$ and $(E)_{aa'} = 0$ otherwise.

The distance function $D(a, a')$ defines the number of edges in the shortest path between agents a and a' . We use the conventions that $D(a, a) = 0$, and that $D(a, a') = \infty$ whenever there is no path between a and a' . The set of direct neighbors to agent a is $S_{a,1} = \{a' : (a, a') \in \mathcal{E}\}$. Moreover, the set of agents at distance $m > 1$ from agent a is $S_{a,m} = \{a' : D(a, a') = m\}$, and the set of agents at distance not further away than m is $R_{a,m} = \cup_{j=1}^m S_{a,m}$. The number of agents at a distance not further away than m from agent a is $V_{a,m} = |R_{a,m}|$. Here, $V_{a,1}$ is the *degree* of agent a , which we also refer to as agent a 's *connectedness*, whereas $V_{a,m}$ is agent a 's *m th order degree*. We use the convention that $V_{a,0} = 0$ for all a . We also define $\Delta V_{a,m} = |S_{a,m}|$.

We define N -vectors V^m , $m = 1, \dots, N$, where the a th element of V^m is $V_{a,m}$. Equivalently,

Definition 1 *The m th order degree vector, $V^m \in \mathbb{R}_+^N$, $m = 1, 2, \dots$, is defined as*

$$V^m = \chi(E^m)\mathbf{1}. \tag{1}$$

Here E^m is the m th power of the adjacency matrix, and $\chi : \mathbb{R}^{N \times N} \rightarrow \{0, 1\}^{N \times N}$ is a matrix

indicator function, such that $(\chi(A))_{i,j} = 0$ if $A_{i,j} = 0$ and $(\chi(A))_{i,j} = 1$ otherwise. Moreover, $\mathbf{1}$ is an N -vector of ones.

First order degree is commonly referred to as *degree centrality*.

Finally, the number of agents within a distance of m from both agents a and a' is $V_{a,a',m} = |R_{a,m} \cap R_{a',m}|$, and the number of neighbors at distance exactly m from both agents is $\Delta V_{a,a',m} = |S_{a,m} \cap S_{a',m}|$

3.1 Information diffusion

At $t = 0$, each agent receives a noisy signal about the asset's value, $s_a = v + \sigma \xi_a$, where $\xi_a \sim N(0, 1)$ are jointly independent across agents, and independent of v . At $T + 1$, the true value of the asset, v , is revealed. It will be convenient to use the precisions $\tau_v = \sigma_v^{-2}$ and $\tau = \sigma^{-2}$.

The graph, \mathcal{G} , determines how agents share information with each other. Specifically, at $t + 1$, agent a shares all signals he has received up until t with all his neighbors. We let $\mathcal{I}_{a,t}$ denote the information set that agent a has received up until t , either directly or via his network.

It is natural to ask why agents would voluntarily reveal valuable information to their neighbors. Of course, in a large economy with an infinite number of agents, sharing signals with ones' (finite number of) neighbors has no cost, since the actions of a finite number of agents will not influence prices. Even with an economy of finite size, as long as signals can be verified ex post, truthful revelation may be optimal in a repeated game setting, since an agent who provides misinformation can be punished by his neighbors, e.g., by being excluded from the network in the future. Even if signals are not ex post verifiable, it may still be possible for an agent to draw inferences about the truthfulness of another agent's signal, by comparing it with other received signals. Again, the threat of future exclusion from the network, could be used to enforce truthful information sharing. We therefore take the truthful information sharing behavior of agents as given. A potentially fruitful area for future research is to better understand in which finite sized financial networks truthful signal sharing can be sustained.

As in Ozsoylev, Walden, Yavuz, and Bildik (2013), we formalize the information sharing role of the network by defining

Definition 2 *The graph \mathcal{G} represents an information network over the signal structure $\{s_a\}_a$, if for all agents $a \in \mathcal{N}$, $a' \in \mathcal{N}$ and times $t = 1, \dots, T$, $s_{a'} \in \mathcal{I}_{a,t}$ if and only if $D(a, a') \leq t$.*

The information about the asset's value that an agent has received through the network up until time t can be summarized (as we shall see) by the sufficient statistic

$$z_{a,t} \stackrel{\text{def}}{=} \frac{1}{V_{a,t}} \sum_{j \in R_{a,t}} s_j = v + \zeta_{a,t},$$

where $\zeta_{a,t} = \frac{V_{a,t}}{\sigma^2} \xi_{a,t}$, and $\xi_{a,t} \sim N(0, 1)$.³ The number of signals agent a receives at t is $\Delta V_{a,t}$, and we therefore expect $\{\Delta V_{a,t}\}_{a \in \mathcal{N}, t \in \mathcal{T}}$ to be important for the dynamics of the economy.

3.2 Market

The market is open between $t = 1$ and $T + 1$. Agents in the information network submit limit orders, and a risk-neutral competitive market maker sets the price such that at each point in time it reflects all publicly available information, $p_t = E_t[v|\mathcal{I}_t^p]$, where \mathcal{I}_t^p is the time- t publicly available information set. At $T + 1$, the asset's value is revealed so $p_{T+1} = v$. Before trading begins, the price is set as the asset's ex ante expected value, $p_0 = \bar{v}$.

To avoid fully revealing prices, we make the standard assumption of stochastic supply of the asset. Specifically, in period t , noise traders submit market orders of u_t per trader in the network, where $u_t \sim N(0, \sigma_u^2)$. In other words, the noise trader demand is defined relative to the size of the population in the information network. As argued elsewhere in the literature, the noise trader assumption need not be taken literally, but is rather a reduced-form representation of unmodeled supply shocks. It could, e.g., represent hedging demand among investors due to unobservable wealth shocks, or other unexpected liquidity shocks. We do not further elaborate on the sources of these shocks. We will use the precision $\tau_u = \sigma_u^{-2}$.

Agents in the network are price takers. At each point in time they submit limit orders to optimize their expected utility of terminal wealth. They thus condition their demand on contemporaneous public information, as well as on their private information. An agent's total demand for the asset at time t is

$$x_{a,t} = \arg \max_x E \left[e^{-\gamma_a W_{a,T+1}} \bar{\mathcal{I}}_{a,t} \right], \quad (2)$$

subject to the budget constraint

$$W_{a,t+1} = W_{a,t} + x_{a,t}(p_{t+1} - p_t), \quad t = 1, \dots, T,$$

and his net time- t demand is $\Delta x_{a,t} = x_{a,t} - x_{a,t-1}$, with the convention that $x_{a,0} = 0$ for all agents. Here, $\bar{\mathcal{I}}_{a,t}$ contains all public and private information available to agent a at time t .

In the linear equilibrium we study, $z_{a,t}$ and p_t are jointly sufficient statistics for an agent's information set, $\bar{\mathcal{I}}_{a,t} = \{z_{a,t}, p_t\}$, leading to the functional form $x_{a,t} = x_{a,t}(z_{a,t}, p_t)$. Of course, an agent's optimal time- t strategy in (2) depends on the (optimal) future strategy. The dynamic problem can therefore be solved by backward induction. The primitives of the economy are summarized by the tuple $\mathcal{M} = (\mathcal{G}, \Gamma, \tau, \tau_u, \tau_v, \bar{v}, T)$.

We note that the assumption that the asset's value is revealed at $T + 1$ means that any residual uncertainty at T of the asset's value is completely mitigated by $T + 1$. We think of this as public information, which becomes available to all agents at $T + 1$. Alternatively, we

³A variation is to let agents receive new private signals in each time period. The analysis in this case is qualitatively similar, but not as clean because of the increased number of signals.

could have assumed that residual uncertainty is gradually incorporated into the market between T and T' for some $T' > T + 1$, keeping the assumption that the information diffusion between agents in the network only occurs until T .

The graph, \mathcal{G} , determines how information diffuses in the network over time, whereas Γ captures agent preferences. We wish to separate dynamics that can be generated solely by heterogeneity in preferences from those that require heterogeneity in network structure. To this end, we define an economy to be *preference symmetric* if $\gamma_a = \gamma$ for all agents, and some constant $\gamma > 0$. There are several symmetry concepts for graphs. The notion we use is so-called distance transitivity.⁴ Informally, symmetry captures the idea that any two vertices can be switched without the network changing its structure. To formalize the concept, we define an automorphism on a graph to be a bijection on the vertices of the graph, $f : \mathcal{N} \leftrightarrow \mathcal{N}$, such that $(f(a), f(a')) \in \mathcal{E}$ if and only if $(a, a') \in \mathcal{E}$. A graph is distance-transitive if for every quadruple of vertices, a, a', b , and b' , such that $D(a, b) = D(a', b')$, there is an automorphism, f , such that $f(a) = a'$ and $f(b) = b'$. An economy is said to be *network symmetric* if its graph is distance-transitive. Preference symmetric economies and network symmetric economies provide useful benchmarks to which the general class of economies can be compared.

We point out that network symmetry does *not* imply that the same amount of information is diffused among agents at each point in time. It does, however, still impose severe restrictions on how information may spread in the economy, as shown by the following lemmas.⁵

Lemma 1 *In a network symmetric economy, $\Delta V_{a,t}$ is the same for all agents at each point in time. That is, for each t , for each a , $\Delta V_{a,t} = \Delta V_t$ for some common ΔV_t .*

Thus, in a network symmetric economy, all agents have an equal precision of information at any point in time, although their signal realizations of course differ.

Lemma 2 *In a network symmetric economy the sequence $\Delta V_1, \Delta V_2, \dots, \Delta V_T$, is unimodal. Specifically, there are times $1 \leq t_1 \leq t_2 \leq T$, such that $\Delta V_{t+1} > \Delta V_t$ for all $t \leq t_1$, $\Delta V_{t+1} = \Delta V_t$ for all $t_1 < t \leq t_2$, and $\Delta V_{t+1} < \Delta V_t$ for all $t > t_2$.*

In other words, the typical behavior of the information diffusion process in a network symmetric economy is “hump-shaped,” initially increasing, after which it reaches a plateau and then decreases.

3.3 Replica network

To justify the assumption that agents are price takers, the number of agents needs to be large. Moreover, as analyzed in Ozsoylev and Walden (2011), restrictions on the distribution of number

⁴Other notions include vertex transitivity, distance regularity, arc-transitivity, t -transitivity, and strong regularity, see Briggs (1993). Distance transitivity is a stronger concept than vertex transitivity, arc-transitivity, and distance regularity, respectively, but neither stronger, nor weaker, than t -transitivity and strong regularity.

⁵The first result follows immediately from the fact that automorphisms preserve distances between nodes, see Briggs (1993), page 118. The second result follows from Taylor and Levingston (1978), where the result is shown for the larger class of distance regular graphs (see also Brouwer, Cohen, and Neumaier 1989, page 167).

of connections are needed, to ensure existence of equilibrium. Ozsoylev and Walden (2011) carry out a fairly general analysis of the restrictions needed on the degree distribution for the existence of equilibrium to be guaranteed. They show that a sufficient condition is that the degree distribution is not too fat-tailed. Compared with their static model, our model has the additional property of being dynamic. Therefore, not only would restrictions on first-order degrees be needed to ensure the existence of equilibrium, but also on degrees of all higher orders. In the dynamic economy, signals spread over longer distances, thereby “fattening” the tail of the distribution of signals among agents over time. We therefore believe that a general analysis would be technically challenging, while adding limited additional economic insight, which is why we choose the simplified approach.

We build on the concept of replica economies, originally introduced by Edgeworth (1881) to study the game theoretic core of an economy (see also Debreu and Scarf 1964). We assume that the full economy consists of a large number, M , of disjoint identical replicas of the network previously introduced, and that agents’ random signals are independent across these replicas. A replica network approach provides the economic and technical advantages of a large economy, namely that price taking behavior is rationalized and that the law of large numbers makes most idiosyncratic signals cancel out in aggregate, while avoiding the issues of signals spreading too quickly among some agents, causing equilibrium to break down.

The total number of agents in the economy is $\bar{N} = N \times M$. Formally, we define the set of agents in an M -replica economy as $\mathcal{A}_m = \mathcal{N} \times \{1, \dots, M\}$, where $a = (i, j) \in \mathcal{A}_m$ represents the i th agent in the j th replica network, in an economy with M replica networks. There is still one asset, one market, and one competitive market maker in the market with \bar{N} agents. We use the enumeration $a = 1, \dots, MN$, of agents, where agent (i, j) maps to $a = (j - 1)N + i$.

Agent (i, j) and (i, j') are thus ex ante identical in their network positions and in their signal distributions, although their signal realizations (typically) differ. We let M increase in a sequence of replica economies, with the natural embedding $\mathcal{A}_1 \subset \mathcal{A}_2 \cdots \subset \mathcal{A}_m \subset \cdots$, and take the limit $\mathcal{A} = \lim_{M \rightarrow \infty} \mathcal{A}_M$, letting \mathcal{A} define our large economy, in a similar manner as in Hellwig (1980). The network \mathcal{G} is thus a representative network in the large economy, \mathcal{A} . Our interpretation is that the network, \mathcal{G} , represents a fairly localized structure, perhaps at the level of a town or municipality in an economy, whereas \mathcal{A} represents the whole economy.

At time t , the market maker observes the average order flow per agent in the network⁶

$$w_t = u_t + \frac{1}{\bar{N}} \sum_{a=1}^{\bar{N}} \Delta x_{a,t}. \quad (3)$$

3.4 Equilibrium

We restrict our attention to linear equilibria in which agents in the same position in different replica networks are (distributionally) identical. Such equilibria are thus characterized by the

⁶Technically, the market maker observes $u_t + \lim_{M \rightarrow \infty} \frac{1}{MN} \sum_{a=1}^{MN} \Delta x_{a,t}$. We avoid such limit notation when this can be done without confusion.

behavior of agents $a = 1, \dots, N$, who are “representative.” Our main existence result is the following theorem, that shows existence of a linear equilibrium in the large economy under general conditions and, furthermore, characterizes this equilibrium.

Theorem 1 *Consider an economy characterized by \mathcal{M} . For $t = 1, \dots, T$, define*

$$\begin{aligned} A_t &= \frac{\tau}{N} \sum_{a=1}^N \frac{V_{a,t}}{\gamma_a}, \\ y_t &= \tau_u (A_t - A_{t-1})^2, \\ Y_t &= \sum_{s=1}^t y_s, \\ C_{a,t} &= \left(\frac{\tau_v + \tau V_{a,t+1} + Y_{t+1}}{\tau_v + \tau V_{a,t} + Y_t} \right) \left(\frac{\tau_v + Y_t}{\tau_v + Y_{t+1}} \right) \left(1 + \tau V_{a,t} \left(\frac{1}{\tau_v + Y_t} - \frac{1}{\tau_v + Y_{t+1}} \right) \right), \\ D_{a,t} &= \prod_{s=t+1}^T C_{a,s}^{-1/2}, \end{aligned}$$

with the convention that $A_0 = 0$, and $Y_{T+1} = \infty$. There is a linear equilibrium, in which prices at time t are given by

$$p_t = \frac{\tau_v}{\tau_v + Y_t} \bar{v} + \frac{Y_t}{\tau_v + Y_t} v + \frac{\tau_u}{\tau_v + Y_t} \sum_{s=1}^t (A_s - A_{s-1}) u_s. \quad (4)$$

In equilibrium, agent a 's time- t demand and expected utility, given wealth $W_{a,t}$ and the realization of signals summarized by $z_{a,t}$, take the form

$$x_{a,t} = \frac{\tau V_{a,t}}{\gamma_a} (z_{a,t} - p_t), \quad (5)$$

$$U_{a,t} = -D_{a,t} e^{-\gamma_a W_{a,t} - \frac{1}{2} \frac{\tau^2 V_{a,t}^2}{\tau_v + Y_t + \tau V_{a,t}} (z_{a,t} - p_t)^2}. \quad (6)$$

Several observations are in place. First, note that the price function (4) has a fairly standard structure. It is determined by the fundamental value (v) and the aggregate supply shocks (u_s , $s = 1, \dots, t$). The weights on these different components are determined by how signals spread through the network. Especially, A_t summarizes how aggressive—and thereby informative—the trades of agents are at time t , consisting of a weighted average of t -degree connectivity of all agents. The variable Y_t corresponds to a cumulative average of squared innovations in A up until time t , and determines how much of the fundamental value that is revealed in the price. The main generalization compared with Vives (1995) is that $V_{a,t}$ varies with agent and over time, depending on the network structure. Moreover, preferences are allowed to vary across agents, through γ_a . This allows us to compare the equilibrium dynamics that may arise because of heterogeneous

preferences with the dynamics that may arise because of heterogeneous information diffusion.

It is notable that Y_t does not only depend on the total amount of information that has been diffused at time t , but also on *how* this information has diffused over time. In other words the price at a specific point in time is information path dependent. For example, consider two economies with 4 agents, all with unit risk aversion ($\gamma_a = 1$), and with parameters $\tau = \tau_v = \tau_u = 1$. The first network, shown in panel A of Figure 1, is tight-knit (it is even complete) with every agent being directly connected to every other agent. It is straightforward to calculate $V_{1,a} = V_{2,a} = 4$, $A_1 = A_2 = 4$, $Y_1 = Y_2 = 16$, via (4) leading to $p_2 - v \sim N\left(0, \frac{1}{17}\right)$. The second network, shown in Panel B of Figure 1, is not as tightly knit, and agents have to wait until $t = 2$ before they have received all signals. In the latter case, $V_{1,a} = 3$, $V_{2,a} = 4$, $A_1 = 3$, $A_2 = 4$, $Y_1 = 9$, $Y_2 = 16$, via (4) leading to $p_2 - v \sim N\left(0, \frac{1}{11}\right)$. Thus, the price at $t = 2$ is less revealing in the second case, even though all agents have the same information at $t = 2$ in both economies. The reason is that in the tight-knit economy, the information revelation is more lumpy, whereas it is more gradual in the less tight-knit economy. Lumpy information diffusion leads to more revealing prices, since it generates more aggressive trading behavior in some periods, in turn making it easier to separate informed trading from supply shocks. This is our first example of how the network structure impacts asset price dynamics.

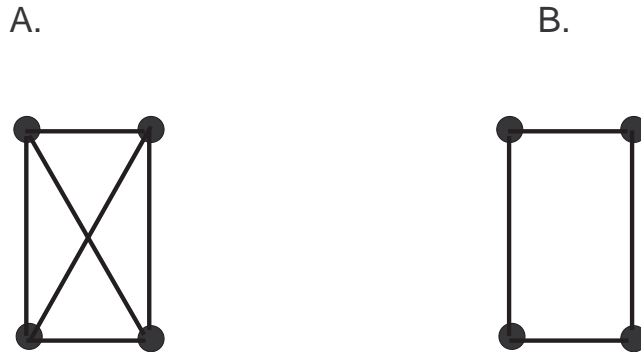


Figure 1: Impact of network structure. *The figure shows two networks with four agents: In Panel A, a tight-knit network is shown, in which every agent is connected with every other agent. In panel B, a less tight-knit network is shown. At $t = 2$, prices are more revealing in the tight-knit network, since the aggregate information arrival has been more lumpy, in turn leading to more revealing trading behavior of informed agents.*

4 Trading, profits, and centrality of individual agents

We study how the trades and performance of agents are determined by their positions in the network.

4.1 Correlation of trades

Feng and Seasholes (2004) studied retail investors in the People’s Republic of China, and found that the geographical position of investors was related to the correlation of their trades: geographically close investors had more positively correlated trades than investors who were farther apart. Colla and Mele (2010) showed that information networks can give rise to such patterns of trade correlations, under the assumption that geographically close agents are also close in the information network. Their analysis was restricted to a cyclical network, but the effect was also shown to arise in general networks in the static model of Ozsoylev and Walden (2011).

In Ozsoylev, Walden, Yavuz, and Bildik (2013), this positive relationship between trades and network position was used to reverse engineer a proxy of the information network in the Istanbul Stock Exchange from individual investor trades. Loosely speaking, agents who repeatedly traded in the same stock, in the same direction, at similar points in time, were assumed to be linked in the market’s information network. Such an approach is justified in the static model of Ozsoylev and Walden (2011), but the situation is more complex in a dynamic setting. Specifically, one may expect a positive relationship between network proximity and portfolio holdings also in the dynamic model, since agents who are close in the network have many overlapping signals and thereby similar information, leading to similar portfolio holdings. However, whereas agents’ trades and portfolio positions are equivalent in the static model, they are not in a dynamic model. Instead, portfolio holdings are equivalent to cumulative period-by-period trades in the dynamics model. For period-by-period trading behavior, the timing of information arrival is also important, and this timing is different even for agents who are close in the network.

Consider, for example, an economy in which one agent, a , is connected to many other agents at a distance t , and therefore receives very precise information about the asset’s value at time t . This agent will at t take a large position in the asset (positive or negative, depending on whether the agent believes that it is under- or over-valued). Moreover, assume that one of agent a ’s neighbors, agent b , does not have many connections within a distance of t but, being connected to a , receives a lot of information at $t + 1$. Now, assume that at $t + 1$ there is also a substantial amount of information incorporated into the asset’s price (driven by many other agents who are also connected to many agents at distance $t + 1$). In this case, agent b will take on a similar position as agent a at time $t + 1$, although less extreme since the price will be closer to the fundamental value at this point. Agent a , however, may actually decrease his position, to realize profits and decrease risk exposure. This argument suggests that the time $t + 1$ trades of the two agents are negatively correlated, although they are neighbors in the network. Thus, the period-by-period relationship between network proximity and correlation of trades seems less clear than that between network proximity and portfolio holdings, suggesting that one needs to be careful in choosing an appropriate window length when inferring network structure from observed trades.

The following theorem characterizes covariances of trades, for an individual agent over time, and between agents at a specific point in time.

Theorem 2 *The covariance between an agent's trade at t and $t + 1$ is*

$$\text{Cov}(\Delta x_{a,t+1}, \Delta x_{a,t}) = \frac{\tau}{\gamma_a} \Delta V_{a,t} \left(\frac{V_{a,t+1}}{\tau_v + Y_{t+1}} - \frac{V_{a,t}}{\tau_v + Y_t} \right). \quad (7)$$

The covariance of agent a and b 's trades at time t is

$$\text{Cov}(\Delta x_{a,t}, \Delta x_{b,t}) = \frac{\tau^2}{\gamma_a \gamma_b} \left(\frac{y_t}{(\tau_v + Y_{t-1})(\tau_v + Y_t)} V_{a,t-1} V_{b,t-1} + \frac{\Delta V_{a,t} \Delta V_{b,t}}{\tau_v + Y_t} + \Delta V_{a,b,t} \right). \quad (8)$$

Equation (7) shows that in the special case when an agent does not receive any new signals (and thus $\Delta V_{a,t} = 0$), the covariance between time t and $t + 1$ trades is also zero. This is natural since the agent's trades in this case depends solely on price changes, and the price process is a martingale. In the more interesting case when the agent does receive signals, the covariance between subsequent trades is determined by the information advantage of the agent over the market at t and $t + 1$, respectively. Specifically, $V_{a,t}$ represents how much information agent a has received at time t , whereas $\tau_v + Y_t$ represents how much aggregate information has been incorporated into prices. A high $V_{a,t}$ relative to $\tau_v + Y_t$, means that the agent has a substantial information advantage at time t . If the information advantage increases between t and $t + 1$, then agent a 's trades will be positively correlated over these two periods, representing a situation where he tends to ramp up investments. If, on the other hand, agent a 's information advantage decreases, his trades will be negatively correlated, representing a situation where he takes home profits and decreases risk exposure by selling stocks (or buying, if in a short position), in line with our previous discussion.

An example of the two different situations is shown in Panels *A* and *B* of Figure 2. In Panel *A*, agent a is at the center of an extended star network and will have a large advantage over the market at $t = 1$. At $t = 2$, the playing field is more even, since information diffusion has made all of agent a 's neighbors well-informed too. Agent a still has an information advantage, now having received all signals from the periphery of the network, but the advantage is lower than in the previous period, and he therefore decreases his asset position, leading to a negative correlation of his trades over the two periods. In Panel *B* of the figure, agent a still receives the same signals period-by-period, but since his neighbors are now more directly connected, his information advantage at $t = 1$ is smaller. In this case, his information advantage may be higher in period 2 than in period 1, so he may tend to gradually ramp up up his position over the two time periods and therefore have positively correlated trades.

We note that this intuition, that relative information advantage over time determines an agent's dynamic trading behavior, which is very clear in the general setting, does not come out clearly if we restrict our attention to symmetric networks. Indeed in a network symmetric economy, $\Delta V_{a,t}$ is the same for all agents, and it can be shown that (7) takes the form

$$\text{Cov}(\Delta x_{a,t+1}, \Delta x_{a,t}) = \frac{\tau}{\gamma_a} \Delta V_{t+1} \Delta V_t \left(\tau_v + Y_t - \frac{\tau}{\bar{\gamma}} V_t \Delta V_{t+1} \right).$$

Thus, all else equal, for low degrees of information diffusion between t and $t + 1$ (i.e., a low ΔV_{t+1}), trades will be positively correlated, whereas they will be negatively correlated when the degree of information diffusion is high (i.e., for a high ΔV_{t+1}). This result is the opposite of what we just showed.

The issue in the symmetric case is that within the class of network symmetric economies, connectivity must increase for all agents at the same time, so there is no way to increase the relative information advantage of one agent. Increasing the connectivity of all agents at the same time has two effects: it increases the total informativeness of their signals and it increases the amount of information that is incorporated into the asset's price, and the second effect always dominates. Thus, increased connectivity at $t + 1$, all else equal, always decreases the information advantage of the agents in a symmetric network setting, making them decrease their portfolio holdings, and thereby potentially leading to negative correlation with their trades in the previous period. This is our first example of how restricting ones' attention to symmetric economies may be misleading—in this case reversing the result.

We next focus on Equation (8), which shows how the time- t trades of two agents are related. The first two terms in the expression represent covariance induced by the fact that two informed agents will tend to trade in the same direction because, both being informed, they will take a similar stand on whether the asset is over-priced or under-priced. This part of the expression increases in the total amount of information the agents have received at $t - 1$ (through $V_{a,t-1}V_{b,t-1}$), as well as in how much additional information they expect to receive between $t - 1$ and t (through $\Delta V_{a,t}\Delta V_{b,t}$). Offsetting these effects is the aggregate informativeness of the market, through the terms $\frac{y_t}{(\tau_v + Y_{t-1})(\tau_v + Y_t)}$ and $\frac{1}{\tau_v + Y_t}$, similarly to what we saw in (7). The third term in the expression provides an additional positive boost to the covariance, and is increasing in the number of common agents at distance t of both agent a and b . This term is zero if the agents are further apart than a distance of $2t$, but will otherwise typically be positive. The term captures the natural intuition that agents who receive identical information signals have more similar trades than agents who receive signals with independent error terms.

The first main implication of (8) is that the covariance is always strictly positive. Thus, the situation with negative correlation between trades—because two nearby agents tend to trade in different directions when one is ramping down portfolio exposure whereas the other one is ramping up—never occurs in the model. To understand why this is the case, we use (5) to rewrite agent a 's time- t demand as

$$\Delta x_{a,t} = \frac{\tau}{\gamma_a} \left(\Delta V_{a,t} \left(\frac{\sum_{j \in \Delta S_{a,t}} s_j}{\Delta V_{a,t}} - p_t \right) - V_{a,t}(p_t - p_{t-1}) \right).$$

The first term in this expression represents the agent's demand because of additional information received between $t - 1$ and t . We note that $\sum_{j \in \Delta S_{a,t}} s_j / \Delta V_{a,t} = v + \zeta^a$, where the error term $\zeta^a \sim N(0, \sigma^2 / \Delta V_{a,t})$ is independent of prices. The second term represents the agent's sloping demand curve, causing him to rebalance portfolio holdings when the price catches up, by selling (buying) stocks when the price increases (decreases) between $t - 1$ and t . For an agent who has

an information advantage at time $t - 1$, but receives no new information between $t - 1$ and t , this second term is the only one present (since $\Delta V_{a,t} = 0$).

Now, agent b 's demand function has the same form as agent a 's and, assuming that agent b receives a lot of new information between $t - 1$ and t , the first term dominates. Negative correlation would then arise if agent b tends to ramp up when agent a ramps down, which is the case if $Cov(v + \zeta^b - p_t, -(p_t - p_{t-1})) < 0$. However, since the market is semi-strong form efficient, $v - p_t$ is independent of $p_t - p_{t-1}$. Furthermore, since ζ^b is independent of aggregate variables, $Cov(\zeta^b, -(p_t - p_{t-1})) = 0$. In other words, since agent a 's rebalancing demand between $t - 1$ and t is publicly known at t , it must be independent of agent b 's time- t demand which in turn is due to informational advantage at time t .

This result is model dependent. It depends on the linear structure of agents' demand functions. However, to a first order approximation we expect the result to hold in more general settings in semi-strong form efficient markets, given that trading for rebalancing purposes mainly depends on price changes, trading for informational purposes depends on the difference between the true value and market price, and the two terms are uncorrelated in a weak-form efficient market.

The second main implication of (8) is that, all else equal, period-by-period covariances (and correlations) between two agents' trades increase in the number of common acquaintances they have (through the $\Delta V_{a,a',t}$ -term). As an example, in Panel *C* of Figure 2, we show a cyclical network. It immediately follows that the correlation between two agents' trades at any time $t < 4$ (at which point all information has reached all agents) is decreasing in the distance between the two agents.

These properties of trade correlations suggest that it may be justified to use trades instead of portfolio holdings to draw inferences about a market's information network as, e.g., done in Ozsoylev, Walden, Yavuz, and Bildik (2013), using short time horizons over which trades are compared.

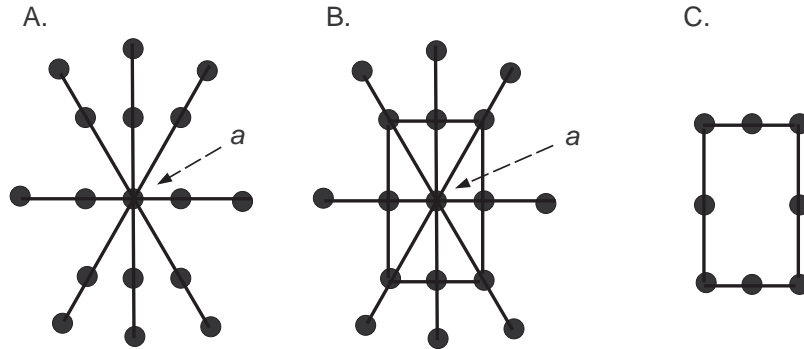


Figure 2: Correlation of trades. The figure shows three networks that lead to different types of trading behavior. Panel A shows a network in which agent a 's trades at time 1 and 2 are negatively correlated, whereas they are positively correlated in Panel B. Panel C shows a network in which the correlation between two agents' trades, at each point in time, is inversely related to the distance between the two agents.

4.2 Profits

Who makes profits in an information network? Our starting point is the following theorem:

Theorem 3 *Define*

$$\begin{aligned}\pi_{a,t} &= \tau V_{a,t} \left(\frac{1}{\tau_v + Y_t} - \frac{1}{\tau_v + Y_{t+1}} \right), & t = 1, \dots, T-1, a = 1, \dots, N \\ \pi_{a,T} &= \tau \frac{V_{a,T}}{\tau_v + Y_T}, & a = 1, \dots, N.\end{aligned}$$

The ex ante certainty equivalent of agent a is

$$U_a = \frac{1}{2\gamma_a} \log(C), \quad \text{where } C = \prod_{t=1}^T (1 + \pi_{a,t}). \quad (9)$$

The expected profit of agent a is

$$\frac{\tau}{\gamma_a} \Pi_a, \quad (10)$$

where

$$\Pi_a = \sum_{t=1}^T (\tau_v + Y_t)^{-1} V_{a,t} \quad (11)$$

is the profitability of agent a .

We focus on profitability, and leave welfare implications of the model for future research. Equation (10) determines (ex ante) expected profits of an agent. It shows that expected profits depends on three components. First, profits are inversely proportional to an agent's risk-aversion, γ_a , because more risk-averse agents take on smaller positions—all else equal. This follows immediately, since an agent's equilibrium trading position is proportional to γ_a , so it corresponds to pure scaling. Therefore, we do not include it in our measure of profitability, as defined by Equation (11). Neither do we include the signal precision, τ , which is constant across agents. Second, expected profits depend on an agent's position in the network through $\{V_{a,t}\}_t, t \in \mathcal{T}$: the higher any given $V_{a,t}$ is, the higher the agent's expected profits. Third, expected profits depend inversely on the amount of aggregate information available in the market, in that at any given point in time, the higher the total amount of aggregate information, the lower the expected profits of any given agent. The third part represents a negative externality of information. Equation (11) thus provides a direct relationship between the properties of a network, local as well as aggregate, and individual agents' profitability.

4.3 Centrality

Equation (11) shows that an agent's profitability is determined by his *centrality*, defined appropriately. Recall that $V_{a,t}$ denotes the number of agents that are within distance t from agent a . So, $V_{a,1}$ is simply the degree of agent a . For $t > 1$, higher order connections are also important in determining profitability. For example, $V_{a,2}$ does not only depend on how connected agent a is, but also on how connected his neighbors are. We use (1) to rewrite (11) on vector form as

$$\Pi = \sum_{t=1}^{\infty} \beta_t \chi(E^t) \mathbf{1}, \quad (12)$$

where $\beta_t = (\tau_v + Y_t)^{-1}$, for $t \in \mathcal{T}$, and $\beta_t = 0$ for $t > T$. Here, Π is an N -vector where the a th element is the profitability of agent a .

We explore the relationship between this profitability measure and standard centrality measures in the literature. Specifically, we define degree centrality, farness, closeness centrality, Katz centrality, and eigenvector centrality (see, e.g., Friedkin (1991) for a detailed discussion of these concepts), and compare (12) with these measures.

Definition 3

- The degree centrality vector is the vector of first-order degrees, V^1 , i.e., the degree centrality of agent a is $V_{a,1}$.
- The farness of agent a is the agent's average distance to all other agents, i.e.,

$$F_a = \frac{\sum_{a' \neq a} D(a, a')}{N - 1}.$$

- The closeness centrality of agent a , \hat{C}_a , is the inverse of that agent's farness,

$$\hat{C}_a = \frac{1}{F_a}.$$

- The Katz centrality vector with parameter $\alpha < 1$ is the vector $K \in \mathbb{R}_+^N$, defined as

$$K = K^\alpha = \sum_{t=1}^{\infty} \alpha^t E^t \mathbf{1}. \quad (13)$$

Here, E^t denotes the t :th power of the adjacency matrix, E , and $\mathbf{1} \in \mathbb{R}^N$ is an N -vector of ones.

- The eigenvector centrality vector is the eigenvector corresponding to the largest eigenvalue of E , i.e., the vector C that solves the equation $C = \lambda EC$, for the largest possible eigenvalue, λ , where we normalize C such that $\sum_{a \in \mathcal{N}} C_a = 1$.

We see that the structures of the profitability measure (12) and Katz centrality (13) are similar. Specifically they are both made up by a weighted sum of powers of the adjacency matrix, multiplied with the vector of ones. The differences are that the weighting is a power of α for Katz centrality but varies more generally with t for profitability, and that the matrix indicator function, χ , operates on the power of the adjacency matrix in the profitability measure. It is a standard result that eigenvector centrality can be viewed as a special case of Katz centrality, since $C = \lim_{\alpha \nearrow \lambda^{-1}} \frac{K^\alpha}{\sum_a K_a^\alpha}$.⁷

The similar structure of the formulas for profitability (12) and Katz centrality (13) suggests that Katz centrality is closely related to profitability, as is eigenvector centrality being a special case of Katz centrality. Therefore, we may expect Katz and eigenvector centrality to dominate the other measures for “an average” network, although there may be exceptions.

To formalize this intuition, we introduce a random network generating process, which allows us to derive a probabilistic ranking of the different measures. We use the classical concept of random graphs, originally studied in Erdős (1947), Erdős and Rényi (1959), and Gilbert (1959), in which links between agents are formed randomly, independently, and with constant probability.⁸

Our focus is on sparse networks—in line with what is observed in practice (see the discussion in Ozsoylev, Walden, Yavuz, and Bildik (2013)). We assume that the expected number of links per agent in a network of size N is $c \log(N)^k + 1$ for some $c > 0$ and $k > 3$.⁹ With this assumption, the connectedness of agents grows with N , but the fraction of expected connections in the network (which is approximately $cN \log(N)^k / 2$) to maximum number of connections (which is approximately $N^2 / 2$) tends to zero for large N , so when N is large the network is indeed sparse. To this end, we define

Definition 4 *The Erdős-Rényi random graph model of size N , and parameters $c > 0$, and $k > 3$, $G(N, c, k)$, is defined so that for each $1 \leq a \leq N$, $1 \leq a' \leq N$, $a \neq a'$, $(a, a') \in \mathcal{E}$ with i.i.d. probability $c \frac{\log(N)^k}{N-1}$.*

We use cross sectional correlation across agents as the metric of similarity.¹⁰ Specifically, for a given network we define the cross sectional correlation between the different centrality measures and profitability,

$$\rho_D = \text{Corr}(D, \Pi), \rho_F = \text{Corr}(-F, \Pi), \rho_{\hat{C}} = \text{Corr}(\hat{C}, \Pi), \rho_{K^\alpha} = \text{Corr}(K^\alpha, \Pi), \rho_C = \text{Corr}(C, \Pi).$$

⁷Uniqueness of the eigenvector centrality measure is not guaranteed, but is almost never an issue in practice.

⁸We focus on the Erdős-Rényi model, since it is the most parsimonious and analytically most tractable. Many other network generating processes have also been suggested in the literature, e.g., the preferential attachment model of Barabasi and Albert (1999), and the model of Watts and Strogatz (1998).

⁹The restriction $k > 3$ is needed for technical reasons.

¹⁰The cross sectional statistics are defined as $EX = \frac{1}{N} \sum_{a \in \mathcal{N}} X_a$, $\sigma^2(X) = E[(X - EX)^2]$, $\text{Cov}(X, Y) = E[(X - EX)(Y - EY)]$, and $\text{Corr}(X, Y) = \text{Cov}(X, Y) / (\sigma(X)\sigma(Y))$.

Here, we use the convention that the correlation between any random variable and a constant is zero. In the random network model, these cross sectional correlations are in turn random, since they depend on the random realization of the network. Note that ρ_F is defined as the correlation between negative farness and profits, since farness is inversely related to connectedness.

For simplicity, we focus on preference symmetric economies, but allow for $\gamma, \tau, \tau_u, \tau_v, \bar{v}$, and T to be arbitrary. The following result shows that Katz centrality dominates degree centrality, farness, and closeness centrality for large random graphs.

Theorem 4 *For any given $c > 0$ and $k > 3$, there is an N_0 , such that for all networks of size $N > N_0$ in the $G(N, c, k)$ model, there is an α , such that*

$$E[\rho_{K^\alpha}] > E[\rho_D] > \max(E[\rho_{\hat{C}}], E[\rho_F]). \quad (14)$$

Thus, for large random networks, profitability is best characterized by Katz centrality. This result provides a strong ranking between different centrality measures for a large class of networks. To the best of our knowledge, this is the first such ranking of centrality measures in equilibrium models of asset pricing with general networks. In contrast, previous applications of network theory to asset pricing typically provide results for very specific networks (e.g., Ozsoylev 2005, Colla and Mele 2010, and Buraschi and Porchia 2012), or is empirically motivated (e.g., Das and Sisk 2005, Adamic, Brunetti, Harris, and Kirilenko 2010, Li and Schurhoff 2012, and Ozsoylev, Walden, Yavuz, and Bildik 2013,). The challenge in proving the result stems from the fact that profitability is determined endogenously in equilibrium.

The intuition behind the result is straightforward, as shown in the proof of the theorem, suggesting that it may extend to more general settings. The profitability equation (12) is made up by a weighted average of the connectedness of different orders (with the weights, β_t , endogenously determined). Degree centrality exclusively focuses on first-order connections, whereas closeness centrality and farness mainly focus on high-order connections (since the vast majority of agents will be quite far away from any given agent in a large network). Katz and eigenvector centrality, in contrast, balance the weights of different orders of connectedness, and are thereby more closely related to profitability.

Of course, our strong ranking of centrality measures holds only for large networks. We also explore the relationship between the different centrality measures in medium-sized networks using simulations, to verify that Katz centrality also works well in such networks. We randomly simulate 1,000 economies of sizes $N = 50, 100, 200, 500,$ and 1000 , respectively, and in each economy we randomly generate links between agents, using the random graph model, so that each agent is expected to have \sqrt{N} links. The growth of the number of links in the network with N is thus slightly faster than in the $G(N, c, k)$ model, although the graph is still asymptotically sparse.

We measure the correlation between profitability (II) and the different centrality measures. The results are shown in Table 1. We see that the ranking is the same as in Theorem 4 and,

Size of network, N	C	K^α	V^1	\hat{C}	F
50	0.96	0.96	0.88	0.0	0.0
100	0.94	0.94	0.82	0.0	0.0
200	0.94	0.94	0.79	0.0	0.0
500	0.95	0.95	0.83	0.0	0.0
1000	0.97	0.97	0.85	0.0	0.0
Mean correlation	0.95	0.95	0.83	0.0	0.0

Table 1: Centrality Measures. The table shows the correlation between profitability, Π , and centrality for several different centrality measures, degree centrality (V^1), eigenvector centrality (C), Katz centrality (K^α), closeness centrality (\hat{C}) and farness (F). The parameters of the economy are $\sigma = \sigma_u = \sigma_v = 10$, $\bar{v} = 0$, $T = 5$, and $\gamma = 1$ for all agents. The number of agents is varied between $N = 50$ and $N = 1000$, and links are randomly drawn between agents, such that on average an agent has \sqrt{N} links. For each N , we simulate 1,000 random networks. The results show that eigenvector and Katz centrality are most closely related to profitability, followed by degree and betweenness centrality. Closeness centrality and farness perform poorly, since there is usually an isolated agent in the network, leading to all agents having closeness centrality of zero and infinite farness centrality.

furthermore, that eigenvector centrality, being a special case of Katz centrality, performs about as well as the Katz centrality measure.

5 Aggregate volatility and trading volume

According to the Mixture of Distributions Hypothesis (MDH), return volatility varies over time, which then leads to heavy-tailed unconditional return distributions. A common explanation for such time varying volatility—as well as trading volume—is lumpy diffusion of information into the market (Clark 1973; Epps and Epps 1976; Andersen 1996).

Rich dynamics of volatility and trading volume have indeed been documented in the literature. First, trading volume is positively autocorrelated over extended periods, i.e., the autocorrelation function of trading volume has “long memory” in that it decreases very slowly (see Bollerslev and Jubinski 1999 and Lobato and Velasco 2000). Roughly speaking, this means that abnormally high trading volume one period predicts abnormally high trading volume for many future periods. Second, return volatility of individual stocks, and markets, also have long memory (e.g., Bollerslev and Jubinski 1999 and Lobato and Velasco 2000). Third, volume and volatility are related (Karpoff 1987; Crouch 1975, Rogalski 1978), in line with the Wall Street wisdom that it takes volume to move markets. Contemporaneously, trading volume and absolute price change are highly positively correlated. The two series also have positive lagged cross-correlations. Using a semi-parametric approach to study returns and trading volume on the NYSE, Gallant, Rossi, and Tauchen (1992) show that large price movements predict large trading volume. In other markets, there is evidence for a reverse casualty, i.e., that large trading volume leads large price movements. For example, Saatcioglu and Starks (1998) find such evidence in several Latin American equity markets.

In our model, agents' preferences (Γ) and the network structure (\mathcal{E}) determine volatility and volume dynamics in the market. It is then natural to ask which type of dynamics can be generated within the model for general Γ and \mathcal{E} , and also what one can infer about the network structure from observed volatility and volume dynamics.

5.1 Volatility

In a general network economy, we would expect price volatility to vary substantially over time. For example, information diffusion may initially be quite limited, with low price volatility as an effect, but eventually reach a hub in the network, at which point substantial information revelation occurs with associated high price volatility. The following result characterizes the price volatility over time, and moreover shows that any volatility structure can be supported in a general economy.

Theorem 5 *For $t = 1, \dots, T$, the variance of prices between $t - 1$ and t , is*

$$\sigma_{p,t}^2 = \frac{y_t}{(\tau_v + Y_t)(\tau_v + Y_{t-1})}, \quad (15)$$

where we use the convention that $Y_0 = 0$, and between T and $T + 1$, it is

$$\sigma_{p,T+1}^2 = \frac{1}{\tau_v + Y_T}. \quad (16)$$

Moreover, given coefficients, k_1, \dots, k_{T+1} , such that $k_t > 0$, and $\sum_{t=1}^{T+1} k_t = 1$ and an arbitrarily small $\epsilon > 0$, there is a preference symmetric economy, such that

$$\left| \sigma_{p,t}^2 - \frac{k_t}{\tau_v} \right| \leq \epsilon, \quad t = 1, \dots, T + 1. \quad (17)$$

From the first part of the theorem, we see that the volatility (the square root of variance) has a general decreasing trend over time because of the increasing denominator in (15), but that it can still have spikes in some time periods because of large values in the numerator. In fact, (17) shows that any structure of time-varying volatility after an information shock can be generated.

Note that the total cumulative variance up until time $t \leq T$ is

$$\sigma_{P,t}^2 = \frac{1}{\tau_v} \frac{Y_t}{\tau_v + Y_t}. \quad (18)$$

This part of the variance that is incorporated until T represents the “information diffusion” component of asset dynamics, whereas the part between T and $T + 1$, i.e., (16), represents the component due to public information sources, in line with the discussion in Section 3.2. The

total price variance between 0 and $T + 1$ is of course equal to the ex ante variance,

$$\sigma_v^2 = \sigma_{P,T}^2 + \sigma_{p,T+1}^2 = \frac{1}{\tau_v} \frac{Y_T}{\tau_v + Y_T} + \frac{1}{\tau_v + Y_T},$$

independently of network structure. But the way the variance is divided period-by-period, and into the information diffusion and public component, depends on the network. It is specifically determined by y_t , $t = 1, \dots, T$.

If we restrict our attention to network symmetric economies, the possible dynamic behavior of volatility is much more restricted. This is not surprising, given the restrictions on information diffusion dynamics described in Lemmas 1 and 2. From (18), it follows that y_t is proportional to $\Delta\eta_t = \eta_t - \eta_{t-1}$, where $\eta_t = \frac{1}{\sigma_v^2 - \sigma_{P,t}^2}$. Since y_t is proportional to ΔV_t^2 (which is the same for all agents in a network symmetric economy), ΔV_t is unimodal, and the square of a nonnegative unimodal function is also unimodal, it follows that the sequence $\Delta\eta$ is also unimodal. Moreover, if the public information component is large compared with the diffusion component, $\sigma_{P,T}^2 \ll \sigma_v^2$, it follows from (18) that

$$\sigma_v^4 \Delta\eta_t \approx \sigma_{p,t}^2, \tag{19}$$

so in this case $\sigma_{p,t}$ is also unimodal. We summarize this in

Corollary 1 *In a network symmetric economy,*

1. $\Delta\eta_t$ is unimodal,
2. if the public information component is high, $\sigma_{P,T}^2 \ll \sigma_v^2$, then $\sigma_{p,t}$ is unimodal.

Of course, the timing of an information shock is typically not known. We have in mind a situation in which information shocks arrive every now and again, and are then gradually incorporated into prices. We do not observe the timing of these shocks, and therefore do not know when $t = 0$ in our model.

However, autocorrelations — which take averages over all time periods — can be calculated even without observing the timing of shocks. Now, there is a close relationship between a function's shape and the shape of its autocorrelation function. For a general sequence, f_0, f_1, \dots, f_T , we define the difference operator $(\Delta f)_k = f_k - f_{k-1}$, $k = 1, \dots, T$, $(\Delta f)_{T+1} = 0 - f_T$, and the higher-order operators $\Delta^{n+1} f = \Delta(\Delta^n f)$. The following lemma relates the monotonicity of a function and its autocorrelation function:

Lemma 3 *Consider a nonnegative sequence f_0, f_1, \dots, f_T , such that $f_0 = 0$. Define the autocorrelation function $R_\tau = \sum_{k=0}^{T-k} f_k f_{k+|\tau|}$, $-T + 1 \leq \tau \leq T - 1$. Then,*

1. if f is unimodal, so is R ,
2. if $\Delta^n f$ does not switch sign, then R has at most $n - 1$ turning points.

In light of Corollary 1, the result immediately leads to:

Corollary 2 *In a network symmetric economy with a high public information component, the autocorrelation function of volatility is unimodal.*

Thus, given the restrictions imposed by network symmetry, nonmonotonicity of the autocorrelation function of volatility suggests that there is asymmetry in the underlying information network.

One way of measuring the degree of monotonicity of an autocorrelation function is by counting its number of turning points for positive τ (since the autocorrelation function is symmetric, it is sufficient to study $\tau > 0$). Another way is to measure its absolute variation, $\sum_{\tau>0} |R_{\tau+1} - R_\tau|$, which will be higher the more nonmonotone the autocorrelation function.

To summarize, network symmetry leads to hump-shaped volatility after an information shock, and also to hump-shaped autocorrelation functions, whereas any dynamics can arise in a preference symmetric economy.

5.2 Volume

Just like with volatility, rich dynamics of trading volume can arise within the network model. Since the model is inherently asymmetric, aggregate trading volume will be made by the heterogeneous trades of many different agents. This contrasts to the uniform behavior in models with a representative informed agent (e.g., Kyle 1985), as well as to the ex ante symmetric behavior in economies with symmetric information structures (e.g., Vives 1995; He and Wang 1995).

We focus on the aggregate period-by-period trading volume of agents in the network, since the stochastic supply is (quite trivially) normally distributed. To this end, we define:

Definition 5 *The time- t aggregate (realized) trading volume is $W_t = \frac{1}{N} \sum_a |\Delta x_{a,t}|$, and the (ex ante) expected trading volume is $X_t = E[W_t]$.*

The following theorem characterizes the expected trading volume, and mirrors our volatility results by showing that any pattern of expected trading volume can be supported in the model.

Theorem 6 *The time- t expected trading volume is*

$$X_t = \frac{\tau}{N} \sum_{a=1}^N \frac{1}{\gamma_a} \sqrt{\frac{2}{\pi} \left(\frac{V_{a,t-1}^2}{\tau_v + Y_{t-1}} - \frac{V_{a,t}^2}{\tau_v + Y_t} + \frac{\Delta V_{a,t}^2}{\tau_v + Y_t} + \frac{\Delta V_{a,t}}{\tau} \right)}. \quad (20)$$

Given positive coefficients, c_1, c_2, \dots, c_{T+1} , and any $\epsilon > 0$, there is an economy such that

$$|X_t - c_t| \leq \epsilon, \quad t = 1, \dots, T + 1.$$

It is clear from (20) that, as is the case for volatility, heterogeneous preferences alone cannot generate such complete generality of trading volume dynamics. In a network symmetric economy, all terms under the square root are identical across agents, and (20) collapses to

$$X_t = \sqrt{\frac{2\tau^2}{\pi\bar{\gamma}^2} \left(\frac{V_{t-1}^2}{\tau_v + Y_{t-1}} - \frac{V_t^2}{\tau_v + Y_t} + \frac{\Delta V_t^2}{\tau_v + Y_t} + \frac{\Delta V_t}{\tau} \right)}. \quad (21)$$

Again, differences in preferences in this case are only important through the effect they have on the (harmonic) average risk aversion coefficient, $\bar{\gamma}$. Moreover, the restrictions on ΔV_t imposed by network symmetry carry over to trading volume. It is easily seen that if the public information component is large compared with the diffusion component, the fourth term under the square root in (21) dominates and $X_t \approx \sqrt{\frac{2\tau^2}{\pi\bar{\gamma}^2} \Delta V_t}$. In this case, we therefore get:

Corollary 3 *In a network symmetric economy, in which the public information component is high:*

- X_t is unimodal, as is its autocorrelation function,
- There is an approximate square root relationship between $\sigma_{p,t}$ and X_t :

$$X_t \sim \sqrt{\sigma_{p,t}}. \quad (22)$$

To summarize, the degree of nonmonotonicity of the autocorrelation functions of volatility and trading volume are related to the degree of asymmetry in the information diffusion process. In the network symmetric case, after an information shock trading volume and price volatility increase, reach a peak and then decrease. In the asymmetric case, the dynamics of volume and volatility after an information shock may be nonmonotone. Also, the instantaneous correlation of volatility and volume is high in the symmetric case.

In the completely general case, with heterogeneity over both preferences (γ_a) and network structure ($V_{a,t}$), we expect the interplay between the two to give rise to quite arbitrary trading volume and volatility dynamics. For example, at a large time, t , almost all information may have diffused among the bulk of agents, leading to a small and decreasing $\Delta V_{a,t}$, and thereby low volatility. A peripheral agent with very low risk aversion, who receives many signals very late, may still generate large trading volume at such a late point in time, despite the low volatility.

The above example captures the important distinction between trading volume driven by high aggregate information diffusion, and by demand from agents with low risk aversion, a

distinction that does not arise in either the preference symmetric or the network symmetric benchmark cases.

5.3 Conditioning on event time

In the previous analysis we studied unconditional relations, assuming that the arrival time of information shocks is unknown. In some cases the arrival time may be observable. Earnings announcements may, for example, constitute well defined information events for which we could use such conditioning. Although the actual announcement is a public event, if agents have additional private information, the inferences they draw from the announcement together with their private information may differ.

If the time of arrival of the information shock is well-identified, we can condition on the time that has passed since arrival and calculate conditional relations between returns and volume. Specifically, we can use t in our information set when calculating relations, e.g., defining $Cov(W_t, W_{t+1}) = E[(W_t - E[W_t])(W_{t+1} - E[W_{t+1}])|t]$. We define the price change, $\mu_t = p_t - p_{t-1}$, and we then have

Theorem 7 *Trading volume and price changes satisfy the following conditional relations:*

- $Corr(W_t, |\mu_t|) > 0$,
- $Corr(W_t, W_{t-1}) > 0$,
- W_t is independent of μ_s for $s < t$.

The first two results state that trading volume is contemporaneously positively related to absolute price changes, as well as positively autocorrelated. These results are in line with the previously discussed empirical literature. The third result states that price movements do not cause lagged trading volume. The result seems to be inconsistent with what is found in Gallant, Rossi, and Tauchen (1992), but consistent with the results in Saatcioglu and Starks (1998). We emphasize though that to test this prediction, one needs to condition on the time elapsed after an information event, which these studies do not do. It is therefore an open question whether the prediction holds empirically.

5.4 Size versus volatility and volume

The results in Sections 5.1 and 5.2 can potentially be used to understand variations in dynamics across different markets and assets. One such variation is between companies of different size, where we may expect more transparent information diffusion for large than for small companies, because of the higher coverage of large companies. The underlying information network structure for large companies may therefore be relatively symmetric, and the public information

component may be large. In contrast, the network structure of smaller companies may be more asymmetric, and the nonpublic information component more significant.

We carry out initial tests of whether the observed dynamics of volatility and volume in the market is consistent with such differences between large and small companies. The tests are suggestive and exploratory. A more extensive empirical investigation, e.g., in the form of reverse engineering the underlying network using maximum likelihood estimation, is outside of the scope of this paper.

5.4.1 Data

Using the Center for Research in Security Prices (CRSP), we collect daily returns, market value, and trading volume over a ten-year period (January 2003-December 2012), for all companies in the Russell 3000 Index. We classify the companies as Large Cap (market capitalization above USD 10 Billion), Mid Cap (market capitalization between USD 1 Billion and 10 Billion), and Small Cap (market capitalization below USD 1 Billion). We require data for at least 98% of the days in the sample (2,466 of the 2,517 trading days) to be available, to include a company in the sample, leaving us with 276 Large Cap, 746 Mid Cap, and 502 Small Cap companies, in total $J = 1524$ companies.

We calculate weekly return volatility and turnover for each stock in the sample, which we then use to calculate weekly autocorrelations of volatility, $R_{j,t}^\sigma$, and turnover, $R_{j,t}^W$, for each stock (j), and lag $t = 1, 2, \dots, 12$ weeks. We use turnover instead of raw volume, to get a normalized trading volume measure.

5.4.2 Results

We calculate the average number of turning points of the autocorrelation functions for volatility and turnover, in the three groups. The results are shown in Table 2. We see that the number of turning points is larger for smaller firms, both for autocorrelation of volatility and of turnover. The differences are statistically significant at the 2.5% level (t -statistic of 2.02, using a one-sided test) when testing for differences of means between Mid Cap and Small Cap companies, whereas it is significant at the 0.1% level for the other three tests (Large Cap versus Mid Cap, and Mid Cap versus Small Cap for volatility, and Large Cap versus Mid Cap for turnover). This is in line with our analysis, that large companies have more symmetric information networks with a larger public component than small companies.

A similar picture emerges for absolute variation of the autocorrelation functions of volatility and volume, for firm j defined as

$$\sum_{t=1}^{12} |R_{j,t}^\sigma - R_{j,t-1}^\sigma| \quad \text{and} \quad \sum_{t=1}^{12} |R_{j,t}^W - R_{j,t-1}^W|,$$

respectively. As seen in Table 2, the average absolute variation is higher for smaller companies, with similar statistical significance as for the number of turning points. Since absolute variation

	Large Cap	Mid Cap	Small Cap
A. Volatility, $R_{j,t}^\sigma$			
Average number of turning points	3.63	4.08	4.46
	(3.76)	(3.81)	
Average absolute variation	0.88	0.92	0.99
	(2.03)	(5.67)	
B. Turnover, $R_{j,t}^W$			
Average number of turning points	3.70	4.18	4.37
	(3.74)	(2.02)	
Average absolute variation	0.71	0.77	0.85
	(3.90)	(6.70)	
C. Correlation	0.57	0.53	0.41
	(-4.00)	(-14.1)	
Number of firms	276	746	502

Table 2: Company size versus dynamics of volatility and turnover. Panel A shows how the autocorrelation function of volatility, $R_{j,t}^\sigma$, is related to company size. Companies are divided into three groups: Large Cap (market capitalization above USD 10 Billion), Mid Cap (market capitalization between USD 1 and 10 Billion), and Small Cap (market capitalization below USD 1 Billion). Average number of turning points, and average absolute variation are shown in rows 1 and 3, respectively. The t -statistic of a difference of means test between number of turning points of Large Cap and Mid Cap companies, and between Mid Cap and Small Cap companies, are shown in columns 2 and 4 of row 2, respectively, and the same difference in means tests for absolute variation in rows 2 and 4 of row 4. Panel B shows the same for the autocorrelation function of turnover. Panel C shows the average correlation between volatility and turnover for the three groups of companies. Columns 2 and 4 show the t -statistics of difference in means test between Large and Mid Cap companies, and between Mid and Small Cap companies, respectively.

provides another measure of nonmonotonicity, this reinforces the view that smaller stocks are associated with more asymmetric information networks than large stocks.

Finally, the cross correlation between volatility and turnover is higher for Large Cap companies than for Small Cap companies (0.57 for Large Cap versus 0.41 for Small Cap, with differences in means strongly statistically significant both between Large Cap and Mid Cap, and between Mid Cap and Small Cap). We argued earlier that in a symmetric network with a large public information component, there will be a close instantaneous link between volatility and volume, so this result is also consistent with less asymmetric networks for large stocks.

One potential issue with the test above is that trading volume is known to be non-stationary, increasing over time (see, e.g., Lo and Wang 2000). Although we use turnover in our test, which did not seem to have an identifiable trend during the sample period (the average weekly turnover was 2.8% in the first year, 2003, peaked around 10% at the height of the financial crises, 2008, and was then down to 3.4% in the last year of the period, 2012), we wish to control for trends in our test, for the sake of robustness. We therefore carry out the same tests as in Table 2, but normalize volatility and turnover by dividing these variables by the average volatility and turnover across all stocks, in any given week. Thus, the sum of normalized volatilities and turnovers in any given week is equal to one. The results (not reported) are virtually identical to those in Table 2, so non-stationarity does not seem to be driving our results.

Another implication of the model is that in network symmetric economies with large public

information components, expected trading volume scales as the square root of volatility (22). In log-coordinates, we can write

$$w_{j,t} = a_j + b_j s_{j,t} + \epsilon_{j,t}, \quad (23)$$

where $w_{j,t} = \log(W_t)$ is the logarithm of trading volume, and $s_{j,t} = \log(\sigma_{p,t})$ is the logarithm of volatility, for firm j at time t , $\epsilon_{j,t}$ are random shocks, and the coefficient $b_j = 0.5$ for firms with symmetric networks and large public information components. We investigate this relation.

For each firm, we regress log-turnover on log-volatility, and we then calculate the average b_j coefficient for firms in the three different groups. The results are shown in Table 3, Panel A. We see that the average coefficients are lower than 0.5, although quite close, for all groups. They are about 0.43 for the Large Cap and Medium Cap groups, and 0.38 for the Small Cap group. This suggests that the relationship between volatility and volume is farthest away from a square root relationship for stocks in the Small Cap group, again consistent with more asymmetry in this group. The difference in means between Mid Cap and Small Cap companies is significant, with a t -stat of -5.67.

The regressions in (23) do not separate between aggregate shocks to volatility and volume, and company specific shocks. If shocks are mainly aggregate, a cross sectional comparison between companies will allow for only limited inferences. We can see how well volatility and turnover are related in the aggregate market, by regressing log-average turnover, $\bar{w}_t = \log\left(\frac{1}{J} \sum_j W_t(j)\right)$, on log-average volatility, $\bar{s}_t = \log\left(\frac{1}{J} \sum_j \sigma_{p,t}(j)\right)$, i.e.,

$$\bar{w}_t = \bar{a} + \bar{b}\bar{s}_t + \epsilon'_t.$$

The \bar{b} coefficient in this regression is insignificantly different from 0.5 (see Panel C of Table 3), suggesting that the square root relationship holds well at the market level.

To focus on excess turnover and volatility, above and beyond market shocks, we regress

$$w_{j,t} - \bar{w}_t = a_j^e + b_j^e(s_{j,t} - \bar{s}_t) + \epsilon_{j,t}^e.$$

The average b_j^e coefficients for companies in the three groups are shown in Panel B of Table 3. The coefficients are lower for all groups compared with those in Panel A—between 0.32 and 0.35—as are the R -squares, suggesting a significant market component that generates both volatility and volume for individual stocks. The lowest average coefficient is again for companies in the Small Cap group, significantly lower than the average of the Mid Cap companies (t -statistic of -2.51). However, the average coefficient for the Large Cap companies is in this case lower than for the Mid Cap companies. Altogether, the results are consistent with the following story: Large companies have symmetric information networks, and the type of information that is relevant for these companies is also closely related to aggregate market information, causing their excess regression coefficient to be low. Mid-sized companies also have symmetric networks, but their information is not as closely related to market information, i.e., it is more idiosyncratic. Finally,

	Large Cap	Mid Cap	Small Cap	Market
A. Log-regression				
Average b	0.43 (0.009)	0.43 (0.005)	0.38 (0.008)	
	(0.35)		(-5.67)	
Average R^2	0.30	0.25	0.15	
B. Excess log-regression				
Average b^e	0.33 (0.011)	0.35 (0.0065)	0.32 (0.0096)	
	(1.64)		(-2.51)	
Average R^2	0.17	0.15	0.11	
C. Market regression				
Market \bar{b}				0.51 (0.022)
R^2				0.52
D. Correlations				
Average $\rho(\bar{s}, s_j)$	0.63	0.60	0.54	
Average $\rho(\bar{w}, w_j)$	0.60	0.54	0.41	

Table 3: Relationship between log-turnover and log-volatility. Panel A shows how the regression coefficient of log-turnover on log-volatility is related to company size. Companies are divided into three groups: Large Cap (market capitalization above USD 10 Billion), Mid Cap (market capitalization between USD 1 and 10 Billion), and Small Cap (market capitalization below USD 1 Billion). Average regression coefficients are shown in row 1, and standard deviations in row 2. The t -statistic of difference of means tests between regression coefficients of Large Cap and Mid Cap companies, and between Mid Cap and Small Cap companies, are shown in columns 2 and 4 of row 3, respectively. The average R-squares are shown in Row 4. Panel B shows the same for excess log-turnover (defined as $w_{j,t} - \bar{w}_t$) regressed on excess log-volatility (defined as $s_{j,t} - \bar{s}_t$). Panel C shows the coefficient when mean log turnover for all firms in the sample, \bar{w}_t , is regressed on mean log volatility, \bar{s}_t . Panel D shows the average (time series) correlations between mean log-volatility and individual company volatility (row 1), and between mean log-turnover and individual company log-turnover (row 2), for the three groups.

small companies have asymmetric networks, and information about these companies is even more idiosyncratic.

The correlations between company volatility and turnover are indeed different for companies of different sizes, as shown in Panel D of Table 3. The panel shows the average (time series) correlation between market and individual firm log-turnover, $Corr(\bar{w}, w_j)$, and between market and individual firm volatility, $Corr(\bar{s}, s_j)$, for the three groups. The average correlations are lower for smaller than for larger companies.

Thus, altogether our tests suggest that price and volume dynamics of small stocks may be especially well explained by information diffusion through asymmetric networks.

6 Concluding remarks

We have introduced a general network model of a financial market with decentralized information diffusion, allowing us to study the effects of heterogeneous preferences and asymmetric diffusion of information among investors, and the interplay between the two. At the individual investor

level, our results show that the trading behavior of investors is closely related to their positions in the network: Closer agents have more positively correlated trades even over short time periods, and standard network centrality measures are closely related to agents' profits.

At the aggregate level, the dynamics of a market's volatility and trading volume is related to—and could therefore be used to draw inferences about—the underlying information network. In an initial empirical investigation, we show that the dynamics for large stocks are consistent with symmetric information networks and large public information components, whereas the dynamics of smaller stocks suggest asymmetric information diffusion through nonpublic channels. We leave to future research to shed further light on the underlying information network in the market.

Another future line of research may be the study of endogenous network formation in financial markets. Given the strong characterization of individual agents' welfare, the model may be extended to include a period before signals are received and trading occurs, during which agents form connections in anticipation of the future value these will generate. It is an open question what type of networks will form in equilibrium in such an extension, and how well those networks could match observed dynamics at the individual and aggregate level.

Proofs

Proof of Theorem 1:

We prove the result using a slightly more general formulation, where the volatility of noise trade demand is allowed to vary over time, so instead of τ_u , we have $\tau_{u_1}, \dots, \tau_{u_T}$. We first state three (standard) lemmas.

Lemma 4 (Projection Theorem) *Assume a multivariate signal $[\tilde{\mu}_x; \tilde{\mu}_y] \sim N([\mu_x; \mu_y], [\Sigma_{xx}, \Sigma_{xy}; \Sigma_{yx}, \Sigma_{yy}])$. Then the conditional distribution is*

$$\tilde{\mu}_x | \tilde{\mu}_y \sim N(\mu_x + \Sigma_{xy} \Sigma_{yy}^{-1} (\tilde{\mu}_y - \mu_y), \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx}).$$

Lemma 5 (Special case of projection theorem) *Assume an K -dimensional multivariate signal $\mathbf{v} = [v; \mathbf{s}] \sim N(\bar{v}\mathbf{1}, \sigma_v^2 \mathbf{1}\mathbf{1}' + \Lambda^2)$, where $\Lambda = \text{diag}(0, \sigma_1, \dots, \sigma_{M-1})$. This is to say that $v \sim N(\bar{v}, \sigma_v^2)$, $\mathbf{s}_i = v + \xi_i$, where $\xi_i \sim N(0, \sigma_i^2)$'s are independent of each other and of v , $i = 1, \dots, K-1$. Then the conditional distribution is*

$$v | \mathbf{s} \sim N\left(\frac{\tau_v}{\tau_v + \tau} \bar{v} + \frac{1}{\tau_v + \tau} \boldsymbol{\tau}' \mathbf{s}, \frac{1}{\tau_v + \tau}\right).$$

Here, $\boldsymbol{\tau} = (\tau_1, \dots, \tau_{K-1})^T$, $\tau_i = \sigma_i^{-2}$, $\tau = \sum_{i=1}^{M-1} \tau_i$, and $\tau_v = \sigma_v^{-2}$.

Lemma 6 (Expectation of exponential quadratic form) *Assume $x \sim N(\mu, \Sigma)$, and that \mathbf{B} is a symmetric positive semidefinite matrix. Then*

$$E\left[e^{-\frac{1}{2}(2\mathbf{a}'\mathbf{x} + \mathbf{x}'\mathbf{B}\mathbf{x})}\right] = \frac{1}{|I + \Sigma\mathbf{B}|^{1/2}} e^{-\frac{1}{2}(\mu'\Sigma^{-1}\mu - (\Sigma^{-1}\mu - \mathbf{a})(\Sigma^{-1} + \mathbf{B})^{-1}(\Sigma^{-1}\mu - \mathbf{a}))}.$$

The structure of the proof is now quite straightforward, the extension compared with previous literature being the heterogeneous information diffusion. We first assume that agents' demand takes a linear form at each point in time, and calculate the market maker's pricing function given observed aggregate demand in (3). This turns out to be linear in a way such that the market maker's information is completely revealed in prices. Thus, p_t and w_t convey the same information. We then close the loop by verifying that given the market maker's pricing function in each time period, each agent when solving their backward induction problem will derive demand and utility according to (5,6), verifying that agents' demand functions are indeed linear.

It will be convenient to use the variables $Q_{a,t} = \tau V_{a,t}$. We enumerate the agents from one-dimensionally from 1 to \bar{N} , so that agent $1, \dots, N$ represents the agents in the first replica network, agents $N+1, \dots, 2N$, the agents in the second replica network, etc. Assume that agent a 's time- t demand function is

$$x_{a,t}(z_{a,t}, p_t) = A_{a,t} z_{a,t} + \eta_{a,t}(p_t).$$

Then the total average agent demand is

$$x_t(v, p_t) = \frac{1}{\bar{N}} \sum_{a=1}^{\bar{N}} A_{a,t} z_{a,t} + \eta_{a,t}(p_t) = A_t v + \eta_t(p_t),$$

where $A_t = \frac{1}{\bar{N}} \sum_{a=1}^{\bar{N}} A_{a,t}$, and $\eta_t = \frac{1}{\bar{N}} \sum_{a=1}^{\bar{N}} \eta_{a,t}(p_t)$, with the convention, $A_0 = 0$, $\eta_0 \equiv 0$. Here, we are using the fact that in our large network $\lim_{M \rightarrow \infty} \frac{1}{M} \sum_{r=0}^{M-1} z_{a+rM,t} = v$ for all a and t (almost surely). The net demand at time t is then the difference between time t and $t-1$ demands,

$$\Delta x_t = x_t(v, p_t) - x_{t-1}(v, p_{t-1}) = (A_t - A_{t-1})v + \eta(p_t) - \eta(p_{t-1}).$$

Now, the market maker observes total time t demands,

$$w_t = \Delta x_t + u_t,$$

and since the functions η_t and η_{t-1} are known, the market maker can back out

$$R_t = (A_t - A_{t-1})v + u_t. \tag{24}$$

This leads to the following pricing formula, which immediately follows from Lemma 2.

Lemma 7 *Given the above assumptions, the time- t price is given by*

$$p_t = \frac{\tau_v}{\tau_v + \hat{\tau}_u^t} \bar{v} + \frac{\hat{\tau}_u^t}{\tau_v + \hat{\tau}_u^t} v + \frac{1}{\tau_v + \hat{\tau}_u^t} \sum_{s=1}^t (A_s - A_{s-1}) \tau_{u_s} u_s, \tag{25}$$

where $\hat{\tau}_u^t = \sum_{s=1}^t (A_s - A_{s-1})^2 \tau_{u_s}$, $\tau_{u_s} = \sigma_{u_s}^{-2}$.
Equivalently,

$$p_t = \lambda_t R_t + (1 - \lambda_t (A_t - A_{t-1})) p_{t-1}, \quad (26)$$

where $\lambda_t = \frac{\tau_{u_t} (A_t - A_{t-1})}{\tau_v + \hat{\tau}_u^t}$, and $p_0 = \bar{v}$.

Proof of Lemma 7: At time t , the market maker has observed R_1, \dots, R_t . We define the vector $\mathbf{s} = (R_1/(A_1 - A_0), R_2/(A_2 - A_1), \dots, R_t/(A_t - A_{t-1}))'$, and it is clear that $\mathbf{s}_i \sim N(\bar{v}, \sigma_v^2 + \sigma_{u_i}^2/(A_i - A_{i-1})^2)$. It then follows immediately from Lemma 2 that

$$v|\mathbf{s} \sim N\left(\frac{\tau_v}{\tau_v + \hat{\tau}_u^t} \bar{v} + \frac{1}{\tau_v + \hat{\tau}_u^t} \sum_{i=1}^t (A_i - A_{i-1})^2 \tau_{u_i} \frac{R_i}{A_i - A_{i-1}}, \frac{1}{\tau_v + \hat{\tau}_u^t}\right),$$

i.e.,

$$v = V_t + \sigma_{V_t} \xi_t^V, \quad (27)$$

where $V_t = \frac{\tau_v}{\tau_v + \hat{\tau}_u^t} \bar{v} + \frac{1}{\tau_v + \hat{\tau}_u^t} \sum_{i=1}^t (A_i - A_{i-1})^2 \tau_{u_i} R_i$, $\sigma_{V_t}^2 = \frac{1}{\tau_{V_t}}$, where $\tau_{V_t} = \tau_v + \hat{\tau}_u^t$.

So, $p_t = V_t = E[v|\mathbf{s}]$ takes the given form in the first expression of the lemma. A standard induction argument, assuming that the second expression is valid up until $t-1$, shows that the expression then is also valid for t .

Note that (27) is the posterior distribution of v given the information R_1, \dots, R_t , so $v \sim N(V_t, \sigma_{V_t}^2)$. We have shown Lemma 7.

Thus, linear demand functions by agents imply linear pricing functions in the market, showing the first part of the proof. We next move to the demand functions and expected utilities of agent a , given the pricing function of the market maker. We have (using lemma 1 for the posterior distribution) at time t , the distribution of value given $\{z_{a,t}, p_t\}$ is

$$v|\{z_{a,t}, p_t\} \sim N\left(\frac{\tau_{V_t}}{\tau_{V_t} + \tau_{a,t}} V_t + \frac{\tau_{a,t}}{\tau_{V_t} + \tau_{a,t}} z_{a,t}, \frac{1}{\tau_{V_t} + \tau_{a,t}}\right).$$

The time- T demand of an agent can now be calculated. Since individual agents condition on prices, they also observe \check{R} , and agent a 's information set is therefore $\{z_a, \check{R}\}$, which via Lemma 1 (with $\Lambda = \text{diag}(0, \sigma_i^2, \sigma_u^2/\bar{A}^2)$) leads to

$$v|\{z_a, \check{R}\} \sim N\left(\underbrace{\frac{\tau_v}{\tau_v + \tau_a + \hat{\tau}_u} \bar{v} + \frac{\tau_a}{\tau_v + \tau_a + \hat{\tau}_u} z_a + \frac{\hat{\tau}_u}{\tau_v + \tau_a + \hat{\tau}_u} \check{R}}_{\mu_a}, \underbrace{\frac{1}{\tau_v + \tau_a + \hat{\tau}_u}}_{\hat{\sigma}_a^2}\right).$$

At time T , given the behavior of the market maker, the asset's value, given agent a 's information set is therefore conditionally normally distributed so given agent a 's CARA utility, the demand for the asset (2) takes the form:

$$x_{a,T}(z_a, p) = \frac{E[v|\mathcal{I}_{a,T}] - p_T}{\gamma_a \sigma^2[v|\mathcal{I}_{a,T}]}, \quad a = 1, \dots, \bar{N}. \quad (28)$$

The demand of agent a is therefore

$$\begin{aligned} x_{a,T}(z_a, p) &= \frac{\mu_a - p_T}{\gamma_a \hat{\sigma}_a^2} \\ &= \frac{1}{\gamma_a} (\tau_v \bar{v} + \tau_a z_a + \hat{\tau}_u \check{R} - (\tau_v + \tau_a + \hat{\tau}_u) p) \\ &= \frac{1}{\gamma_a} \left(\tau_v \bar{v} + \tau_a z_a + \hat{\tau}_u \check{R} - \left(1 + \frac{\tau_a}{\tau_v + \hat{\tau}_u}\right) (\tau_v \bar{v} + \hat{\tau}_u \check{R}) \right) \\ &= \frac{1}{\gamma_a} (\tau_a z_a - \tau_a p) \\ &= \frac{\tau_{a,T}}{\gamma_a} (z_{a,T} - p_T). \end{aligned}$$

It follows that $A_T = \frac{1}{\bar{N}} \sum_{a=1}^{\bar{N}} \frac{N_{a,T}}{\sigma^2 \gamma_a}$.

Since $v - p_T \sim N(0, \sigma_{V_t}^2)$, and $z_{a,T} - p_T = \zeta_{a,T} + (v - p_T)$, where $\zeta_{a,T}$ is independent of $v - p_T$, it follows that $\zeta_{a,T} | (z_{a,T} - p_T) \sim N(\frac{\tau_{V_T}}{\tau_{V_T} + \tau_{a,T}}(z_{a,T} - p_T), \frac{1}{\tau_{V_T} + \tau_{a,T}})$. The expected utility of the agent at time T (with time- T wealth of zero), given $z_{a,T} - p_T$, is then

$$\begin{aligned}
U_{a,T} &= -E \left[e^{-\gamma_a x_{a,T}(v-p_T)} \middle| z_{a,T} - p_T \right] \\
&= -E \left[e^{-\tau_{a,T}(z_{a,T}-p_T)(z_{a,T}-p_T-\zeta_{a,T})} \middle| z_{a,T} - p_T \right] \\
&= -e^{-\tau_{a,T}(z_{a,T}-p_T)^2} E \left[e^{-\tau_{a,T}(z_{a,T}-p_T)(-\zeta_{a,T})} \middle| z_{a,T} - p_T \right] \\
&= -e^{-\tau_{a,T}(z_{a,T}-p_T)^2} e^{\frac{\tau_{a,T}(z_{a,T}-p_T)}{\tau_{V_T} + \tau_{a,T}}(z_{a,T}-p_T) - \frac{1}{2}\tau_{a,T}^2(z_{a,T}-p_T)^2 \frac{1}{\tau_{V_T} + \tau_{a,T}}} \\
&= -e^{-\frac{\tau_{a,T}}{\tau_{V_t} + \tau_{a,T}}(z_{a,T}-p_T)^2 ((\tau_{V_t} + \tau_{a,T} - \tau_{V_t}) - \frac{1}{2}\tau_{a,T})} \\
&= -e^{-\frac{1}{2}\frac{\tau_{a,T}^2}{\tau_{V_t} + \tau_{a,T}}(z_{a,T}-p_T)^2}.
\end{aligned}$$

This shows the result at T .

It is easy to check that the unconditional expected utility is $-E_0[e^{-\gamma_a x_{a,T}(v-p_T)}] = -\sqrt{\frac{\tau_{V_T}}{\tau_{V_T} + \tau_{a,T}}}$, using lemma 3.

We define $Y_t = \hat{\tau}_u^t = \sum_{i=1}^t y_i$, where $y_i = (A_i - A_{i-1})^2 \tau_{u_i}$, and recall that $Q_{a,t} = \tau_{a,t} = \frac{V_{a,t}}{\sigma^2} = \sum_{i=1}^t q_{a,i}$, where $q_{a,i} = \frac{V_{a,i} - V_{a,i-1}}{\sigma^2} = \frac{\Delta V_{a,i}}{\sigma^2}$. With this notation we have

$$\begin{aligned}
U_{a,T} &= -e^{-\frac{1}{2}\frac{Q_{a,T}^2}{\tau_v + Y_T + Q_T}(z_{a,T} - p_T)^2}, \\
x_{a,T} &= \frac{Q_{a,T}}{\gamma_a}(z_{a,T} - p_T),
\end{aligned}$$

and $-E_0[e^{-\gamma_a x_{a,T}(v-p_T)}] = -\sqrt{\frac{\tau_v + Y_T}{\tau_v + Y_T + \tau_{V_{a,T}}}} = -\frac{1}{\sqrt{c_{a,T}}} = -D_{a,T}$.

We proceed with an induction argument: We show that given that (5,6) is satisfied at time t , then it is satisfied at time $t-1$. As already shown, p_{t-1} , and $z_{a,t-1}$ sufficiently summarizes agent a 's information at time $t-1$ (given the linear pricing function). From the law of motion, $W_{a,t} = W_{a,t-1} + x_{a,t-1}(p_t - p_{t-1})$, an agent's optimization at time $t-1$ is then

$$\begin{aligned}
U_{a,t} &= \arg \max_{x_{a,t-1}} -E_{a,t-1} \left[e^{-\gamma_a W_{a,t-1} - \gamma_a x_{a,t-1}(p_t - p_{t-1})} D_{a,t} e^{-\frac{1}{2}\frac{Q_t^2}{\tau_v + Y_t + Q_t}(z_{a,t} - p_t)^2} \middle| z_{a,t-1}, p_{t-1} \right] \\
&= \arg \max_{x_{a,t-1}} -D_{a,t} e^{-\gamma_a W_{a,t-1}} E_{a,t-1} \left[e^{-\gamma_a x_{a,t-1}(p_t - p_{t-1}) - \frac{1}{2}\frac{Q_t^2}{\tau_v + Y_t + Q_t}(z_{a,t} - p_t)^2} \middle| z_{a,t-1}, p_{t-1} \right] \\
&\stackrel{\text{def}}{=} \arg \max_b -D_{a,t} e^{-\gamma_a W_{a,t-1}} E_{a,t-1} \left[e^{-b(p_t - p_{t-1}) - \frac{1}{2}\frac{Q_t^2}{\tau_v + Y_t + Q_t}(z_{a,t} - p_t)^2} \middle| z_{a,t-1}, p_{t-1} \right]. \tag{29}
\end{aligned}$$

Thus, we need to calculate the distributions of $p_t - p_{t-1}$ and $z_{a,t} - p_t$ given $z_{a,t-1}$ and p_{t-1} . From the signal structure, we have the following relationship

$$z_{a,t-1} = v + \xi_{t-1}, \quad \xi_{t-1} \sim N\left(0, \frac{1}{Q_{t-1}}\right), \tag{30}$$

$$z_{a,t} = v + \xi_t = v + \frac{Q_{t-1}}{Q_t} \xi_{t-1} + \frac{q_t}{Q_t} e_t, \quad e_t \sim N\left(0, \frac{1}{q_t}\right), \tag{31}$$

where e_t and ξ_{t-1} jointly independent and independent of all other variables. In the new notation, from (4), we have

$$p_t = \frac{\tau_v}{\tau_v + Y_t} \bar{v} + \frac{Y_t}{\tau_v + Y_t} v + \frac{1}{\tau_v + Y_t} \sum_{s=1}^t (A_s - A_{s-1}) \tau_{u_s} u_s, \tag{32}$$

$$p_{t-1} = \frac{\tau_v}{\tau_v + Y_{t-1}} \bar{v} + \underbrace{\frac{Y_{t-1}}{\tau_v + Y_{t-1}} v}_{A_4} + \frac{1}{\tau_v + Y_{t-1}} \sum_{s=1}^{t-1} (A_s - A_{s-1}) \tau_{u_s} u_s, \tag{33}$$

so

$$\begin{aligned}
p_t - p_{t-1} &= \left(\frac{\tau_v}{\tau_v + Y_t} - \frac{\tau_v}{\tau_v + Y_{t-1}} \right) \bar{v} + \underbrace{\left(\frac{Y_t}{\tau_v + Y_t} - \frac{Y_{t-1}}{\tau_v + Y_{t-1}} \right)}_{A_1} v + \frac{1}{\tau_v + Y_t} \underbrace{(A_t - A_{t-1})\tau_{u_t}}_{\sqrt{y_t \tau_{u_t}}} u_t \\
&+ \underbrace{\left(\frac{1}{\tau_v + Y_t} - \frac{1}{\tau_v + Y_{t-1}} \right)}_{B_1} \sum_{s=1}^{t-1} (A_s - A_{s-1})\tau_{u_s} u_s,
\end{aligned} \tag{34}$$

and also

$$\begin{aligned}
z_{a,t} - p_t &= \frac{Q_{t-1}}{Q_t} \xi_{t-1} + \frac{q_t}{Q_t} e_t - \frac{\tau_v}{\tau_v + Y_t} \bar{v} + \underbrace{\frac{\tau_v}{\tau_v + Y_t}}_{A_2} v \\
&- \frac{1}{\tau_v + Y_t} \underbrace{(A_t - A_{t-1})\tau_{u_t}}_{\sqrt{y_t \tau_{u_t}}} u_t - \frac{1}{\tau_v + Y_t} \sum_{s=1}^{t-1} (A_s - A_{s-1})\tau_{u_s} u_s.
\end{aligned} \tag{35}$$

This leads to the unconditional distribution:

$$\begin{bmatrix} p_t - p_{t-1} \\ s_t - p_t \\ s_{t-1} \\ p_{t-1} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ \bar{v} \\ \bar{v} \end{bmatrix}, \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma'_{XY} & \Sigma_{YY} \end{bmatrix} \right), \tag{36}$$

Here,

$$\begin{aligned}
\Sigma_{XX} &= \begin{bmatrix} \frac{A_1^2}{\tau_v} + \frac{y_t}{(\tau_v + Y_t)^2} + B_1^2 Y_{t-1} & \frac{A_1 A_2}{\tau_v} - \frac{y_t}{(\tau_v + Y_t)^2} - \frac{B_1 Y_{t-1}}{\tau_v + Y_t} \\ \frac{A_1 A_2}{\tau_v} - \frac{y_t}{(\tau_v + Y_t)^2} - \frac{B_1 Y_{t-1}}{\tau_v + Y_t} & \frac{1}{Q_t} + \frac{A_2^2}{\tau_v} + \frac{y_t}{(\tau_v + Y_t)^2} + \frac{Y_{t-1}}{(\tau_v + Y_t)^2} \end{bmatrix} \\
&= \begin{bmatrix} \frac{y_t}{(\tau_v + Y_t)(\tau_v + Y_{t-1})} & 0 \\ 0 & \frac{1}{Q_t} + \frac{1}{\tau_v + Y_t} \end{bmatrix},
\end{aligned}$$

$$\begin{aligned}
\Sigma_{YY} &= \begin{bmatrix} \frac{1}{\tau_v} + \frac{1}{Q_{t-1}} & \frac{A_4}{\tau_v} \\ \frac{A_4}{\tau_v} & \frac{A_4^2}{\tau_v} + \frac{Y_{t-1}}{(\tau_v + Y_{t-1})^2} \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{\tau_v} + \frac{1}{Q_{t-1}} & \frac{Y_{t-1}}{\tau_v(\tau_v + Y_{t-1})} \\ \frac{Y_{t-1}}{\tau_v(\tau_v + Y_{t-1})} & \frac{Y_{t-1}}{\tau_v(\tau_v + Y_{t-1})} \end{bmatrix},
\end{aligned}$$

$$\begin{aligned}
\Sigma_{XY} &= \begin{bmatrix} \frac{A_1}{\tau_v} & \frac{A_4 A_1}{\tau_v} + \frac{B_1}{\tau_v + Y_{t-1}} Y_{t-1} \\ \frac{1}{Q_t} + \frac{A_2}{\tau_v} & \frac{A_2 A_4}{\tau_v} - \frac{1}{\tau_v + Y_{t-1}} \frac{1}{\tau_v + Y_t} Y_{t-1} \end{bmatrix} \\
&= \begin{bmatrix} \frac{y_t}{(\tau_v + Y_t)(\tau_v + Y_{t-1})} & 0 \\ \frac{1}{Q_t} + \frac{1}{\tau_v + Y_t} & 0 \end{bmatrix}.
\end{aligned}$$

We use the projection theorem to write $[p_t - p_{t-1}; z_{a,t} - p_t] \sim N(\mu, \hat{\Sigma})$, where $\mu = \Sigma_{XY} \Sigma_{YY}^{-1} [z_{a,t-1} - \bar{v}; p_{t-1} - \bar{v}]$, and $\hat{\Sigma} = \Sigma_{XX} - \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma'_{XY}$. It follows that

$$\mu = \begin{bmatrix} \frac{Q_{t-1} y_t}{(\tau_v + Y_t)(\tau_v + Q_{t-1} + Y_{t-1})} \\ \frac{Q_{t-1}(\tau_v + Y_{t-1})(\tau_v + Q_t + Y_t)}{Q_t(\tau_v + Q_{t-1} + Y_{t-1})(\tau_v + Y_t)} \end{bmatrix} (z_{a,t-1} - p_{t-1}).$$

We rewrite (29) as

$$U_{a,t} = \arg \max_q -D_{a,t} e^{-\gamma_a W_{a,t-1}} E \left[e^{-\mathbf{a}\mathbf{x}_1 - \frac{1}{2}\mathbf{x}'\mathbf{B}\mathbf{x}} \right],$$

where $\mathbf{x} = [p_t - p_{t-1}; z_t - p_t]$, $\mathbf{B} = \left[0, 0; 0, \frac{Q_t^2}{\tau_v + Y_t + Q_t} \right]$, $\mathbf{a} = [q; 0]$, and $q = \gamma_a x_{a,t-1}$. From Lemma 3, it follows directly that this maximization problem is equivalent to

$$U_{a,t} = \arg \max_q \frac{-D_{a,t} e^{-\gamma_a W_{a,t-1}}}{|I + \hat{\Sigma}\mathbf{B}|^{1/2}} e^{-\frac{1}{2}(\mu' \hat{\Sigma}^{-1} \mu - (\hat{\Sigma}^{-1} \mu - \mathbf{a})\mathbf{Z}(\hat{\Sigma}^{-1} \mu - \mathbf{a}))},$$

where $\mathbf{Z} = (\hat{\Sigma}^{-1} + \mathbf{B})^{-1}$. Clearly, the optimal solution is given by

$$\arg \max_q q(\mathbf{Z}\hat{\Sigma}^{-1}\mu)_1 - \frac{1}{2}\mathbf{Z}_{11}q^2,$$

leading to $q^* = \frac{(\mathbf{Z}\hat{\Sigma}^{-1}\mu)_1}{\mathbf{Z}_{11}}$. It is easy to verify that $\hat{\Sigma}^{-1}\mu = \frac{Q_t Q_{t-1}}{Q_t - Q_{t-1}} [1; 1](z_{a,t-1} - p_{t-1})$, and some further algebraic manipulations shows that indeed $q^* = Q_{t-1}(z_{a,t-1} - p_{t-1})$, leading to the stated demand function at $t-1$, (5).

Given the form of q , it then follows that

$$\mu' \hat{\Sigma}^{-1} \mu - (\hat{\Sigma}^{-1} \mu - \mathbf{a})\mathbf{Z}(\hat{\Sigma}^{-1} \mu - \mathbf{a}) = \frac{Q_{t-1}^2}{\tau_v + Q_{t-1} + Y_{t-1}} (z_{a,t-1} - p_{t-1})^2, \quad (37)$$

leading to the form of the utility stated in the theorem (6), with $C_{a,t-1} = |I + \hat{\Sigma}\mathbf{B}|^{1/2}$. It is easy to check that $C_{a,t-1}$ takes the prescribed form, as does then $D_{a,t-1} = C_{a,t-1}^{-1/2} D_{a,t}$.

Thus, given a linear pricing function, agents' demand take a linear form and, moreover, the coefficients take the functional forms shown in the Theorem, as do agents' expected utility. We are done. \blacksquare

Proof of Theorem 2: The proof follows immediately from (5), (30-31), and (35). \blacksquare

Proof of Theorem 3: The certainty equivalent satisfies $-e^{-\gamma_a CE} = E_0[-e^{-W_{T+1}}] = -E_0 \left[\prod_{t=2}^T C_{t-1}^{-1} e^{-\frac{1}{2} \frac{Q_t^2}{\tau_v + Q_t + Y_t} (z_{a,1} - p_1)^2} \right]$.

It is easy to see that

$$\prod_{t=2}^T C_{t-1}^{-1} = \left(\frac{\tau_v + Q_T + Y_T}{\tau_v + Q_1 + Y_1} \right) \left(\frac{\tau_v + Y_1}{\tau_v + Y_T} \right) \prod_{t=2}^T \left(1 + \frac{Q_{t-1}(Y_t - Y_{t-1})}{(\tau_v + Y_{t-1})(\tau_v + Y_t)} \right).$$

Moreover, since at $t=0$, $z_{a,1} - p_1 \sim N(0, \frac{1}{Q_1} + \frac{1}{\tau_v + Y_1})$, it follows that

$$E_0 \left[e^{-\frac{1}{2} \frac{Q_1^2}{\tau_v + Q_1 + Y_1} (z_{a,1} - p_1)^2} \right] = \frac{\tau_v + Q_1 + Y_1}{\tau_v + Y_1},$$

and the result for the ex ante certainty equivalent follows.

The ex ante expected profits between t and $t+1$ are $E_0[(z_{a,t} - p_t)(p_{t+1} - p_t)]$. Plugging in the form (34,35) yields the result. Using a similar approach for expected profits, we get that the expected total, time T trading profit of agent a 's trade in time t is

$$\frac{1}{\gamma_a} \frac{Q_{a,t}}{\tau_v + Y_t},$$

and the total expected trading profit over time therefore is

$$\frac{1}{\gamma_a} \sum_{t=1}^T \frac{Q_{a,t}}{\tau_v + Y_t}, \quad (38)$$

We are done. ■

Proof of Theorem 4:

Proof of Theorem 4: We focus on the Erdős-Renyi random graph $G(N, p_N)$ model, where

$$p_N = c \log^k(N)/(N-1), \quad c > 0, \quad k > 3, \quad (39)$$

$$q_N = p(N-1) = c \log^k(N). \quad (40)$$

Throughout the proof we will often suppress the dependence of variables on N , e.g., writing p and q instead of p_N and q_N . Also, ϵ denotes a constant arbitrarily close to zero, and C, c, c_1 , etc. denote strictly positive (possibly large) constants.

The proof is based on the fact that for arbitrary constants, $\alpha > 0, \beta > 0$, it follows that $q^\alpha \ll N^\beta$, for large N . We shall see that the variables of interest can be given as well-behaved functions of $q^{-\alpha}$ with error terms of order $N^{-\beta}$, which then leads to tight bounds on the expected correlations.

We first introduce some convenient notation. For a sequence of random variables w_N , we write $w = O_s(f(N))$, where it is assumed that $\lim_{N \rightarrow \infty} f(N) = 0$, if for large N

$$\mathbb{P}(|w| \geq C f(N)) \leq C f(N)$$

for some constant, $C > 0$. If $w_N - r = O_s(f(N))$, for some function (or constant) r , we also write $w_N = r + O_s(f(N))$. We will also use standard $O()$ (Big-O) and $o()$ (little-o) notation. Finally, $a_N \sim b_N$ means that $0 < c \leq a_N/b_N \leq C < \infty$ for large N (i.e., that $a_N = \Theta(b_N)$).

We use the notation $Var(x)$, $Cov(x, y)$, and $Corr(x, y)$ for the population variance, covariance and correlation, respectively, of random variables x and y , whereas we use $V[X]$, $C[X, Y]$, and $\rho[X, Y]$ for the sample (cross sectional) variance, covariance and correlation of vectors of realization of random variables, X and Y .

We are interested in the expectations of the cross sectional correlations

$$\begin{aligned} \rho_D &\stackrel{\text{def}}{=} \rho[D_N, \Pi_N], \\ \rho_F &\stackrel{\text{def}}{=} \rho[-F, \Pi_N], \\ \rho_{\hat{C}} &\stackrel{\text{def}}{=} \rho[\hat{C}, \Pi_N], \\ \rho_K &\stackrel{\text{def}}{=} \rho[K^\alpha, \Pi_N]. \end{aligned}$$

We begin with deriving properties of y_t and $\zeta_t = \beta_t/\beta_1$. To do this, we start with the following Lemma that shows the distributional properties of $V_{a,t}$ for large N .

Lemma 8 Consider constants $K > 14$ (arbitrarily large), and $\epsilon > 0$ (arbitrarily close to zero), and an integer $T > 1$ (arbitrarily large). Then there is an N_0 , such that for all $N > N_0$, for all $a = 1, \dots, N$, for all $t = 1, \dots, T$,

$$\mathbb{P}(|V_{a,t} - V_{a,t-1} - q^t| \geq \epsilon q^t) \leq N^{-K}. \quad (41)$$

Proof: We apply Lemma 10.7 in Bollobas (2001). Given that $\alpha_t = (K-2)\sqrt{\log(N)/(c \log^k(N))^t}$, $\beta_t = (c \log^k(N))^t/N$, and $\gamma_t = 2(c \log^k(N))^{t-1}/N$, it follows that $\alpha_1 \leq C \log(N)^{\frac{1-k}{2}}$, and

$$\sum_{i=1}^t \alpha_i + \beta_i + \gamma_i = \alpha_1(1 + o(1)),$$

where α_i, β_i and γ_i are defined in Lemma 10.6 of Bollobas (2001). Lemma 10.7 in Bollobas (2001) now implies that,

$$\mathbb{P}(|V_{a,t} - V_{a,t-1} - q^t| \geq \delta_t q^t) \leq N^{-K}, \quad (42)$$

where

$$\delta_t \leq C' \log(N)^{\frac{1-k}{2}}, \quad t = 1, \dots, T, \quad (43)$$

and since $\delta_t \rightarrow 0$ when $N \rightarrow \infty$, the result follows. \blacksquare

Bounds on y_t :

We recall that $V_{a,t}$ is the number of nodes within a distance of t from node a , and derive asymptotic properties of the averages

$$\bar{V}_t \stackrel{\text{def}}{=} \frac{1}{N} \sum_{a=1}^N V_{a,t}, \quad t = 1, \dots, T. \quad (44)$$

We note that $A_t = \frac{\tau}{\gamma} \bar{V}_t$, so all asymptotic results we derive for \bar{V}_t will, up to a constant, also hold for A_t , and thereby be important for y_t . We have

$$y_t = \alpha (\bar{V}_t - \bar{V}_{t-1})^2,$$

where we define $\bar{V}_0 = 0$, and $\alpha = \frac{\tau u \tau^2}{\gamma}$ (the risk aversion coefficient being constant since we assume a preference symmetric economy). Lemma (8) immediately implies that

$$\mathbb{P}(|\bar{V}_t - \bar{V}_{t-1} - q^t| \geq \epsilon q^t) \leq N^{-K},$$

for any $\epsilon > 0$, for large N , and therefore

$$\begin{aligned} |\bar{V}_t - \bar{V}_{t-1} - q^t| \leq \epsilon q^t &\Rightarrow |y_t - \alpha q^{2t}| \\ &= \alpha |\bar{V}_t - \bar{V}_{t-1} - q^t| \times |\bar{V}_t - \bar{V}_{t-1} + q^t| \\ &\leq \alpha \epsilon q^t \times (1 + \epsilon) q^t \\ &= \epsilon' q^{2t}. \end{aligned}$$

Since ϵ' is also arbitrarily close to zero, we get:

$$\mathbb{P}(|y_t - \alpha q^{2t}| \geq \epsilon q^{2t}) \leq N^{-K}. \quad (45)$$

Bounds on ζ_t :

We can write the expected profit of agent a as

$$\Pi_a = \beta_1 \left(V_{a,1} + \sum_{t=2}^T \zeta_t V_{a,t} \right),$$

where $\beta_1 = (\tau_v + y_1)^{-1}$ and

$$\zeta_t = \frac{\tau_v + y_1}{\tau_v + \sum_{k=1}^t y_k}.$$

The coefficients, ζ_t (which do not depend on a) represent the relative value for agents of having many links at distance t .

The bound on y_t (45) now immediately imply that

$$\mathbb{P}\left(|\zeta_t \times q^{2(t-1)} - 1| \geq \epsilon\right) \leq N^{-K}. \quad (46)$$

for any $\epsilon > 0$, $K > 14$, for large enough N .

Asymptotic behavior of $E[\rho_D]$, and $E[\rho_K]$:

First, we restrict Ω_N to only contain events for which $\zeta_t q^{2(t-1)} \in [1 - \epsilon, 1 + \epsilon]$, $t = 1, \dots, T$, for some fixed $\epsilon \ll 1$. From (46) we know that this set of events satisfies

$$\mathbb{P}(\Omega_N) \geq 1 - O(q^{-K}) \quad (47)$$

for arbitrarily large K .

We have

$$\begin{aligned}
\rho[D_N, \Pi_N] &= \rho[V_{a,1}^N, \beta_1(V_{a,1}^N + \zeta_2 V_{a,2}^N + \zeta_3 V_{a,3}^N + \cdots + \zeta_T^T V_{a,T}^N)] \\
&= \frac{S_1^2 + \zeta_2 C[V_{a,1}, V_{a,2}] + \cdots + \zeta_T C[V_{a,1}, V_{a,2}]}{S_1 \sqrt{V[V_{a,1} + \zeta_2 V_{a,2} + \cdots + \zeta_T V_{a,T}]}} \\
&= \frac{S_1^2 + \sum_{i=2}^T \zeta_i S_{i,1}}{S_1 \sqrt{S_1^2 + \sum_{i=2}^T \zeta_i^2 S_i^2 + \sum_{i=1, j>i}^T 2\zeta_i \zeta_j S_{j,i}}}, \tag{48}
\end{aligned}$$

where we have defined $\zeta_1 = 1$, and

$$S_t^2 = V[V_{a,t}], \quad t = 1, 2, \dots, T, \tag{49}$$

$$S_{t,s} = C[V_{a,t}, V_{a,s}], \quad s, t = 1, 2, \dots, T. \tag{50}$$

We define $a_t = \zeta_i q^{2(t-1)}$, $t = 2, \dots, T$, where the coefficients a_i can be chosen arbitrarily close to 1 by letting N be sufficiently large and focusing on events in Ω defined in (47), we have good control over the behavior of ζ_t . We also need to control S_T^2 , and $S_{t,s}$, $t > s$.

The following Lemma is helpful

Lemma 9 *Given the definitions (49), (50), we have*

$$S_t^2 = \sum_{i=1}^t Z_i^2 + 2 \sum_{i=1}^t \sum_{j=i+1}^t q^{j-i} Z_i + O_s(N^{-1/3+\epsilon}), \quad 1 \leq t \leq T, \tag{51}$$

$$S_{t,s} = S_s^2 + \sum_{i=1}^s \sum_{j=s+1}^t q^{j-i} Z_i + O_s(N^{-1/3+\epsilon}), \quad 1 \leq s, t \leq T. \tag{52}$$

Here,

$$\begin{aligned}
Z_t^2 &= q^t \sum_{i=0}^{t-1} q^i, \\
&= q^{2t-1} \sum_{i=0}^{t-1} q^{-i}, \quad t = 1, \dots, T,
\end{aligned}$$

Proof of Lemma 9:

For large N , the distributions of $V_{a,t}$ for small t will resemble those of a branching process, the difference being that branching processes do not choose from the same nodes when going from t to $t+1$. For large N , the difference will be small. We therefore begin with studying the branching process defined as $Z_{a,1} = \xi_{a,0}$, $Z_{a,t+1} = \sum_{i=1}^{V_{a,t}} \xi_{a,t,i}$, $a = 1, \dots, N$, $t = 1, \dots, T-1$. Here, $\xi_{a,t,i} \sim \text{Bin}(N-1, p)$, and are independent across a , t , and i . Using iterated expectations, and the law of total variance, it is straightforward to show the following (unconditional) moment formulas

$$\begin{aligned}
E[Z_{a,t}] &= E[\xi]^t, \\
\text{Var}[Z_{a,t}] &= E[\xi]^t \text{Var}(\xi) + E[\xi]^{2t} \text{Var}[Z_{a,t-1}], \\
\text{Cov}[Z_{a,t}, Z_{a,s}] &= E[\xi]^{t-s} \text{Var}(Z_{a,s}), \quad t > s.
\end{aligned}$$

In the specific case where $E[\xi] = q$, and $\text{Var}(\xi) = q + O(N^{-1})$ —i.e., our case—a recursive argument shows that

$$\begin{aligned}
E[Z_{a,t}] &= q^t, \\
\text{Var}[Z_{a,t}] &= q^t \sum_{i=0}^{t-1} q^i + O(N^{-1+\epsilon}), \tag{53}
\end{aligned}$$

$$\text{Cov}[Z_{a,t}, Z_{a,s}] = q^t \sum_{i=0}^{s-1} q^i + O(N^{-1+\epsilon}), \quad t > s. \tag{54}$$

We next show that the moments for $V_{a,t+1} - V_{a,t}$ are similar to the moments of $Z_{a,t+1}$. Of course, $V_{a,1}$ has identical moments as $Z_{a,1}$. For $V_{a,2}$, we have, given that node a has $V_{a,1}$ links, the probability that node b which is not a neighbor of a is not a neighbor to one of a 's neighbors is $(1-p)^{V_{a,1}}$. Therefore, the conditional distribution of the number of nodes at distance 2 from a , given $V_{a,1}$ is

$$V_{a,2} - V_{a,1} \sim \text{Bin}(N - V_{a,1}, 1 - (1-p)^{V_{a,1}}),$$

leading to

$$\begin{aligned} E[V_{a,2} - V_{a,1} | V_{a,1}] &= (N - V_{a,1})(1 - (1-p)^{V_{a,1}}) \\ &= -N(1 - V_{a,1}/N) \left(\sum_{i=1}^{V_{a,1}} \binom{V_{a,1}}{i} (-1)^i p^i \right) \\ &= -N(1 - V_{a,1}/N) \left(\sum_{i=1}^{V_{a,1}} \binom{V_{a,1}}{i} \left(\frac{-q}{N-1} \right)^i \right), \end{aligned} \quad (55)$$

and

$$\begin{aligned} \text{Var}[V_{a,2} - V_{a,1} | V_{a,1}] &= (N - V_{a,1})(1 - (1-p)^{V_{a,1}})(1-p)^{V_{a,1}} \\ &= -N(1 - V_{a,1}/N)(1-p)^{V_{a,1}} \left(\sum_{i=1}^{V_{a,1}} \binom{V_{a,1}}{i} (-1)^i p^i \right) \\ &= -N(1 - V_{a,1}/N)(1-p)^{V_{a,1}} \left(\sum_{i=1}^{V_{a,1}} \binom{V_{a,1}}{i} \left(\frac{-q}{N-1} \right)^i \right). \end{aligned} \quad (56)$$

The law of total variance now implies that

$$\text{Var}[V_{a,2} - V_{a,1}] = E[\text{Var}[V_{a,2} - V_{a,1} | V_{a,1}]] + \text{Var}[E[V_{a,2} - V_{a,1} | V_{a,1}]], \quad (57)$$

and since (as is easily shown) $E[V_{a,t}^i / (N-1)] = O(N^{-1+\epsilon})$ for any fixed i and t , and furthermore, $E\left[\sum_{i=2}^{\infty} V_{a,t}^i / (N-1)^i\right] = O(N^{-1})$, (55,56) immediately imply, when plugged into (57), that

$$\text{Var}[V_{a,2} - V_{a,1}] = q^2 + q^3 + O(N^{-1+\epsilon}) = \text{Var}[Z_{a,1}] + O(N^{-1+\epsilon}).$$

The same argument for

$$V_{a,3} - V_{a,2} \sim \text{Bin}(N - (V_{a,2} - V_{a,1}), 1 - (1-p)^{V_{a,2} - V_{a,1}}),$$

leads to

$$\text{Var}[V_{a,3} - V_{a,2}] = q^3 + q^2(q^2 + q^3) + O(N^{-1+\epsilon}) = \text{Var}[Z_{a,2}] + O(N^{-1+\epsilon}),$$

and for higher orders to

$$\text{Var}[V_{a,t} - V_{a,t-1}] = \text{Var}[Z_{a,t}] + O(N^{-1+\epsilon}). \quad (58)$$

A similar argument for the covariances leads to,

$$\text{Cov}[V_{a,t} - V_{a,t-1}, V_{a,s} - V_{a,s-1}] = \text{Cov}[Z_{a,t}, Z_{a,s}] + O(N^{-1+\epsilon}), \quad t > s. \quad (59)$$

Now, since $V_{a,t} = V_{a,1} + (V_{a,2} - V_{a,1}) + \dots + (V_{a,t} - V_{a,t-1})$, from (58,59) it immediately follows that

$$\text{Var}[V_{a,t}] = \sum_{i=1}^t Z_i^2 + 2 \sum_{i=1}^t \sum_{j=i+1}^t q^{j-i} Z_i + O_s(N^{-1+\epsilon}), \quad \text{and} \quad (60)$$

$$\text{Cov}[V_{a,t}, V_{a,s}] = \text{Var}[V_{a,s}] + \sum_{i=1}^s \sum_{j=s+1}^t q^{j-i} Z_i + O_s(N^{-1+\epsilon}), \quad t > s. \quad (61)$$

We next show that covariances of $V_{a,t}$ and $V_{b,t}$ when $a \neq b$ (i.e., across agents) is low. First, $V_{a,1}^m - 1$, is of course binomially distributed,

$$V_{a,1} - 1 \sim \text{Bin}(N, p).$$

We recall that the moment generating function of a binomial distribution is $M(t) = (1 - p + pe^t)^n$, and it therefore follows that for any fixed m ,

$$\begin{aligned} \mu_m &\stackrel{\text{def}}{=} E[V_{a,1}^m] = M^{(m)}(0) = q^m(1 + O(q^{-1})), \\ \text{Var}[V_{a,1}^m] &= M^{(2m)}(0) - (M^{(m)}(0))^2 = q^{2m-1}(1 + O(q^{-1})). \end{aligned}$$

The covariance between $V_{a,1}^m$ and $V_{b,1}^n$, $a \neq b$ is on the form

$$\text{Cov}(V_{a,1}^m, V_{b,1}^n) = E[(w_a + I)^m (w_b + I)^n] - E[(w_a + I)^m] E[(w_b + I)^n],$$

where I is Bernoulli distributed, $I \sim \text{Ber}(p)$, representing the possible common link between a and b , $w_a \sim \text{Bin}(N-1, p)$, $w_b \sim \text{Bin}(N-1, p)$ represent all the other links, and I , w_a , and w_b are independent.

Moreover, since the Bernoulli distribution is idempotent, $I^k = I$, $k \geq 1$, it follows that

$$\begin{aligned} \text{Cov}(V_{a,1}^m, V_{b,1}^n) &= E[(w_a + I)^m (w_b + I)^n] - E[(w_a + I)^m] E[(w_b + I)^n] \\ &= E \left[\left(w_a^m + \sum_{i=0}^{m-1} \binom{m}{i} w_a^i I \right) \left(w_b^n + \sum_{j=0}^{n-1} \binom{n}{j} w_b^j I \right) \right] \\ &= E \left[w_a^m + \sum_{i=0}^{m-1} \binom{m}{i} w_a^i I \right] E \left[w_b^n + \sum_{j=0}^{n-1} \binom{n}{j} w_b^j I \right] \\ &= E \left[\left(\sum_{i=0}^{m-1} \binom{m}{i} w_a^i I \right) \left(\sum_{j=0}^{n-1} \binom{n}{j} w_b^j I \right) \right] - E \left[\sum_{i=0}^{m-1} \binom{m}{i} w_a^i I \right] E \left[\sum_{j=0}^{n-1} \binom{n}{j} w_b^j I \right] \\ &= \left(\sum_{i=0}^{m-1} \binom{m}{i} \mu_i \right) \left(\sum_{j=0}^{n-1} \binom{n}{j} \mu_j \right) E[I] - \left(\sum_{i=0}^{m-1} \binom{m}{i} \mu_i \right) \left(\sum_{j=0}^{n-1} \binom{n}{j} \mu_j \right) E[I]^2 \\ &= \left(\sum_{i=0}^{m-1} \binom{m}{i} \mu_i \right) \left(\sum_{j=0}^{n-1} \binom{n}{j} \mu_j \right) \text{Var}(I) \\ &= O(q^{m+n-2} p(1-p)) \\ &= O(N^{-1+\epsilon}). \end{aligned}$$

Together with the bounds on $\text{Var}(V_{a,1}^m)$, this then implies that

$$\begin{aligned} \text{Corr}(V_{a,1}^m, V_{b,1}^n) &= \frac{\text{Cov}(V_{a,1}^m, V_{b,1}^n)}{\sqrt{\text{Var}(V_{a,1}^m) \text{Var}(V_{b,1}^n)}} \\ &= \frac{O(N^{-1+\epsilon})}{\sqrt{q^{2m-1} q^{2n-1} (1 + O(q^{-1}))}} \\ &= O(N^{-1+\epsilon}). \end{aligned}$$

An identical argument as that above shows that

$$\text{Cov}((V_{a,1} - V_{b,1})^m, (V_{c,1} - V_{d,1})^n) = O(q^{m+n-2} N^{-1+\epsilon}) = O(N^{-1+\epsilon}),$$

when a, b, c and d are different, and that

$$\text{Corr}((V_{a,1} - V_{b,1})^m, (V_{c,1} - V_{d,1})^n) = O(N^{-1+\epsilon}),$$

and similarly for general $T \geq t \geq s \geq 1$, $n \geq 1$, $m \geq 1$

$$\text{Cov}(V_{a,t}^m, V_{b,s}^n) = O(N^{-1+\epsilon}), \quad (62)$$

$$\text{Corr}(V_{a,t}^m, V_{b,s}^n) = O(N^{-1+\epsilon}), \quad (63)$$

$$\text{Cov}((V_{a,t} - V_{b,t})^m, (V_{c,s} - V_{d,s})^n) = O(N^{-1+\epsilon}), \quad (64)$$

$$\text{Corr}((V_{a,t} - V_{b,t})^m, (V_{c,s} - V_{d,s})^n) = O(N^{-1+\epsilon}). \quad (65)$$

To go from these bounds on population variances and covariances to bounds on sample variances and covariances, we use the following lemmas:

Lemma 10 *Assume x_1, \dots, x_N are identically distributed (but not necessarily independent) random variables, with mean μ and (finite) variance σ^2 . Also, assume that $\text{Corr}(x_i, x_j) \leq CN^{-\alpha}$, $\alpha > 0$. Then the sample mean*

$$\bar{x}_N \stackrel{\text{def}}{=} \frac{1}{N} \sum_{i=1}^N x_i$$

satisfies

$$\bar{x}_N = \mu + O_s(\max(\sigma, 1)N^{-\beta/3}),$$

where $\beta = \min(\alpha, 1)$. Here, μ and σ are allowed to depend on N . Moreover, if σ is independent of N , the formula reduces to

$$\bar{x}_N = \mu + O_s(N^{-\beta/3}).$$

Proof of Lemma 10: It is clear that $\text{Var}(\bar{x}_N) = \frac{\sigma^2}{N^2} (N + \sum_{i,j \neq i} \text{Corr}(x_i, x_j)) \leq C^2 \sigma^2 N^{-\beta}$. Chebyshev's inequality then immediately implies that

$$\mathbb{P}(|\bar{x}_N - \mu| \geq C\sigma N^{-\beta/2} K) \leq K^{-2},$$

and by choosing $K = N^{\beta/6}$ the result follows. ■

Lemma 11 *If*

$$\bar{x}_N = \mu + O_s(N^{-\alpha+\epsilon})$$

for all $\epsilon > 0$, then, for any fixed $m > 1$, such that $\mu^{m-1} = o(N^\epsilon)$ for all $\epsilon > 0$,

$$\bar{x}_N^m = \mu^m + O_s(N^{-\alpha+\epsilon}).$$

Proof of Lemma 11: The proof is by induction. Assume that the result has been proved for all $m = 1, 2, \dots, M$, i.e.,

$$\mathbb{P}(|\bar{x}_N^m - \mu^m| \geq C_m N^{-\alpha+\epsilon}) \leq C_m N^{-\alpha+\epsilon}, \quad m = 1, 2, \dots, M.$$

We expand

$$|\bar{x}_N^{M+1} - \mu^{M+1}| = |\bar{x}_N - \mu| \left| \sum_{i=0}^M a_i \mu^i \bar{x}_N^{M-i} \right|,$$

for some constants a_i , $i = 1, \dots, M$. Now, clearly under our induction assumption,

$$\mathbb{P}(|\bar{x}_N^m| \geq 2\mu^m) \leq N^{-\alpha+\epsilon}, \quad m = 1, \dots, M$$

for large enough N . Therefore,

$$\mathbb{P} \left(\left| \sum_{i=0}^M a_i \mu^i \bar{x}_N^{M-i} \right| \geq C' \mu^M \right) \leq C' N^{-\alpha+\epsilon},$$

for some $C' > 0$. Moreover,

$$\mathbb{P}(|\bar{x}_N - \mu| \geq C_1 N^{-\alpha+\epsilon}) \leq C_1 N^{-\alpha},$$

and since if $\left| \sum_{i=0}^M a_i \mu^i \bar{x}_N^{M-i} \right| \leq C' \mu^M$ and $|\bar{x}_N - \mu| \leq C_1 N^{-\alpha}$, then

$$|\bar{x}_N^{M+1} - \mu^{M+1}| = |\bar{x}_N - \mu| \left| \sum_{i=0}^M a_i \mu^i \bar{x}_N^{M-i} \right| \leq C_1 N^{-\alpha+\epsilon} C' \mu^M \leq C'' N^{-\alpha+2\epsilon}.$$

Finally,

$$\mathbb{P} \left(\left| \sum_{i=0}^M a_i \mu^i \bar{x}_N^{M-i} \right| \leq C' \mu^M \cap |\bar{x}_N - \mu| \leq C_1 N^{-\alpha} \right) \geq 1 - (C_1 + C') N^{-\alpha} = 1 - C''' N^{-\alpha}.$$

Setting $C_{M+1} = \max(C'', C''')$, and recalling that $\epsilon > 0$ was arbitrary, the lemma follows. \blacksquare

Lemma 12 *Assume x_1, \dots, x_N are identically distributed (but not necessarily independent) random variables, with mean μ , variance σ^2 , and finite fourth moments. Assume that $|\text{Corr}(x_i, x_j)| \leq CN^{-\alpha}$ and that $|\text{Corr}((x_i - x_j)^2, (x_k - x_n)^2)| \leq CN^{-\alpha}$, where $\alpha > 0$, and i, j, k, n are different indexes. Then the sample variance*

$$s_N^2 \stackrel{\text{def}}{=} \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x}_N)^2$$

satisfies

$$s_N^2 = \sigma^2 + O_s(N^{-2\beta/3}),$$

where $\beta = \min(\alpha, 1)$.

Proof of Lemma 12: We rewrite

$$s_N^2 \stackrel{\text{def}}{=} \frac{1}{N(N-1)} \sum_{i,j} \frac{1}{2} (x_i - x_j)^2.$$

We have

$$E[(x_i - x_j)^2] = E[((x_i - \mu) - (x_j - \mu))^2] = \text{Var}(x_i) + \text{Var}(x_j) - 2\text{Cov}(x_i, x_j),$$

implying that

$$E[s_N^2] = \sigma^2 - \overline{\text{Cov}(x_i, x_j)} = \sigma^2(1 - r) \stackrel{\text{def}}{=} \bar{\sigma}^2, \quad r = O(N^{-\alpha}),$$

in turn leading to

$$\text{Var}(s_N^2) = \left(\frac{1}{N(N-1)} \right)^2 E \left[\left(\sum_{i,j} \left(\frac{1}{2} (x_i - x_j)^2 - \bar{\sigma}^2 \right) \right)^2 \right].$$

We expand the square, and separate

- $N(N-1)$ terms on the form:

$$\left(\frac{1}{N(N-1)} \right)^2 E \left[\left(\frac{1}{2} (x_i - x_j)^2 - \bar{\sigma}^2 \right)^2 \right].$$

The sum of these terms will thus be of $O(N^{-2})$.

- $N(N-1)(N-2)$ terms on the form:

$$\left(\frac{1}{N(N-1)} \right)^2 E \left[\left(\frac{1}{2} (x_i - x_j)^2 - \bar{\sigma}^2 \right) \left(\frac{1}{2} (x_i - x_k)^2 - \bar{\sigma}^2 \right) \right].$$

The sum of these terms will thus be of $O(N^{-1})$.

- $N^4 - 3N^3 - N^2 + N$ terms on the form:

$$\left(\frac{1}{N(N-1)} \right)^2 E \left[\left(\frac{1}{2} (x_i - x_j)^2 - \bar{\sigma}^2 \right) \left(\frac{1}{2} (x_k - x_n)^2 - \bar{\sigma}^2 \right) \right].$$

Since

$$\begin{aligned} \left| E \left[\left(\frac{1}{2} (x_i - x_j)^2 - \bar{\sigma}^2 \right) \left(\frac{1}{2} (x_k - x_n)^2 - \bar{\sigma}^2 \right) \right] \right| &= \left| \text{Corr} \left(\frac{1}{2} (x_i - x_j)^2, \frac{1}{2} (x_k - x_n)^2 \right) \right| \\ &\times \text{Var} \left(\frac{1}{2} (x_i - x_j)^2 \right) \\ &= O(N^{-\alpha}), \end{aligned}$$

the sum of these terms will be of $O(N^{-\alpha})$.

Thus, altogether, $\text{Var}(s_N^2) = O(N^{-\beta})$. As in Lemma 10, the result then follows from Chebyshev's inequality. \blacksquare

Lemma 13 *Assume x_1, \dots, x_N , and y_1, \dots, y_N , are identically distributed random variables, with means μ_X, μ_Y , variances σ_X^2, σ_Y^2 , covariance $\sigma_{XY} = \text{Cov}(x_i, y_i)$ for all i , and finite fourth moments. Assume that $|\text{Corr}(x_i, y_j)| \leq CN^{-\alpha}$, and that $|\text{Corr}((x_i - x_j)^2, (y_k - y_n)^2)| \leq CN^{-\alpha}$, where $\alpha > 0$, and i, j, k, n are different indexes. Then the sample covariance*

$$s_{XY} \stackrel{\text{def}}{=} \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x}_N)(y_i - \bar{y}_N)$$

satisfies

$$s_{XY} = \sigma_{XY} + O_s(N^{-2\beta/3}),$$

where $\beta = \min(\alpha, 1)$.

Proof of Lemma 13: Identical to the Proof of Lemma 12. \blacksquare

Lemma 12, together with (60) and (63) therefore implies (51), and Lemma 13 together with (61) and (65) implies (52). We have shown Lemma 9. \blacksquare

Plugging (51-52) into (48), we get

$$\begin{aligned} \rho[D_N, \Pi_N] &= \frac{S_1^2 + \sum_{i=2}^T \zeta_i S_{i,1}}{S_1 \sqrt{S_1^2 + \sum_{i=2}^T \zeta_i^2 S_i^2 + \sum_{i=1, j>i}^T 2\zeta_i \zeta_j S_{j,i}}} + O_s(N^{-1/3+\epsilon}) \\ &= \frac{q(1 + a_2(q^{-1} + q^{-2}) + a_3(q^{-2} + q^{-3}) + a_4q^{-3} + q^{-4}P_1(q^{-1}))}{q(1 + a_2^2(q^{-2} + 3q^{-3}) + 2a_2(q^{-1} + q^{-2}) + 2a_3(q^{-2} + q^{-3}) + 2a_4q^{-3} + 2a_2a_3q^{-3} + q^{-4}Q(q^{-1}))^{1/2}} \\ &+ O_s(N^{-1/3+\epsilon}), \\ &= \frac{1 + a_2(q^{-1} + q^{-2}) + a_3(q^{-2} + q^{-3}) + a_4q^{-3} + q^{-4}P_1(q^{-1})}{1 + a_2^2(q^{-2} + 3q^{-3}) + 2a_2(q^{-1} + q^{-2}) + 2a_3(q^{-2} + q^{-3}) + 2a_4q^{-3} + 2a_2a_3q^{-3} + q^{-4}Q(q^{-1})^{1/2}} \\ &+ O_s(N^{-1/3+\epsilon}), \end{aligned} \tag{66}$$

where P and Q are polynomials finite orders, and we define $a_1 = 1$. A Taylor expansion of (66) around $x = 0$, where $x = q^{-1}$ now yields,

$$\frac{1 + a_2(q^{-1} + q^{-2}) + a_3(q^{-2} + q^{-3}) + a_4q^{-3} + q^{-4}P_1(q^{-1})}{1 + a_2^2(q^{-2} + 3q^{-3}) + 2a_2(q^{-1} + q^{-2}) + 2a_3(q^{-2} + q^{-3}) + 2a_4q^{-3} + 2a_2a_3q^{-3} + q^{-4}Q(q^{-1})^{1/2}} = 1 - cq^{-3} + O(q^{-4})$$

altogether leading to

$$\rho_D = 1 - cq^{-3} + O_s(q^{-4}), \tag{67}$$

where $c > 0$ depends on a_2, \dots, a_T , and when $a_t \rightarrow 1$, $i = 2, \dots, a_T$, then $c \rightarrow 1/2$. Of course, this immediately implies that

$$\mathbb{P}(|1 - \rho_D - cq^{-3}| \geq Cq^{-4}) \leq Cq^{-4}, \tag{68}$$

for large N for some bounded constant $C > 0$.

Now, since correlations are bounded between -1 and 1, for general sequences of random variables, a_N , and b_N , if

$$\mathbb{P}(|1 - \rho[a_N, b_N] - f(N)| \geq o(f(N))) = o(f(N)),$$

where $\lim_{N \rightarrow \infty} f(N) = 0$, then

$$1 - E(\rho[a_N, b_N]) = f(N)(1 + o(1)).$$

Thus, from (68), it follows that

$$1 - E[\rho_D] = cq^{-3}(1 + o(1)) \sim q^{-3}. \tag{69}$$

A similar expansion of $V_N^1 + \zeta_2 V_N^2$ gives

$$\rho[V_{a,1} + \zeta_2 V_{a,2}, \Pi_N] = 1 - cq^{-5} + O_s(q^{-6}),$$

where $c = 1/2$ when $a_i \rightarrow 1$, $i = 2 \dots, a_T$, which then leads to

$$1 - E[\rho_K] \sim q^{-5}. \quad (70)$$

This establishes the relationship $E[\rho_K] > E[\rho_D]$ for large N .

We next study $F = \frac{1}{C}$, the average distance from an agent to all other agents, and show that

$$1 - E[\rho[D, F]] \sim q^{-1}. \quad (71)$$

We note that (53,54) imply that for $i > 1$, and fixed α ,

$$\begin{aligned} \text{Corr}(Z_1, Z_1 + \alpha Z_i) &= \frac{1 + \alpha \frac{\text{Cov}(Z_1, Z_i)}{\text{Var}(Z_1)}}{\sqrt{1 + 2\alpha \frac{\text{Cov}(Z_1, Z_i)}{\text{Var}(Z_1)} + \alpha^2 \frac{\text{Var}(Z_i)}{\text{Var}(Z_1)}}} \\ &= \frac{1 + \alpha q^{i-1}}{\sqrt{1 + 2\alpha q^{i-1} + \alpha^2 (q^{2i-2} + q^{2i-3} + O(q^{2i-4}))}} \\ &= \frac{1 + (\alpha q)^{-1}}{\sqrt{1 + 2(\alpha q)^{-1} + q^{-1} + O(q^{-2})}} \\ &= 1 - \frac{1}{2}q^{-1} + O(q^{-2}), \end{aligned}$$

in turn implying that for $\alpha_1, \dots, \alpha_R$, such that $\alpha_1 / (\sum_i \alpha_i) \leq 1 - \epsilon$, $\epsilon > 0$,

$$\text{Corr}\left(Z_1, \sum_{i=1}^R Z_i\right) = 1 - \frac{1}{2}q^{-1} + O(q^{-2}). \quad (72)$$

Now, from Theorem 10.10 in Bollobas (2001), it follows that if we define $R = \lceil \log(N) / \log(c) \rceil + 1$, with probability greater than $1 - N^{-k}$ for arbitrary k , the maximum distance between any two agents is either $R - 1$ or R . We will therefore have that the average distance between an agent, a , and all other agents is

$$\begin{aligned} F_a &= \frac{1}{N} \left(V_{a,1} + \sum_{i=2}^{R-1} i(V_{a,i} - V_{a,i-1}) + R \left(N - (V_{a,1} + \sum_{i=2}^{R-1} (V_{a,i} - V_{a,i-1})) \right) \right) \\ &= R - \frac{R-1}{N} V_{a,1} - \sum_{i=2}^{R-1} \frac{R-i}{N} (V_{a,i} - V_{a,i-1}). \end{aligned}$$

Now, from (72), if we replace $V_{a,i} - V_{a,i-1}$ with $Z_{a,i}$,

$$\begin{aligned} \tilde{F}_a &\stackrel{\text{def}}{=} \frac{1}{N} \left(Z_{a,1} + \sum_{i=2}^{R-1} (i Z_i) + R \left(N - \left(Z_{a,1} + \sum_{i=2}^{R-1} Z_i \right) \right) \right) \\ &= R - \frac{R-1}{N} Z_{a,1} - \sum_{i=2}^{R-1} \frac{R-i}{N} Z_{a,i} \end{aligned}$$

we have

$$\text{Corr}(V_{a,1}, -\tilde{F}_a) = 1 - \frac{1}{2}q^{-1} + O(q^{-2}).$$

A similar argument as that leading to (67) therefore implies that

$$\text{Corr}(V_{a,1}, -F_a) = 1 - \frac{1}{2}q^{-1} + O(q^{-2}),$$

and

$$\rho[D, -F] = 1 - \frac{q}{2} + O_s(q^{-2}), \quad (73)$$

and therefore, following a similar argument as for ρ_D , that

$$E[\rho[D, -F]] = 1 - \frac{q}{2} + O(q^{-2}).$$

We use the following triangle-inequality like lemma:

Lemma 14 *Assume $\rho(a, b) = 1 - \alpha$, and $\rho(a, c) = 1 - \beta$, with $\beta > \alpha$. Then $\rho(b, c) \leq 1 - (\sqrt{\beta} - \sqrt{\alpha})^2$.*

Proof of Lemma 14: Because of the simple renormalizations, $a \mapsto (a - E[a])/\sigma(a)$, $b \mapsto (b - E[b])/\sigma(b)$, and $c \mapsto (c - E[c])/\sigma(c)$, we can without loss of generality assume that a, b , and c have zero expectations and unit variances. All correlations can then be expressed as, $\rho(a, b) = E[ab]$, etc. We introduce the metric $d(a, b) = \sqrt{E[(a - b)^2]}$, and we then have the triangle inequality $d(a, c) \leq d(a, b) + d(b, c)$, leading to $d(b, c) \geq d(a, c) - d(a, b)$. We have $d(a, b)^2 = E[(a - b)^2] = E[a^2] + E[b^2] - 2E[ab] = 2 - 2(1 - \alpha) = 2\alpha$, and similarly $d(a, c)^2 = 2\beta$. Finally, $d(b, c)^2 = 2 - 2\rho(b, c)$ which via the triangle inequality leads to $2 - 2\rho(b, c) \geq (\sqrt{2\beta} - \sqrt{2\alpha})^2 = 2(\sqrt{\beta} - \sqrt{\alpha})^2$, leading to the result. We are done. \blacksquare

Equations (68) and (73), together with Lemma 14, where $a = D$, $b = -F$, $c = \Pi$, now implies that if $1 - \rho[D, \Pi] \leq cq^{-3}$, and $1 - \rho[D, -F] \geq cq^{-1}$, then

$$1 - \rho[-F, \Pi] = 1 - \rho_F > c(1 - \epsilon)q^{-1}$$

for large N , and thus,

$$1 - E[\rho_F] \geq c(2 - \epsilon)q^{-1},$$

implying the third part of the Theorem, $E[\rho_D] > E[\rho_F]$.

Finally, for \hat{C} , we first note that

$$\begin{aligned} E[V_{a,1}] &= q + O(N^{-1+\epsilon}), \\ \text{Var}[V_{a,1}] &= q + O(N^{-1+\epsilon}), \\ E[F_a] &= R(1 + O(q^{-1})), \\ \text{Var}[F_a] &= \frac{R^2}{q}(1 + O(q^{-1})), \end{aligned}$$

and more generally that

$$E[(1 - F/R)^i] = q^{-i}(c_i + O(q^{-1}))$$

for any fixed $i \geq 1$, where c_i is a (finite) constant, and $c_2 = 1$. We have

$$\begin{aligned} \text{Corr}(V_{a,1}, \hat{C}) &= \text{Corr}\left(V_{a,1}, \frac{1}{F}\right) \\ &= \text{Corr}\left(V_{a,1}, \frac{1}{R - (R - F)}\right) \\ &= \text{Corr}\left(V_{a,1}, \frac{1}{1 - \left(1 - \frac{F}{R}\right)}\right) \\ &= \text{Corr}\left(V_{a,1}, -\frac{F}{R} + \sum_{i \geq 2} (-1)^i \left(1 - \frac{F}{R}\right)^i\right) \\ &= 1 - \frac{q}{2} + O(q^{-2}) + \sum_{i \geq 2} d_i q^{-i} \\ &= 1 - \frac{q}{2} + O(q^{-2}). \end{aligned}$$

A similar argument as for ρ_F , again using Lemma 14, then implies the final part of the theorem, $E[\rho_D] > E[\rho_{\hat{C}}]$. We are done. \blacksquare

Proof of Theorem 5: For the first part, we note that since $(p_t - p_{t-1})$ is independent of p_{t-1} (given publicly available information), it follows that the price volatility between $t-1$ and t is equal to $(\Sigma_{XX})_{11}$, $\sigma_{p,t}^2 = (\Sigma_{XX})_{11}|_{t-1} = \frac{y_t}{(\tau_v + Y_t)(\tau_v + Y_{t-1})}$. with the convention $Y_0 = 0$. Also, the final period volatility of $v - p_T$ is $\sigma_{p,T+1}^2 = \frac{1}{\tau_v + Y_T}$. This proves the first part of the theorem.

For the second part, we first note that from the definition of y_1 , it follows that $y_1 = \tau_v \frac{k_1}{1-k_1}$ and—defining $K_t = \sum_{i=1}^t k_i$ —a simple induction argument further shows $y_t = \tau_v \frac{k_t}{(1-K_{t-1})(1-K_t)}$, $t = 1, \dots, T-1$, and $y_T = \tau_v \frac{1-K_T}{K_T(1-K_{T-1})}$. We note that all the y_1, \dots, y_T are all well defined.

Next, we back out the connectedness that is needed to be consistent with the y 's. We have $A_1 = \sqrt{\frac{y_1}{\tau_u}}$, $A_t = A_{t-1} + \sqrt{\frac{y_t}{\tau_u}}$, leading to $\bar{V}_1 \stackrel{\text{def}}{=} \frac{\sum_a V_{a,1}}{N} = \frac{\gamma}{\tau} \sqrt{\frac{y_1}{\tau_u}}$, $\Delta \bar{V}_t \stackrel{\text{def}}{=} \bar{V}_t - \bar{V}_{t-1} = \frac{\gamma}{\tau} \sqrt{\frac{y_t}{\tau_u}}$. Thus, if we can replicate, arbitrarily closely, any sequence of diffusions, through which the average number of signals, \bar{V}_t , increases over time, then we can generate any y_t , and thereby any volatility structures. We note that γ is a free parameter that allows us to scale the network to arbitrary sizes. The result now follows from the following lemma:

Lemma 15 *For any T , there are networks of size N , such that $\bar{V}_{T'} = (1 + o(1))\bar{V}_T$ for $T' > T$, and $\bar{V}_{T'} = \frac{1}{N(1+o(1))}\bar{V}_T$ for $T' < T$.*

This lemma thus states that we can always find a (possibly large) network such that very little happens before and after time T , with respect to information diffusion. The result follows immediately: For $T = 1$, a tightly-knit network would have these properties. For $T = 2$, a large star network. For $T = 3$, a star-like network with $N^2 + N$ nodes, in which there are N tightly-knit nodes in the center, each connected to N peripheral agents. For even $T \geq 4$, adding longer distance to the $T = 2$ (star) network, and for odd $T \geq 5$, adding longer distances to the $T = 3$ network will generate these properties. Let's call such a network a T -network.

Finally, any sequence of $\frac{V_{t+1}}{V_t}$, $t = 1, \dots, T$ can be generated by choosing a network with many disjoint $1-$, $2-$, \dots , $T-$ networks in such a way so that the relative sizes of the networks match the fractions.

We are done. ■

Proof of Lemma 3:

i): Let us study the function g , defined by $g_i = f_i$, $i = 0, \dots, N$, and $g_{N+1} = 0$, which has the same autocorrelation function as f , for $\tau \leq N$, and which always has a maximum at an interior point when f is unimodal, nonnegative, and $f_0 = 0$. Let us denote that maximum by m , i.e., $g_m \geq \max(g_{m-1}, g_{m+1})$. Of course, the result follows trivially if $f \equiv 0$, so we assume that $f_i > 0$ for some i .

We use the following lemma

Lemma 16 *Consider a sequence f_0, f_1, \dots, f_{N+1} , such that $f_0 = f_{N+1} = 0$. Then*

$$\sum_{i=0}^N f_i(f_{i+1} - f_i) - (f_{k+1} - f_k)(f_k - f_{k-1}) \leq 0$$

for any $1 \leq k \leq N$.

Proof of Lemma 16: The summation by parts rule implies that $\sum_{i=0}^N f_i(f_{i+1} - f_i) = -\sum_{i=0}^N f_{i+1}(f_{i+1} - f_i)$, in turn implying that $\sum_{i=0}^N f_i(f_{i+1} - f_i) = -\frac{1}{2} \sum_{i=0}^N (f_{i+1} - f_i)^2$. We now have

$$\sum_{i=0}^N f_i(f_{i+1} - f_i) = -\frac{1}{2} \sum_{i=0}^N (f_{i+1} - f_i)^2 \leq -\frac{1}{2} ((f_k - f_{k-1})^2 + (f_{k+1} - f_k)^2) \leq \frac{1}{2} 2(f_{k+1} - f_k)(f_k - f_{k-1}),$$

where the second inequality follows from the inequality $(x+y)^2 = x^2 + y^2 + 2xy \geq 0$. Thus, $\sum_{i=0}^N f_i(f_{i+1} - f_i) + (f_{k+1} - f_k)(f_k - f_{k-1}) \leq 0$, as claimed. This completes the proof of Lemma 16. ■

Since R_τ is symmetric around $\tau = 0$, we restrict our attention to the region $\tau \geq 0$, showing that R_τ is nonincreasing in this region. This will immediately imply i). We prove the result by contradiction: Assume that there is a τ , such that $R_{\tau+1} - R_\tau > 0$. Given the definition of R_τ , this means that

$$\sum_{i=0}^{N+1-\tau} g_i(g_{i+\tau} - g_i) > 0 \tag{74}$$

for some τ . By zero padding (adding extra zeros to) g , we can without loss of generality assume that the maximum occurs at $m = n\tau$ for some integer n , and that $N+1-\tau = M\tau - 1$, i.e., by studying the function $g'_i = 0$, $i < v$, $g'_i = g_{i-v}$, $0 \leq i-v \leq N+1-\tau$, $g'_i = 0$, $N+1-\tau < i-v \leq M\tau - 1$. Thus, it must be that

$$\left(\sum_{i=0}^{(m-1)\tau-1} g'_i(g'_{i+\tau} - g'_i) \right) + \left(\sum_{i=(m-1)\tau}^{m\tau-1} g'_i(g'_{i+\tau} - g'_i) \right) + \left(\sum_{i=m\tau}^{M\tau-1} g'_i(g'_{i+\tau} - g'_i) \right) > 0.$$

Now, because g'_i is increasing for $i \leq m\tau$, and decreasing for $i \geq m\tau$, we can replace g'_i with $g'_{r_i\tau}$, where $r_i = \tau(\lfloor i/\tau \rfloor + 1)$ in the first and third term, and by $g'_{(m-1)\tau}$ for the middle term, leading to

$$\begin{aligned} \sum_{i=0}^{m-2} g'_{i\tau} \sum_{j=0}^{\tau-1} (g'_{i\tau+j+\tau} - g'_{i\tau+j}) &\geq \sum_{i=0}^{(m-1)\tau-1} g'_i (g'_{i+\tau} - g'_i), \\ g'_{(m-1)\tau} \sum_{j=0}^{\tau-1} (g'_{m\tau+j+\tau} - g'_{m\tau+j}) &\geq \sum_{i=(m-1)\tau}^{m\tau-1} g'_i (g'_{i+\tau} - g'_i), \\ \sum_{i=m}^{M-1} g'_{i\tau} \sum_{j=0}^{\tau-1} (g'_{i\tau+j+\tau} - g'_{i\tau+j}) &> \sum_{i=m\tau}^{M\tau-1} g'_i (g'_{i+\tau} - g'_i). \end{aligned}$$

But since $\sum_{j=0}^{\tau-1} g_{k+j+1} - g_{k+j} = g_{k+\tau} - g_k$, this implies that

$$\left(\sum_{i=0}^{m-2} g'_{i\tau} (g'_{(i+1)\tau} - g'_{i\tau}) \right) - \left(g'_{(m-1)\tau} (g'_{m\tau} - g'_{(m+1)\tau}) \right) + \left(\sum_{i=m}^{M-1} g'_{i\tau} (g'_{(i+1)\tau} - g'_{i\tau}) \right) > 0,$$

in turn leading to

$$\left(\sum_{i=0}^{M-1} g'_{i\tau} (g'_{(i+1)\tau} - g'_{i\tau}) \right) + (g'_{(m-1)\tau} - g'_{m\tau})(g'_{(m+1)\tau} - g'_{m\tau}) > 0.$$

Now, defining $h_i = g'_{i\tau}$, $0 \leq i \leq M$, it follows that $h_0 = h_M = 0$, that h is positive and unimodal, and that this in turn implies that

$$\left(\sum_{i=0}^{M-1} h_i (h_{i+1} - h_i) \right) - (h_m - h_{m-1})(h_{m+1} - h_m) > 0.$$

However, from Lemma 16 no such sequence can exist, and thus neither can a sequence g satisfying (74). So, R is nonincreasing on the positive axis, and it must therefore be unimodal (with maximum at 0). We have shown i).

ii): It is easy to show that $(\Delta^n R)(\tau) = \sum_{i=1}^{N-\tau} f_i (\Delta^n f)_i$. Thus, if $\Delta^n f$ is positive (negative), so is R_τ . This, in turn, implies that R_τ has at most $n - 1$ turning points. \blacksquare

Proof of Theorem 6: The proof is based on the following standard lemma:

Lemma 17 Assume a normally distributed random variable, $y \sim N(\mu, \sigma^2)$. Then $E[|y|] = \sigma \sqrt{\frac{2}{\pi}} e^{-\frac{\mu^2}{2\sigma^2}} + \mu(1 - 2\Phi(-\mu/\sigma))$, where Φ is the cumulative normal distribution of a standard normal variable.

We note that from (5) and given that $v = \bar{v} + \eta$, it follows that agent a 's net time- t demand is

$$\begin{aligned} \gamma_a \Delta x_{a,t} &= Q_{a,t}(z_{a,t} - p_t) - Q_{a,t}(z_{a,t-1} - p_{t-1}) \\ &= Q_{a,t} \left(\bar{v} + \eta + \frac{Q_{a,t-1}}{Q_{a,t}} \xi_{a,t-1} + \frac{q_t}{Q_{a,t}} e_{a,t} - \left(\frac{\tau_v}{\tau_v + Y_t} \bar{v} + \frac{Y_t}{\tau_v + Y_t} (\bar{v} + \eta) + \frac{1}{\tau_v + Y_t} \sum_{i=1}^t (A_i - A_{i-1}) \tau_{u_i} u_i \right) \right) \\ &\quad - Q_{a,t-1} \left(\bar{v} + \eta + \xi_{a,t-1} + - \left(\frac{\tau_v}{\tau_v + Y_{t-1}} \bar{v} + \frac{Y_{t-1}}{\tau_v + Y_{t-1}} (\bar{v} + \eta) + \frac{1}{\tau_v + Y_{t-1}} \sum_{i=1}^{t-1} (A_i - A_{i-1}) \tau_{u_i} u_i \right) \right) \\ &= \underbrace{q_{a,t} e_{a,t}}_{\sim N(0, q_{a,t})} - \underbrace{\left(\frac{Q_{a,t}}{\tau_v + Y_t} - \frac{Q_{a,t-1}}{\tau_v + Y_{t-1}} \right) \left(\tau_v \eta - \sum_{i=1}^{t-1} (A_i - A_{i-1}) \tau_{u_i} u_i \right) - \frac{Q_{a,t}}{\tau_v + Y_t} (A_t - A_{t-1}) \tau_{u_t} u_t}_{\sim N(0, \hat{r}_{a,t}^2)} \end{aligned}$$

where

$$\hat{r}_t^2 = \left(\frac{Q_{a,t}}{\tau_v + Y_t} - \frac{Q_{a,t-1}}{\tau_v + Y_{t-1}} \right)^2 (\tau_v + Y_{t-1}) + \left(\frac{Q_{a,t}}{\tau_v + Y_t} \right)^2 y_t^2 = \frac{Q_{a,t-1}^2}{\tau_v + Y_{t-1}} - \frac{Q_{a,t}^2}{\tau_v + Y_t} + \frac{q_{a,t}^2}{\tau_v + Y_t}.$$

Recalling that $q_{a,t} = \tau \Delta V_{a,t}$ and $Q_{a,t} = \tau V_{a,t}$, this leads to

$$\gamma_a \Delta x_{a,t} = \sqrt{\tau \Delta V_{a,t}} \left(\omega_{a,t} + \frac{r_{a,t} \sqrt{\tau}}{\sqrt{\Delta V_{a,t}}} \xi_t \right),$$

where $\omega_{a,t} \sim N(0, 1)$ is independent of $\xi_t \sim N(0, 1)$ and across agents,

$$r_t^2 = \frac{V_{a,t-1}^2}{\tau_v + Y_{t-1}} - \frac{V_{a,t}^2}{\tau_v + Y_t} + \frac{\Delta V_{a,t}^2}{\tau_v + Y_t},$$

when $\Delta V_{a,t} > 0$, and

$$\gamma_a \Delta x_{a,t} = r_{a,t} \tau \xi_t,$$

when $\Delta V_{a,t} = 0$. Assuming that $\Delta V_{a,t} > 0$, we note that $\gamma_a \frac{\Delta x_{a,t}}{\tau \Delta V_{a,t}} \Big| \xi_t \sim N\left(\frac{r_{a,t} \sqrt{\tau}}{\sqrt{\Delta V_{a,t}}} \xi_t, 1\right)$. The law of large numbers, together with Lemma 17, then in turn implies that, conditioned on ξ_t ,

$$\gamma_a \frac{1}{M} \sum_a |\Delta x_{a,t}| \rightarrow_{a.s.} \sqrt{\tau \Delta V_{a,t}} \left(\sqrt{\frac{2}{\pi}} e^{-\frac{\tau r_{a,t}^2 \xi_t^2}{2 \Delta V_{a,t}}} + \frac{r_{a,t} \sqrt{\tau}}{\sqrt{\Delta V_{a,t}}} \xi_t \left(1 - 2\Phi\left(-\frac{r_{a,t} \sqrt{\tau}}{\sqrt{\Delta V_{a,t}}} \xi_t\right) \right) \right).$$

It follows immediately that the unconditional expectation of the first term (not conditioning on ξ_t) is $\sqrt{\frac{2}{\pi}} \frac{\sqrt{\tau \Delta V_{a,t}}}{\sqrt{\frac{\tau r_{a,t}^2}{\Delta V_{a,t}} + 1}} = \sqrt{\frac{2}{\pi}} \frac{\sqrt{\tau \Delta V_{a,t}}}{\sqrt{\tau r_{a,t}^2 + \Delta V_{a,t}}}$. For the second term, we use the fact that $E[y\Phi(ay)] = \sqrt{\frac{1}{2\pi}} \frac{a}{\sqrt{a^2+1}}$, for a random variable $y \sim N(0, 1)$, to get

$$\sqrt{\tau \Delta V_{a,t}} E \left[\frac{r_{a,t} \sqrt{\tau}}{\sqrt{\Delta V_{a,t}}} \xi_t \left(1 - 2\Phi\left(-\frac{r_{a,t} \sqrt{\tau}}{\sqrt{\Delta V_{a,t}}} \xi_t\right) \right) \right] = \sqrt{\frac{2}{\pi}} \sqrt{\tau \Delta V_{a,t}} \frac{\frac{r_{a,t}^2 \tau}{\Delta V_{a,t}}}{\sqrt{1 + \frac{r_{a,t}^2 \tau}{\Delta V_{a,t}}}} = \sqrt{\frac{2}{\pi}} \sqrt{\tau} \frac{\tau r_{a,t}^2}{\sqrt{\tau r_{a,t}^2 + \Delta V_{a,t}}}.$$

Summing the two terms together, we get

$$E \left[\frac{1}{M} \sum_a |\Delta x_{a,t}| \right] \rightarrow_{a.s.} \frac{\tau}{\gamma_a} \sqrt{\frac{2}{\pi} \left(r_{a,t}^2 + \frac{\Delta V_{a,t}}{\tau} \right)}.$$

We note that this formula also holds when $\Delta V_{a,t} = 0$, since $E|r_{a,t} \tau \xi_t| = \tau \sqrt{\frac{2}{\pi} r_{a,t}^2}$. This finally leads to (20)

$$\begin{aligned} X_t &= \frac{\tau}{N} \sum_{a=1}^N \frac{1}{\gamma_a} \sqrt{\frac{2}{\pi} \left(r_{a,t}^2 + \frac{\Delta V_{a,t}}{\tau} \right)} \\ &= \frac{\tau}{N} \sum_{a=1}^N \frac{1}{\gamma_a} \sqrt{\frac{2}{\pi} \left(\frac{V_{a,t-1}^2 y_t}{(\tau_v + Y_{t-1})(\tau_v + Y_t)} - \frac{2V_{a,t-1} \Delta V_{a,t}}{\tau_v + Y_t} + \frac{\Delta V_{a,t}}{\tau} \right)} \\ &= \frac{\tau}{N} \sum_{a=1}^N \frac{1}{\gamma_a} \sqrt{\frac{2}{\pi} \left(\frac{V_{a,t-1}^2}{\tau_v + Y_{t-1}} - \frac{V_{a,t}^2}{\tau_v + Y_t} + \frac{\Delta V_{a,t}^2}{\tau_v + Y_t} + \frac{\Delta V_{a,t}}{\tau} \right)}. \end{aligned}$$

Now, if one of the $\gamma_a \rightarrow 0$, then agent $V_{a,t}$ will determine A_t , y_t , and Y_t . In this case, we get

$$X_t = \frac{\tau}{N} \sum_{a=1}^N \frac{1}{\gamma_a} \sqrt{\frac{2}{\pi} \left(\frac{V_{t-1}^2 \frac{\tau_u \tau^2}{\gamma_a^2} \Delta V_t^2}{\left(\tau_v + \frac{\tau_u \tau^2}{\gamma_a^2} \sum_{i=1}^{t-1} \Delta V_i^2 \right) \left(\tau_v + \frac{\tau_u \tau^2}{\gamma_a^2} \sum_{i=1}^t \Delta V_i^2 \right)} - \frac{2V_{t-1} \Delta V_t}{\tau_v + \frac{\tau_u \tau^2}{\gamma_a^2} \sum_{i=1}^t \Delta V_i^2} + \frac{\Delta V_t}{\tau} \right)}.$$

Using the fact that $V_t = \sum \Delta V_t$, leading to the inequality $V_t^2 \leq t \sum_{i=1}^t \Delta V_i^2$ (from $E[x^2] \geq E[x]^2$), it follows that for large ΔV_t , the third term will be dominant, and therefore any sequence of X_t can be generated by choosing ΔV_t appropriately. This shows the second part of the theorem.

We are done. \blacksquare

Proof of Theorem 7: We first show the second result: We use (24,26) to define

$$\hat{R}_t = (A_t - A_{t-1})\tau_{u_t} R_t - y_t \bar{v} = y_t \eta + (A_t - A_{t-1})\tau_{u_i} u_i,$$

and, of course, $\hat{R}_1, \dots, \hat{R}_t$ can be backed out of p_1, \dots, p_t . Now, it is straightforward to use (26) to show that

$$\begin{aligned} p_t &= \bar{v} + \frac{1}{\tau_v + Y_t} \left(Y_t \eta + \sum_{i=1}^t (A_i - A_{i-1}) \tau_{u_i} u_i \right) \\ &= \bar{v} + \frac{1}{\tau_v + Y_t} \sum_{i=1}^t \hat{R}_t = \bar{v} + \frac{1}{\tau_v + Y_t} \left(Y_t \eta + \sum_{i=1}^t (A_i - A_{i-1}) \tau_{u_i} u_i \right). \end{aligned} \quad (75)$$

Using this relation, we get

$$\begin{aligned} \gamma_a \Delta x_{a,t} &= Q_{a,t}(z_{a,t} - p_t) - Q_{a,t-1}(z_{a,t-1} - p_{t-1}) \\ &= Q_{a,t} \left(\bar{v} + \eta + \frac{Q_{t-1}}{Q_{a,t}} \xi_{a,t-1} + \frac{q_t}{Q_{a,t}} e_{a,t} - \left(\bar{v} + \frac{1}{\tau_v + Y_t} \sum_{i=1}^t \hat{R}_t \right) \right) \\ &\quad - Q_{a,t-1} \left(\bar{v} + \eta + \xi_{a,t-1} - \left(\bar{v} + \frac{1}{\tau_v + Y_t} \sum_{i=1}^t \hat{R}_t \right) \right) \\ &= q_{a,t} e_{a,t} + Q_{a,t} \left(\eta - \frac{1}{\tau_v + Y_t} \left(Y_t \eta + \sum_{i=1}^t (A_i - A_{i-1}) \tau_{u_i} u_i \right) \right) \\ &\quad - Q_{a,t-1} \left(\eta - \frac{1}{\tau_v + Y_{t-1}} \left(Y_{t-1} \eta + \sum_{i=1}^{t-1} (A_i - A_{i-1}) \tau_{u_i} u_i \right) \right) \\ &= \underbrace{q_{a,t} e_{a,t}}_{\sim N(0, q_t)} + \frac{Q_{a,t}}{\tau_v + Y_t} \left(\tau_v \eta - \sum_{i=1}^t (A_i - A_{i-1}) \tau_{u_i} u_i \right) \\ &\quad - \frac{Q_{a,t-1}}{\tau_v + Y_{t-1}} \left(\tau_v \eta - \sum_{i=1}^{t-1} (A_i - A_{i-1}) \tau_{u_i} u_i \right) \\ &= \underbrace{q_{a,t} e_{a,t}}_{\sim N(0, q_t)} - \left(\frac{Q_{a,t}}{\tau_v + Y_t} - \frac{Q_{a,t-1}}{\tau_v + Y_{t-1}} \right) \underbrace{\left(\tau_v \eta - \sum_{i=1}^{t-1} (A_i - A_{i-1}) \tau_{u_i} u_i \right)}_{\sim N(0, \tau_v + Y_{t-1})} - \frac{Q_{a,t}}{\tau_v + Y_t} \underbrace{(A_t - A_{t-1}) \tau_{u_t} u_t}_{\sim N(0, y_t)} \\ &= q_{a,t} e_{a,t} - \left(\frac{Q_{a,t}}{\tau_v + Y_t} - \frac{Q_{a,t-1}}{\tau_v + Y_{t-1}} \right) f_{t-1} - \frac{Q_{a,t}}{\tau_v + Y_t} g_t, \end{aligned}$$

where $f_1 = \tau_v \eta \sim N(0, \tau_v)$, $f_t = f_{t-1} + g_t$, $g_t \sim N(0, y_t)$, $f_t | f_{t-1} \sim N(f_{t-1}, y_t)$.

We note that the coefficient in front of f_{t-1} is strictly positive. The total trading volume at times t and $t+1$ then have the form:

$$\begin{aligned} W_t &= \sum_a |\alpha_{a,t}^1 \xi_{a,t} - \alpha_{a,t}^2 f_{t-1} - \alpha_{a,t}^3 g_t|, \\ W_{t+1} &= \sum_a |\alpha_{a,t+1}^1 \xi_{a,t+1} - \alpha_{a,t+1}^2 (f_{t-1} + g_{t+1}) - \alpha_{a,t+1}^3 g_{t+1}|, \end{aligned}$$

where $\xi_{a,t}$, $\xi_{a,t+1}$, f_{t-1} , g_t , and g_{t+1} are all jointly independent, and all α 's are positive.

The result now follows from the following lemma:

Lemma 18 *Assume that \tilde{A} , \tilde{B} and \tilde{z} are independent random variables with standardized normal distributions. Then, for any $a > 0$ and $b > 0$, $Cov(|a\tilde{A} + \tilde{z}|, |b\tilde{B} + \tilde{z}|) > 0$.*

Proof of Lemma 18: First, we note that if $Cov(\tilde{X}, \tilde{Y} | \tilde{Z}_1, \dots, \tilde{Z}_n) > 0$ for all realizations of the random variables $\tilde{Z}_1, \dots, \tilde{Z}_n$, then it must be that $Cov(\tilde{X}, \tilde{Y}) > 0$ unconditionally. This follows from the law of iterated expectations, since

$$Cov(\tilde{X}, \tilde{Y}) = E[\tilde{X}\tilde{Y}] - E[\tilde{X}]E[\tilde{Y}] = E \left[E[\tilde{X}\tilde{Y}] - E[\tilde{X}]E[\tilde{Y}] | \tilde{Z} \right] = E[Cov(\tilde{X}, \tilde{Y}) | \tilde{Z}] > 0.$$

Now, define $Z_1 = a|\tilde{A}|$, $Z_2 = b|\tilde{B}|$. Then it follows that $E[|a\tilde{A} + \tilde{z}| | Z_1] = E[\max(Z_1, |\tilde{z}|)]$, $E[|b\tilde{B} + \tilde{z}| | Z_2] = E[\max(Z_2, |\tilde{z}|)]$, and $E[|a\tilde{A} + \tilde{z}||b\tilde{B} + \tilde{z}| | Z_1, Z_2] = E[\max(Z_1, |\tilde{z}|)\max(Z_2, |\tilde{z}|)]$, so

$$Cov(|a\tilde{A} + \tilde{z}|, |b\tilde{B} + \tilde{z}| | Z_1, Z_2) = Cov(\max(Z_1, |\tilde{z}|), \max(Z_2, |\tilde{z}|) | Z_1, Z_2) > 0$$

where the inequality follows from $\max(c, x)$ being a monotone transformation of x , so $\max(Z_1, |\tilde{z}|)$, and $\max(Z_2, |\tilde{z}|)$ are comonotonic. Given the argument, the positivity must then also hold unconditionally, and the lemma therefore follows.

Now, this lemma implies that all the terms that make up W_t and W_{t+1} have pairwise positive covariances, and it is indeed the case that $Cov(W_t, W_{t+1}) > 0$, showing the second result.

For the third result, we proceed as follows: It follows from (75) that

$$v - p_t = \frac{1}{\tau_v + Y_t} f_t = \eta - \frac{1}{\tau_v + Y_t} \sum_{i=1}^t \hat{R}_i. \quad (76)$$

Therefore, since $v - p_{t-1}$ is independent of p_{t-1} , the same holds for f_{t-1} (and, of course for g_t), trading volume is indeed independent of p_{t-1} , and further of p_{t-i} for all $i > 1$, showing the third result.

For the first result, from (76), the relationship

$$|p_t - p_{t-1}| = \left| \frac{1}{\tau_v + Y_{t-1}} f_t - \frac{1}{\tau_v + Y_t} f_{t-1} \right| = \left| \left(\frac{1}{\tau_v + Y_{t-1}} - \frac{1}{\tau_v + Y_t} \right) f_{t-1} - \frac{1}{\tau_v + Y_t} g_t \right|$$

follows, and a similar argument as for trading volume over time then implies that $Cov(|p_t - p_{t-1}|, W_t) > 0$, showing the first result.

We are done. ■

References

- ADAMIC, L., C. BRUNETTI, J. HARRIS, AND A. KIRILENKO (2010): “Trading networks,” Working Paper, University of Michigan.
- ADMATI, A., AND P. PFLEIDERER (1988): “A Theory of Intraday Patterns: Volume and Price Volatility,” *Review of Financial Studies*, 1, 3–40.
- ANDERSEN, T. G. (1996): “Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility,” *Journal of Finance*, 51, 169–204.
- ANDREI, D. (2012): “Information Percolation Driving Volatility,” Working Paper, UCLA.
- BABUS, A., AND P. KONDOR (2013): “Trading and Information Diffusion in Over-the-Counter Markets,” Working Paper, Imperial College.
- BARABASI, A. L., AND R. ALBERT (1999): “Emergence of Scaling in Random Networks,” *Science*, 286, 509–512.
- BLUME, L., D. EASLEY, AND M. O’HARA (1994): “A Theory of Intraday Patterns: Volume and Price Volatility,” *Journal of Finance*, 49, 153–181.
- BOLLERSLEV, T., AND D. JUBINSKI (1999): “Equity Trading Volume and Volatility: Latent Information Arrivals and Common Long-Run Dependencies,” *Journal of Business and Economic Statistics*, 17, 9–21.
- BOLLOBAS, B. (2001): *Random Graphs*. Cambridge University Press, 2nd edn.
- BREON-DRISH, B. M. (2010): “Asymmetric Information in Financial Markets: Anything Goes,” Working Paper, Stanford University.
- BRIGGS, N. (1993): *Algebraic Graph Theory*. Cambridge University Press.
- BROUWER, A. E., A. M. COHEN, AND A. NEUMAIER (1989): *Distance-Regular Graphs*. Springer-Verlag.
- BURASCHI, A., AND P. PORCHIA (2012): “Dynamic Networks and Asset Pricing,” Working Paper, Imperial College.
- CALVET, L., J. CAMPBELL, AND P. SODINI (2007): “Down or Out: Assessing the Welfare Costs of Household Investment Mistakes,” *Journal of Political Economy*, 115(5), 707–747.
- (2009): “Fight or Flight? Portfolio Rebalancing by Individual Investors,” *Quarterly Journal of Economics*, 124(1), 301–348.
- CLARK, P. K. (1973): “A subordinated Stochastic Process Model with Finite Variance for Speculative Prices,” *Econometrica*, 41, 135–155.
- COLLA, P., AND A. MELE (2010): “Information linkages and correlated trading,” *Review of Financial Studies*, 23(1), 203–246.
- CROUCH, R. L. (1975): “The volume of transactions and price changes on the New York Stock Exchange,” *Financial Analysts Journal*, 65, 104–109.
- 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.
- DEBREU, G., AND H. SCARF (1964): “A limit theorem on the core of an economy,” *International Economic Review*, 4, 235–256.
- DUFFIE, D., S. MALAMUD, AND G. MANSO (2009): “Information percolation with equilibrium search dynamics,” *Econometrica*, 77, 1513–1574.

- DUFFIE, D., AND G. MANSO (2007): “Information percolation in large markets,” *American Economic Review*, 97, 203–209.
- EDGEWORTH, F. Y. (1881): *Mathematical Psychics: An essay on the mathematics to the moral sciences*. Wiley.
- EPPS, T. W., AND M. L. EPPS (1976): “The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture-of-Distributions Hypothesis,” *Econometrica*, 44, 305–321.
- ERDÖS, P. (1947): “Some remarks on the theory of graphs,” *Bulletin of the American Mathematical Society*, 53, 292–294.
- ERDÖS, P., AND A. RÉNYI (1959): “On random graphs I,” *Publ. Math. Debrecen*, 6, 290–297.
- FAIR, R. C. (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, D., AND S. VISWANATHAN (1995): “Can speculative trading explain the volume-volatility relation?,” *Journal of Business and Economics Statistics*, 13, 379–396.
- FRIEDKIN, N. (1991): “Theoretical Foundations for Centrality Measures,” *American Journal of Science*, 96, 1471–1504.
- GABAIX, X., P. GOPIKRISHNAN, V. PLEROU, AND H. E. STANLEY (2003): “A theory of power-law distributions in financial market fluctuations,” *Nature*, 423, 267–270.
- GALLANT, A. R., P. E. ROSSI, AND G. TAUCHEN (1992): “A subordinated Stochastic Process Model with Finite Variance for Speculative Prices,” *Review of Financial Studies*, 5, 199–242.
- GILBERT, E. N. (1959): “Random Graphs,” *Annals of Mathematical Statistics*, 30, 1141–1144.
- GLOSTEN, L., AND P. MILGROM (1985): “Bid, ask and transaction prices in a specialist market with heterogeneously informed traders,” *Journal of Financial Economics*, 14, 71–100.
- GROSSMAN, S. (1976): “On the Efficiency of Competitive Stock Markets Where Traders Have Diverse Information,” *Journal of Finance*, 31, 573–584.
- HAN, B., AND L. YANG (2013): “Social Networks, Information Acquisition, and Asset Prices,” *Management Science*, 59, 1444–1457.
- HE, H., AND J. WANG (1995): “Differential Information and Dynamic Behavior of Stock Trading Volume,” *Review of Financial Studies*, 8, 919–972.
- HEIMER, R. Z., AND D. SIMON (2012): “Facebook Finance: How Social Interaction Propagates Active Investing,” Working Paper, Brandeis University.
- HELLWIG, M. F. (1980): “On the aggregation of information in competitive markets,” *Journal of Economic 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.
- 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., AND Y. ZENOU (2012): “Games on Networks,” *Handbook of Game Theory*, Vol. 4, forthcoming.
- JACKSON, M. O. (2005): “A survey of models of network formation: stability and efficiency,” in *Group Formation in Economics: Networks, Clubs, and Coalitions*, ed. by G. Demange, and M. Wooders. Cambridge University Press.

- KARPOFF, J. M. (1987): “The Relation Between Price Changes and Trading Volume: A Survey,” *Journal of Financial and Quantitative Analysis*, 22, 109–126.
- KYLE, A. S. (1985): “Continuous Auctions and Insider Trading,” *Econometrica*, 53, 1315–1336.
- LI, D., AND N. SCHURHOFF (2012): “Dealer Networks,” Working Paper, University of Lausanne.
- LO, A. W., AND J. WANG (2000): “Trading Volume: Definitions, Data Analysis and Implications of Portfolio Theory,” *Review of Financial Studies*, 13, 257–300.
- LOBATO, I. N., AND C. VELASCO (2000): “Long Memory in Stock-Market Trading Volume,” *Journal of Business and Economic Statistics*, 18, 410–427.
- MASSA, M., AND A. SIMONOV (2006): “Hedging, Familiarity, and Portfolio Choice,” *Review of Financial Studies*, 19(2), 633–685.
- OZSOYLEV, H. (2005): “Asset pricing implications of social networks,” Working paper, University of Oxford.
- OZSOYLEV, H., AND J. WALDEN (2011): “Asset Pricing in Large Information Networks,” *Journal of Economic Theory*, 146, 2252–2280.
- OZSOYLEV, H., J. WALDEN, D. YAVUZ, AND R. BILDIK (2013): “Investor Networks in the Stock Market,” *Review of Financial Studies*, Forthcoming.
- PAREEK, A. (2012): “Information networks: Implications for mutual fund trading behavior and stock returns,” Working Paper, Rutgers University.
- ROGALSKI, R. J. (1978): “The dependence of prices and volume,” *Review of Economics and Statistics*, 60, 268–274.
- SAATCIOGLU, K., AND L. T. STARKS (1998): “The Stock Price-Volume Relationship in Emerging Stock Markets: The Case of Latin America,” *International Journal of Forecasting*, 14, 215–225.
- SCHNEIDER, J. (2009): “A Rational Expectations Equilibrium with Informative Trading Volume,” *Journal of Finance*, 64, 2783–2805.
- SHILLER, R. J., AND J. POUND (1989): “Survey evidence on the diffusion of interest and information among investors,” *Journal of Economic Behavior*, 12, 47–66.
- TAYLOR, D. E., AND R. LEVINGSTON (1978): “Distance-Regular Graphs,” in *Combinatorial Mathematics*, ed. by D. A. Holton, and J. Seberry, Berlin.
- VIVES, X. (1995): “Short term investment and the informational efficiency of the market,” *Review of Financial Studies*, 8, 125–160.
- WATTS, D. J., AND S. H. STROGATZ (1998): “Collective dynamis of ‘small world’ networks,” *Nature*, 393, 440–442.