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

Preliminary and incomplete

Johan Walden†

December 5, 2012

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 profitability 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 how central the agent is, using a centrality measure that is closely related to so-called Katz centrality. The model generates rich dynamics of aggregate trading volume and volatility, beyond what can be generated by heterogeneous preferences in a symmetric setting. Casual observations suggest that price and volume dynamics of small stocks in the market may be especially well explained by such asymmetric information diffusion. The model could potentially be used to study individual investor behavior and performance, and to analyze endogenous network formation in financial markets.

*I thank seminar participants at the Duisenberg School of Finance, Maastricht University, Rochester (Simon School) and the Oxford-Man Institute, Nicolae Garleanu, Brett Green, Pierre de la Harpe, Svante Jansson, Ron Kaniel, Anders Karlsson, Martin Lettau, Dmitry Livdan, Hanno Lustig, Gustavo Manso, Christine Parlour, and Han Ozsoylev for valuable comments and suggestions.

†Haas School of Business, University of California at Berkeley, 545 Student Services Building #1900, CA 94720-1900. E-mail: walden@haas.berkeley.edu. Phone: +1-510-643-0547. Fax: +1-510-643-1420.
1 Introduction

There is extensive evidence that decentralized and heterogeneous 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 et al. (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 are hedging motives and heterogeneous preferences, but several studies indicate that there are limitations to how well such motives can explain observed heterogeneous investor behavior.1 With decentralized information diffusion, however, it is unsurprising if significantly heterogeneous behavior of investors is observed in the market.

Second, stock markets are known to experience large price movements that are unrelated to public news, as documented in Cutler et al. (1989), and Fair (2002). These studies find that over two thirds of major stock market movements cannot be attributed to public news events, suggesting that there are other channels through which information is incorporated into asset prices.

Third, the dynamics of trading volume and asset prices are known to be very rich. Returns and trading volume in many markets are heavy-tailed, time varying, show “long-term memory,” and are related to each other in a complex way (see Gabaix et al. (2003), Karpoff (1987), Gallant et al. (1992), Bollerslev and Jubinski (1999), and Lobato and Velasco (2000)). Lumpy information diffusion provides a potential explanation for such behavior. In periods when more information arrives, volatility is higher, as is trading volume (see, Clark (1973), Epps and Epps (1976), and Andersen (1996)). To generate rich trading volume dynamics, however, such information diffusion must necessarily be heterogeneous.

So, given the potentially important role of decentralized, heterogeneous, information diffusion in markets, how can it be modeled? In this paper, we follow a recent strand of literature that uses information networks to model such diffusion, see Colla and Mele (2010), Ozsoylev and Walden (2011), Han and Yang (2011), and Ozsoylev et al. (2012). Agents who are directly linked in a network share information with each other, and over time information therefore diffuses among the population in a well-specified manner. This literature has made important observations about effects of investor networks, but several key questions still remain open. How will the network structure in a market determine the dynamic trading behavior of its agents, and their performance? How does the network structure influence aggregate properties of the market, e.g., price volatility and trading volume? What does heterogeneous information diffusion “add” compared with, e.g., what can be generated by

---

1For example, Massa and Simonov (2006) find that hedging motives for human capital risk—a fundamental source of individual investor risk—does not explain heterogeneous investment behavior among individual investors well. Similarly, Calvet et al. (2007) and Calvet et al. (2009) find that diversification and portfolio rebalancing motives do not explain investors’ portfolio holdings well.
heterogeneous preferences alone? These questions are obviously fundamental for our understanding of the impact of information networks on financial markets and, in the extension, for how well such information networks can explain observed investor and market behavior. In this exploratory paper we analyze these questions.

We introduce a dynamic noisy rational expectations model in which agents in a network share information with their neighbors. Briefly, 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, agents share with their neighbors all the information they have received up until that point, leading to gradual diffusion of the private signals. The structure of the network is completely general.

We show the existence of a noisy rational expectations equilibrium, and derive closed-form expressions for all variables of interest. Theorem 1 provides the main existence and characterization result for a price taking equilibrium in a large economy. To define the large economy equilibrium, we use the concept of replica networks, assuming that there is a local network structure (e.g., at the level of a municipality) and that there are many similar such network structures in the economy. This technical trick 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. Throughout the paper, we use standard results from deterministic and random graph theory.

The structure of the network is crucial for asset pricing dynamics. For example, we show in a simple example that price informativeness and volatility at any given point in time does not only depend on the specific information agents in the model have obtained at that point, but also on how the information has diffused through the network. The equilibrium outcome thus depends on high dimensional properties of the network, beyond, e.g., 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 different agents. A priori, one may expect that portfolio holdings of agents who are close in the network should be positively related, 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. 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 their connectedness. This results justifies the use of data on trades to draw inferences about network structure.

It is argued in Ozsoylev and Walden (2011) that some type of centrality measure should be important in determining agent profitability. Centrality captures the concept that it is not only who your neighbors are that matters, but also who your neighbors’ neighbors are, who your neighbors’ neighbors’ neighbors are, etc. Specifically, investors who are centrally placed in a network should tend to receive information signals early and therefore perform better, whereas peripheral investors should tend to receive them later and perform worse. Ozsoylev and Walden (2011) do not show which specific centrality measure—among many—that captures profitability. In an empirical study, Ozsoylev et al. (2012) 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 cen-
trality measure is somewhat arbitrary. Several other finance papers discuss and use various centrality measures without a general theoretical justification, e.g., Das and Sisk (2005), Adamic et al. (2010), and Buraschi and Porchia (2012).

We show that profitability is determined by a centrality measure closely related to eigenvector centrality or, more generally, to so-called Katz centrality. To the best of our knowledge, this is the first complete characterization of the relationship between agent centrality and performance in information networks. For large random networks, Katz centrality dominates other common centrality measures, i.e., degree centrality, closeness centrality, and betweenness centrality. We also show in simulations that the same ranking of centrality measures holds in medium-sized random networks. Thus, the use of centrality measures when studying individual investor behavior, e.g., in Ozsoylev et al. (2012), is theoretically justified. Our analysis also leads to closed form expressions for the welfare of agents in the economy. This opens up for the study of endogenous network formation in financial markets in future work. Our main results that characterize trading behavior, profitability and welfare of agents, and the relationship between centrality and profitability, are Theorems 2-4.

We next derive several aggregate results regarding the dynamic behavior of price volatility and trading volume. The network structure in an economy is important for the distribution of return volatility and trading volume in the time series, as well as predicts relationships between the two. Several stylized properties naturally arise in the model. For example, lead-lag relationships between trading volume and return volatility, as well as persistent volatility and volume dynamics, are natural properties in the information network setting.

We also show that heterogeneous preferences among agents alone, in the form of different risk aversion in a symmetric network setting, cannot generate as rich dynamics of volatility and volume, as a general information network. We furthermore look at the behavior of stock prices of some large and small companies, and find that whereas symmetric networks may be consistent with the behavior of large stocks, the dynamic behavior of small stocks is better explained by asymmetric networks, in which some agents are better positioned than others. Thus, the richness of the general information network model may be key to understanding the rich dynamics of volatility and trading volume. Our theoretical results are summarized in Theorems 5-7.

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

2 Related Literature

Our paper is most closely related to the recent strand on 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 “cycle,” 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 static model is not appropriate for studying dynamic information diffusion, however, and is therefore not well suited for several of the questions analyzed in this paper, e.g., the relationship between agent profitability and centrality, and the short-term correlation between agents’ trade.

Han and Yang (2011) 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 et al. (2012) 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 (2009) studies how information networks—proxied by the commonality in stock holdings—among mutual is related to return momentum.

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

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

Our paper further explores the richness of the dynamics of volatility and volume that arises when agents share their signals, and there is 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 et al. (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 et al. (1994), Schneider (2009), and Breon-Drish (2010). Our model does not explore this potential informational role of trading volume.

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

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

3 Model

In this section, we introduce the model and characterize equilibrium. Theorem 1 will be the main work-horse result that we build upon in the rest of the paper.

3.1 Network

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

\[
U_a = E[-e^{-\gamma_a \tilde{W}_{a,T+1}}].
\]

We summarize agents’ risk aversion coefficients in the vector \( \Gamma = (\gamma_1, \ldots, \gamma_N) \).

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

Agents are connected in a network. A network is described by a graph \( \mathcal{G} = (\mathcal{N}, \mathcal{E}) \). The relation \( \mathcal{E} \subset \mathcal{N} \times \mathcal{N} \) describes which agents (vertices) are connected in the network. Specifically, \( (a, a') \in \mathcal{E} \), if and only if there is a connection (edge) 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 herself, \((a, a) \in E\) for all \(a \in N\), i.e., \(E\) is reflexive. We also assume that connections are bidirectional, i.e., that \(E\) 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 \((a, a') \in E\) 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 agents adjacent to agent \(a\) is \(S_{a, 1} = \{a' : (a, a') \in E\}\). Further, 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 \(R_{a, m} = \bigcup_{j=1}^{m} S_{a, m}\). The number of agents at a distance not further away than \(m\) from agent \(a\) is \(V_{a, m} = |R_{a, m}|\) (here \(|A|\) denotes the number of elements in the set \(A\)). Here, \(V_{a, 1}\) is the degree of agent \(a\), which we also refer to as agent \(a\)'s connectedness. We use the convention that \(V_{a, 0} = 0\) for all \(a\). We also define \(\Delta V_{a, m} = |S_{a, m}|\). We note that if we define the \(N\)-vectors \(V^m\), where the \(a\)th element of \(V^m\) is \(V_{a, m}\), (agent \(a\)'s \(m\)-th order degree) then we can define

**Definition 1** The \(m\)th order degree vector, \(V^m \in \mathbb{R}^N_+\), \(m = 1, 2, \ldots\), is defined as

\[
V^m = \chi(E^m)1. \tag{1}
\]

Here \(E^m\) is the \(m\)th power of the adjacency matrix, and \(\chi: \mathbb{R}^{N \times N} \rightarrow \{0, 1\}^{N \times N}\) is a matrix indicator function, such that \((\chi(A))_{i,j} = 0\) if \(A_{i,j} = 0\) and \((\chi(A))_{i,j} = 1\) otherwise. Further, \(1\) is an \(N\)-vector of ones.

First order degree is commonly referred to as degree centrality. Finally, the number of common agents within a distance of \(m\) from agents \(a\) and \(a'\) is \(V_{a,a', m} = |R_{a, m} \cap R_{a', m}|\), and the number of common neighbors at a distance of (exactly) \(m\) from both agents is \(\Delta V_{a,a', m} = |S_{a, m} \cap S_{a', m}|\).

### 3.2 Information diffusion

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

The graph, \(G\), 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 \(\tilde{I}_{a,t}\) denote the information set that agent \(a\) has received up until \(t\), either directly or via his network.

A natural question is to ask why agents would voluntarily reveal valuable information to their neighbors. In a large economy, sharing signals with ones’ neighbors has no cost, since the actions of a finite number of agents will not influence prices. Even with a finite size economy, as long as signals can be verified ex post, truthful revelation may be optimal in a repeated game setting, since an agent’s misinformation can be punished by his neighbors by future exclusion from the network. 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.

Following Ozsoylev et al. (2012), we formalize the information sharing role of the network, by defining

**Definition 2** The graph $G$ 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 $t = 1, \ldots, T$, $s_{a'} \in \hat{I}_{a,t}$ if and only if $D(a, a') \leq t$.

The information about the asset value that an agent has received through the network up until time $t$ can be summarized by the sufficient statistic

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

where $\zeta_{a,t} = \frac{V_{a,t}}{\sigma^2} \xi_{a,t}$, and $\xi_{a,t} \sim N(0,1)$.

### 3.3 Market

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

To avoid revealing prices, we make the standard assumption of stochastic supply of the asset. Specifically, in period $t$ “noise traders” submit market orders of $u_t$ per trader in the network, where $u_t \sim N(0, \sigma_u^2)$. In other words, the noise trader demand is defined relative to the size of the population in the information network. As argued elsewhere in the literature, the noise trader assumption need not be taken literally, but is rather a reduced-form representation of stochastic supply, or of unmodeled demand shocks. It could, e.g., represent hedging demand among investors, due to unobservable wealth 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. Their total demand at time $t$ is

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

subject to the budget constraint

$$\hat{W}_{a,t+1} = \hat{W}_{a,t} + x_{a,t}(p_{t+1} - p_t), \quad t = 1, \ldots, T,$$
and their 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, $I_{a,t}$ contains all public and private information available to agent $a$ at time $t$. In the linear equilibrium we shall study, $z_{a,t}$ and $p_t$ will be jointly sufficient statistics for an agent’s information set, $I_{a,t} = \{z_{a,t}, p_t\}$, 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 her (optimal) future strategies. The dynamic problem can therefore be solved by backward induction. The primitives of the economy are summarized by the tuple $\mathcal{M} = (G, \Gamma, \tau, \tau_u, \tau_v, \bar{v}, T)$.

The graph, $G$, determines how information diffuses in the network over time, whereas $\Gamma$ captures differences in preferences. We will 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 a market 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 will use is so-called distance transitivity.² 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, $f : N \leftrightarrow N$, such that $(f(a), f(a')) \in E$ if and only if $(a, a') \in E$. A graph is distance-transitive if for every pair of vertices, $(a, a')$, at distance $d$, there is a distance preserving automorphism, i.e., an automorphism, $f$, such that the distance between $f(a)$ and $f(a')$ is also $d$. An economy is said to be network symmetric if its graph is distance-transitive. Network symmetric economies and preference symmetric economies, provide useful benchmarks to which the general class of economies can be compared.

We note that network symmetry does not imply that the same amount of information is incorporated into prices at each point in time. It does, however, still put severe restrictions on how information spreads in the economy, as shown by the following lemmas:³

**Lemma 1** In a network symmetric economy, $\Delta V_{a,t}$ is the same for all agents for all $t$, i.e., for each $t$, $\Delta V_{a,t} = \Delta V_t$ for some $\Delta V_t$ that is common across agents.

Thus, all agents have an equal amount of information at any point in time in a network symmetric economy, although the signal realizations will typically differ.

**Lemma 2** In a network symmetric economy the sequence $\Delta V_1, \Delta V_2, \ldots, \Delta V_T$, is unimodal, i.e., there are times $1 \leq t_1 \leq t_2 \leq t_3 \leq T - 1$, such that $\Delta V_{t+1} > \Delta V_t$ for all $t \leq t_1$, $\Delta V_{t+1} = \Delta V_t$ for all $t_1 < t \leq t_2$, and $\Delta V_{t+1} < \Delta V_t$ for all $t > t_2$.

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

²Other 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.

³The 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 et al. (1989), page 167).
3.4 Replica network

To justify the assumption that agents are price takers, \( N \) needs to be large. Further, as analyzed in Ozsoylev and Walden (2011), restrictions on the distribution of number of links that agents may have are needed, to ensure existence of equilibrium. Ozsoylev and Walden (2011) carry out a fairly general analysis of the restrictions needed on the degree distribution for the existence of equilibrium to be guaranteed. They show that a sufficient condition is that the degree distribution is not too heavy-tailed. Compared with their static model, our model has the additional property of being dynamic. Therefore, not only would restrictions on first-order degrees be needed, but also on degrees of all higher orders. In the dynamic economy, signals spread over longer distances, thereby “fattening” the tail of the distribution of signals among agents over time. We believe that a general analysis would be technically challenging, while adding limited additional economic insight, and therefore choose a simplified approach compared with Ozsoylev and Walden (2011).

We build on the concept of replica economies, originally introduced by Edgeworth (1881) to study the core of an economy, and then further formalized in 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 the 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 make most idiosyncratic signals cancel out in aggregate, while avoiding the issues of signals spreading too quickly among some agents in a large network so that equilibrium breaks 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 \( A_m = N \times \{ 1, \ldots, M \} \), where \( a = (i, j) \in A_m \) represents the \( i \)th agent in the \( j \)th replica network, in a market with \( M \) replica networks. There is still one asset, one market, and a competitive market maker. We also 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 \( A_1 \subset A_2 \subset \cdots \subset A_m \subset \cdots \), and take the limit \( A = \lim_{M \to \infty} A_M \), letting \( A \) 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\(^4\)

\[
wt = \frac{1}{N} \sum_{a=1}^{N} \Delta x_{a,t} + ut. 
\]  \hfill (3)

\(^4\)Technically, the market maker observes \( \lim_{M \to \infty} \frac{1}{MN} \sum_{a=1}^{MN} \Delta x_{a,t} + ut \). We avoid such limit notation when this can be done without confusion.
3.5 Equilibrium

We focus on 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, further, characterizes this equilibrium.

**Theorem 1** Consider a network economy characterized by \(M\). For \(t = 1, \ldots, T\), define

\[
A_t = \frac{\tau}{N} \sum_{a=1}^{N} V_{a,t},
\]

\[
y_t = \tau u(A_t - A_{t-1})^2,
\]

\[
Y_t = \sum_{s=1}^{t} y_s,
\]

\[
C_{a,t} = \left( \frac{\tau_v + \tau V_{a,t+1} + Y_{t+1}}{\tau_v + \tau V_{a,t} + Y_t} \right) \left( 1 + \tau V_{a,t} \left( \frac{1}{\tau_v + Y_t} - \frac{1}{\tau_v + Y_{t+1}} \right) \right),
\]

\[
D_{a,t} = \prod_{s=t+1}^{T} C_{a,s}^{-1/2},
\]

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

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

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

\[
x_{a,t} = \frac{\tau V_{a,t}}{\gamma_a} (z_{a,t} - p_t), \tag{5}
\]

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

Several observations are in place. First, note that the price function (4) has a fairly standard structure. It is determined by the fundamental value \((\bar{v})\) 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 investors 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. Further,
preferences are allowed to vary across agents, through $\gamma_a$. This allows us to compare which variations in equilibrium dynamics that may arise solely from heterogeneous preferences, and which that can only arise from heterogeneous information diffusion.

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

Figure 1: Impact of network structure. The figure shows two networks with four agents: In Panel A, a tight-knit network is shown, in which every agent is connected with every other agent. In panel B, a less tight-knit network is shown.

4 Trading, profitability, and centrality of individual agents

This section focuses on the behavior and performance of individual agents, and how these relate to the agents’ positions in the network.
4.1 Correlation of trades

We are interested in the relationship between network structure and agents’ trading behavior, to understand whether the model can explain observed investors behavior, and also whether inferences about the underlying network structure can be drawn from observed trading behavior. As a possible example, 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 geographically far apart. Colla and Mele (2010) showed that information networks can give rise to such behavior of trade correlations, under the assumption that geographically close agents are also close in the information network. Their analysis was restricted to a cyclical network, but the effect was also shown to arise more generally in the static model of Ozsoylev and Walden (2011).

In Ozsoylev et al. (2012), this relationship 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 tended to trade in the same stock, in the same direction, at similar points in time, were assumed to be linked in the market’s investor network. Such an approach is justified in the static model of Ozsoylev and Walden (2011). The situation is more complicated in a dynamic setting, however, and it is a priori unclear whether their approach is justified. 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. This leads to similar portfolio holdings. In the static model agents’ trades and portfolio positions are equivalent, but they are not in the dynamic model: portfolio holdings in the dynamic model are equivalent to cumulative period-by-period trades.

For period-by-period trading behavior an additional complication of timing of information arrival arises. This timing will not be identical, even for agents who are close in the network. Consider, for example, a situation where one agent, $a$, has 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 that point take a large position in the asset (positive or negative, depending on whether the agent believes it is under- or over-valued). Further, assume that one of this agent’s neighbors, $b$, does not have many neighbors within a distance of $t$ but, being connected to $a$, receives a lot of information at time $t + 1$. Now, assume that at $t + 1$ there is also a substantial amount of information incorporated into prices (driven by many other agents who also have 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 prices will be closer to fundamental value at this point. Agent $a$, however, may actually decrease his/her position, to take home 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 very careful in choosing an appropriate window length when inferring network structure from observed trades.

The following theorem characterizes the covariance of agents’ trade, for an individual agent across time, and between agents at a specific point in time.
Theorem 2. The covariance between an agent’s trade at time $t$ and $t+1$ is

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

(7)

The covariance of trades between agent $a$ and $b$ at time $t$ is

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

(8)

Equation (7) shows that in the special case when the agent does not receive any new signals, and $\Delta V_{a,t} = 0$ therefore is 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 time $t$ 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 the agent takes home profits and decreases risk exposure by selling off stocks, 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 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 leaves in 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 two than in period 1, so he may tend to gradually scale 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 this setting, does not come out clearly if we restrict our attention to symmetric networks. Indeed in a network symmetric economy, $\Delta V_{a,t}$ is the same for all agents, and it can be shown that (7) takes the form

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

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

The issue here is that within the class of network symmetric economies, connectivity must be increased for all agents at the same time, so there is no way to increase the relative information advantage of only a few agents. 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 example shows that restricting one’s attention to symmetric economies may be misleading—in this case reversing the result. We will see further examples in Section 5.

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

The first main implication of (8) is that the covariance is always strictly positive. Thus, the situation where negative correlation arises 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. 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 $\sum_{j \in \Delta S_{a,t}} s_{j} / \Delta V_{a,t} = \tilde{v} + \zeta$, where the error term $\zeta \sim N(0, \sigma^{2}/\Delta V_{a,t})$ is independent of prices. The second term represents the agent’s sloping demand curve, leading the agent to ramp down investments when the price catches up, e.g., by selling stocks after the asset price was undervalued at $t − 1$ but increases at $t$. 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 and, assuming that agent $b$ receives a lot of
new information between $t - 1$ and $t$, the first term will be dominant. Negative correlation would then arise if agent $b$ tends to ramp up when agent $a$ ramps down, i.e., if $\text{Cov} \left( \tilde{v} + \zeta^b - p_t, -(p_t - p_{t-1}) \right) < 0$. However, since the market is semi-strong form efficient, $\tilde{v} - p_t$ is independent of $p_t - p_{t-1}$. Furthermore, since $\zeta^b$ is independent of aggregate variables, this correlation is zero. In other words, since agent $a$’s hedging demand between $t - 1$ and $t$ is publicly known at $t$, it must be independent of agent $b$’s time $t$ demand, which is due to informational advantage at time $t$. This strong result is model dependent. It depends on the linear structure of agents’ demand functions. However, to a first order approximation, we expect the result to hold in more general settings in semi-strong form efficient markets, given that trading for rebalancing purposes mainly depends on price changes, trading for informational purposes depends on the difference between the true value and market price, and the two terms are uncorrelated in a weak-form efficient market.

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

These properties of trade correlations suggest that it may be justified to use trades instead of portfolio holdings to draw inferences about a market’s information network as, e.g., done in Ozsoylev and Walden (2011), and that the time horizon over which trades are compared may be quite short.

![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 Panel B shows a network in which the trades are positively correlated. 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.](image)

4.2 Profitability

Our next question concerns who is profitable in an information network. Our starting point is the following theorem:
Theorem 3  Define

\[ \pi_{a,t} = \tau V_{a,t} \left( \frac{1}{\tau v + Y_t} - \frac{1}{\tau v + Y_{t+1}} \right), \quad t = 1, \ldots, T - 1, \ a = 1, \ldots, N \]

\[ \pi_{a,T} = \tau \frac{V_{a,T}}{\tau v + Y_T}, \quad a = 1, \ldots, N. \]

The ex ante certainty equivalent of agent \( a \) is

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

The total expected trading profit of agent \( a \) is

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

where

\[ \Pi_a = \sum_{t=1}^{T} (\tau v + Y_t)^{-1} V_{a,t} \quad (11) \]

is the profitability of agent \( a \).

We focus on profitability, and leave welfare implications of the model for future research. Equation (10) determines ex ante expected profits of an agent. It shows that expected profits depends on three important components, which are all intuitive. First, profits are inversely proportional to an agent’s risk-aversion, \( \gamma_a \), 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 \( \gamma_a \), so it corresponds to pure scaling. Therefore, we do not include it in our measure of profitability, as defined by Equation 11. Second, expected profits depend on an agent’s position in the network through \( V_{a,t} \), for \( t > 1 \): 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 so 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: Any given agent’s profitability decreases when other agents become more informed. Equation (11) thus provides a direct relationship between the properties of a network, local as well as aggregate, and individual agents’ profitability.

4.3 Centrality

Equation (11) shows that an agent’s profitability is determined by the agent’s centrality, defined appropriately. Recall that \( V_{a,t} \) denotes the number of agents that are within distance \( t \) from agent \( a \). So, \( V_{a,1} \) is simply the degree of agent \( a \). For \( t > 1 \), higher order links are also important in determining profitability. Specifically, \( V_{a,2} \) does not only depend on how connected an agent is, but also how connected the neighbors of that agent are, and similarly for higher orders.
We use (1) to rewrite (11) as

\[
\Pi = \sum_{t=1}^{\infty} \beta_t \chi(E^t)1,
\]

where \(\beta_t = (\tau + Y_t)^{-1}\), for \(t = 1, \ldots, T\), and \(\beta_t = 0\) for \(t > T\). Here, \(\Pi\) is an \(N\)-vector where the \(a\)th element is the profitability of agent \(a\). We explore the relationship between this measure and standard centrality measures in the literature. Specifically, we review the concepts of degree centrality, Katz centrality, eigenvector centrality, closeness centrality, and betweenness centrality (see, e.g., Friedkin (1991) for a detailed discussion of these concepts), and compare (12) with these measures.

**Definition 3** The degree centrality vector is the vector of first order degrees, \(V^1\).

**Definition 4** The Katz centrality vector with parameter \(\alpha < 1\) of a network with matrix representation \(E\) is the vector \(K \in \mathbb{R}_+^N\), defined as

\[
K = \sum_{t=1}^{\infty} \alpha^t E^t 1.
\]

(13)

Here, \(E^t\) denotes the \(t\)th power of the adjacency matrix, \(E\), and \(1 \in \mathbb{R}^N\) is an \(N\)-vector of ones. When stressing the value of \(\alpha\), we write \(K^\alpha\).

The interpretation is that the \(a\)th element of the vector \(K\) represents agent \(a\)’s Katz centrality. We see that the structures of the profitability measure (12) and Katz centrality (13) are similar. Specifically they are both made up by a weighted sum of powers of the neighborhood matrix, multiplied with the vector of ones. The differences are that the weighting is a power of \(\alpha\) for Katz centrality but varies more generally with \(t\) for profitability, and that the matrix indicator function, \(\chi\), is taken on the power of the adjacency matrix in the profitability measure.

**Definition 5** The eigenvector centrality vector is the eigenvector corresponding to the largest eigenvalue of \(E\), i.e., the vector \(C\) that solves the equation \(C = \lambda EC\), for the largest possible eigenvalue, \(\lambda\), where, for uniqueness, we normalize \(C\) such that \(\sum_a C_a = 1\).

Similarly, to Katz centrality, the interpretation is that \(C_a\) is agent \(a\)’s eigenvector centrality. It is a standard result that eigenvector centrality can be viewed as a special case of Katz centrality, since \(C = \lim_{\alpha \to \lambda^{-1}} \frac{K^\alpha}{\sum_a K^\alpha_a}\).

Another commonly used centrality measure is closeness centrality.

**Definition 6** The closeness centrality of agent \(a\), \(\hat{C}_a\), is defined as

\[
\hat{C}_a = \frac{N - 1}{\sum_{a' \neq a} D(a, a')},
\]

\[\tag{5}\]
The closeness centrality is thus defined as the inverse of the average distance from an agent to all other agents. We note that for any two agents, $a$ and $a'$, who are not connected at any order, $\hat{C}_a = \hat{C}_{a'} = 0$ since $D(a, a') = \infty$ for two such agents.

A final centrality measure is given by betweenness centrality. Given that $D(i, j) < \infty$ between two nodes, $i \neq j$, define $f_{ij}$ to be the number of paths between $i$ and $j$ of length $D(i, j)$. Further, define $f_{ij}^a$ to be the number of such paths that pass through the third node, $a$, and finally the set $S_a \overset{\text{def}}{=} \{(i, j) \in \mathbb{N}^2 : 1 \leq i \leq N, 1 \leq j \leq N, i \neq j, i \neq a, j \neq a\}$. We then have

**Definition 7** Agent $a$’s betweenness centrality is defined as

$$B_a = \sum_{(i,j) \in S_a} \frac{f_{ij}^a}{f_{ij}}.$$  

Thus, the betweenness centrality of agent $a$ measures how often, on average, $a$ belongs to shortest paths between other pairs of agents.

The similar structures of the formulas for profitability (12) and Katz centrality (13) suggest that Katz centrality will be closely related to profitability, as will eigenvector centrality, being a special case of Katz centrality. The formulas are not identical, so we may expect to be able to find specific networks where any of the other centrality measures are more closely related to profitability, but we would expect Katz and eigenvector centrality to dominate for the “average” network. To formalize this intuition, we will use the concept of random graphs, originally studied in Erdös (1947), Erdös and Rényi (1959), and Gilbert (1959). Random graph theory provides a convenient method of statistically analyzing large networks, in which links between any pair of nodes are equally probable independent events.

Our focus is on sparse networks—in line with what is observed in practice (see the discussion in Ozsoylev et al. (2012)). We therefore assume that the expected number of links per agent in a network of size $N$ is $c \log(N)$ for some $c > 0$. With this assumption, the connectedness of agents grows with $N$, but the fraction of expected existing links (which is approximately $cN \log(N)/2$) to total potential number of links (which is approximately $N^2/2$) tends to zero for large $N$. Thus, large networks will indeed be sparse. To this end, we define

**Definition 8** The Erdös-Rényi random graph model of size $N$, and parameter $c > 0$, and $k > 0$, $G(N, c, k)$, 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 $c \frac{\log(N)}{N}$.

We use cross sectional correlation across agents as the measure of what it means for two variables in a given network to be closely related. Specifically, for a given network, we define the cross sectional (across agents) correlation between the different centrality measures and profitability,

$$\rho_{K^a} = \text{Corr}(K^a, \Pi), \quad \rho_C = \text{Corr}(C, \Pi), \quad \rho_D = \text{Corr}(D, \Pi), \quad \rho_{\hat{C}} = \text{Corr}(\hat{C}, \Pi).$$

Here, we use the convention that the correlation between any random variable and a constant is zero.
The following asymptotic result shows that Katz centrality dominates degree centrality, closeness centrality, and betweenness centrality for large random graphs.

**Theorem 4** For any given \( c > 0 \) and \( k > 1 \), there is a \( N_0 \), such that for all networks of size \( N > N_0 \) in the \( G(N, c, k) \) model, there is an \( \alpha \), such that the expected correlation between different centrality measures and profitability satisfies

\[
E[\rho_K(N)] > E[\rho_D(N)] > \max\left(E[\rho_B(N)], E[\rho_C(N)]\right).
\]

Thus, for large networks, profitability is best characterized by Katz centrality. This result provides a strong ranking between different centrality measures for a large class of networks. This contrasts with most applications of network theory to asset pricing, which typically provide results for very specific networks, e.g., Ozsoylev (2005), Colla and Mele (2010), and Buraschi and Porchia (2012), or do not provide a theoretical justification for the use of centrality, e.g., Ozsoylev et al. (2012), Das and Sisk (2005), and Adamic et al. (2010).

Of course, our strong ranking of centrality measures holds only asymptotically. 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 draw links between agents, 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 \( G(N, c) \) model, although the graph is still asymptotically sparse.

We measure the cross sectional correlation between profitability (\( \Pi \)) and the different centrality measures. The results are shown in Table 1. We see that the ranking is the same as in the theorem and, further, that eigenvector centrality, being a special case of Katz centrality, performs about as well as the Katz centrality measure.

### 5 Aggregate volatility and trading volume

A rich dynamic structure of volatility and trading volume has been documented in the literature. According to the Mixture of Distributions Hypothesis, this is because of lumpy information diffusion into the market. Along this line of argument, our model allows for general, decentralized, heterogeneous information diffusion.

The following stylized facts of volatility and trading volume, have been verified over a large range of time horizons (intra-daily, daily, weekly, and monthly):

1. Trading volume is positively autocorrelated over long time periods, i.e., the autocorrelation function of trading volume has long memory, i.e., decreases very slowly (see Bollerslev and Jubinski (1999) and Lobato and Velasco (2000)). Roughly speaking, this means that abnormally high trading volume one period predicts abnormally high trading volume for many future periods.
Table 1: Centrality Measures. The table shows the correlation between profitability, $\Pi$, and centrality for several different centrality measures, degree centrality ($V^1$), eigenvector centrality ($C$), Katz centrality ($K^\alpha$), closeness centrality ($\hat{C}$) and betweenness centrality ($B$). 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 performs poorly, since there is usually an isolated agent in the network, leading to all agents having closeness centrality of zero.

2. Similarly, return volatility of individual stocks and markets have long memory (see Bollerslev and Jubinski (1999) and Lobato and Velasco (2000)).

3. Volume and volatility are related (see 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 shows highly positive correlation. The two series also have positive lagged cross-correlations. Using a semi-parametric approach to study returns and trading volume on the NYSE, Gallant et al. (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.

If these rich volume and volatility dynamics represent lumpy information diffusion, they may be informative about how information diffuses in the market. Specifically the dynamics of volatility and trading volume in our model is determined by agents’ preferences ($\Gamma$) and by the network structure ($E$). A natural question is then which type of dynamics these two structural parameters can generate. In other words, which type of dynamics can be generated in a preference symmetric economy, a network symmetric economy, and in a completely general economy, respectively?

To gain some further intuition, we study the behavior of trading volume and price movements over time for some randomly chosen stocks. The objective of this example is exploratory and suggestive. A more extensive empirical study is outside the scope of this paper. In Figure 3, the dynamics of trading volume and stock price movements is shown for some randomly chosen large firms (black) with market capitalization above USD 10 billion, and for some small firms (red) with market capitalization between USD 100-500 million.

The left panel shows the correlogram of daily trading volume, defined as number of dollars traded in a day, normalized by the average number of dollars traded per day over the previous one-year
period. The normalization is introduced because of the non-stationarity of trading volume (see Lo and Wang (2000)). The thick red line represents the average correlogram of ten small stocks, and the thick black line represents the average correlogram of ten large stocks.\(^6\)

![Figure 3: Volume and volatility](image)

The figure shows autocorrelations of daily trading volume (left), weekly return volatility (middle), and cross-correlation between daily trading volume and squared returns (right). The solid black line represents an average of 10 large stocks, whereas the solid red line represents an average of 10 small stocks. Two examples of stocks within each group are also shown (PFE and MSFT, and OXM and RDN, respectively). The time period used is January 1, 2000 — December 31, 2011. Source: CRSP

We see that both the average series are decreasing from a peak of about 0.5 correlation between trading volume and one-day lagged trading volume, for both series. From day 5, the average autocorrelation of small cap stocks is significantly higher than for large cap stocks, decreasing down to about 0.14 after three weeks (15 trading days), versus 0.08 for large cap stocks. We also show the correlogram of two of the large cap stocks (PFE and MSFT), and two of the small cap stocks (OXM and RDN). We see that the two series for large cap companies are in general decreasing, whereas they are nonmonotonic for the small cap stocks.

\(^6\)The tickers for large cap stocks are MMM, DIS, MSFT, IBM, PG, MCD, PFE, JNJ, HD, and CSCO, and the tickers for the small cap stocks are AIR, NDN, OME, AZZ, OXM, ADC, QTM, RDN, ALK, and CPK. The stocks were chosen randomly, with the only constraint that data for the whole period be available.
A similar pattern emerges in the central panel, which shows the correlogram of weekly return volatility for the same stocks, between one and ten weeks. The average autocorrelation is in general higher for the small cap stocks, and there is significant nonmonotonicity for the individual stocks, whereas the trend is is more monotone for the large cap stocks.

Finally, the right panel shows the daily cross-correlogram between trading volume and squared returns for the same stocks. We see that the cross correlation is in general higher for the large cap stocks. Specifically, the cross-correlations are positive over the whole time period, showing a positive relation between trading volume and lagged price movements (negative lags), as well as between price movements and lagged trading volume (positive lags). In contrast, for the small stocks the correlation becomes negative after a few days, both with positive and negative lags.

Now, there is a close relationship between a function’s shape and the shape of its autocorrelation function. For example, the following standard result shows that the sign of higher-order derivatives are preserved by the autocorrelation function:

**Lemma 3** Consider an \( n \geq 2 \) times nonnegative continuously differentiable function, \( f \), with support on \([0,r]\), \( r < 1 \). Define the autocorrelation function \( R(\tau) = \int_{1-\tau}^{1} (f(x) - q)(f(x + \tau) - q)dx, 0 \leq \tau \leq 1 \), where \( q = \int_{0}^{1} f(x)dx \). Suppose that \( f^{(n)}(x) \) is nonnegative (nonpositive). Then \( R^{(n)}(\tau) \) is also nonnegative (nonpositive).

The result thus implies that a significantly nonmonotone autocorrelation function—whose higher-order derivatives must necessarily have switching signs—cannot arise from an underlying function that has a monotone behavior.

### 5.1 Volatility

How does the previous example fit with our network model? In a general network, we would expect that price volatility may 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 also shows that any volatility structure can be supported in a general economy.

**Theorem 5** For \( t = 1, \ldots, T \), the volatility of prices between \( t - 1 \) and \( t \), is

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

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

\[
\sigma_{p,T+1}^2 = \frac{1}{\tau_v + Y_T}.
\]

Further, 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

$$\left| \frac{\sigma^2_{P,t}}{\tau_v} - \frac{k_t}{\tau_v} \right| \leq \epsilon, \quad t = 1, \ldots, T + 1.$$  

From the first part of the theorem, we see that the volatility has a general decreasing trend over time through the denominator of (16), but that it can still have spikes in some time periods through big changes in the numerator. Note that the total price volatility up until time $t \leq T$ is

$$\sigma^2_{P,t} = \frac{1}{\tau_v} \frac{Y_t}{\tau_v + Y_t},$$  

and that the total price volatility between 0 and $T + 1$ is independent of the network structure—it is equal to the ex ante uncertainty of the asset’s value—but the way the volatility is divided period-by-period depends on the network through the dynamics of $y_t$ over time.

If we focus on network symmetric economies, the generality of the possible volatility structure does not hold. This is unsurprising, since in such networks we know from Lemmas 1 and 2 that there are severe restrictions on how information spreads over time. Specifically, Lemma 2 implies that the change in information precision that an agent has is unimodal, e.g., making it impossible for the information diffusion process to “slow down for a while” and then “speed up.” Since information diffusion drives price volatility in the model, we would expect similar implications for price volatility, i.e., it should be decreasing over time after a (potential) initial “hump.” Indeed, it is easy to show that in a network symmetric economy, $y_t$ is proportional to the square of the (common) information diffusion coefficient, $\Delta V_t$, regardless of the heterogeneity in preferences among agents, $\Gamma$.

A unimodality condition does indeed also hold for volatility in any network symmetric economy, when defined over a transformed volatility. We define the increasing function

$$\eta_t(\sigma^2_{P,t}) = \frac{1}{1 - \sigma^2_{P,t}/\sigma^2_v} - 1.$$  

The term $\sigma^2_{P,t}/\sigma^2_v$ denotes the fraction of total volatility ($\sigma^2_{v,t}$) realized at time $t$. Here, $\eta_t$ denotes a convex, increasing, transformation such that $\eta_0 = 0$, and $\eta_{T+1} = \infty$. Finally, we define $\Delta \eta_t = \eta_t - \eta_{t-1}$. It then follows immediately that $\Delta \eta_t \propto (\Delta V_t)^2$ and therefore, since the square of a unimodal nonnegative sequence is also unimodal, we arrive at

**Corollary 1** In a network symmetric economy, there is a $t^*$, $1 \leq t^* \leq T + 1$, such that $\Delta \eta_t$ is (weakly) increasing for $1 \leq t \leq t^*$, and (weakly) decreasing for $t^* \leq t \leq T + 1$.

It also follows that in the region where $\Delta V_t$ is decreasing, so is volatility:

**Corollary 2** For $t \geq t^*$, $\sigma_{p,t}$ decreasing.

Thus, the possible dynamics of price volatility is quite restricted in a network symmetric economy, regardless of the heterogeneity in preferences; it is typically “hump”-shaped, consistent with the behavior of large stocks in Figure 3, but not with the severely nonmonotone behavior shown by the small stocks in the same figure. Also, the faster decrease of volatility shown by large cap stocks in the same figure suggests that the information networks of large stocks are more connected.
Altogether, the result are in line with the view that information spreads fast and symmetrically among investors in large stocks, whereas it spreads more slowly and asymmetrically for small stocks. The result for smaller stocks is in line with what is found in Ozsoylev et al. (2012). The authors identify several information events in small stocks on the Istanbul Stock Exchange, for which it is plausible that there was heterogeneous information diffusion before information about the event reached public media.\footnote{For example, one such event was that a small pharmaceutical company announced that it would construct new facilities to satisfy increasing customer demands. The event was accompanied by a stock return of 15.8% beginning a few days before the company’s announcement.}

7 The paper finds that agents who were more central, as defined using an empirical proxy based on observed trading behavior, on average traded earlier than peripheral agents around these events.

5.2 Trading volume

Similar to volatility, rich structures of trading volume can arise within the network model. Since the model is inherently asymmetric, aggregate trading volume will be made up of the behavior of many different agents. This contrasts to the uniform behavior in representative agent models (e.g., Kyle (1985)), as well as to the ex ante symmetric behavior in network symmetric economies (e.g., Vives (1995) and He and Wang (1995)).

We focus on the aggregate period-by-period trading volume of investors in the network—the stochastic supply is trivially normally distributed. To this end, we define:

**Definition 9** The time-\(t\) aggregate trading volume is \(W_t = \frac{1}{N} \sum_a |\Delta x_{a,t}|\), and the expected trading volume is \(X_t = E[W_t]\).

The following theorem characterizes the expected trading volume over time, and mirrors our volatility results by showing that any expected trading volume can be supported.

**Theorem 6** The time-\(t\) expected trading volume is

\[
X_t = \frac{\tau}{N} \sum_{a=1}^{N} \frac{1}{\gamma_a} \sqrt{\frac{2}{\pi}} \left( \frac{V_{a,t-1}^2}{\tau_v + Y_{t-1}} - \frac{V_{a,t}^2}{\tau_v + Y_t} + \frac{\Delta V_{a,t}^2}{\tau_v + Y_t} + \frac{\Delta V_{a,t}}{\tau} \right).
\] (19)

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

\[|X_t - c_t| \leq \epsilon, \quad t = 1, \ldots, T+1.\]

It is clear from (19) that, as was the case for volatility, preferences alone cannot generate such generality of trading volume dynamics. In a network symmetric economy, all terms under the square root are identical across agents, and (19) collapses to

\[
X_t = \sqrt{\frac{2\tau^2}{\pi \gamma^2}} \left( \frac{V_{t-1}^2}{\tau_v + Y_{t-1}} - \frac{V_t^2}{\tau_v + Y_t} + \frac{\Delta V^2}{\tau_v + Y_t} + \frac{\Delta V_t}{\tau} \right).
\] (20)
Again, differences in preferences in this case are only important through the effect they have on the (harmonic) average risk aversion. Further, the restrictions on $\Delta V_t$ imposed by network symmetry carry over to trading volume. For example, after the bulk of an information shock has been incorporated into prices and $\Delta V_t$ is decreasing, so is $X_t$. This follows since (as is straightforward to show) the fourth term under the square root in (20) dominates in this region, and thus $X_t \approx \sqrt{\frac{2\tau^2}{\pi \gamma^2}} \Delta V_t$. We therefore have

**Corollary 3** In a network symmetric economy, for large $t$, $X_t$ is positively related to $\Delta V_t$ and decreases over time.

We can also rewrite (20) to get

$$X_t = \sqrt{\frac{2\tau^2}{\pi \gamma^2}} \left( (1 - c_t) V_{t-1}^2 \sigma_{p,t}^2 + \Delta V_t \right),$$

where $c_t = \frac{\gamma^2}{\tau^2} \times \frac{\tau u_{t+1}}{V_{t+1} \Delta V_{t+1}}$. For small $t$, we expect $V_{t-1}$ to be relatively small, and quickly increasing in $t$, as information spreads, implying that $0 < c_t << 1$. In this case, trading volume is positively related to volatility and the speed of information diffusion. We thus have

**Corollary 4** In a network symmetric economy, for small $t$, $X_t$ is positively related to $\Delta V_t$ and $\sigma_{p,t}$.

To summarize, the network symmetric case, again, leads to quite restricted dynamics. Right after an information shock, trading volume increases, driven by price volatility and increasing information diffusion. After the information diffusion has reached its peak, volatility decreases, as does trading volume, now tightly linked to the decreasing information diffusion. Again, such dynamics is consistent with the behavior of large stocks in the left panel of Figure 3, but less so with the small stocks. Moreover, it is consistent with a high average contemporaneous correlation between price movements and trading volume. Indeed, from the right panel of the Figure 3, we see that the average correlation for the large stocks is much larger than for the small stocks (0.36 versus 0.15), again suggesting that network symmetry provides a better fit for large stocks than for small.

### 5.3 The general economy

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

The above example captures the important distinction between trading volume driven by high aggregate information diffusion, and trading volume driven 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. With a disconnect between the points in time when information is incorporated into
prices—and price movements are large—and when agents with low risk-aversion trade—and volumes are high—it is not surprising if the cross-correlogram between the two variables quickly becomes negative, as in the right panel of Figure 3 for small stocks.

So, are there any unambiguous relations that must hold in the general case, given the rich dynamics that may arise when heterogeneous preferences and network structures interact? Indeed, there are, if we focus on conditional relations. So far, we have studied unconditional (over time) expectations under the assumption that an information shock arrives in the market, and then drives volatility and trading volume until the information has been incorporated into prices.

If the time of arrival of the information shock is well-identified, we could also condition on the time that has surpassed since arrival, and calculate conditional relations between volume and volatility. Earnings announcements may, for example, constitute well defined information events for which we could use such conditioning. Specifically, for such events, we can use $t$ in our information set when calculating relations, e.g., defining $\text{Cov}(W_t, W_{t+1}) = E[(W_t - E[W_t])(W_{t+1} - E[W_{t+1}])|t]$. Defining the price return $\mu_t = p_t - p_{t-1}$, we then have

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

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

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

### 6 Concluding remarks

We have introduced a general network model of a financial market with decentralized information diffusion, allowing us to study the effects of heterogeneous preferences and asymmetric diffusion of information among agents, and the interplay between the two. At the individual investor level, our results show that the trading behavior of individual investors is closely related to their position in the network: Closer agents have more positively correlated trades, even over short time periods, and standard network centrality measures are closely related to an agent’s profitability.

At the aggregate level, the dynamics of a market’s volatility and trading volume is related to—and could therefore be used to draw inferences about—the underlying information network. In a small random sample of large and small stocks, the dynamics for large stocks are consistent with fast
symmetric information diffusion among shareholders after information shocks, whereas the dynamics of smaller stocks suggest slower, asymmetric information diffusion. Future research may shed further light on the underlying information network in the market.

Another future line of research may be the study of endogenous network formation in financial markets. Given the strong characterization of individual agents’ welfare, it would be relatively straightforward to extend the model to include a period before signals are received and trading occurs, during which agents form links in anticipation of the future value these will generate. Investors’ actual trading behavior may then be used to empirically compare predicted and observed agent behavior in such markets.
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 \( \tau_a \), we have \( \tau_{a_1}, \ldots, \tau_{a_T} \). We first state three (standard) lemmas

Lemma 4 (Projection Theorem) Assume a multivariate signal \([\hat{\mu}_2; \hat{\mu}_y] \sim N([\mu_2; \mu_y], [\Sigma_{xx}, \Sigma_{xy}, \Sigma_{yx}, \Sigma_{yy}]) \). Then the conditional distribution is

\[
\hat{\mu}_2 | \hat{\mu}_y \sim N \left( \mu_2 + \Sigma_{xy} \Sigma_{yy}^{-1} (\mu_y - \mu_y), \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx} \right).
\]

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

\[
\tilde{v} | \mathbf{s} \sim N \left( \frac{\tau_v}{\tau_v + \tau} \tilde{v} + \frac{1}{\tau_v + \tau} \tau \mathbf{s}, \frac{1}{\tau_v + \tau} \right).
\]

Here, \( \tau = (\tau_1, \ldots, \tau_{K-1})^T \), \( \tau_i = \sigma_i^{-2} \), \( \tau = \sum_{i=1}^{M-1} \tau_i \), 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

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

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, so that agent 1

\[
\text{to } \bar{v}, \text{ and the market maker's pricing function in each time period, agents' demand functions are indeed linear.}
\]

We first assume that agents' demand takes a linear form at each point in time, so that agent 1 demands, \( A_a \) for all \( a \) and \( t \) (almost surely). The net demand at time \( t \) is then the difference between time \( t \) and \( t - 1 \) demands,

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

Now, the market maker observes total time \( t \) demands,

\[
w_t = \Delta x_t + u_t,
\]

and since the functions \( \eta_t \) and \( \eta_{t-1} \) are known, the market maker can back out

\[
R_t = (A_t - A_{t-1})\tilde{v} + u_t.
\]
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

\[
p_t = \frac{\tau_0}{\tau_0 + \hat{\tau}_u} \tilde{v} + \frac{\hat{\tau}_u}{\tau_0 + \hat{\tau}_u} \bar{v} + \frac{1}{\tau_0 + \hat{\tau}_u} \sum_{s=1}^{t} (A_s - A_{s-1}) \tau_{u,s} u_s, \tag{22}
\]

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

Equivalently,

\[
p_t = \lambda_t R_t + (1 - \lambda_t (A_t - A_{t-1})) p_{t-1}, \tag{23}
\]

where \( \lambda_t = \frac{\tau_u (A_t - A_{t-1})}{\tau_u + \hat{\tau}_u} \), 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 \( s = (R_t/(A_1 - A_0), R_2/(A_2 - A_1), \ldots, R_t/(A_t - A_{t-1}))' \), and it is clear that \( s_i \sim N(\bar{v}, \sigma_v^2 + \sigma_u^2/(A_i - A_{i-1})^2) \). It then follows immediately from Lemma 2 that

\[
\tilde{v} | s \sim N \left( \frac{\tau_0}{\tau_0 + \hat{\tau}_u} \bar{v} + \frac{1}{\tau_0 + \hat{\tau}_u} \sum_{i=1}^{t} (A_i - A_{i-1})^2 \tau_{u,i} R_i, \frac{1}{\tau_0 + \hat{\tau}_u} \right),
\]

i.e.,

\[
\tilde{v} = V_t + \sigma_V \xi^V_t, \tag{24}
\]

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

So, \( p_t = V_t = E[\tilde{v}|s] \) 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 (24) is the posterior distribution of \( \bar{V} \) given the information \( R_1, \ldots, R_t \), so \( v \sim N(V_t, \sigma_V^2) \). 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

\[
\tilde{v} | \{z_{a,t}, p_t\} \sim N \left( \frac{\tau_0}{\tau_0 + \hat{\tau}_u} V_t + \frac{\tau_u}{\tau_0 + \tau_u} z_{a,t}, \frac{1}{\tau_0 + \tau_u} \right).
\]

The time-\( T \) demand of an agent can now be calculated. Since individual investors condition on prices, they also observe \( \bar{R} \) and agent \( a \)'s information set is therefore \( \{z_a, \bar{R}\} \), which via Lemma 1 (with \( \Lambda = \text{diag}(0, \sigma_1^2, \sigma_2^2/\bar{A}^2) \)) leads to

\[
\tilde{v} | \{z_a, \bar{R}\} \sim N \left( \frac{\tau_0}{\tau_0 + \hat{\tau}_u} V_r + \frac{\tau_u}{\tau_0 + \tau_u} z_a + \frac{\tau_u}{\tau_0 + \tau_u} \bar{R}, \frac{1}{\tau_0 + \tau_u} \right).
\]

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:

\[
x_{a,T}(z_a, p) = \frac{E[\tilde{v}|I_a,T] - p_T}{\gamma_a \sigma^2[\tilde{v}|I_a,T]}, \quad a = 1, \ldots, \bar{N}, \tag{25}
\]
The demand of agent $a$ is therefore

$$x_{a,T}(z_{a,T},pt) = \frac{\mu_a - pt}{\gamma_a \sigma_a^2}$$

$$= \frac{1}{\gamma_a} (\tau_v \tilde{v} + \tau_a z_a - (\tau_v + \tau_a + \tilde{v}) p)$$

$$= \frac{1}{\gamma_a} (\tau_v \tilde{v} + \tau_a z_a + \tilde{v} R - (1 + \frac{\tau_v}{\tau_v + \tilde{v}}) (\tau_v \tilde{v} + \tilde{v} R))$$

$$= \frac{1}{\gamma_a} (\tau_a z_a - \tau_a p)$$

$$= \frac{\tau_a T}{\gamma_a} (z_{a,T} - pt).$$

It follows that $A_T = \frac{1}{N} \sum_{a=1}^N \frac{N_a}{\sigma_a^2 z_{a,T}}$.

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

$$U_{a,T} = -E \left[ e^{-\gamma_a x_{a,T}(\tilde{v} - pt) \mid z_{a,T} - pt} \right]$$

$$= -E \left[ e^{-\gamma_a x_{a,T}(z_{a,T} - pt) \mid z_{a,T} - pt} \right]$$

$$= -e^{-\gamma_a x_{a,T}(z_{a,T} - pt)} E \left[ e^{-\gamma_a x_{a,T}(z_{a,T} - pt) \mid z_{a,T} - pt} \right]$$

$$= -\frac{\tau_{v}^a}{\gamma_{v,T} + \gamma_{a,T}} (z_{a,T} - pt)^2$$

$$= -e^{\frac{\tau_{v}^a}{\gamma_{v,T} + \gamma_{a,T}} (z_{a,T} - pt)^2}.$$

This shows the result at $T$.

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

We define $Y_t = \tilde{v}_t = \sum_{i=1}^{p} y_i$, where $y_i = (A_i - A_{i-1})^2 \tau_{a,i}$, and recall that $Q_{a,t} = \tau_{a,t} = \frac{\tau_{v}^a}{\gamma_{v,T} + \gamma_{a,T}} = \sum_{i=1}^{t} q_{a,i}$, where $q_{a,i} = \frac{\gamma_{a,i} - \gamma_{a,i-1}}{\gamma_{a} - \gamma_{a}}$. With this notation we have

$$U_{a,T} = -e^{\frac{Q_{a,T}^2}{\gamma_{v,T} + \gamma_{a,T}} (z_{a,T} - pt)^2},$$

$$x_{a,T} = \frac{Q_{a,T}}{\gamma_{a}} (z_{a,T} - pt),$$

and $-E_0 \left[ e^{-\gamma_a x_{a,T}(\tilde{v} - pt)} \right] = -\sqrt{\frac{Q_{a,T}^2}{\tau_{v} + \gamma_{a,T}}} = \frac{1}{\sqrt{\gamma_{a,T}}} = -D_{a,T}$.

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, $W_{a,t} = W_{a,t-1} + \alpha_{a,t-1}(p_t - p_{t-1})$, an agent’s optimization at time $t - 1$ is then

$$U_{a,t} = \arg \max_{x_{a,t-1}} -E_{a,t-1} \left[ e^{-\gamma_a x_{a,t-1} - (p_t - p_{t-1}) D_{a,t} e^{\frac{Q_{a,t}^2}{\gamma_{v,T} + \gamma_{a,T}} (z_{a,t} - pt)^2} \mid z_{a,t-1}, p_{t-1}} \right]$$

$$= \arg \max_{x_{a,t-1}} \frac{D_{a,t} e^{\frac{Q_{a,t}^2}{\gamma_{v,T} + \gamma_{a,T}} (z_{a,t} - pt)^2} \mid z_{a,t-1}, p_{t-1}}}$$

$$= \arg \max_{b} \frac{D_{a,t} e^{\frac{Q_{a,t}^2}{\gamma_{v,T} + \gamma_{a,T}} (z_{a,t} - pt)^2} \mid z_{a,t-1}, p_{t-1}}}$$

$$= \frac{D_{a,t} e^{\frac{Q_{a,t}^2}{\gamma_{v,T} + \gamma_{a,T}} (z_{a,t} - pt)^2} \mid z_{a,t-1}, p_{t-1}}}.$$ (26)
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

\[
\begin{align*}
z_{a,t-1} &= \bar{v} + \xi_{t-1}, \quad \xi_{t-1} \sim N \left(0, \frac{1}{Q_{t-1}}\right), \\
z_{a,t} &= \bar{v} + \xi_t = \bar{v} + \frac{Q_{t-1}}{Q_t} \xi_{t-1} + \frac{q_t}{Q_t} \varepsilon_t, \quad \varepsilon_t \sim N \left(0, \frac{1}{q_t}\right),
\end{align*}
\]

(27)

(28)

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

\[
\begin{align*}
p_t &= \frac{\tau_v}{\tau_v + Y_t} \bar{v} + \frac{Y_t}{\tau_v + Y_t} \bar{v} + \frac{1}{\tau_v + Y_t} \sum_{s=1}^{t} (A_s - A_{s-1}) \tau_{a,s} u_s, \\
p_{t-1} &= \frac{\tau_v}{\tau_v + Y_{t-1}} \bar{v} + \frac{Y_{t-1}}{\tau_v + Y_{t-1}} \bar{v} + \frac{1}{\tau_v + Y_{t-1}} \sum_{s=1}^{t-1} (A_s - A_{s-1}) \tau_{a,s} u_s,
\end{align*}
\]

(29)

(30)

so

\[
\begin{align*}
p_t - p_{t-1} &= \left(\frac{\tau_v}{\tau_v + Y_t} - \frac{\tau_v}{\tau_v + Y_{t-1}}\right) \bar{v} + \left(\frac{Y_t}{\tau_v + Y_t} - \frac{Y_{t-1}}{\tau_v + Y_{t-1}}\right) \bar{v} + \frac{1}{\tau_v + Y_t} \left(A_t - A_{t-1}\right) \tau_{a,t} u_t \\
&\quad + \frac{1}{\tau_v + Y_t} \sum_{s=1}^{t-1} (A_s - A_{s-1}) \tau_{a,s} u_s,
\end{align*}
\]

(31)

and also

\[
\begin{align*}
z_{a,t} - p_t &= \frac{Q_{t-1}}{Q_t} \xi_{t-1} + \frac{q_t}{Q_t} \varepsilon_t - \frac{\tau_v}{\tau_v + Y_t} \bar{v} + \frac{\tau_v}{\tau_v + Y_t} \bar{v} \\
&\quad - \frac{1}{\tau_v + Y_t} \left(A_t - A_{t-1}\right) \tau_{a,t} u_t - \frac{1}{\tau_v + Y_t} \sum_{s=1}^{t-1} (A_s - A_{s-1}) \tau_{a,s} u_s.
\end{align*}
\]

(32)

This leads to the unconditional distribution:

\[
\left[ \begin{array}{c} p_t - p_{t-1} \\ s_t - p_t \\ s_{t-1} \\ p_{t-1} \end{array} \right] \sim N \left( \left[ \begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \end{array} \right], \left[ \begin{array}{cccc} \Sigma_{XX} & \Sigma_{XY} & 0 & 0 \\ \Sigma_{XY} & \Sigma_{YY} & 0 & 0 \\ 0 & 0 & \Sigma_{XX} & \Sigma_{XY} \\ 0 & 0 & \Sigma_{XY} & \Sigma_{YY} \end{array} \right] \right),
\]

(33)

Here,

\[
\begin{align*}
\Sigma_{XX} &= \left[ \begin{array}{cccc} A^2_t & 0 & 0 & 0 \\ 0 & B^2_{Y_t} & 0 & 0 \\ 0 & 0 & \frac{A^2_t A_{Y_{t-1}}}{\tau_v + Y_{t-1}} & \frac{A_t A_{Y_{t-1}}}{\tau_v + Y_{t-1}} \\ 0 & 0 & \frac{A_t A_{Y_{t-1}}}{\tau_v + Y_{t-1}} & \frac{B^2_{Y_{t-1}}}{\tau_v + Y_{t-1}} \end{array} \right],
\end{align*}
\]

\[
\begin{align*}
&= \left[ \begin{array}{cccc} \frac{A^2_t}{\tau_v + Y_t} & \frac{A_t}{\tau_v + Y_t} & 0 & 0 \\ 0 & B^2_{Y_t} & 0 & 0 \\ 0 & 0 & \frac{A^2_t}{\tau_v + Y_t} & \frac{A_t}{\tau_v + Y_t} \\ 0 & 0 & \frac{A_t}{\tau_v + Y_t} & \frac{B^2_{Y_t}}{\tau_v + Y_t} \end{array} \right].
\end{align*}
\]
\[\Sigma_{YY} = \begin{bmatrix} \frac{1}{\tau_e} + \frac{1}{Q_{t-1}} & \frac{1}{\tau_e} \\ \frac{1}{\tau_e} & \frac{1}{\tau_e} + Y_{t-1} \end{bmatrix} \]

\[\Sigma_{XY} = \begin{bmatrix} \frac{1}{Q_t} + \frac{1}{\tau_e} & -\frac{1}{\tau_e} + \frac{1}{Q_t} Y_{t-1} \\ \frac{1}{Q_t} (\tau_e + Y_{t-1}) & 0 \end{bmatrix} \]

We use the projection theorem to write \([p_t - p_{t-1}; z_{a,t} - p_t] \sim N(\mu, \hat{\Sigma})\), where \(\mu = \Sigma_{XY}^{-1} \Sigma_{Y}^{-1} [z_{a,t-1} - \bar{v}; p_{t-1} - \bar{v}]\), and \(\hat{\Sigma} = \Sigma_{XX} - \Sigma_{XY} \Sigma_{Y}^{-1} \Sigma_{XY}\). It follows that

\[\mu = \begin{bmatrix} \frac{Q_t}{\tau_e + Y_{t-1}} (\tau_e + Q_{t-1} + Y_{t-1}) \\ \frac{Q_t (\tau_e + Y_{t-1}) (\tau_e + Q_{t-1} + Y_{t-1})}{\tau_e + Y_{t-1}} \end{bmatrix} (z_{a,t-1} - p_{t-1}).\]

We rewrite (26) as

\[U_{a,t} = \arg \max_q -D_{a,t} e^{-\gamma_a W_{a,t-1}} E\left[e^{-ax^t - \frac{1}{2}x^t B x}\right],\]

where \(x = [p_t - p_{t-1}; z_t - p_t], B = \begin{bmatrix} 0 & 0 & \frac{Q_t}{\tau_e + Y_{t-1}} \\ 0 & 0 & \frac{Q_t (\tau_e + Y_{t-1}) (\tau_e + Q_{t-1} + Y_{t-1})}{\tau_e + Y_{t-1}} \end{bmatrix}\), \(a = [q; 0]\), and \(q = \gamma_a x_{a,t-1}\). From Lemma 3, it follows directly that this maximization problem is equivalent to

\[U_{a,t} = \arg \max_q -D_{a,t} e^{-\gamma_a W_{a,t-1}} e^{-\frac{1}{2}(\mu' \Sigma^{-1} a - (\Sigma^{-1} a) Z (\Sigma^{-1} a))},\]

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

\[\arg \max_q q (Z \Sigma^{-1} a) = \frac{1}{2} Z_{11} q^2,\]

leading to \(q^* = (Z \Sigma^{-1} a)\). It is easy to verify that \(\hat{\Sigma}^{-1} \mu = \frac{Q_t}{\tau_e + Q_{t-1} + Y_{t-1}} [1; 1] (z_{a,t-1} - p_{t-1})\), 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

\[\mu' \hat{\Sigma}^{-1} \mu - (\hat{\Sigma}^{-1} \mu - a) Z (\Sigma^{-1} a) = \frac{Q_t^2}{\tau_e + Q_{t-1} + Y_{t-1}} (z_{a,t-1} - p_{t-1})^2,\]

leading to the form of the utility stated in the theorem (6), with \(C_{a,t-1} = |I + \hat{\Sigma} B|^{1/2}\). 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 3:** The certainty equivalent satisfies

\[-e^{-\gamma_a c E} = E \left[ -e^{-W_T} \right] = -E \left[ \prod_{t=2}^T C_{t-1}^{-1} e^{-\frac{1}{2} \sigma_j^2 (z_{a,t-1} - p_t)^2} \right].\]

32
It is easy to see that
\[
\prod_{t=2}^{T} C_{t-1}^{-1} = \left(\frac{\tau_0 + Q_T + Y_T}{\tau_0 + Q_1 + Y_1}\right) \left(\frac{\tau_0 + Y_1}{\tau_0 + Y_T}\right) \prod_{t=2}^{T} \left(1 + \frac{Q_{t-1}(Y_t - Y_{t-1})}{(\tau_0 + Y_{t-1})(\tau_0 + Y_t)}\right).
\]

Moreover, since at \( t = 0, z_{a,1} - p_1 \sim N(0, \overline{Q_1} + \frac{1}{\tau_0 + Y_1}) \), it follows that
\[
E_0\left[ e^{-\frac{1}{2} \frac{Q_T^2}{\tau_0 + Y_T}(z_{a,1} - p_1)^2}\right] = \frac{\tau_0 + Q_1 + Y_1}{\tau_0 + Y_1},
\]
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 (31.32) 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
\[
\frac{1}{\gamma_a} Q_{a,t}
\]
and the total expected trading profit over time therefore is
\[
\frac{1}{\gamma_a} \sum_{t=1}^{T} \frac{Q_{a,t}}{\tau_0 + Y_t}, \tag{35}
\]
We are done.

**Proof of Theorem 2:** The proof follows immediately from (5), (27-28), and (32).

**Preliminary Proof of Theorem 4:**
We use the standard order notations, \( O(f(n)), o(f(n)), O_p(f(n)), \) and \( o_p(f(n)) \). For example, \( y_n = O(f(n)) \) denotes that the sequence \( y_n \) satisfies \( \lim_{n \to \infty} |y_n/f(n)| \leq C \) for some \( C > 0 \). We also write \( y_n = x_n + O(f(n)) \) if \( y_n - x_n = O(f(n)) \) and similarly for the other order notations. Finally, we write \( y_n \sim x_n \) if \( y_n = x_n + o_n(x_n) \).

Throughout the proof, we will adopt the standard convention in asymptotic analysis of using generic constants, \( C, C' \), etc., that may represent different values in different parts of the proof.

We begin with the inequality: \( E[\rho_D(N)] > E[\rho_C(N)] \). We rewrite
\[
\Pi = \beta_1 \left( D + \infty \sum_{t=2}^{\infty} \zeta_t \chi(E'1) \right), \quad \zeta_t = \frac{\beta_n}{\beta_1},
\]
where it follows from the definition of \( Y_t \) that \( \zeta_T << \zeta_{T-1} << \cdots << \zeta_2 << 1 \) for large \( N \) (since \( V_{a,t} \sim c \log(N)^k \)). We next note that, since \( A_t \sim (c \log(N)^k)^t \), whereas \( Y_t \sim (c \log(N)^k)^{2t} \), it follows that the \( D \) term dominates in the sum above for large \( N \), and thus \( E[\rho_D] \to 1 \) for large \( N \). More generally, if \( \zeta_p = \frac{\beta_n}{\beta_1} = \frac{\tau_0 + Y_{p-1}}{\tau_0 + Y_p} \sim \frac{(c \log(N)^k)^2}{(c \log(N)^k)^{2p}} = \frac{1}{(c \log(N)^k)^{2(p-1)}}, \) for any fixed \( 2 \leq p \leq T \).

Next, we note that a standard property of Erdos-Renyi random networks (see Bollobas (2001)) is that the degree, \( D_N \), for large \( N \) is close to \( c \log(N)^k \), in the sense that \( Z_N \equiv \frac{D_N - c \log(N)^k}{\sqrt{c \log(N)^k}} \to_d Z \sim N(0, 1) \). The result follows from the de Moivre-Laplace formula for binomial distributions, since \( pN = c \log(N)^k \to \infty \) as \( N \to \infty \). Further, it follows that the second order degree distribution \( V_N^{\tilde{Z}} \) satisfies \( \tilde{Z}_N \equiv \frac{\sqrt{2} - (c \log(N)^k)^2}{c \log(N)^k} \to_d \tilde{Z} \sim N(0, 1) \).
Calculating the cross sectional correlation between $D_N$ and $\Pi_N$, we therefore have

$$
\text{Corr}(D_N, \Pi_N) = \text{Corr}(D_N, \beta_1(D_N + \zeta_2 V_N^i)) + O_p \left( (\log(N)^k)^{-4} \right) 
= \text{Corr} \left( \sqrt{\log(N)^k} Z_N, \sqrt{\log(N)^k} Z_N + \frac{1}{(\log(N)^k)^{3/2}} \log(N)^k \hat{Z}_N \right) + O_p \left( (\log(N)^k)^{-4} \right) 
= \text{Corr} \left( Z_N, Z_N + \frac{1}{(\log(N)^k)^{3/2}} \hat{Z} \right) + O_p \left( (\log(N)^k)^{-4} \right).
$$

We have

$$
\text{Corr} \left( Z_N, Z_N + \frac{1}{(\log(N)^k)^{3/2}} \hat{Z} \right) = \frac{\text{Cov} \left( Z_N, Z_N + \frac{1}{(\log(N)^k)^{3/2}} \log(N)^k \hat{Z}_N \right)}{\sigma(Z_N) \sigma \left( Z_N + \frac{1}{(\log(N)^k)^{3/2}} \log(N)^k \hat{Z}_N \right)} = \frac{\sigma^2(Z_N) \left( 1 + \frac{1}{(\log(N)^k)^{3/2}} \text{Cov}(Z_N, \hat{Z}_N) \right)}{\sigma^2(Z_N) \left( 1 + \frac{2}{(\log(N)^k)^{3/2}} \text{Cov}(Z_N, \hat{Z}_N) \right) \sigma^2(Z_N) \left( \frac{1}{(\log(N)^k)^{3/2}} \text{Cov}(Z_N, \hat{Z}_N) \right)^{3/2}} \times \left( 1 - \frac{1}{(\log(N)^k)^{3/2}} \text{Cov}(Z_N, \hat{Z}_N) \right) + O_p \left( (\log(N)^k)^{-3} \right) = 1 - O_p \left( (\log(N)^k)^{-3} \right).
$$

Altogether, we therefore get $1 - E[\rho_D] = O \left( (\log(N)^k)^{-3} \right)$. We note that as long as $0 < C < \text{Corr}(Z_N, \hat{Z}_N) < C' < 1$, the argument can also be used to obtain an upper bound,

$$
E[\rho_D] \leq 1 - C'' \left( \log(N)^k \right)^{-3}
$$

(36) for some positive constant $C'' > 0$.

Now, to prove the result, it is sufficient to show that $E[\text{Corr}(D, \hat{C})] \leq 1 - C((\log(N)^k)^{-4})$, for constants $q < 3$, and $C > 0$, since this together with $E[\rho_D] = 1 - O((\log(N)^k)^{-3})$, implies that $E[\rho_D(N)] \leq 1 - C''((\log(N)^k)^{-4})$, $C' > 0$, and therefore $E[\rho_D(N)] > E[\rho_C(N)]$ for large $N$. This follows immediately from the following triangle inequality-like lemma:

**Lemma 8** Assume $\rho(a, b) = 1 - a$, and $\rho(a, c) = 1 - \beta$, with $\beta > \alpha$. Then $\rho(b, c) \leq 1 - (\beta - \alpha)^2$.

**Proof of Lemma 8:** Because of the simple renormalizations, $a \mapsto (a - E[a])/\sigma(a)$, $b \mapsto (b - E[b])/\sigma(b)$, and $c \mapsto (c - E[c])/\sigma(c)$, we can without loss of generality assume that $a, b$, and $c$ have zero expectations and unit variances. All correlations can then be expressed as, $\rho(a, b) = E[ab]$, etc. We introduce the metric $d(a, b) = \sqrt{E[(a - b)^2]}$, and we then have the triangle inequality $d(a, c) \leq d(a, b) + d(b, c)$, leading to $d(b, c) \geq d(a, c) - d(b, c)$. 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 $2 - 2\rho(a, c) \geq (\sqrt{\beta - \alpha})^2 = 2(\sqrt{\beta} - \sqrt{\alpha})^2$, leading to the result. We are done.

We next introduce the following lemma:

**Lemma 9** Given two random variables, $X$, and $Y$, with finite variances. Define the random variable $f = f(x) = E[Y | X = x]$, and assume that $f$ is an increasing function on the range of $X$. Then $\text{Corr}(X, Y) \leq \text{Corr}(X, f)$.

**Proof of Lemma 9:** We have $\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X) \sigma(Y)}$, and $\text{Corr}(X, f) = \frac{\text{Cov}(X, f)}{\sigma(X) \sigma(f)}$. From the law of total variance (Eve’s law), $\sigma_f^2 = E(\sigma^2(Y | X)) + \sigma^2(f) \geq \sigma_f^2$. Moreover, $\text{Cov}(X, Y) = E[X Y] - E[X] E[Y] = E[E[Y | X]] - E[X] E[Y | X] = E[X f] - E[X] E[f] = \text{Cov}(X, f)$. Finally, since $f$ is an increasing function, we follow from a standard monotonicity argument that $\text{Cov}(X, f) \geq 0$ (since $\text{Cov}(X, f) = E[(X - E[X])(f - E[f])] = E[(X - E[X])(f - f(E[X]))]$, and the term within this expectation is nonnegative). Altogether we then have $\text{Corr}(X, Y) = \frac{\text{Cov}(X, f)}{\sigma(X) \sigma(Y)} \leq \frac{\text{Cov}(X, f)}{\sigma(X) \sigma(f)} = \text{Corr}(X, f)$. We are done.

Finally, the following lemma describes the asymptotic behavior of expected closeness centrality:

\[ \text{34} \]
Lemma 10 \[ E[C_{N}(Z)-b'_{N}] = Z - \frac{1}{2\sqrt{c\log(N)}} Z^2 + O\left(\frac{\log(N)^{1/2-k}}{\log(N)}\right), \]
where \( a'_{N} = \frac{\log(c\log(N)^k)}{\log(N)^{2/\sqrt{c\log(N)^k}}} \), and \( b'_{N} = \frac{\log(c\log(N)^k)}{\log(N)} \left( 1 + \frac{\log(c\log(N)^k)}{\log(N)} \right). \]

Proof of Lemma 10:
We first note that for a given (integer) degree, \( D = pN + Z\sqrt{pN} = (c\log(N)^k) + Z\sqrt{c\log(N)^k} \), the asymptotic normality of the degree distribution implies that for large \( N \) there will typically be approximately \( \frac{N}{\sqrt{c\log(N)^k}} \) vertices with \( D \) neighbors for large \( N \) (where \( \varphi(x) = (2\pi)^{-1/2}e^{-x^2/2} \) is the standardized normal distribution function). In a large random graph, the probability that no vertex has \( D \) neighbors therefore very quickly becomes negligible.

Now, in a random graph, \( G \in \mathcal{G}(N,c,k) \), consider a vertex, \( a \), with \( D = (c\log(N)^k) + Z\sqrt{c\log(N)^k} \) neighbors. Define
\[ d = \frac{\log(N) - \log(D)}{\log(c\log(N)^k)} = d_0 + z_0, \]
where \( d_0 = \lfloor d \rfloor \) is the integer part of \( d \), and \( 0 \leq z_0 < 1 \) is the fractional part. For simplicity, assume that \( z_0 > 0 \) (the case with \( z_0 = 0 \) can be handled in a similar manner). Note that \( (c\log(N))^d = N/D \).

Following the same procedure as Bollobas (2001), Lemma 10.7, page 257, it follows that for arbitrary \( K > 12 \), with probability at least \( 1 - N^{-K} \),
\[ \left| \frac{V_{a,t} - D(c\log(N)^k)^t}{D(c\log(N)^k)^t} \right| \leq C_K \log(N)^{1/2-k}, \]
for all \( 1 \leq t \leq d_0 \), and that
\[ \sum_{t=1}^{d_0} V_{a,t} - N \leq C_K \log(N)^{1/2-k}. \]

Here, the constant \( C_K \) is allowed to depend on \( K \), but not on \( N \). For the time being, we focus on graphs that satisfy these conditions for some constants \( K, C_K \). We therefore define the event \( \Omega_L(N) \) to contain the set of graphs that satisfy (37,38).

The mean distance from \( a \) to all other vertices is
\[ \frac{1}{C} = \frac{1}{N-1} \sum_{t=1}^{\infty} tV_{a,t}. \]

The number of vertices at distance larger than \( d_0 + 1 \) from \( a \) is vanishingly small in relative terms, but since the sum in (39) is weighted by \( i \), a few vertices very far away from \( a \) may potentially have a disproportionate effect, so we need to ensure that no such vertices exist. Since \( k > 1 \), it follows that the graph is asymptotically connected, so the distances between all nodes is bounded (see Bollobas (2001), Theorem 7.3, page 164). Indeed, the graph’s diameter is close to \( d_0 \). For example, for \( k \geq 4 \), Theorem 10.9 in Bollobas (2001) implies that almost all graphs have a diameter not larger than \( r = d_0 + 2 \). A similar argument implies that for a weaker constraint, e.g., \( r \leq d_0(1 + o(1)) \), the probability that the distance between any \( a \) and any other vertex is greater than \( r \) decreases extremely quickly. Thus, we need not worry about such disproportionate effects of a few distant vertices.

Moreover, the geometric structure of \( V_{a,t} \) as functions of \( t \) implies that the number of vertices at distances less than \( d_0 \) are also very few compared with the number of vertices at distances \( d_0 \) and \( d_0 + 1 \). Altogether, this implies that we can focus on two dominant terms in (39) and write
\[ \frac{1}{C} = \frac{1}{N-1} \left( d_0 V_{a,d_0} + (d_0 + 1)V_{a,d_0+1} \right) + O(d\log(N)^{1/2-k}) \]
\[ = \frac{1}{N} \left( (d-z)V_{a,d_0} + (d+1-z)V_{a,d_0+1} \right) + O(d\log(N)^{1/2-k}) \]
\[ = d \frac{V_{a,d_0} + V_{a,d_0+1}}{N} + \frac{V_{a,d_0+1} - z_0(V_{a,d_0} + V_{a,d_0+1})}{N} + O(d\log(N)^{1/2-k}) \]
\[ = d(1 + O(\log(N)^{1/2-k})). \]

So, \( \hat{C} = \frac{1}{d} (1 + O(\log(N)^{1/2-k})). \)

Now,
\[ \frac{1}{d} = \frac{\log(c\log(N)^k)}{\log(N) - \log(D)} = \frac{\log(c\log(N)^k)}{\log(N)} \left( 1 + \frac{\log(D)}{\log(N)} + O(\log(N)^{-2}) \right), \]

35
and since \( D = \sqrt{c \log(N)} Z + c \log(N) \), we get

\[
\frac{1}{d} = \log(c \log(N)^k) \left(1 + \frac{\log(\sqrt{c \log(N)^k} Z + c \log(N)^k)}{\log(N)} + O(\log(N)^{-2})\right)
\]

\[
= \frac{\log(c \log(N)^k)}{\log(N)} \left(1 + \frac{\log(c \log(N)^k) + \log\left(1 + \frac{Z}{\sqrt{c \log(N)^k}}\right)}{\log(N)} + O(\log(N)^{-2})\right)
\]

\[
= \frac{\log(c \log(N)^k)}{\log(N)} \left(1 + \frac{1}{\log(N)} \left(\frac{Z}{\sqrt{c \log(N)^k}} \left(1 - \frac{1}{2} \frac{Z}{\sqrt{c \log(N)^k}}\right)\right) + O(\log(N)^{-2})\right)
\]

\[
= b_N'(1 + O(\log(N)^{-2})) + a_N' \left(Z - \frac{1}{2c \log(N)^k} Z^2\right).
\]

Altogether, we therefore get

\[
\hat{C} = \frac{1}{d} (1 + O(\log(N)^{1/2-k}))
\]

\[
= \left( b_N'(1 + O(\log(N)^{-2})) + a_N' \left(Z - \frac{1}{2c \log(N)^k} Z^2\right)\right) (1 + O(\log(N)^{1/2-k}))
\]

\[
= \left( b_N' + a_N' \left(Z - \frac{1}{2c \log(N)^k} Z^2\right)\right) (1 + O(\log(N)^{1/2-k})). \tag{40}
\]

Equation (40) holds for \( \Omega_K(N) \). To calculate the unconditional expectation we use the law of total expectation, together with the bound \( 0 \leq \hat{C} \leq 1 \), to get

\[
E[\hat{C}|Z] = E[\hat{C}|Z, \Omega_K]P(\Omega_K) + E[\hat{C}|Z, \Omega_K^C]P(\Omega_K^C)
\]

\[
= \left( b_N' + a_N' \left(Z - \frac{1}{2c \log(N)^k} Z^2\right)\right) (1 + O(\log(N)^{1/2-k}))(1 - N^{-K}) + O(N^{-K})
\]

\[
= \left( b_N' + a_N' \left(Z - \frac{1}{2c \log(N)^k} Z^2\right)\right) (1 + O(\log(N)^{1/2-k})), \tag{41}
\]

where the last inequality follows since \( K > 12 \) can be chosen arbitrarily. By subtracting \( b_N' \) and dividing by \( a_N' \), we arrive at

\[
\frac{E[\hat{C}|Z] - b_N'}{a_N'} = Z - \frac{1}{2c \log(N)^k} Z^2 + O \left(\log(N)^{1/2-k}\right). \tag{42}
\]

We have shown Lemma 10.

We note that another way of writing (42) is as

\[
\frac{E[\hat{C}|Z] - b_N'}{a_N'} = Z - \frac{1}{2c \log(N)^k} Z^2 + \log(N)^{1/2-k} W_N,
\]

where \( W_N \) is a random variable with tails that are bounded independently of \( N \). Lemma 10, together with an identical argument as that leading to (36) then implies that

\[
E \left[ \text{Ccorr}[Z_N, E[\hat{C}_N|Z_N]] \right] = E \left[ \text{Ccorr}\left[Z_N, Z_N - \frac{1}{2c \log(N)^k} Z^2 + \log(N)^{1/2-k} W_N\right]\right]
\]

\[
\leq 1 - C(\log(N)^k)^{-1},
\]

36
for a constant $C > 0$. Lemma 9 can then be applied to show that

$$E \left[ Corr[D, C] \right] \leq 1 - C(c \log(N)^k)^{-1},$$  \hspace{1cm} (43)$$

and finally we use Lemma 8: Equation (43) states that $E[Corr(D, C)] \leq 1 - C((c \log(N))^{-q})$, for constants $q < 3$ ($q = 1$) and $C > 0$. Together with $E[\rho_B] = 1 - O((c \log(N)^k)^{-3})$, this implies that $E[\rho_C] \leq 1 - C''((c \log(N)^k)^{-q})$, $C'' > 0$, and therefore $E[\rho_D(N)] > E[\rho_C(N)]$ for large $N$.

For the inequality $E[\rho_D(N)] > E[\rho_B(N)]$, we use the following lemma:

**Lemma 11** $\frac{E[B_N | Z] - b_N}{a_N} = Z + \frac{1}{2 \sqrt{\log(N^k)}} Z^2 + O((c \log(N)^k)^{-1})$, where $a_N = \frac{2 \log(N)}{(c \log(N)^k)^{1/2} N \log(c \log(N)^k)}$, and $b_N = \frac{\log(N)}{N \log(c \log(N)^k)}$.

**Proof of Lemma 11**: Similarly to Lemma 10, define $d = \frac{\log(N)}{\log(c \log(N)^k)} = d_0 + z_0$. From Bollobas (2001), Lemma 10.9, page 259, it follows that with probability at least $1 - N^{-K}$, for any vertex, the fraction of vertices that are at distance $d_v \in V \overset{d_0}{=} \{d_0 - 1, d_0, d_0 + 1\}$ from that vertex, is at least $1 - C_K(c \log(N)^k)^{-1}$. Thus, with very high probability, the vast majority of pairwise vertices will be at distance very close to $d$. We denote this event by $\Omega_k$.

Focusing on a vertex, $a$, with $D = (c \log(N)^k) + Z \sqrt{(c \log(N)^k)^2}$ neighbors, we study two vertices $i, j$, at distance $d_v \in V$ from each other, and choose an arbitrary shortest path between the two. The likelihood of that such a shortest path will go through one of $a$’s neighbors, is then

$$p = \frac{D}{N} d_v = \frac{D \log(N)}{N \log(c \log(N)^k)} (1 + O(N^{-1})).$$

Given that the path goes through one of $a$’s neighbors, $a'$, and that $a'$ has $(c \log(N)^k) + \sqrt{(c \log(N)^k)^2} Z'$ neighbors, it can either go through $a$ and one of $a$’s $D$ neighbors, or through one of $a$’s $D' - 1$ other neighbors, and one of the $(c \log(N)^k)^2 (1 + (c \log(N)^k)^{-1/2} Z' + O((c \log(N)^k)^{-1}))$ neighbors to that neighbor. The conditional probability for the former scenario is then

$$p' = \frac{D}{D} = \frac{D}{(c \log(N)^k)^2 (1 + (c \log(N)^k)^{-1/2} Z' + O((c \log(N)^k)^{-1}))} = \frac{D}{(c \log(N)^k)^2} \left(1 - \frac{D}{(c \log(N)^k)^2} - (c \log(N)^k)^{-1/2} Z' + O((c \log(N)^k)^{-1})\right) = \frac{D}{(c \log(N)^k)^2} \left(1 - (c \log(N)^k)^{-1/2} Z' + O((c \log(N)^k)^{-1})\right).$$

The total probability for the path to go through $D$ is then

$$pp' = \frac{D}{(c \log(N)^k)^2 N \log(c \log(N)^k)} (1 + O(N^{-1})) \left(1 - (c \log(N)^k)^{-1/2} Z' + O((c \log(N)^k)^{-1})\right) = \frac{D^2 \log(N)}{(c \log(N)^k)^2 N \log(c \log(N)^k)} \left(1 - (c \log(N)^k)^{-1/2} Z' + O((c \log(N)^k)^{-1})\right).$$

To calculate $B_N$, we need to sum over all pairwise vertices. We note that the fraction of all shortest paths that go through one of $a$’s neighbors, that go through the specific vertex $a'$ is

$$1 + Z' (c \log(N)^k)^{-1/2} \frac{D + O(1)}{D}.$$
The sum then takes the form
\[
B_N = \sum_{a' \in S_{a,1}} \frac{D^2 \log(N)}{(c \log(N))^k 2^N \log(c \log(N))} \left( 1 - (c \log(N)^k)^{-1/2} Z_{a'} + O((c \log(N)^k)^{-1}) \right) \frac{1 + Z_{a'}(c \log(N)^k)^{-1/2}}{D + O_p(1)}
\]
\[
= \frac{1}{D} \sum_{a' \in S_{a,1}} \frac{D^2 \log(N)}{(c \log(N))^k 2^N \log(c \log(N))} \left( 1 + O_p((c \log(N)^k)^{-1}) \right)
\]
\[
= \frac{D^2 \log(N)}{(c \log(N))^k 2^N \log(c \log(N))} \left( 1 + O_p((c \log(N)^k)^{-1}) \right).
\]

The second equality uses the fact that \( \sum_{a' \in S_{a,1}} Z_{a'} = O_p((c \log(N)^k)^{1/2}) \), since the \( Z_{a'} \)'s are very close to independent across \( a^8 \).

We therefore have
\[
E[B_N | Z] = \frac{D^2 \log(N)}{(c \log(N))^k 2^N \log(c \log(N))} \left( 1 + O((c \log(N)^k)^{-1}) \right)
\]
\[
= \left( \frac{\log(N)}{N \log(c \log(N))} + \frac{2 \log(N)}{(c \log(N)^k)^{1/2} N \log(c \log(N))} \left( Z + \frac{1}{2 \sqrt{c \log(N)^k}} Z^2 \right) \right) \left( 1 + O((c \log(N)^k)^{-1}) \right)
\]
\[
= \left( b'_N + a'_N \left( Z + \frac{1}{2 \sqrt{c \log(N)^k}} Z^2 \right) \right) \left( 1 + O((c \log(N)^k)^{-1}) \right).
\]

Subtracting \( b'_N \), and dividing by \( a'_N \), we arrive at
\[
\frac{E[B_N | Z] - b'_N}{a'_N} = Z + \frac{1}{2 \sqrt{c \log(N)^k}} Z^2 + O \left( (c \log(N)^k)^{-1} \right). \tag{44}
\]

So far we have conditioned on \( \Omega_K' \), but an identical argument as that in (41) implies that (44) also holds unconditionally. We have shown Lemma 11.

Now, an identical argument as for \( \hat{C} \) implies that \( \text{Corr}(\Pi, B_N) \leq 1 - C(c \log(N)^k)^{-1} \) for some \( C > 0 \), and the inequality \( E[\rho_0(N)] > E[\rho_d(N)] \) therefore follows.

For the final inequality, \( E[\rho_k(N)] > E[\rho_d(N)] \), we proceed as follows: We note that \( \chi(E_1)1 \equiv E_11 \). Also, as noted, for large \( N \), each vertex has approximately \( c \log(N)^k \) neighbors, and each of these neighbors has approximately \( c \log(N)^k \) neighbors. As a consequence, the expected number of vertices to which a vertex has two paths is approximately \( (c \log(N)^k)^3 / N \), and the expected number of vertices to which there are three or more paths is approximately \( (c \log(N)^k)^4 / N^2 \). Therefore, there are expected to be in total about \( (c \log(N)^k)^3 / N \) vertices with two paths to a vertex within distance two, and no (in expectation \( (c \log(N)^k)^4 / N \to 0 \) vertices with more than two paths. Especially, the probability that there are more than \( 2(c \log(N)^k)^3 / N \) vertices with two paths quickly becomes extremely small. Indeed, we can focus on the event \( \Omega'' \), which contains the graphs that satisfy these bounds for the number of double and triple-or-higher paths, and use the fact that \( P(\Omega'') = 1 - O(N^{-K}) \) for any \( K > 0 \). An identical argument as in (41) then implies that if the asymptotic result holds for \( \Omega'' \), it holds unconditionally.

For \( \Omega'' \), \( \chi(E')1 = E'1 - r \), where \( r \) is a vector with at most \( 2(c \log(N)^k)^3 \) ones and the remaining elements zero. Choosing \( \alpha = \frac{1}{(c \log(N)^k)^3} \), and using a similar argument as when determining the asymptotic behavior of \( D \), we therefore

\footnote{Indel, with probability at least \( 1 - O((c \log(N)^k)^4 / N) \), there are no shared neighbors of any two vertices in \( S_{a,a'} \), except for \( a \), in which case all of \( Z_{a'} \) are jointly independent.}

38
arrive at

\[
\text{Corr}(\Pi, K^{n}) = \text{Corr}(\Pi, D + \alpha E^{2}1) + O_{p}\left( (c \log(N)^{k})^{-4} \right)
\]

\[
= \text{Corr}(\Pi, \Pi - \alpha r) + O_{p}\left( (c \log(N)^{k})^{-4} \right)
\]

\[
= \text{Corr}\left( \sqrt{c \log(N)^{k}}Z_{N}, \sqrt{c \log(N)^{k}}Z_{N} - \frac{1}{(c \log(N)^{k})^{1/2}}r \right) + O_{p}\left( (c \log(N)^{k})^{-5/2} \right)
\]

\[
= \text{Corr}\left( Z_{N}, Z_{N} - \frac{1}{(c \log(N)^{k})^{1/2}}r \right) + O_{p}\left( (c \log(N)^{k})^{-4} \right)
\]

\[
= 1 - O_{p}(c \log(N)^{k})^{-4}).
\] (45)

Taking expectations (again using the boundedness of correlation for the rare events not covered by (45)), we arrive at

\[E[\rho_{N}] = E[\text{Corr}(\Pi, K^{n})] = 1 - O((c \log(N)^{k})^{-4}).\]

So, since \(E[\rho_{D}] \leq 1 - C(c \log(N)^{k})^{-3}\), it is indeed the case that \(E[\rho_{K^{n}}(N)] > E[\rho_{D}(N)]\) for large \(N\). We are done. \(\blacksquare\)

**Proof of Theorem 5:** For the first part, we note that since \((p_{1} - p_{T-1})\) is independent of \(p_{T-1}\) (given publicly available information), it follows that the price volatility between \(t - 1\) and \(t\) is equal to \((\Sigma_{XX})_{11}, \sigma_{p_{t},t} = (\Sigma_{XX})_{11}\mid_{t-1} = (\tau_{t+1} \mid_{t-1} \tau_{t+1})\), with the convention \(Y_{0} = 0\). Also, the final period volatility of \(\tilde{\nu} - \nu_{T}\) is \(\sigma_{p_{T},T+1} = c + \nu_{T} \tau_{T+1} \nu_{T} \tau_{T+1} \nu_{T} \tau_{T+1}\).

For the second part, we first note that from the definition of \(y_{t}\), it follows that \(y_{t} = \tau_{c} \frac{k_{t}}{1-k_{t}}\) and—defining \(K_{t} = \sum_{i=1}^{t} k_{i}\)—a simple induction argument further shows \(y_{t} = \tau_{c} \frac{k_{t}}{1-K_{t-1}}\), \(t = 1, \ldots, T-1\), and \(y_{T} = \tau_{c} \frac{1-K_{T}}{1-K_{T-1}}\). We note that all the \(y_{1}, \ldots, y_{T}\) are all well defined.

Next, we back out the connectedness that is needed to be consistent with the \(y_{t}\)'s. We have \(A_{1} = \frac{\tau_{c}}{\tau_{c}}, A_{t} = \frac{\tau_{c}}{\tau_{c}} + \frac{\tau_{c}}{\tau_{c}}, \text{ leading to } V_{1} = \frac{\tau_{c} V_{t-1}}{N}, \text{ and } \Delta V_{1} = \frac{\Delta V_{t-1}}{N}. \text{ Thus, if we can replicate arbitrarily closely arbitrary sequences of diffusions, through which the average number of signals, } V_{t}, \text{ increases over time, then we can generate any } y_{t}, \text{ and thereby any volatility structures. We note that } \gamma \text{ is a free parameter that allows us to scale the network to arbitrary sizes. The result now follows from the following lemma:}

**Lemma 12** For any \(T\), there are networks of size \(N\), such that \(\tilde{V}_{T'} = (1 + o(1))\tilde{V}_{T}\) for \(T' > T\), and \(\tilde{V}_{T'} = \frac{1}{N(t + o(1))} \tilde{V}_{T}\) 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. 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.

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

We are done. \(\blacksquare\)

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

**Lemma 13** 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 \(\Phi\) is the cumulative normal distribution of a standard normal variable.
We note that from (5) and given that \( \tilde{v} = \bar{v} + \eta \), it follows that agent \( a \)'s net time-\( t \) demand is

\[
\gamma_a \Delta x_{a,t} = Q_{a,t}(z_{a,t} - p_t) - Q_{a,t}(z_{a,t-1} - p_{t-1})
\]

\[
= Q_{a,t} \left( \tilde{v} + \eta + \frac{Q_{t-1}}{Q_{a,t}} \xi_{a,t-1} + \frac{q_t}{Q_{a,t}} e_{a,t} - \left( \frac{\tau_v}{\tau_v + Y_t} \tilde{v} + \frac{Y_{t-1}}{\tau_v + Y_{t-1}} (\tilde{v} + \eta) + \frac{1}{\tau_v + Y_t} \sum_{i=1}^{t} (A_i - A_{i-1}) \tau_{u_i} u_{t} \right) \right)
\]

\[
= Q_{a,t-1} \left( \tilde{v} + \eta + \xi_{a,t-1} + - \left( \frac{\tau_v}{\tau_v + Y_{t-1}} \tilde{v} + \frac{Y_{t-1}}{\tau_v + Y_{t-1}} (\tilde{v} + \eta) + \frac{1}{\tau_v + Y_{t-1}} \sum_{i=1}^{t-1} (A_i - A_{i-1}) \tau_{u_i} u_{t-1} \right) \right)
\]

\[
= \mathbb{E}(a,t) \left( \frac{Q_{a,t}}{\tau_v + Y_t} - \frac{Q_{a,t-1}}{\tau_v + Y_{t-1}} \right) \left( \tau_v \eta - \sum_{i=1}^{t-1} (A_i - A_{i-1}) \tau_{u_i} u_{t} \right) \sim N(0, q_{a,t}^2)
\]

where

\[
r_{t}^2 = \left( \frac{Q_{a,t}}{\tau_v + Y_t} - \frac{Q_{a,t-1}}{\tau_v + Y_{t-1}} \right) \left( \tau_v + Y_t \right)
\]

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

\[
\gamma_a \Delta x_{a,t} \sim \frac{q_{a,t} \sqrt{r}}{\sqrt{\Delta V_{a,t}}} \xi_t
\]

where \( \xi_t \sim N(0, 1) \) is independent of \( \xi_t \sim N(0, 1) \), and also independent across different agents, and

\[
r_{t}^2 = \frac{V_{a,t}^2}{\tau_v + Y_t - 1} = \frac{V_{a,t}^2}{\tau_v + Y_t} + \Delta V_{a,t}^2
\]

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

\[
\gamma_a \Delta x_{a,t} \sim \frac{r_{a,t}}{\tau} \xi_t
\]

when \( \Delta V_{a,t} = 0 \). Assuming \( \Delta V_{a,t} > 0 \), we note that \( \frac{\Delta x_{a,t}}{\Delta V_{a,t}} \xi_t \sim N \left( \frac{\tau \Delta V_{a,t}}{\tau \Delta V_{a,t}} \xi_t, 1 \right) \). The law of large numbers, together with Lemma 13 then in turn implies that, conditioned on \( \xi_t \),

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

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

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

Summing the two terms together, we get

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

40
We note that this formula also holds when \( \Delta V_{a,t} = 0 \), since \( E[r_{a,t}\tau\xi_i] = \tau \sqrt{\frac{1}{\pi} \frac{\Delta V_{a,t}}{\tau}} \), which finally leads to (19)

\[
X_t = \frac{\tau}{N} \sum_{a=1}^{N} \frac{1}{\gamma_a} \left[ \frac{2}{\pi} \left( \frac{V_{a,t-1}^2}{\tau_0 + \tau} \right) - \frac{2V_{a,t-1}\Delta V_{a,t}}{\tau_0 + \tau} + \frac{\Delta V_{a,t}}{\tau} \right]
\]

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

\[
X_t = \frac{\tau}{N} \sum_{a=1}^{N} \frac{1}{\gamma_a} \left[ \frac{2}{\pi} \left( \frac{V_{a,t-1}^2}{\tau_0 + \tau} \right) - \frac{2V_{a,t-1}\Delta V_{a,t}}{\tau_0 + \tau} + \frac{\Delta V_{a,t}}{\tau} \right] \]

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

We are done.

**Proof of Theorem 7:** We first show the second result: We use (21,23) to define

\[
\hat{R}_t = (A_t - A_{t-1})\tau_u \hat{R}_t - y_t \bar{v} = y_t \eta + (A_t - A_{t-1})\tau_u u_t,
\]

and, of course, \( \hat{R}_1, \ldots, \hat{R}_t \) can be backed out of \( p_1, \ldots, p_t \). Now, it is straightforward to use (23) to show that

\[
p_t = \bar{v} + \frac{1}{\tau_0 + Y_t} \left( Y_t \eta + \sum_{i=1}^{t} (A_i - A_{i-1})\tau_u u_i \right)
\]

\[
= \bar{v} + \frac{1}{\tau_0 + Y_t} \sum_{i=1}^{t} \hat{R}_i = \bar{v} + \frac{1}{\tau_0 + Y_t} \left( Y_t \eta + \sum_{i=1}^{t} (A_i - A_{i-1})\tau_u u_i \right).
\]
Using this relation, we get

\[ \gamma_n \Delta x_{a,t} = Q_{a,t}(z_{a,t} - p_t) - Q_{a,t-1}(z_{a,t-1} - p_{t-1}) \]

\[ = Q_{a,t}\left( \tilde{v} + \eta + \frac{Q_{t-1}}{Q_{a,t}} \xi_{a,t-1} + \frac{q_t}{Q_{a,t}} e_{a,t} - \left( \tilde{v} + \frac{1}{\tau_0 + Y_t} \sum_{i=1}^t \tilde{R}_i \right) \right) \]

\[ = Q_{a,t-1}\left( \tilde{v} + \eta + \xi_{a,t-1} - \left( \tilde{v} + \frac{1}{\tau_0 + Y_{t-1}} \sum_{i=1}^{t-1} \tilde{R}_i \right) \right) \]

\[ = q_{a,t} e_{a,t} + Q_{a,t} \left( \eta - \frac{1}{\tau_0 + Y_t} \left( Y_t \eta + \sum_{i=1}^t (A_i - A_{i-1}) \tau_{u_i} u_i \right) \right) \]

\[ = Q_{a,t-1} \left( \eta - \frac{1}{\tau_0 + Y_{t-1}} \left( Y_{t-1} \eta + \sum_{i=1}^{t-1} (A_i - A_{i-1}) \tau_{u_i} u_i \right) \right) \]

\[ = q_{a,t} e_{a,t} \sim N(0, q_t) \left( \tau_0 \eta - \sum_{i=1}^{t-1} (A_i - A_{i-1}) \tau_{u_i} u_i \right) \]

\[ \sim N(0, q_t) \left( \tau_0 \eta - \sum_{i=1}^{t-1} (A_i - A_{i-1}) \tau_{u_i} u_i \right) - \frac{Q_{a,t}}{\tau_0 + Y_t} \left( A_t - A_{t-1} \right) \tau_{u_t} u_t \]

\[ \sim N(0, q_t) \]

where \( f_t = \tau_0 \eta \sim N(0, \tau_0) \), \( f_t = f_{t-1} + g_t, g_t \sim N(0, y_t) \), \( f_t | f_{t-1} \sim N(f_{t-1}, y_t) \).

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

\[ W_t = \sum_a |\alpha_{a,t} \xi_{a,t} - \alpha_{a,t+1} f_{t-1} - \alpha_{a,t}^3 g_t|, \]

\[ W_{t+1} = \sum_a |\alpha_{a,t+1} \xi_{a,t+1} - \alpha_{a,t+1}^2 f_{t-1} + g_{t+1} - \alpha_{a,t+1}^3 g_{t+1}|, \]

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

The result now follows from the following lemma:

**Lemma 14** Assume that \( \tilde{A}, \tilde{B} \) and \( \tilde{z} \) are independent with standardized normal distributions. Then, for any \( a > 0 \) and \( b > 0 \), \( \text{Cov}(|a \tilde{A} + \tilde{z}|, |b \tilde{B} + \tilde{z}|) > 0 \).

**Proof of Lemma 14:** First, we note that if \( \text{Cov}(\tilde{X}, \tilde{Y} | \tilde{Z}_1, \ldots, \tilde{Z}_n) > 0 \) for all realizations of the random variables \( \tilde{Z}_1, \ldots, \tilde{Z}_n \), then it must be that \( \text{Cov}(\tilde{X}, \tilde{Y}) > 0 \) unconditionally. This follows since

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

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

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

where the inequality follows from \( \text{max}(c, x) \) being a monotone transformation of \( x \), so \( \text{max}(Z_1, |\tilde{z}|) \), and \( \text{max}(Z_2, |\tilde{z}|) \) are therefore comonotonic. Given the argument, the positivity must then also hold unconditionally, and the lemma therefore follows.
Now, this lemma implies that all the terms that make up $W_t$ and $W_{t+1}$ have pairwise positive covariances, and it is indeed the case that $\text{Cov}(W_t, W_{t+1}) > 0$, showing the second result.

For the third result, we proceed as follows: It follows from (46) that

$$\tilde{v} - p_t = \frac{1}{\tau_v + Y_t} f_t = \eta - \frac{1}{\tau_v + Y_t} \sum_{i=1}^{t} R_i.$$  \hfill (47)

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

For the first result, from (47), the relationship

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

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

We are done.
References


Andrei, D., 2012, Information percolation driving volatility, Swiss Finance Institute, unpublished manuscript.


Edgeworth, F. Y., 1881, Mathematical Psychics: An essay on the mathematics to the moral sciences (Wiley).


Han, B., and L. Yang, 2011, Social networks, information acquisition, and asset prices, Working paper, University of Texas at Austin.


