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

Johan Walden†

April 4, 2016

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 eigenvector centrality. In line with the Mixture of Distributions Hypothesis, the market's network structure influences aggregate trading volume and price volatility. Volatility after an information shock is more persistent in less central networks, and in markets with a higher degree of private information. Similar results hold for trading volume. The shape of the autocorrelation functions of volatility and volume are related to the degree of asymmetry of the information network. Altogether, our results suggest that these dynamics contain important information about the underlying information diffusion process in the market.

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

*I thank seminar participants at the 2014 Meetings of the American Finance Association, the 6th International Finance Conference in Santiago, Chile, the Duisenberg School of Finance, FGV/EPGE, Insper, Koc University, Maastrict University, the Oxford-Man Institute, Rochester (Simon School), Stockholm School of Economics, Universidad Torcuato di Tella, Uppsala University, and USC, Fernando Alvarez, Bradyn Breon-Drish, Laurent Calvet, Nicolae Gârleanu, Brett Green, Pierre de la Harpe, Philipp Illeditsch, Matt Jackson, Svante Jansson, Ron Kaniel, Anders Karlsson, Peter Kondor, Martin Lettau, Dmitry Livdan, Hanno Lustig, Gustavo Manso, Han Ozsoylev, Christine Parlour, Matas Šileikis, Elvira Sojli, Henghsiu Tsai, and Deniz Yavuz for valuable comments and suggestions. I am grateful to Tim Bollerslev and Dan Jubinski for sharing their data.

†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 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, have "long memory" in that shocks are very persistent, 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 2014). 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 the 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 returns and trading volume? What does heterogeneous information diffusion "add" compared with, e.g., what can can be generated by heterogeneous preferences alone? The answers to these questions

[1]For example, Massa and Simonov (2006a) 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 (2007a, 2009) find that diversification and portfolio rebalancing motives do not explain investors' portfolio holdings well.

are obviously fundamental for our understanding of the impact of information networks on financial markets, as well as for any general analysis of endogenous network formation, and ultimately for how well information networks can explain observed behavior of investors and financial markets. 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 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. We believe that the introduction of such a work-horse model is valuable in itself, by allowing for further study of the general relation between information networks and asset pricing.

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 network 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, in turn. 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 short-horizon 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 makes profits in the network. It is argued in Ozsoylev and Walden (2011) and elsewhere in the literature that some type of centrality measure should determine agent profitability.[2] Centrality—a fundamental concept in network theory—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 a priori unclear, however, which is the appropriate definition of centrality in this context.

We show that profitability is determined by a centrality measure that is closely related to eigenvector 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. The result is quite intuitive: the eigenvector centrality measure provides a good balance between connections at all distances compared with other measures. Degree centrality, for example, focuses exclusively on direct neighbors, excluding agents at farther distances which are also important for an agent's performance. Our main results that characterize trading behavior, profitability, and welfare of agents are Theorems 2 and 3.

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, in line with the Mixture of Distributions Hypothesis, the information diffusion process generated by the network determines volatility and trading volume in the time series. We show that very rich dynamics of volatility and volume can be generated in the general information network model, in contrast to the benchmark case with symmetric networks and heterogeneous preferences, suggesting that the information network part of the model is indeed a crucial component in explaining these dynamics. Moreover, volatility after an information shock is more persistent in less central networks, and in markets with a higher degree of private information. The autocorrelation function of volatility depends on the structure of the network in that its nonmonotonicity is related to network asymmetry. Similar results hold for trading volume. Using simulations, we explore these implications of the model, using the preferential attachment network formation model of Barabasi and Albert (1999).

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. In Section 6, we discuss the empirical predictions of the model, and in Section 7 we discuss potential extensions. Finally, Section 8 concludes. All proofs are delegated to the appendix.

[2] In an empirical study, Ozsoylev, Walden, Yavuz, and Bildik (2014) study the trades of all investors on the 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 (several different centrality measures exist) 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).

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. Manela (2014) analyzes how the speed of information diffusion affects the welfare of agents, showing that the value is hump-shaped in the diffusion speed. Again, the diffusion process is quite specific.

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 (2014) 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 funds 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. Andrei (2012), shows that persistent price volatility can arise in such a model, and Andrei and Cujean (2014) analyze momentum and reversal in a similar setting. 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 leads to variation in return volatility and trading volume, as well as to 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 literature 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. 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,

$a = 1, \ldots, N$,

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

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

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 *mth 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, $m = 1, \ldots, N$, where the ath element of V^m is $V_{a,m}$ Equivalently,

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

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

Here E^m is the mth power of the adjacency matrix, and $\chi : \mathbb{R}^{N \times N} \to \{0,1\}^{N \times N}$ 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 (2014), 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 $t = 1, \ldots, T$, $s_{a'} \in \mathcal{I}_{a,t}$ if and only if $D(a, a') \le t$.*

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 $\zeta_{a,t} = \frac{\sigma}{\sqrt{V_{a,t}}}\xi_{a,t}$, 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

[3]A 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.

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 needs 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

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

subject to the budget constraint

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

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 at $T+1$. We think of this as public information, which becomes available to all agents at $T+1$. Alternatively, we 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. Especially, models with symmetric information structures typically fall into the class of network symmetric economies (e.g., the model in Vives (1995)).

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 $\Delta V_1, \Delta V_2, \ldots, \Delta V_T$, is unimodal. Specifically, there are times $1 \le t_1 \le t_2 \le T$, such that $\Delta V_{t+1} > \Delta V_t$ for all $t \le t_1$, $\Delta V_{t+1} = \Delta V_t$ for all $t_1 < t \le t_2$, 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 of connections agents have are needed, to ensure existence of equilibrium. Ozsoylev and Walden (2011) carry out a fairly general analysis of the restrictions needed for the existence of equilibrium to be guaranteed. They show that a sufficient condition is that the distribution of number of connections 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 connections be needed to ensure the existence of

[4] Other 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.

[5] The 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).

equilibrium, but also on connections 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 $\mathcal{A}_m = \mathcal{N} \times \{1, \ldots, M\}$, where $a = (i, j) \in \mathcal{A}_m$ represents the ith agent in the jth 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 $a = 1, \ldots, MN$, 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 $\mathcal{A} = \lim_{M\to\infty} \mathcal{A}_M$, 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 network[6]

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

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 behavior of agents $a = 1, \ldots, N$, 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.

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

Theorem 1 *Consider an economy characterized by $\mathcal{M}$. For $t = 1, \ldots, T$, 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 unique linear equilibrium, in which prices at time t are given by

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

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

$$
\begin{aligned}
x_{a,t} &= \frac{\tau V_{a,t}}{\gamma_a} (z_{a,t} - p_t), &(5) \\
U_{a,t} &= -D_{a,t} e^{-\gamma_a W_{a,t} - \frac{1}{2} \frac{\tau^2 V_{a,t}^2}{\tau_v + Y_t + \tau V_{a,t}} (z_{a,t} - p_t)^2}. &(6)
\end{aligned}
$$

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, $s = 1, \ldots, t$). 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 through the network 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 in the network have received 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 for the market maker to separate informed trading from supply shocks. This is our first example of how the network structure impacts asset price dynamics.

We define

$$h_t = \frac{1}{N} \sum_{a=1}^{N} \left(\frac{\bar{\gamma}}{\gamma_a}\right) \Delta V_{a,t}, \qquad t = 1, \ldots, T, \tag{7}$$

which measures the (weighted) average number of agents at distance t from any agent in the network. Here, $\bar{\gamma}$ is the (harmonic) average risk aversion coefficient of agents in the economy, $\bar{\gamma} = (\sum_a \gamma_a^{-1})^{-1}$, and weights are chosen such that agents with lower risk aversion get higher weights, which is natural since they will be more active in the market. With this definition, we can write

$$y_t = \frac{\tau_u \tau^2}{\bar{\gamma}^2} h_t^2, \tag{8}$$

showing that y_t is closely related to this average.

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

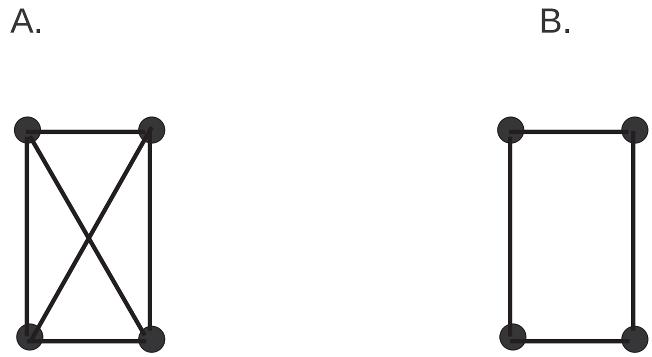

Figure 1: Impact of network structure. *The figure shows two networks with four agents: In Panel A, a tight-knit network is shown, in which every agent is connected with every other agent. In panel B, a less tight-knit network is shown. At $t = 2$, prices are more revealing in the tight-knit network,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

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 (2014), 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*

$$Cov(\Delta x_{a,t+1}, \Delta x_{a,t}) = \frac{\tau^2}{\gamma_a^2}\Delta V_{a,t}\left(\frac{V_{a,t+1}}{\tau_v + Y_{t+1}} - \frac{V_{a,t}}{\tau_v + Y_t}\right). \tag{9}$$

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

$$Cov(\Delta x_{a,t}, \Delta x_{b,t}) = \frac{\tau^2}{\gamma_a\gamma_b}\left(\frac{y_t}{(\tau_v + Y_{t-1})(\tau_v + Y_t)}V_{a,t-1}V_{b,t-1} + \frac{\Delta V_{a,t}\Delta V_{b,t}}{\tau_v + Y_t} + \Delta V_{a,b,t}\right). \tag{10}$$

Equation (9) 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 (9) takes the form

$$Cov(\Delta x_{a,t+1}, \Delta x_{a,t}) = \frac{\tau^2}{\gamma_a^2} \Delta V_{t+1} \Delta V_t \left(\tau_v + Y_t - \frac{\tau}{\bar{\gamma}} V_t \Delta V_{t+1} \right).$$

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 (10), 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 (9). 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 (10) 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 $Cov\left(v + \zeta^b - p_t, -(p_t - p_{t-1})\right) < 0$. 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, $Cov\left(\zeta^b, -(p_t - p_{t-1})\right) = 0$. 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.[7]

The second main implication of (10) 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 over fairly short time horizons, rather than portfolio holdings, to draw inferences about a market's information network as, e.g., done in Ozsoylev, Walden, Yavuz, and Bildik (2014).

4.2 Profits and centrality

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

[7] We do not expect these assumptions to hold in markets where agents' price impact introduces strategic motives to hide their trades, as in Colla and Mele (2010) where negatively correlated trades may arise in equilibrium. Of course, in their model informed investors are risk neutral so hedging motives are absent.

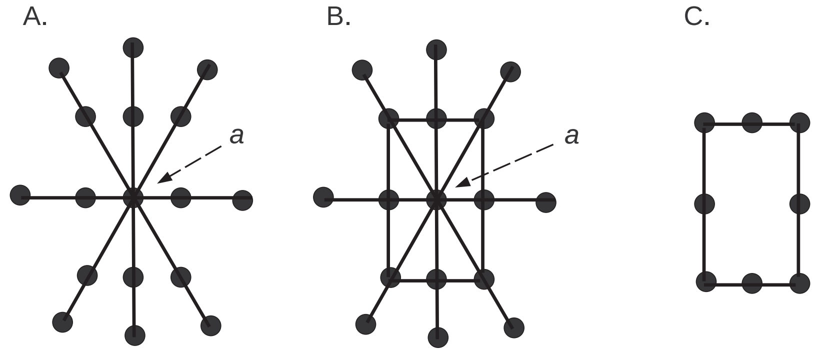

Figure 2: Correlation of trades. *The figure shows three networks that lead to different types of trading behavior. Panel A shows a network in which agent a's trades at time 1 and 2 are negatively correlated, whereas they are positively correlated in Panel B. Panel C shows a network in which the correlation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

Theorem 3 *Define*

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

The ex ante certainty equivalent of agent a is

$$U_a = \frac{1}{2\gamma_a}\log(C), \qquad \text{where } C = \prod_{t=1}^{T}(1 + \pi_{a,t}). \tag{11}$$

The expected profit of agent a is

$$\frac{\tau}{\gamma_a}\Pi_a, \tag{12}$$

where

$$\Pi_a = \sum_{t=1}^{T}(\tau_v + Y_t)^{-1}V_{a,t} \tag{13}$$

is the profitability *of agent a.*

Our focus in this study is on profitability, but we have also stated the formula for welfare, to allow for future research on endogenous network formation. Equation (12) 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 (13). 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 (13) thus provides a direct relationship between the properties of a network, local as well as aggregate, and individual agents' profitability.

Equation (13) shows that an agent's profitability is determined by his *centrality*, appropriately defined. 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 (13) on vector form as

$$\Pi = \sum_{t=1}^{\infty} \beta_t \chi(E^t)\mathbf{1}, \tag{14}$$

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 ath element is the profitability of agent a. The functional form of (14) is close to the standard Katz and eigenvector centrality measures. Recall that the *Katz centrality* vector with parameter $\alpha < 1$ is the vector $K \in \mathbb{R}_+^N$, defined as

$$K = K^\alpha = \sum_{t=1}^{\infty} \alpha^t E^t \mathbf{1}. \tag{15}$$

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. Moreover, the *eigenvector centrality* vector is the eigenvector corresponding to the largest eigenvalue of E, i.e., the vector C that solves the equation $C = \lambda E C$, for the largest possible eigenvalue, λ, where we normalize C such that $\sum_{a \in \mathcal{N}} C_a = 1$. It is a standard result that eigenvector centrality can be viewed as a special case of Katz centrality, since $C \propto K^{\lambda^{-1}}$ in the sense that $C = \lim_{\alpha \nearrow \lambda^{-1}} \frac{K^\alpha}{\sum_a K_a^\alpha}$.[8]

The structures of the profitability measure (14) and equation (15) 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 and eigenvector 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. Importantly, both expressions highlight that an agent's centrality depends on direct as well as higher-order connections, in contrast to the degree measure which only measures direct connections. We therefore take (14) as the definition of agent centrality within our model, immediately implying that more central agents make higher profits.

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

In Appendix B, we provide further evidence that the centrality measure (14) is indeed closely related to eigenvector and Katz centrality, both asymptotically in large scale networks and in mid-size networks, by studying random networks generated by the classical Erdös-Rényi model (see Erdös (1947), Erdös and Rényi (1959), and Gilbert (1959)), in which links between agents are formed randomly, independently, and with constant probability.

5 Aggregate volatility and trading volume

According to the Mixture of Distributions Hypothesis (MDH), return volatility varies over time, which 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 shocks to trading volume are very persistent in that abnormally high trading volume in one period predicts abnormally high trading volume for many future periods. Second, return volatility of individual stocks, and markets, also have long memory (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 4 *For $t = 1, \cdots, T$, the variance of prices between $t-1$ and t, is*

$$\sigma_{p,t}^2 = \frac{y_t}{(\tau_v + Y_t)(\tau_v + Y_{t-1})}, \tag{16}$$

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

$$\sigma_{p,T+1}^2 = \frac{1}{\tau_v + Y_T}. \tag{17}$$

Moreover, given coefficients, $k_1, \ldots, k_{T+1}$, 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

$$\left| \sigma_{p,t}^2 - \frac{k_t}{\tau_v} \right| \leq \epsilon, \qquad t = 1, \ldots, T+1. \tag{18}$$

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 (16), but that it can still have spikes in some time periods because of large values in the numerator. In fact, (18) 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

$$\sigma_{P,t}^2 = \frac{1}{\tau_v} \frac{Y_t}{\tau_v + Y_t} = \frac{1}{\tau_v} \frac{1}{1 + \frac{\tau_v}{Y_t}}. \tag{19}$$

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., (17), 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,

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

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, $t = 1, \ldots, T$.

Now, it is clear from (19) that volatility after an information shock will be more persistent in economies with lower Y_t, for all times $t = 1, \ldots, T$, in that at each t there is more residual volatility left to be incorporated into prices, the lower is Y_t. We can use this fact to link the persistence of volatility to the centrality of the network. Specifically, recall that h_t describes the (weighted) average number of nodes at distance t from any node in the network (7). It is reasonable to call a network with sequence h_t more central that one with sequence h'_t, $t = 1, \ldots, T$, if $h_t > h'_t$ for all t, since the average number of nodes at distance t is higher in the former network than in the latter, for *all* t.[9] From (8), it then follows that

Corollary 1 *Volatility is less persistent in more central networks than in less central networks, all else equal.*

Corollary 1 thus provides a direct link between the persistence of volatility and network structure, a link that we will explore further in Section 5.3.

Network structure is also important for equilibrium volatility patterns over time. If we restrict our attention to network symmetric economies, the possible dynamic behavior of volatility is quite restricted. This is not surprising, given the restrictions on information diffusion dynamics described in Lemmas 1 and 2. From (19), 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, $\sigma_{P,T}^2 << \sigma_v^2$, it follows from (19) that

$$\sigma_v^4 \Delta\eta_t \approx \sigma_{p,t}^2, \tag{20}$$

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

Corollary 2 *In a network symmetric economy,*

1. *$\Delta\eta_t$ is unimodal,*
2. *if the public information component is high, $\sigma_{P,T}^2 << \sigma_v^2$, 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, $g_0, g_1, \ldots, g_T$, we define the difference operator $(\Delta g)_k = g_k - g_{k-1}$, $k = 1, \ldots, T$, $(\Delta g)_{T+1} = 0 - g_T$, and the higher-order operators $\Delta^{n+1} g = \Delta(\Delta^n g)$. The following lemma relates the monotonicity of a function and its autocorrelation function:

[9] Note that this concept, which provides a measure of the average centrality of agents in the network, is distinct from network *centralization*, which measures the difference between the most and least central nodes in a network.

Lemma 3 *Consider a nonnegative sequence $g_0, g_1, \ldots, g_T$, such that $g_0 = 0$. Define the autocorrelation function $R_\tau = \sum_{k=0}^{T-k} g_k g_{k+|\tau|}$, $-T+1 \leq \tau \leq T-1$. Then,*

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

In light of Corollary 2, Lemma 3 immediately leads to the following result, showing that nonmonotonicity of volatility is associated with information networks outside of the symmetric class:

Corollary 3 *In a network symmetric economy with a high public information component, the autocorrelation function of volatility is unimodal.*

5.2 Volume

Just like with volatility, rich dynamics of trading volume can arise within the network model. Aggregate trading volume will be made up 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 3 *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 5 *The time-t expected trading volume is*

$$X_t = \frac{\tau}{N}\sum_{a=1}^{N}\frac{1}{\gamma_a}\sqrt{\frac{2}{\pi}\left(\frac{V_{a,t-1}^2}{\tau_v+Y_{t-1}} - \frac{V_{a,t}^2}{\tau_v+Y_t} + 2\frac{\Delta V_{a,t}V_{a,t}}{\tau_v+Y_t} + \frac{\Delta V_{a,t}}{\tau}\right)}. \tag{21}$$

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

$$|X_t - c_t| \leq \epsilon, \qquad t = 1, \ldots, T+1.$$

It is clear from (21) 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 (21) collapses to

$$X_t \;=\; \sqrt{\frac{2\tau^2}{\pi\bar{\gamma}^2}\left(\frac{V_{t-1}^2}{\tau_v+Y_{t-1}}-\frac{V_t^2}{\tau_v+Y_t}+2\frac{\Delta V_t V_t}{\tau_v+Y_t}+\frac{\Delta V_t}{\tau}\right)}. \tag{22}$$

Again, differences in preferences in this case are only important through the effect they have on the 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 (22) dominates and $X_t \approx \sqrt{\frac{2\tau}{\pi\bar{\gamma}^2}\Delta V_t}$. In this case, we therefore get:

Corollary 4 *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*:*

$$X_t \sim \sqrt{\sigma_{p,t}}. \tag{23}$$

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.

The difference between the dynamics in symmetric and asymmetric networks is exemplified by the two networks in Figure 3. The left panel shows the symmetric so-called Heawood network with 14 agents, whereas the right panel shows an asymmetric subnetwork of the Heawood network, in which two agents are more connected than the others. The average numbers of connections in these networks, as a function of time, are shown in the left panel of Figure 4. In line with Lemma 2, for the symmetric network the average number of connections is unimodal, whereas it is not for the asymmetric network. This carries over to the same behavior of volatility and trading volume after an information shock, as shown in the middle and right panels of Figure 4, in line with our previous arguments.

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

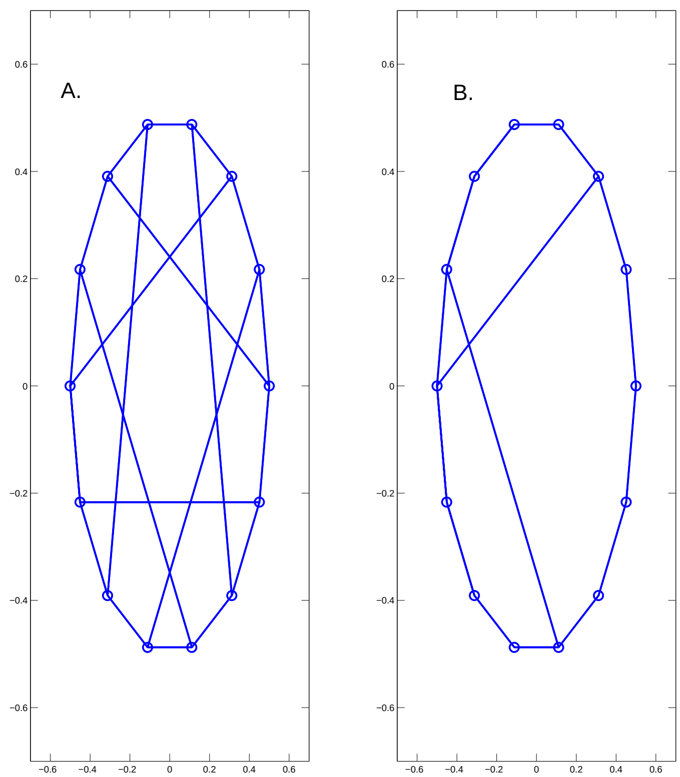

Figure 3: Symmetric and asymmetric network. *The left panel (A.) shows the symmetric so-called Heawood network, with 14 agents. The right panel (B.) shows an asymmetric subnetwork of the Heawood network.*

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. This argument 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 Dynamics in simulated networks

Given of our strong characterization of equilibrium, with closed form expressions for all variables of interest, it is straightforward to simulate a class of larger networks. Our goal is to gain further insight about how persistence is informative about the underlying information diffusion process. Specifically, we are interested in how network structure and the channels through which information propagate (the network versus public channel) relate to the long memory of volatility and volume documented in Bollerslev and Jubinski (1999).

We recall that for a general series, K_t, with the coefficient of fractional integration, $0 < d <$

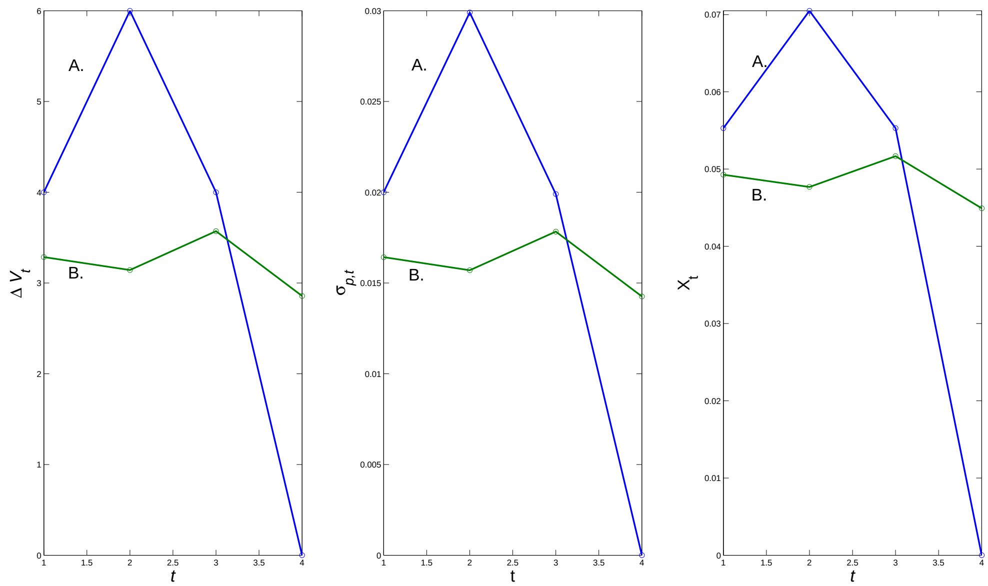

Figure 4: Dynamics of different networks. *The figure shows the dynamics of average number of connections (left panel), volatility (middle panel) and trading volume (right panel) of the two networks in Figure 3. The dynamics of the symmetric network (A.) is unimodal in contrast to the dynamics of the asymmetric network (B.), in all panels. Parameters:* $\sigma_v = 2$, $\sigma_u = 1$, $\sigma = 1/10$, $\gamma = 10$.

0.5, the process $(1-L)^d K_t$ has a bounded spectrum across all frequencies, K_t is stationary and its autocorrelation function function decays like $|t|^{2d-1}$ for large t. Here, L is the lag operator. Thus, the higher d is, the longer the memory of the process, the higher the persistence after a shock, and the slower the decay of the autocorrelation function. As a comparison, for stationary ARMA processes the decay is much faster; it is exponential.

Studying the constituents of the S&P 100 index, Bollerslev and Jubinski (1999) find coefficients of fractional integration for absolute returns, $d_{|\mu|}$, between 0.15 and 0.57, and between 0.07 and 0.74 for trading volume, d_V, with most coefficients lying between 0.2 and 0.5. (see Figure 5 in Bollerslev and Jubinski 1999). Our previous analysis suggests that stocks with more central networks should have lower coefficients, which we now wish to verify via simulations.

To generate networks randomly, we choose the preferential attachment model introduced in Barabasi and Albert (1999), which has been extensively used in various applications. The underlying idea in their model is that agents are more likely to form links with other agents who already have many links. This generates a snowball effect where some agents end up having a disproportionately large number of links, and thereby leads to a heavier-tailed distribution of number of connections. For large networks, the fraction of agents with n links is roughly proportional to n^{-3}. In other words, the network has a power-law distributed degree distribution with a tail exponent of 3. As discussed in Ozsoylev and Walden (2011), such power-law distributed networks may provide reasonable approximations of real world financial networks.

We simulate a large number of networks, each with $N = 1,000$ agents. Following Barabasi and Albert (1999), we assume that a fraction, f, of these agents are initially connected in the network (with an an average of 4 connections per agent), and that the rest of the agents arrive sequentially (before the asset is traded), and connect with those already in the network. As in Barabasi and Albert (1999), the probability for a new agent to connect with any given existing agent in the network is proportional to the number of connections the existing agent already has.

The initial agents will on average be more connected and centrally placed than those who arrive sequentially, and thereby more informed. Our interpretation is that the original investors represent well-informed—perhaps institutional—investors, with established information channels, whereas those who arrive sequentially represent less informed—perhaps retail—investors, who receive information only once it has passed through the established part of the network.

By varying the fraction of well-informed agents, f, we generate different network structures. For example, in Figure 5 we show two realizations, one with a low fraction, $f = 5\%$ (left panel), and one with a higher fraction, $f = 20\%$ (right panel). For expositional reasons, we chose a smaller network with $N = 400$ agents in the figure. We see that the network with the high fraction of well-informed agents is more tight-knit than the network with a low fraction, and we would expect this to lead to higher centrality, less persistence, and lower coefficients of fractional integration.

We would also expect the ratio of information propagating through the private and public channel to be important. We can interpret τ_v as the precision of public information, i.e., the information precision after all the information agents have from public channels has been revealed, whereas τ is the precision of the private signal. This ratio then takes the form $g = \frac{\tau}{\tau_v}$. A higher g, all else equal, makes the information network relatively more important as a channel of information diffusion.

We study how f and g impact the long-memory properties of volatility and volume in the network. We vary f and g (by varying τ, keeping τ_v fixed). For each value of the pair (f, g), we simulate a large number (10,000) of economies with $T = 60$ time periods, using the preferential attachment network generation process described above. For each simulation, we calculate equilibrium and then estimate the coefficients of fractional integration for volatility, d_σ, and trading volume, d_V, using the GPH estimator (see Geweke and Porter-Hudak 1983).[10]

The average estimated coefficients for each choice of f and g are shown in Figure 6. Each line in the figure represents a constant f, where g varies between 0.1 and 0.001. Higher values of g's lie on the upper left part of each line. The main implications are two-fold. First, in line with our expectations, higher values of f lead to lower estimates of d_V and d_σ. Second, since the higher values of g are on the upper left part of each line, economies with relatively more private information—all else equal—have higher estimated d_σ's but lower d_V's. Again, the intuition for d_σ is straightforward: Diffusion through

[10]Since T is finite (although potentially large), our processes are technically not "true" long-memory processes. From a practical perspective, however, it is the behavior over a limited number of periods that matters, which also determines the estimated coefficients.

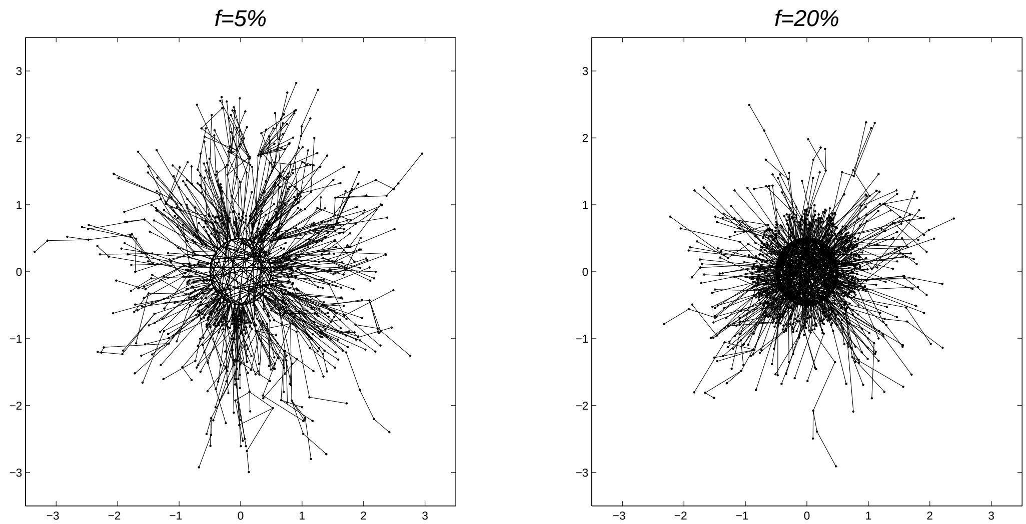

Figure 5: Examples of networks. *The figure shows two networks generated by the preferential attachment model of Barabasi and Albert (1999). Each network contains $N = 400$ agents. In the left panel, 5% of the agents are initially connected, and make up the well-informed part of the network, whereas 20% of the agents are well-informed in the right panel. The network in the right panel is more centralized than that in the left panel.*

the network is a slower process, so when a higher fraction of the information diffuses through this channel, i.e., when g is higher, shocks last longer, thereby leading to a higher estimated value of d_σ.

It is a priori less clear why the effect on d_V of relatively more private information is actually the opposite, i.e., trading volume shocks are shorter-lived with more private information. However, from our previous discussion we know that the link between information diffusion and trading volume may be more complex than that between information diffusion and volatility. Even with a precise public signal, trading volume can be high although the information advantage of the informed agents is low, since there is little uncertainty about the fundamental value and the (risk averse) agents therefore take on large positions. Moreover, trading volume also reflects the different risk aversion of agents: agents with low risk aversion who receive information late—at which point profit opportunities are low—may still generate substantial trading volume by taking on large positions. In our example, these effects dominate, leading to a negative relation between g and d_V. This further highlights the disconnect between volume and volatility dynamics in general economies, discussed in Section 5.2.

6 Empirical Implications

As discussed in the introduction, our model has far reaching consequences in suggesting that decentralized information diffusion among the investor population is an important determinant of investor behavior and performance, as well as of the dynamics of asset prices and trading volume. Especially,

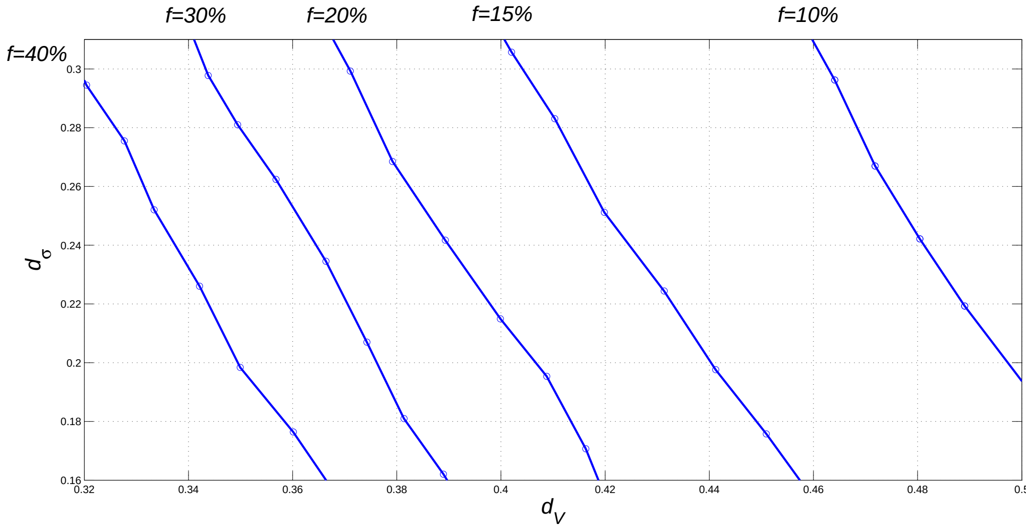

Figure 6: Long memory of volatility and trading volume. *The figure shows how the estimated coefficient of fractional integration of the autocorrelation functions of trading volume (d_V) and volatility (d_σ) depend on the fraction of well-informed agents (f) and the relative degree of public information (g), using the preferential attachment network formation model. For each line, g varies and f is constant; the higher the value of g, the further to the upper left part of each line. Higher values of f leads to lower d_V and d_σ. Higher values of g leads to higher d_V and lower d_σ. Parameters: $T = 60$, $N = 1000$, $\tau_u = 1$, $\tau_v = 1$, γ_a i.i.d. uniformly distributed between 1 and 2.*

the centrality of individual agents and of the network in aggregate are important for the equilibrium outcome. An empirical test of the model will therefore require two steps: first one needs to find a good proxy for the information network, then the model's predictions can be tested. We discuss these steps below.

6.1 Identification of the information network

Ozsoylev and Walden (2011) discuss two approaches for identifying a market's information network. The first approach would use information about individual households to build a proxy for network connections. An example is the Swedish dataset used by Calvet, Campbell, and Sodini (2007b), containing household information such as residential address, education level, employer, and demographic data, as well as financial data, such as holdings of financial securities and bank accounts, for the entire Swedish population. Other examples include the Swedish LINDA dataset, used in Massa and Simonov (2006b), and the Finnish Central Securities Depositary dataset used in Grinblatt and Keloharju (2000) and Grinblatt and Keloharju (2001). Proximity of household address, a common employer or the same educational institution would, for example, suggest the presence of an information link. These datasets have the advantage of covering a large portion of, or even the whole, population. Other datasets may also be used. For example, the dataset in Cohen, Frazzini, and Malloy (2008) contains information about educational linkages between portfolio managers and corporate board members, and could be used to construct the information network for this subpopulation.

The second approach discussed in Ozsoylev and Walden (2011) identifies information networks directly from trades. The idea is that since the correlations of individual agents' trades are closely related to their position in the information network (as shown in Theorem 2), these trades, if observed at the individual level, can be used to reverse-engineer the network. Datasets that contain account level information about all trades in a market are, e.g., used in Barber, Lee, Liu, and Odean (2009) (for the Taiwan Stock Exchange) and in Ozsoylev, Walden, Yavuz, and Bildik (2014) (for the Istanbul Stock Exchange). Indeed, Ozsoylev, Walden, Yavuz, and Bildik (2014) uses this approach to identify agents as being connected in an information network if they trade in the same direction in the same stock within a short time interval (up to one day), a pre-specified number of times within a one-year period. Pareek (2009), in a somewhat similar approach, uses correlations between mutual fund managers' portfolio holdings to identify information networks.

A third approach would be to use news events, possibly in conjunction with data on individual trades, to draw inferences about the information network. As mentioned in Ozsoylev, Walden, Yavuz, and Bildik (2014), although one interpretation of the information network is that it represents gradual diffusion of private information through word-of-mouth communication, another possible interpretation is that it represents the gradual diffusion of information that is technically public, e.g., through different media outlets. That is, even though a piece of information may have been publicly announced, it may still reach agents and affect their trading in a sequential fashion, because of various costs of information gathering, agents' limited processing skills, etc. A similar argument is made in Manela (2014). The dataset in Tetlock (2010), which contains 2.2 million news events in the Dow Jones news archive, may serve as an example. Alternatively, Internet discussion boards may be used to draw inferences about the information network (see, e.g., Das and Sisk 2005). For example, the time it takes for information to spread from one discussion board to another could be used as a proxy for how tight-knit the information network is.

A fourth approach would be to draw inferences about the network structure of shareholders from information about investor holdings. For example, 13(f) filings of institutional common stock holdings, provided by Thomson Reuters (formerly known as CDA/Spectrum database), may be used to draw inferences about how central a stock's investor network is. The Herfindahl index of institutional holdings could, e.g., serve as a proxy for network centrality, under the natural assumption that institutional investors correspond to well-informed agents in the preferential attachment model of Section 5.3.

6.2 Model predictions

The model provides two types of predictions about the behavior of individual agents, which if supported suggest that decentralized information diffusion is indeed important for the behavior and performance of individual agents:

Prediction 1 (Individual investors)

a) The closer two agents are in the network, the more similar are their trades.

b) *The more central an agent is in the network, the better his performance.*

Of course, for prediction 1a) to be testable, the network identification method cannot be based on trades. However, it would be testable if the network was identified by any of the other approaches discussed. Prediction 2b) was explored in Ozsoylev, Walden, Yavuz, and Bildik (2014), who found support for a positive relation between performance and centrality, using a network identification methodology based on trades. The authors discuss the issue of using trades both to identify the network and to calculate investor performance. They try to control for it, but a stronger test would, again, identify the network from other sources.

The model also provides predictions about aggregate dynamics:

Prediction 2 (Aggregate dynamics)

a) *The more central the network, the more short-lived are shocks to volatility and trading volume.*

b) *The higher the public information component, the more short-lived are shocks to volatility.*

c) *The more symmetric the network, the more monotone are volatility and trading volume over time after an information shock.*

To test prediction 2a), a natural approach would be to study the estimated coefficients of fractional integration, d_σ and d_V for different stocks (e.g., using the approach in Bollerslev and Jubinski 1999). Stocks with more central networks should then have lower coefficients than stocks with less central networks. The test could either be carried out in the cross section of stocks, or for individual stocks over time given time-series network data. In either case, the test would need to control for stock-specific characteristics, like liquidity and firm size.

To test prediction 2b), analyst coverage, obtained through IBES (Institutional Brokers Estimates System), may be used to draw inferences about the degree of information propagating through the information network channel (i.e., the variable g in Section 5.3). Under the assumption that more analyst coverage corresponds to a higher degree of public information, high analyst coverage would correspond to low g for a stock.

Prediction 2c) relates asymmetry of the information network to nonmonotonicity of (the autocorrelation functions of) volatility and trading volume, R_t^σ and R_t^V. To measure the nonmonotonicity of R_t^σ and R_t^V, one may proceed by counting their number of turning points (the number of times that $R_t - R_{t-1}$ shifts signs). Another potential measure of nonmonotonicity would be the absolute variation, $\sum_t |R_{t+1} - R_t|$. As the measure of network asymmetry, we propose using the cross sectional variance of centrality across traders.

Finally, the window length over which volatility and volume is measured needs to be matched with the speed at which information diffuses through the network. In Ozsoylev, Walden, Yavuz, and Bildik (2014), it is argued that a natural horizon for complete diffusion of information after an information shock is about a month. This is between the short horizons (less than a day) influenced by market microstructure dynamics, and longer-term (at least six months) factors, like momentum. With a one-month horizon, it would be reasonable to use daily volatility and trading volume in the tests.

7 Extensions

Several extensions of the model would be fairly straightforward. In its current version, signals are assumed to be perfectly communicated from agent to agent. A more realistic extension may be to assume that some "miscommunication" noise is introduced each time an agent shares information with another agent. Agents who are farther apart would then not only receive each others' signals later, but also with more noise than agents who are close. The most straightforward way to incorporate such an effect would be to make the signals an agent receives increasingly noisy over time, i.e., to make the signal precision in Theorem 1 decreasing over time, $\tau_1 > \tau_2 > \cdots > \tau_T$. The main effect of this extension would be to offset the increased information agents gain over time, by decreasing the signal informativeness. For example, in the expression for agents' expected profits, (12,13), the signal precision for each agent at higher t would be lower than in the economy with perfect information sharing, acting as a force to lower profits. On the other hand, the aggregate informativeness of the market, represented by Y_t, would also be lower, increasing the total opportunity for informational rents. It is a priori unclear whether agents would be better or worse off in total with more noise, similar to the hump-shaped behavior of equilibrium welfare in signal precision discussed in Ozsoylev and Walden (2011) and Manela (2014).

The extension just described implicitly assumes that the noisiness of the signal agent a_1 receives from agent a_2 depends on the distance, $D(a_1, a_2)$, but not on the number of paths between the two agents. This corresponds to a situation where the signal is identical, regardless of the path it travels. As a consequence, the only effect is to make the aggregate signal precision, τ, time dependent. Alternatively, one could assume that the miscommunication noise terms introduced when a signal propagates on two different paths are independent. For example, if agent a_1 is indirectly connected to agent a_2 via agents a_3 and a_4, then at $t = 2$ agent a_1 would be able to infer more about agent a_2's signal by receiving a "report" from both a_3 and a_4 than if he was only connected to a_3, if miscommunication noise is independent. Having multiple paths between agents would thus partially offset the additional noise effect.

Existence of equilibrium (Theorem 1) in this latter extension would follow from similar arguments as before, but the formulas would be somewhat different. Especially, the aggregate variables in Theorem 1, as well as the formulas for individual agent performance in Theorem 3 would not simply depend on the number of agents within a distance of a specific agent, represented by $\chi(E^t)$, but also on the number of paths between the agents. The number of paths of length t from an agent to any of the other agents is measured by the t:th power of the adjacency matrix, E^t, which would therefore be important for equilibrium. For example, agent performance would no longer just depend on the number of agents within a certain distance, as in (14), but also on the number of paths between agents, and therefore be more similar to the Katz/eigenvector measure in (15).

A perhaps more interesting extension would take a step toward endogenous network formation, by introducing a cost of sustaining a link between any two agents and studying which networks could be sustained with such a cost present. As we discussed before, in a large network with price-taking agents, there are only advantages for an agent of being linked to another agent. A cost of sustaining links would lead to a trade-off. In the simplest version of this extension, there would be a one-time cost—the same for any link—imposed at time $t = 0$ on both agents. Only links valuable enough for both agents (as measured in welfare terms, (3)) would be sustainable in equilibrium, suggesting that:

- Agents are less likely to sever links to other agents who are central, and to agents whose information is not obtained through other links (redundant information).
- There may be a threshold centrality such that agents below the threshold cannot sustain links with any other agents, since the value of their information does not outweigh the cost of linking to them. Such agents would thus end up isolated.
- There may be a threshold centrality, above which is not worthwhile for an agent to sustain any further link, since the marginal benefit of additional information is decreasing, but the marginal cost of sustaining each link is the same.
- All else equal, an agent with higher risk aversion will be less likely to sustain a link than an agent with low risk aversion, since risk averse agents trade less aggressively on private information.

The first and third points describe two forces pulling in opposite directions: On the one hand, it is very valuable for an agent—especially one with low centrality—to be connected to a central agent. On the other hand, the value for the central agent of sustaining a link to an agent with low centrality is low. Altogether, we expect such considerations to have interesting implications for which information networks are sustainable in equilibrium, complementing other studies, e.g., Han and Yang (2013), who focus on endogenously determined information acquisition.

Further extensions along these lines are possible. A more general form of the cost function would introduce a period-by-period cost of sustaining links. In this case we may expect the price discovery process to decrease over time—and possibly even come to a halt—as profit opportunities are mitigated when the price becomes more informative, and agents therefore do not find it worthwhile to sustain links. Heterogeneous cost functions across agents may also be assumed, where some agents are more "talented" than others in sustaining links. By assuming increasing marginal costs of sustaining a link in the number of links an agent already has, an interesting trade-off between the costs of having a high degree (many direct links) and the benefits of having a high centrality (many higher order links) is introduced. We leave these interesting extensions for future research.

8 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 more central agents make higher profits. At the aggregate level, network structure affects the dynamics of a market's volatility and trading volume, as shown theoretically and in simulations. These dynamics could therefore potentially be used to draw inferences about the underlying information network in a market. We view this as an exciting area for future research.

Endogenous network formation in financial markets provides another interesting and feasible area for future research, given our strong characterization of agent welfare in the model. As discussed, a first step in this direction would be to add costs of links in the model and study which networks could be sustained in equilibrium when such costs are present. A more ambitious approach would be to introduce a network formation game, played before the market opens, in which agents add and sever links among themselves, knowing what the consequences will be for equilibrium welfare. It is an open question what type of networks would form in equilibrium in such an extension, and how well those networks could match observed dynamics at the individual and aggregate level, in practice.

A 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*

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

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* $\Lambda = diag(0, \sigma_1, \dots, \sigma_{M-1})$. *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 $A_t = \frac{1}{N}\sum_{a=1}^N A_{a,t}$, and $\eta_t = \frac{1}{N}\sum_{a=1}^N \eta_{a,t}(p_t)$, with the convention, $A_0 = 0$, $\eta_0 \equiv 0$. Here, we are using the fact that in our large network $\lim_{M\to\infty}\frac{1}{M}\sum_{r=0}^{M-1} z_{a+rM,t} = v$ for all a and t (almost surely). This allows us to use the L.L.N. for each node, and collapse the sum from $\bar{N}$ to N. 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 net 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,

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

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 $R_1, \ldots, R_t$. We define the vector $\mathbf{s} = (R_1/(A_1 - A_0), R_2/(A_2 - A_1), \ldots, R_t/(A_t - A_{t-1}))'$, 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.,

$$v = V_t + \sigma_{V_t} \xi_t^V, \tag{27}$$

where $V_t = \frac{\tau_v}{\tau_v + \hat{\tau}_u^t}\bar{v} + \frac{1}{\tau_v + \hat{\tau}_u^t}\sum_{i=1}^t (A_i - A_{i-1})\tau_{u_i} R_i$, $\sigma_{V_t}^2 = \frac{1}{\tau_{V_t}}$, 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 $R_1, \ldots, R_t$, 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 $\Lambda = diag(0, \sigma_i^2, \sigma_u^2/\bar{A}^2)$) 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:

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

The demand of agent a is therefore

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

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

Since $v-p_T \sim N(0,\sigma^2_{V_t})$, 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

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

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 $Y_t = \hat{\tau}^t_u = \sum_{i=1}^t y_i$, 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

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

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

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

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

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

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

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

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

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

so

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

and also

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

This leads to the unconditional distribution:

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

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

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

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

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

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. ∎

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

Proof of Theorem 3: The certainty equivalent satisfies $-e^{-\gamma_a CE} = E_0[-e^{-W_{T+1}}] = -E_0\left[\prod_{t=2}^T C_{t-1}^{-1} e^{-\frac{1}{2}\frac{Q_1^2}{\tau_v + Q_1 + Y_1}(z_{a,1} - p_1)^2}\right]$. 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

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

We are done. ∎

Proof of Theorem 4: 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 $(\Sigma_{XX})_{11}$, $\sigma^2_{p,t} = (\Sigma_{XX})_{11}|_{t-1} = \frac{y_t}{(\tau_v+Y_t)(\tau_v+Y_{t-1})}$. with the convention $Y_0 = 0$. Also, the final period volatility of $v - p_T$ is $\sigma^2_{p,T+1} = \frac{1}{\tau_v+Y_T}$. 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 8 *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-nit 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 $\frac{\bar{V}_{t+1}}{\bar{V}_t}$, $t = 1,\dots,T$ 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 9 *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 9: 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

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

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 9. ∎

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

$$\sum_{i=0}^{N+1-\tau} g_i(g_{i+\tau} - g_i) > 0 \tag{39}$$

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 9 no such sequence can exist, and thus neither can a sequence g satisfying (39). 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. ∎

Proof of Theorem 5: The proof is based on the following standard lemma:

Lemma 10 *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

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

where

$$\begin{aligned}
\hat{r}^2_{a,t} &= \left(\frac{Q_{a,t}}{\tau_v + Y_t} - \frac{Q_{a,t-1}}{\tau_v + Y_{t-1}}\right)^2(\tau_v + Y_{t-1}) + \left(\frac{Q_{a,t}}{\tau_v + Y_t}\right)^2 y_t \\
&= \left(\frac{Q^2_{a,t}}{(\tau_v + Y_t)^2} + \frac{Q^2_{a,t-1}}{(\tau_v + Y_{t-1})^2} - 2\frac{Q_{a,t}Q_{a,t-1}}{(\tau_v + Y_{t-1})(\tau_v + Y_t)}\right)(\tau_v + Y_{t-1}) + \left(\frac{Q_{a,t}}{\tau_v + Y_t}\right)^2 y_t \\
&= \frac{Q^2_{a,t}}{\tau_v + Y_t} + \frac{Q^2_{a,t-1}}{\tau_v + Y_{t-1}} - 2\frac{Q_{a,t}(Q_{a,t} - q_{a,t})}{\tau_v + Y_t} \\
&= \frac{Q^2_{a,t-1}}{\tau_v + Y_{t-1}} - \frac{Q^2_{a,t}}{\tau_v + Y_t} + 2\frac{q_{a,t}Q_{a,t}}{\tau_v + Y_t}.
\end{aligned}$$

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,

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

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 10, then in turn implies that, conditioned on ξ_t,

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

It follows immediately that the unconditional expectation of the first term (not conditioning on ξ_t) is $\sqrt{\frac{2}{\pi}}\frac{\sqrt{\tau\Delta V_{a,t}}}{\sqrt{\frac{\tau r^2_{a,t}}{\Delta V_{a,t}}+1}} = \sqrt{\frac{2}{\pi}}\frac{\sqrt{\tau}\Delta V_{a,t}}{\sqrt{\tau r^2_{a,t}+\Delta V_{a,t}}}$.

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

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

Summing the two terms together, we get

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

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 (21)

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

∎

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

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

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. ∎

B Centrality and profits in large and medium size random networks

We study the asymptotic relationship between agent profitability and several common centrality measures in large networks. To this end, 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 Erdös (1947), Erdös and Rényi (1959), and Gilbert (1959), in which links between agents are formed randomly, independently, and with constant probability.[11]

Our focus is on sparse networks—in line with what is observed in practice (see the discussion in Ozsoylev, Walden, Yavuz, and Bildik (2014)). 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 $k > 3$.[12] 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 Erdös-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 $c\frac{\log(N)^k}{N-1}$.*

Our goal is to explore the relationship between this profitability measure and standard centrality measures in the literature, for large networks. 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 (14) with these measures.

Definition 5

- *The* degree centrality *vector is the vector of first-order degrees, V^1, i.e., the degree centrality of agent a is $V_{a,1}$.*

[11] We focus on the Erdös-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).

[12] The restriction $k > 3$ is needed for technical reasons.

- *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*

$$K = K^\alpha = \sum_{t=1}^{\infty} \alpha^t E^t \mathbf{1}.$$

  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 $C = \lambda E C$, for the largest possible eigenvalue, λ, where we normalize C such that $\sum_{a \in \mathcal{N}} C_a = 1$.*

We use cross sectional correlation across agents as the metric of similarity.[13] Specifically, for a given network we define the cross sectional correlation between the different centrality measures and profitability, $\rho_D = Corr(V^1, \Pi)$, $\rho_F = Corr(-F, \Pi)$, $\rho_{\hat{C}} = Corr(\hat{C}, \Pi)$, $\rho_{K^\alpha} = Corr(K^\alpha, \Pi)$, and $\rho_C = Corr(C, \Pi)$. 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 γ, τ, τ_u, τ_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 6 *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*

$$E\left[\rho_{K^\alpha}\right] > E\left[\rho_D\right] > \max\left(E\left[\rho_{\hat{C}}\right], E\left[\rho_F\right]\right). \tag{40}$$

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; Buraschi and Porchia 2012), or is empirically motivated (e.g., Das and Sisk 2005; Adamic, Brunetti, Harris, and Kirilenko 2010; Li and Schurhoff 2012; Ozsoylev, Walden, Yavuz, and Bildik 2014,). 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 (14) 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 (Π) and the different centrality measures. The results are shown in Table 1. We see that the ranking is the same as in Theorem 6 and, furthermore, that eigenvector centrality, being a special case of Katz centrality, performs about as well as the Katz centrality measure.

Proof of Theorem 6:

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

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.

We focus on the Erdös-Renyi random graph $G(N, p_N)$ model, where

$$p_N = c\log^k(N)/(N-1), \qquad c > 0, \qquad k > 3, \tag{41}$$

$$q_N = p(N-1) = c\log^k(N). \tag{42}$$

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 $q^\alpha << N^\beta$, 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 $\lim_{N\to\infty} f(N) = 0$, if for large N

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

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

$$\rho_D \stackrel{\text{def}}{=} \rho[D_N, \Pi_N],$$

$$\rho_F \stackrel{\text{def}}{=} \rho[-F, \Pi_N],$$

$$\rho_{\hat{C}} \stackrel{\text{def}}{=} \rho[\hat{C}, \Pi_N],$$

$$\rho_K \stackrel{\text{def}}{=} \rho[K^\alpha, \Pi_N].$$

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 11 *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 $a = 1, \ldots, N$, for all $t = 1, \ldots, T$,*

$$\mathbb{P}\left(\left|V_{a,t} - V_{a,t-1} - q^t\right| \geq \epsilon q^t\right) \leq N^{-K}. \tag{43}$$

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,

$$\mathbb{P}\left(\left|V_{a,t} - V_{a,t-1} - q^t\right| \geq \delta_t q^t\right) \leq N^{-K}, \tag{44}$$

where

$$\delta_t \leq C' \log(N)^{\frac{1-k}{2}}, \qquad t = 1, \ldots, T, \tag{45}$$

and since $\delta_t \to 0$ when $N \to \infty$, the result follows. ∎

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

$$\bar{V}_t \stackrel{\text{def}}{=} \frac{1}{N}\sum_{a=1}^{N} V_{a,t}, \qquad t = 1, \ldots, T. \tag{46}$$

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 (11) immediately implies that

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

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

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

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

$$\mathbb{P}\left(\left|y_t - \alpha q^{2t}\right| \geq \epsilon q^{2t}\right) \leq N^{-K}. \tag{47}$$

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 (47) now immediately imply that

$$\mathbb{P}\left(\left|\zeta_t \times q^{2(t-1)} - 1\right| \geq \epsilon\right) \leq N^{-K}. \tag{48}$$

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 $\zeta_t q^{2(t-1)} \in [1-\epsilon, 1+\epsilon]$, $t = 1, \ldots, T$, for some fixed $\epsilon << 1$. From (48) we know that this set of events satisfies

$$\mathbb{P}(\Omega_N) \geq 1 - O(q^{-K}) \tag{49}$$

for arbitrarily large K.

We have

$$\begin{aligned}
\rho[D_N, \Pi_N] &= \rho[V_{a,1}^N, \beta_1(V_{a,1}^N + \zeta_2 V_{a,2}^N + \zeta_3 V_{a,3}^N + \cdots + \zeta_N^T V_{a,T}^N)] \\
&= \frac{S_1^2 + \zeta_2 C[V_{a,1}, V_{a,2}] + \cdots + \zeta_T C[V_{a,1}, V_{a,2}]}{S_1 \sqrt{V[V_{a,1} + \zeta_2 V_{a,2} + \cdots + \zeta_T V_{a,T}]}} \\
&= \frac{S_1^2 + \sum_{i=2}^T \zeta_i S_{i,1}}{S_1 \sqrt{S_1^2 + \sum_{i=2}^T \zeta_i^2 S_i^2 + \sum_{i=1,j>i}^T 2\zeta_i \zeta_j S_{j,i}}},
\end{aligned} \tag{50}$$

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

$$S_t^2 = V[V_{a,t}], \qquad t = 1, 2, \ldots T, \tag{51}$$

$$S_{t,s} = C[V_{a,t}, V_{a,s}], \qquad s, t = 1, 2, \ldots, T. \tag{52}$$

We define $a_t = \zeta_i q^{2(t-1)}$, $t = 2, \ldots, T$, 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 (49), 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 12 *Given the definitions (51), (52), we have*

$$S_t^2 = \sum_{i=1}^t Z_i^2 + 2 \sum_{i=1}^t \sum_{j=i+1}^t q^{j-i} Z_i + O_s(N^{-1/3+\epsilon}), \qquad 1 \leq t \leq T, \tag{53}$$

$$S_{t,s} = S_s^2 + \sum_{i=1}^s \sum_{j=s+1}^t q^{j-i} Z_i + O_s(N^{-1/3+\epsilon}), \qquad 1 \leq s, t \leq T. \tag{54}$$

Here,

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

Proof of Lemma 12:

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 $Z_{a,1} = \xi_{a,0}$, $Z_{a,t+1} = \sum_{i=1}^{V_{a,t}} \xi_{a,t,i}$, $a = 1, \ldots, N$, $t = 1, \ldots, T-1$. Here, $\xi_{a,t,i} \sim Bin(N-1, p)$, 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

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

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

$$\begin{aligned} E[Z_{a,t}] &= q^t, \\ Var[Z_{a,t}] &= q^t \sum_{i=0}^{t-1} q^i + O(N^{-1+\epsilon}), && (55) \\ Cov[Z_{a,t}, Z_{a,s}] &= q^t \sum_{i=0}^{s-1} q^i + O(N^{-1+\epsilon}), \qquad t > s. && (56) \end{aligned}$$

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

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

leading to

$$\begin{aligned} E[V_{a,2} - V_{a,1} | V_{a,1}] &= (N - V_{a,1})(1 - (1-p)^{V_{a,1}}) \\ &= -N(1 - V_{a,1}/N) \left(\sum_{i=1}^{V_{a,1}} \binom{V_{a,1}}{i} (-1)^i p^i \right) \\ &= -N(1 - V_{a,1}/N) \left(\sum_{i=1}^{V_{a,1}} \binom{V_{a,1}}{i} \left(\frac{-q}{N-1} \right)^i \right), && (57) \end{aligned}$$

and

$$\begin{aligned} Var[V_{a,2} - V_{a,1} | V_{a,1}] &= (N - V_{a,1})(1 - (1-p)^{V_{a,1}})(1-p)^{V_{a,1}} \\ &= -N(1 - V_{a,1}/N)(1-p)^{V_{a,1}} \left(\sum_{i=1}^{V_{a,1}} \binom{V_{a,1}}{i} (-1)^i p^i \right) \\ &= -N(1 - V_{a,1}/N)(1-p)^{V_{a,1}} \left(\sum_{i=1}^{V_{a,1}} \binom{V_{a,1}}{i} \left(\frac{-q}{N-1} \right)^i \right). && (58) \end{aligned}$$

The law of total variance now implies that

$$Var[V_{a,2} - V_{a,1}] = E[Var[V_{a,2} - V_{a,1} | V_{a,1}]] + Var[E[V_{a,2} - V_{a,1} | V_{a,1}]], \qquad (59)$$

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, $E\left[\sum_{i=2}^{\infty} V_{a,t}^i/(N-1)^i\right] = O(N^{-1})$, (57,58) immediately imply, when plugged into (59), that

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

The same argument for

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

leads to

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

and for higher orders to

$$Var[V_{a,t} - V_{a,t-1}] = Var[Z_{a,t}] + O(N^{-1+\epsilon}). \qquad (60)$$

A similar argument for the covariances leads to,

$$Cov[V_{a,t}-V_{a,t-1},V_{a,s}-V_{a,s-1}] \quad = \quad Cov[Z_{a,t},Z_{a,s}]+O(N^{-1+\epsilon}), \qquad t>s. \tag{61}$$

Now, since $V_{a,t}=V_{a,1}+(V_{a,2}-V_{a,1})+\ldots+(V_{a,t}-V_{a,t-1})$, from (60,61) it immediately follows that

$$Var[V_{a,t}] \quad = \quad \sum_{i=1}^{t} Z_i^2+2\sum_{i=1}^{t}\sum_{j=i+1}^{t} q^{j-i}Z_i+O_s(N^{-1+\epsilon}), \qquad \text{and} \tag{62}$$

$$Cov[V_{a,t},V_{a,s}] \quad = \quad Var[V_{a,s}]+\sum_{i=1}^{s}\sum_{j=s+1}^{t} q^{j-i}Z_i+O_s(N^{-1+\epsilon}), \qquad t>s. \tag{63}$$

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,

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

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,

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

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

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

where I is Bernoulli distributed, $I\sim Ber(p)$, representing the possible common link between a and b, $w_a\sim Bin(N-1,p)$, $w_b\sim Bin(N-1,p)$ 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

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

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

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

An identical argument as that above shows that

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

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

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

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

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

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

$$Cov((V_{a,t}-V_{b,t})^m, (V_{c,s}-V_{d,s})^n) = O(N^{-1+\epsilon}), \tag{66}$$

$$Corr((V_{a,t}-V_{b,t})^m, (V_{c,s}-V_{d,s})^n) = O(N^{-1+\epsilon}). \tag{67}$$

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

Lemma 13 *Assume $x_1, \ldots, x_N$ are identically distributed (but not necessarily independent) random variables, with mean μ and (finite) variance σ^2. Also, assume that $Corr(x_i, x_j) \leq CN^{-\alpha}$, $\alpha > 0$. 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 13: It is clear that $Var(\bar{x}_N) = \frac{\sigma^2}{N^2}(N + \sum_{i,j\neq i} Corr(x_i, x_j)) \leq C^2\sigma^2 N^{-\beta}$. 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 14 *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 14: The proof is by induction. Assume that the result has been proved for all $m = 1, 2, \ldots, M$, i.e.,

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

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, $i = 1, \ldots, M$. Now, clearly under our induction assumption,

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

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,

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

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

Lemma 15 *Assume $x_1, \ldots, x_N$ are identically distributed (but not necessarily independent) random variables, with mean μ, variance σ^2, and finite fourth moments. Assume that $|Corr(x_i, x_j)| \leq CN^{-\alpha}$ and that $|Corr((x_i - x_j)^2, (x_k - x_n)^2)| \leq CN^{-\alpha}$, 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 15: 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

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

implying that

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

in turn leading to

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

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

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

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

Thus, altogether, $Var(s_N^2)=O(N^{-\beta})$. As in Lemma 13, the result then follows from Chebyshev's inequality. ∎

Lemma 16 *Assume $x_1,\ldots,x_N$, and $y_1,\ldots,y_N$, are identically distributed random variables, with means μ_X, μ_Y, variances σ_X^2, σ_Y^2, covariance $\sigma_{XY}=Cov(x_i,y_i)$ for all i, and finite fourth moments. Assume that $|Corr(x_i,y_j)|\leq CN^{-\alpha}$, and that $|Corr((x_i-x_j)^2,(y_k-y_n)^2)|\leq CN^{-\alpha}$, 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 16: Identical to the Proof of Lemma 15. ∎

Lemma 15, together with (62) and (65) therefore implies (53), and Lemma 16 together with (63) and (67) implies (54). We have shown Lemma 12. ∎.

Plugging (53-54) into (50), we get

$$\begin{aligned}\rho[D_N,\Pi_N] &= \frac{S_1^2+\sum_{i=2}^{T}\zeta_i S_{i,1}}{S_1\sqrt{S_1^2+\sum_{i=2}^{T}\zeta_i^2S_i^2+\sum_{i=1,j>i}^{T}2\zeta_i\zeta_jS_{j,i}}}+O_s(N^{-1/3+\epsilon}) \\ &= \frac{q(1+a_2(q^{-1}+q^{-2})+a_3(q^{-2}+q^{-3})+a_4q^{-3}+q^{-4}P_1(q^{-1}))}{q(1+a_2^2(q^{-2}+3q^{-3})+2a_2(q^{-1}+q^{-2})+2a_3(q^{-2}+q^{-3})+2a_4q^{-3}+2a_2a_3q^{-3}+q^{-4}Q(q^{-1}))^{1/2}} \\ &+ \; O_s(N^{-1/3+\epsilon}), \\ &= \frac{1+a_2(q^{-1}+q^{-2})+a_3(q^{-2}+q^{-3})+a_4q^{-3}+q^{-4}P_1(q^{-1})}{1+a_2^2(q^{-2}+3q^{-3})+2a_2(q^{-1}+q^{-2})+2a_3(q^{-2}+q^{-3})+2a_4q^{-3}+2a_2a_3q^{-3}+q^{-4}Q(q^{-1})^{1/2}} \\ &+ \; O_s(N^{-1/3+\epsilon}),\end{aligned} \tag{68}$$

where P and Q are polynomials finite orders, and we define $a_1=1$. A Taylor expansion of (68) 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

$$\rho_D \quad = \quad 1 - cq^{-3} + O_s(q^{-4}), \tag{69}$$

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

$$\mathbb{P}\left(|1 - \rho_D - cq^{-3}| \geq Cq^{-4}\right) \leq Cq^{-4}, \tag{70}$$

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

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

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

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

Thus, from (70), it follows that

$$1 - E[\rho_D] = cq^{-3}(1 + o(1)) \sim q^{-3}. \tag{71}$$

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

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

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

$$1 - E[\rho_K] \sim q^{-5}. \tag{72}$$

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

$$1 - E[\rho[D, F]] \sim q^{-1}. \tag{73}$$

We note that (55,56) imply that for $i > 1$, and fixed α,

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

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

$$Corr\left(Z_1, \sum_{i=1}^{R} Z_i\right) \quad = \quad 1 - \frac{1}{2} q^{-1} + O(q^{-2}). \tag{74}$$

Now, from Theorem 10.10 in Bollobas (2001), it follows that if we define $R = \lfloor \log(N)/\log(c) \rfloor + 1$, 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 (74), if we replace $V_{a,i}-V_{a,i-1}$ with $Z_{a,i}$,

$$\begin{aligned} \tilde{F_a} &\stackrel{\text{def}}{=} \frac{1}{N}\left(Z_{a,1}+\sum_{i=2}^{R-1}(iZ_i)+R\left(N-\left(Z_{a,1}+\sum_{i=2}^{R-1} Z_1\right)\right)\right) \\ &= R-\frac{R-1}{N}Z_{a,1}-\sum_{i=2}^{R-1}\frac{R-i}{N}Z_{a,i} \end{aligned}$$

we have

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

A similar argument as that leading to (69) therefore implies that

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

and

$$\rho[D,-F] = 1-\frac{q}{2}+O_s(q^{-2}), \tag{75}$$

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

Proof of Lemma 17: 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 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 $d(a,c)\le d(a,b)+d(b,c)$, leading to $d(b,c)\ge d(a,c)-d(b,c)$. 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, $d(a,c)^2=2-2\rho(a,c)$ which via the triangle inequality leads to $2-2\rho(a,c)\ge(\sqrt{2\beta}-\sqrt{2\alpha})^2=2(\sqrt{\beta}-\sqrt{\alpha})^2$, leading to the result. We are done. ∎

Equations (70) and (75), together with Lemma 17, where $a=D$, $b=-F$, $c=\Pi$, now implies that if $1-\rho[D,\Pi]\le cq^{-3}$, and $1-\rho[D,-F]\ge cq^{-1}$, then

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

for large N, and thus,

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

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

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

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

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

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

A similar argument as for ρ_F, again using Lemma 17, then implies the final part of the theorem, $E[\rho_D] > E[\rho_{\hat{C}}]$. 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.

Andrei, D., and J. Cujean (2014): "Information Percolation, Momentum and Reversal," 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.

Barber, B., Y. T. Lee, Y. J. Liu, and T. Odean (2009): "Just how much do individual investors lose by trading?," *Review of Financial Studies*, 22, 609–632.

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 (2007a): "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes," *Journal of Political Economy*, 115(5), 707–747.

Calvet, L., J. Campbell, and P. Sodini (2007b): "Down or Out: Assessing the welfare costs of household investment mistakes," *Journal of Political Economy*, 115, 707–747.

Calvet, L., J. Campbell, and P. Sodini (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.

Cohen, L., A. Frazzini, and C. Malloy (2008): "The small world of investing: Board connections and mutual fund returns," *Journal of Political Economy*, 116, 951–979.

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.

Geweke, J., and S. Porter-Hudak (1983): "The Estimation and Application of Long-Memory Time Series Models," *Journal of Time Series Analysis*, 4, 221–238.

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.

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

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

Grossman, 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.

Lobato, I. N., and C. Velasco (2000): “Long Memory in Stock-Market Trading Volume,” *Journal of Business and Economic Statistics*, 18, 410–427.

Manela, A. (2014): “The Value of Diffusing Information,” Journal of Financial Economics, forthcoming.

Massa, M., and A. Simonov (2006a): “Hedging, Familiarity, and Portfolio Choice,” *Review of Financial Studies*, 19(2), 633–685.

Massa, M., and A. Simonov (2006b): “Hedging, familiarity and portfolio choice,” *Review of Financial Studies*, 19, 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 (2014): “Investor Networks in the Stock Market,” *Review of Financial Studies*, 27, 1323–1366.

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

——— (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.

Tetlock, P. C. (2010): "Does Public Financial News Resolve Asymmetric Information," *Review of Financial Studies*, 23, 3520–3557.

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.