

Information Linkages and Correlated Trading

Unabridged Version*

Paolo Colla
Università Bocconi

Antonio Mele
London School of Economics

First draft: September 13, 2004

This version: October 6, 2008

Abstract

In a market with informationally connected traders, the dynamics of volume, price informativeness, price volatility, and liquidity are severely affected by the information linkages every trader experiences with his peers. We show that in the presence of information linkages among traders, volume and price informativeness increase. Moreover, we find that information linkages improve or damage market depth, and lower or boost the traders' profits, according to whether these linkages convey positively or negatively correlated signals. Finally, our model predicts patterns of trade correlation consistent with those identified in the empirical literature: trades generated by "neighbor" traders are positively correlated and trades generated by "distant" traders are negatively correlated.

*We wish to thank Matthew Spiegel (the editor), an anonymous referee, Henry Cao, Marco Pagano, Andrea Prat, Jean-Charles Rochet, Randi Rosenblatt, Mark Seasholes, Hyun Shin, Dimitri Vayanos, Paolo Vitale, Johan Walden and seminar participants at the Centre for Studies in Economics and Finance (Salerno), Ente Luigi Einaudi (Rome), ESSEC Business School (Paris), ISEG (Technical University of Lisbon), Instituto de Empresa Business School (Madrid), FMG-LSE, Norwegian School of Economics and Business Administration, Norwegian School of Management, Stockholm School of Economics, Universitat Pompeu Fabra, University of Toulouse, USI Lugano (Institute of Finance), the 2008 American Finance Association, CEPR/Studienzentrum Gerzensee ESSFM (2005), CORE Summer School (2005), the 12th Mitsui Life Symposium on Financial Markets (Information in Trading) at the Ross School of Business, Ann-Arbor (2006), and the Society for Advancement in Economic Theory meeting in Vigo (2005) for very useful comments and suggestions. The first author acknowledges financial support provided through the European Community's Human Potential Programme under contract HPRN-CT-2002-00232. The second author thanks the British EPSRC for financial support via grant EP/C522958/1. The usual disclaimer applies. Address correspondence to Antonio Mele, The London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom, or email: a.mele@lse.ac.uk.

One pervasive feature in financial markets is the existence of information linkages among market participants. Traders and investors are socially connected and have access to comparable sources of information. Many writers describe the financial community as one of overlapping groups of people who share similar opinions, either because they are endowed with comparable signals about the fundamentals and/or communicate regularly with one another [e.g., Shiller (1984, 2005), Shiller and Pound (1989), Hertz (1998)], or simply because they are exposed to similar cultural biases [e.g., Guiso, Sapienza and Zingales (2006)]. Many information-based explanations of asset price movements hinge upon the assumption that investors do not experience information linkages at all. In this paper, we relax this assumption and explore the resulting implications along several dimensions: market efficiency, liquidity, trading volume, correlation among trades and volumes generated by heterogenous traders, and gains from informed trading.

Our notion of information linkages is closely related to the recent empirical literature on the value of local information, social interactions and information networks in financial markets. For example, portfolio decisions are known to be related to social networks, be they cultural or linguistic [Kelly and O’Grada (2000), Grinblatt and Keloharju (2001), Cohen, Frazzini and Malloy (2007)]. Locality effects matter as well. Coval and Moskowitz (1999) provide strong evidence that geographical proximity influences managers’ portfolio choices. Hong, Kubik and Stein (2005) document that US fund managers located in the same city commit to correlated investment decisions. The authors argue that such correlated choices arise either through peer-to-peer communication or because fund managers in a given area base their decisions upon common sources of information - such as a local newspaper or TV station. Similarly, Feng and Seasholes (2004) find that in the Chinese stock market, trades are positively correlated for geographically close investors, but *negatively* correlated for distant investors.

A rational explanation of these findings must necessarily rely on a pronounced heterogeneity in the investors’ information endowments. Thus, at the heart of our analysis is the idea that in asset markets, there are groups of traders whose signals and beliefs are more correlated with some and less correlated with other groups of traders. A natural measure of informational distance between any two traders is the amount of information they share. To generate heterogeneity in informational distance, we consider a model with strategic traders who are locally connected to common sources of information about the long term value of an asset. We call these local connections “information linkages.” Accordingly, we define close traders as those who are connected through these linkages, and distant traders as those who are not. Local connections give rise

to overlapping networks of traders, which may include only one's closest neighbors or the entire market. Indeed, there are no obvious arguments suggesting whether information connections should be best thought of as local or global. Our framework is kept as general as possible to account for a wide spectrum of possibilities.

Our model builds on the seminal papers of Holden and Subrahmanyam (1992), Foster and Viswanathan (1996) and Back, Cao and Willard (2000), who develop multi-trader generalizations of the Kyle's (1985) model. The point of departure is the introduction of information linkages among traders. Our main goal is to uncover patterns of trade correlations among traders. In equilibrium, every trader makes use of the information available at the linkages he has with his local peers. He also knows, however, that by trading aggressively, he reveals part of this information to distant peers. What is the ultimate effect on the correlation between "close" and "distant" trades?

The central prediction of our multiperiod model is that in a market where strategic traders have access to information linkages, the correlation among trades is heterogeneous, both temporally and spatially. More in detail, the correlation among trades is very high at the beginning of the trading period. The same correlation decreases over time, and exhibits different patterns, depending upon the informational distance among traders:

- For traders who are sufficiently close (close neighbors), the correlation among opinions and trades decreases over the trading period, although it remains persistently high.
- Traders' opinions and trades diverge with the informational distance. Eventually, the correlation between trades is negative for relatively distant traders. A significant and persistent divergence in trades occurs even when the number of information linkages is large enough to make any two traders' opinions quite similar at the beginning of the trading period.

The economic interpretation for the *positive* correlation between close trades is intuitive: the linkages among traders raise the correlation of the information endowments and, hence, the trade correlation. This property, although intuitive, is in sharp contrast with well-known features of markets without information linkages, where the correlation among trades eventually becomes negative [see Foster and Viswanathan (1996)]. At the same time, our model also matches the empirical evidence on the *negative* correlation between distant trades documented in the literature [see our previous discussion of Feng and Seasholes (2004) and the papers we mention below].

The economic mechanism at work in the model is the following. Over time, and as in Foster and Viswanathan (1996), the equilibrium asset price conveys more and more information about the traders' average opinion of the asset value, not the single private opinions every trader has about this value. Therefore, over time, traders stand on opposite sides of the market, on average, which makes the correlation between trades decrease. In our model, this correlation tends to become negative, especially for distant traders. However, the presence of information linkages in the market "kicks in" for close traders. In particular, we find that for close traders, the information effects induced by the linkages dominate the negative correlation arising from the tendency of each trader to stay on the opposite side of the market. Therefore, the cross sectional properties of correlated trading we unveil in this paper originate from the combined effects of both the information linkages and the market maker's learning process.

Although these properties are consistent with the available empirical evidence, there remains a number of empirical issues called for by our model. Importantly, the current empirical literature on correlated trading relies on the unconditional correlation among trades in a given period [as in Lakonishok, Shleifer and Vishny (1992) or, more recently, Feng and Seasholes (2004), Hong, Kubik and Stein (2005), Bae, Yamada and Ito (2006), Barber, Odean and Zhu (2006), and Dorn, Huberman and Sengmueller (2008)]. Our theoretical analysis suggests that a more interesting concept is that of trade correlation occurring prior to a corporate event such as an earnings announcement. The model implications on this correlation are quite strong. For example, they can be used to test whether correlated trading in financial markets is induced by an alternative and, perhaps, more standard mechanism, based on herding behavior. Herding would indeed lead to a positive correlation among trades initiated by close traders. However, herding-based explanations for correlated trading rely on agents' sequential moves [e.g., Bikhchandani and Sharma (2001)], and should be consistent with a trade correlation that increases prior to a corporate earnings event. As noted, our information linkages mechanism predicts just the opposite: correlation of trades for close traders is high and positive, but it decreases over the trading period.

Our model also predicts that information linkages among traders have implications on traditional market variables such as volume, liquidity, and efficiency, as well as on the gains from informed trading. Linkages do affect all these variables. As we shall show, even a small number of information linkages among traders can induce a quite large effect on the equilibrium price and trading activity. The predictions of our model can be streamlined as follows:

- Compared to a market without information linkages, a market with information linkages is characterized by higher volume, efficiency, and in general, higher liquidity.
- Information linkages boost the expected gains from informed trading if they convey negatively correlated signals. However, if the signals available at the information linkages are positively correlated, the mere existence of these linkages damages the traders' profits.

The economic interpretation for the first property is as follows. Consider the prediction about volume. Heterogeneity in private information is a source of monopolistic power for traders. But information linkages destroy part of this monopolistic power. Consequently, every trader trades aggressively in order to preempt his peers, which makes market-wide volume increase. (The model predictions about efficiency and liquidity can be understood in a similar vein.)

The model implications on the gains from informed trading can be explained as follows. In our model, information linkages affect the traders' profits through two channels. On the one hand, information linkages damage the traders' monopolistic power. On the other hand, the very same linkages improve the quality of the traders' inference about the fundamental value of the asset. If the signals available at the information linkages are positively correlated, the losses generated by the first effect are larger than the gains stemming from the second effect. This property holds under a wide range of conditions on the initial traders' beliefs and the market structure, as summarized by the number of traders and batch auctions, as well as the initial correlation of the signals made available at the traders' locations and the number of information linkages. These results, however, are reversed if the signals available at the information linkages are negatively correlated.

Although Section 3 further qualifies these findings, the model implications about volume and liquidity are in general quite clear: stocks traded in markets with many informationally connected traders are those with large volume and liquidity. To date, the available empirical literature provides only indirect support to these predictions. For example, Dorn, Huberman and Sengmueller (2008) show that correlated trading tends to be greater in heavily traded stocks, a property our model generates through the information linkages channel. Similarly, Hong, Lim and Stein (2000) supply evidence that stocks with lower analyst coverage are those for which firm-specific information flows more slowly to the market; Brennan and Subrahmanyam (1995) find that stocks with higher analyst coverage are those with lower adverse selection costs of trading.

While analyst coverage might proxy for the information linkages in the market, there exist

alternative measures of information connectedness devised in the more recent literature, which purportedly capture locality and network effects [see, e.g., Coval and Moskowitz (1999), Grinblatt and Keloharju (2001), Ivković and Weisbenner (2005), or the empirical networks methods surveyed in Goyal (2007) and Jackson (2008)]. These measures are well suited to investigate how trading volume and liquidity can possibly relate to the presence of information linkages among traders. Similarly, the empirical literature lacks a systematic analysis of the impact information linkages have on trading profits. One exception is Hau (2001), who provides evidence that local interactions between traders and financial intermediaries might adversely affect trading profits. Our theoretical work sheds new light on this finding. At the same time, it calls for further investigations in which such preliminary evidence about information linkages and traders’ profits could be expanded and related to the other testable implications of our model discussed so far.

The remainder of the article is as follows. In the next section, we develop the information structure of the model. In Section 2, we derive the equilibrium while in Section 3, we discuss in detail the testable implications of the model. Section 4 contains a succinct discussion of related work. Section 5 concludes. The Appendix contains technical details omitted in the main text.

1 Information structure

1.1 The asset market and traders’ location

We consider a market for one risky asset in which trading takes place in $N \geq 1$ batch auctions. The asset pays a random payoff $f \sim N(0, \sigma_f^2)$ at the end of the trading period. The crucial feature of the model is that a number of imperfectly competitive traders experience information linkages related to the asset payoff. Precisely, we assume that the traders are physically located around a circle. By convention, they are ordered clockwise, such that trader i has trader $i + 1$ as his clockwise neighbor and trader $i - 1$ as his counterclockwise neighbor (see Figure 1). There are M such traders and, for reasons developed below, we assume M is an odd number.

Signals about the fundamental value f are available at each trader’s location. Let $\mathbf{s}_0 = [s_{1,0}, \dots, s_{i,0}]^\top$ be the $M \times 1$ vector of the signals in the market. We assume that each signal $s_{i,0}$ is available at the i -th trader’s location and is observed by some of the i -th trader’s neighbors, on both sides. Hence, we allow for “double-sided” information linkages. Precisely, we assume

that the signal available at any trader's location is observed by G clockwise neighbors and G counterclockwise neighbors of any given trader, and we take G to be exogenous in the model. The i -th trader's information set is, then, $\mathbf{s}_{i,0} = [s_{i-G,0}, \dots, s_{i,0}, \dots, s_{i+G,0}]^\top$, $G \in [0, (M-1)/2]$.¹ For example, for $G = 1$, the signal available at the i -th trader location is also observed by traders $i-1$ and $i+1$, as in Figure 1. In this case, traders $i-1$, i and $i+1$ observe $s_{i,0}$; trader i observes $s_{i-1,0}$, $s_{i,0}$ and $s_{i+1,0}$, and so forth.

To ease notation, we let $\hat{G} = 2G + 1$ be the number of signals every trader has access to. Let $\bar{s}_{i,0}$ denote the average signal available to the i -th trader,

$$\bar{s}_{i,0} = \hat{G}^{-1} \sum_{k=-G}^G s_{i+k,0}. \quad (1)$$

In the absence of information linkages, we have that $\hat{G} = 1$ and, hence, $\mathbf{s}_{i,0} = \bar{s}_{i,0} = s_{i,0}$ for all i . In principle, the maximum number of information linkages is $\hat{G} - 1 = M - 1$, in which case $\mathbf{s}_{i,0} = \mathbf{s}_0$ for all i . However, this market may fail to have a linear equilibrium as the number of auctions N gets large and the uncertainty related to liquidity trades (to be introduced later) gets small, as initially conjectured by Holden and Subrahmanyam (1992) and shown by Back, Cao and Willard (2000). Therefore, we shall limit ourselves to analyze cases in which $\hat{G} < M$.

Our information structure can be interpreted in a variety of ways. For example, every signal $s_{i,0}$ can be thought of as being broadcast to the i -th trader's location through a local newspaper or TV station, as suggested in the empirical work of Hong, Kubik and Stein (2005) and Ivković and Weisbenner (2005), who argue that locality effects are likely to be related to information linkages among investors. In our model, informationally linked traders are those who have access to some common information source. In the limiting case in which $\hat{G} = 1$, every trader gathers information from a unique local news source, and there are no information linkages among them. As \hat{G} increases, these sources of news overlap across traders, and the number of information linkages every trader experiences with his peers, $2G$, is interpreted as the media coverage of information providers.

The information network we consider can also be interpreted as one in which each trader observes the average signal of his neighbors with some error. To illustrate, consider again Figure

¹In the language of the networks literature, our information structure is a regular ring lattice with M vertices (the traders), with each of the vertices being connected to $2G$ neighbors through undirected edges: see, for example, the seminal work by Watts and Strogatz (1998).

1. In this example, the i -th trader observes, among other things, $\bar{s}_{i+1,0} + \eta_{i,0}$, where $\bar{s}_{i+1,0} = \frac{1}{3}(s_{i,0} + s_{i+1,0} + s_{i+2,0})$ is the average signal available to the $(i+1)$ -th trader, and the error term is $\eta_{i,0} = -\frac{1}{3}s_{i+2,0}$. This interpretation is important: as we shall show, each trader's strategy is linear in his average signal, in equilibrium. Therefore, the information structure in this market is such that in equilibrium, each trader observes the trade of his neighbors with some error.

Another interpretation of our information structure is that of social proximity among traders. Social connections can relate to geographical, cultural, demographic or linguistic distance, and translate into differences in beliefs among traders. For example, in an early contribution, Shiller and Pound (1989) argue that the presence of social networks might explain portfolio decisions among US investors. More recently, Grinblatt and Keloharju (2001) find that language, culture and geographical distance affect portfolio choice in the Finnish stock market; Cohen, Frazzini and Malloy (2007) document that education networks affect the trading behavior of portfolio managers connected to senior officers of publicly traded companies. In general, repeated interactions among traders can lead them to sharing close opinions and views. At the other extreme, socially distant traders are less likely to interact and, hence, more likely to have relatively more independent opinions and views. Then, our model can be understood as one that focuses on the asset pricing implications of a given architecture of opinion formation. Stein (2007) lays down the foundations for an honest exchange of ideas to arise among strategic players. This exchange stems from complementarities in the production of ideas. For example, to come up with a useful idea about an asset payoff, it may be necessary to have access to the peers previous ideas on the same topic. In our model, information linkages can be thought of as the result of such fruitful conversations among close traders.

1.2 The distribution of signals

Next, we describe the distribution of the signals available at the traders' location. We follow Foster and Viswanathan (1996) and assume that all the signals are jointly normally distributed with mean zero and variance-covariance matrix equal to $\Psi_0 = E([s_{1,0}, \dots, s_{M,0}]^\top [s_{1,0}, \dots, s_{M,0}])$. The unconditional distribution of the signals is symmetric in that it is invariant to a permutation

of the indices $1, \dots, M$. The joint distribution of the vector $[f \ \mathbf{s}_0]^\top$ is given by:

$$\begin{bmatrix} f \\ \mathbf{s}_0 \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \sigma_f^2 & c_0 \mathbf{1}^\top \\ c_0 \mathbf{1} & \Psi_0 \end{bmatrix} \right), \quad \Psi_0 = \begin{bmatrix} \Lambda_0 & \Omega_0 & \cdots & \Omega_0 \\ & \Lambda_0 & & \Omega_0 \\ & & \ddots & \vdots \\ & & & \Lambda_0 \end{bmatrix}, \quad (2)$$

where we assume that Ψ_0 is invertible, which it is, provided $\Lambda_0 > -(M-1)\Omega_0$, a restriction we maintain throughout the paper. Such restriction ensures that the average of the average signals, $\bar{s} = M^{-1} \sum_{i=1}^M \bar{s}_{i,0}$, is a sufficient statistic for the full information liquidation value,

$$E(f | \mathbf{s}_0) = \theta \bar{s}, \quad (3)$$

where $\theta = c_0 M (\Lambda_0 + (M-1)\Omega_0)^{-1}$.

The average signal every trader has access to, $\bar{s}_{i,0}$, plays a key role for each trader forecasting problem and, hence, for the trading strategies, as we shall show below. We now derive the distribution of the average signals $(\bar{s}_{i,0})_{i=1}^M$, under the assumptions made so far.

We denote the unconditional variance-covariance matrix of the average signals with $\bar{\Psi}_0 = E([\bar{s}_{1,0}, \dots, \bar{s}_{M,0}]^\top [\bar{s}_{1,0}, \dots, \bar{s}_{M,0}])$. The elements of this matrix depend on the number of information linkages in the market. Accordingly, we set $\bar{\Psi}_0 = \bar{\Psi}_0(G)$, where

$$\bar{\Psi}_0(G) = \begin{bmatrix} \bar{\Lambda}_0(G) & \bar{\Omega}_0(1, G) & \cdots & \bar{\Omega}_0(\frac{M-1}{2}, G) & \cdots & \bar{\Omega}_0(-1, G) \\ & \bar{\Lambda}_0(G) & & & & \bar{\Omega}_0(-2, G) \\ & & & \ddots & & \vdots \\ & & & & & \bar{\Lambda}_0(G) \end{bmatrix}, \quad (4)$$

and

$$\begin{aligned} \bar{\Lambda}_0(G) &= \text{var}(\bar{s}_{i,0}) \\ \bar{\Omega}_0(k, G) &= \text{cov}(\bar{s}_{i+k,0}, \bar{s}_{i,0}), \quad k = \mp 1, \mp 2, \dots, \mp \frac{M-1}{2} \end{aligned}$$

denote the unconditional variance of the average signals available to any trader ($\bar{\Lambda}_0(G)$), and the unconditional covariance between the average signals of any two traders who are located k positions apart ($\bar{\Omega}_0(k, G)$) ($k \neq 0$). We have that $\bar{\Omega}_0(k, G) = \bar{\Omega}_0(-k, G)$, which follows by both the circular information structure and the double-sided nature of the information linkages in this

market. Furthermore, $\bar{\Lambda}_0(G)$ is the same for each trader and given by

$$\bar{\Lambda}_0(G) = \frac{\Lambda_0 + 2G\Omega_0}{\hat{G}}. \quad (5)$$

due to the symmetric unconditional distribution of the signals in Eq. (2).

Consider, next, any two traders i and $j = i + k$ with $k \neq 0$, and the unconditional covariance between the average signals these two traders have access to, i.e. the off-diagonal element $\bar{\Omega}_0(k, G)$ in Eq. (4). The covariance between the average signals depends on: (i) k , the distance between the i -th trader and the $(i + k)$ -th trader; and (ii) the number of information linkages in the market, G . This dependence on k arises because the number of signals every trader shares with the other peers depends on their relative position on the circle. For example, assume that $2G < (M - 1)/2$. In this case, trader i shares $2G$ signals with trader $i + 1$, $2G - 1$ signals with trader $i + 2$ and in general $2G + 1 - k$ signals with trader $i + k$. As Figure 2 illustrates, eventually, trader i shares no signals with trader $i + 2G + 1$ and beyond.

To derive explicitly the covariance between any two traders' information endowments, we need to distinguish between two situations that may alternatively arise, depending on information linkages being "large" or "small." Intuitively, if the number of information linkages, $2G$, is larger than the number of traders each trader has on either side, $(M - 1)/2$, some pieces of information might "go through" over and above the dimension of the information network any trader belongs to. Then, we need to analyze the following two cases.

- (i) *Small number of information linkages:* $2G \leq (M - 1)/2$. In this case, any two traders can not observe common signals, provided they are located sufficiently apart, as in Figure 1 with, e.g., $M = 7$. However, if two traders are sufficiently close, they have access to some common information. In Appendix A, we show that,

$$\text{For } 2G \leq \frac{M-1}{2}, \quad \bar{\Omega}_0(k, G) = \begin{cases} \bar{\Lambda}_0(G) - \hat{G}^{-2}(\Lambda_0 - \Omega_0)k, & \text{for } k \in [1, 2G + 1] \\ \Omega_0, & \text{for } k \in [2G + 1, \frac{M-1}{2}] \end{cases} \quad (6)$$

- (ii) *Large number of information linkages:* $2G \geq (M - 1)/2$. In this case, information linkages are such that any two traders might have access to common sources of information even when each of them does not observe the signals available at the location of the other, a property we label "double overlap." This feature of the model results in a higher information

clustering of traders. It parallels a similar “small worlds” property analyzed in the networks literature, by which a network may be both highly clustered and have a low average distance among its nodes (our traders) [see, e.g., Watts and Strogatz (1998)]. Consider, for example, Figure 3. In this example, trader i has access to the signal available at the location of trader $i - \ell$, although he can not observe the signal available at the location of trader $i + k_1$. However, traders $i + k_1$ and $i - \ell$ are informationally connected, as their relative distance is less than G . Therefore, traders i and $i + k_1$ do observe some common signal (at least the signal $s_{i-\ell,0}$), even if they are not connected to the same information linkage. In Appendix A, we show that the “double overlap” modifies the correlation structure in (6) as follows:

$$\text{For } 2G \geq \frac{M-1}{2}, \quad \bar{\Omega}_0(k, G) = \begin{cases} \bar{\Lambda}_0(G) - \hat{G}^{-2}(\Lambda_0 - \Omega_0)k, & \text{for } k \in [1, 2(\frac{M-1}{2} - G)] \\ 2\bar{\Lambda}_0(G) - \hat{G}^{-2}M(\Lambda_0 - \Omega_0) - \Omega_0, & \text{for } k \in [2(\frac{M-1}{2} - G), \frac{M-1}{2}] \end{cases} \quad (7)$$

To summarize, the variance-covariance matrix of the average signals available to all traders, $\bar{\Psi}_0(G)$ in Eq. (4), depends on the number of information linkages, which makes $\bar{\Psi}_0(G)$ not invariant to a permutation of the indices $i = 1, \dots, M$, unlike Ψ_0 in Eq. (2). By Eq. (5), the elements on the main diagonal, $\bar{\Lambda}_0(G)$, are the same. The off-diagonal elements, instead, are decreasing in the traders’ relative distance k , according to the pattern in Eqs. (6)-(7). Thus, the correlation between the average signals of any two traders located k positions apart, $\bar{\rho}_0(k, G) = \bar{\Lambda}_0(G) / \bar{\Omega}_0(k, G)$, varies with their relative distance.

Despite its asymmetry, the matrix $\bar{\Psi}_0(G)$ preserves a useful property of symmetric matrices. Define the unconditional covariance between the average signal available to any trader i , $\bar{s}_{i,0}$, with the sum of the remaining traders’ average signals,

$$\bar{\Gamma}_0(G) = \text{cov} \left(\sum_{k \neq i} \bar{s}_{k,0}, \bar{s}_{i,0} \right) = \sum_{k \neq i} \bar{\Omega}_0(k, G).$$

Due to the information structure in this market, $\bar{\Gamma}_0(G)$ is the same for each trader. In Appendix A, we prove that,

$$\bar{\Gamma}_0(G) = (M - 1)\Omega_0 + \frac{2G}{\hat{G}}(\Lambda_0 - \Omega_0), \quad \text{for all } G \in [0, \frac{M-1}{2}]. \quad (8)$$

The fact that both $\bar{\Lambda}_0(G)$ and $\bar{\Gamma}_0(G)$ do not depend on location implies that each trader’s forecasts of (i) the asset value and (ii) the *sum* of all remaining traders’ average signals, are independent of k , as we shall show below.

2 Equilibrium price and trades

This section develops a dynamic model that relies on the information structure described in the previous section. We derive the equilibrium price and traders' strategies (Proposition 1) following the methods set forth in Foster and Viswanathan (1996), in which the equilibrium outcome arises out of a set of possibly off-equilibrium strategies and prices. Then, we provide a characterization of the trading strategies (in Proposition 2), which we shall use to understand the correlation of trades and volumes we uncover in Section 3.

2.1 Equilibrium characterization

Let $(x_{i,n})_{n=1}^N$ be the sequence of orders submitted by the i -th trader over the trading period. Trades are chosen so as to maximize the expected profits, viz

$$W_{i,n} \equiv \max_{(x_{i,t})_{t=n}^N} E \left[\sum_{t=n}^N (f - p_t) x_{i,t} \middle| F_{i,n} \right], \quad n = 1, \dots, N, \quad (9)$$

where $F_{i,n}$ is the trader i information set at the n -th batch auction. On top of these informed orders, there is a sequence of liquidity trades $(u_n)_{n=1}^N$, where u_n is independent and identically distributed as a standard normal variable, with mean zero and variance σ_u^2 . Thus, the aggregate order flow is given by:

$$y_n = \sum_{i=1}^M x_{i,n} + u_n, \quad n = 1, \dots, N. \quad (10)$$

The $(M + 1)$ -th market participant is the market maker, who commits himself to offset the order flow according to the Semi-Strong efficiency rule:

$$p_n = E(f | F_n), \quad n = 1, \dots, N. \quad (11)$$

where $F_n = \{(y_t)_{t=1}^n\}$ denotes the market maker's information set at the n -th batch auction. Upon observing the aggregate order flow, the market maker updates his estimate of the average signal available to any trader i as

$$t_n = E(\bar{s}_{i,0} | F_n) = \hat{G}^{-1} \sum_{k=-G}^G E(s_{i+k,0} | F_n) = E(s_{i,0} | F_n), \quad (12)$$

which is independent of the i -th trader's specific location since the variance-covariance matrix Ψ_0 in Eq. (2) is symmetric.

We focus on equilibria in which both traders' strategies and the pricing function are linear with respect to the information structure, that is, (i) each trader strategy is linear in his informational advantage relative to the market maker,

$$x_{i,n} = \hat{G}\beta_n (\bar{s}_{i,0} - t_{n-1}), \quad (13)$$

and (ii) the market maker's learning about the asset value satisfies,

$$p_n = p_{n-1} + \lambda_n y_n, \quad (14)$$

for some deterministic sequences β_n and λ_n to be determined in equilibrium. We shall come back later to the economic interpretation of the linear strategy $x_{i,n}$ in Eq. (13) (see Proposition 2).

Foster and Viswanathan (1996), Back, Cao and Willard (2000) and Bernhardt and Miao (2004) restrict attention to the same equilibrium conditions summarized by Eqs. (13) and (14). Moreover, Foster and Viswanathan (1996) explicitly prove that any trader i can manipulate the remaining traders' beliefs about the asset value only through the aggregate order flow. The existence of information linkages does not destroy this property, as we show in Appendix B (see Lemma 3). Consequently, the residual order flow, $(y_{i,t} - x_{i,t})_{t=1}^{n-1}$, constitutes a redundant piece of information in the information set of the i -th trader at the n -th batch auction, and we set $F_{i,n} = \{\bar{s}_{i,0}, (y_t)_{t=1}^{n-1}\}$. It follows that $(\bar{s}_{i,0} - t_{n-1})$ is sufficient for any trader i to (i) forecast the asset value and (ii) forecast (the sum of) the other traders' informational advantage, before submitting his order at time n , i.e.

$$E(f - p_{n-1} | F_{i,n}) = \frac{\theta (\bar{\Gamma}_{n-1}(G) + \bar{\Lambda}_{n-1}(G))}{M \bar{\Lambda}_{n-1}(G)} (\bar{s}_{i,0} - t_{n-1}), \quad (15)$$

and

$$E\left(\sum_{j \neq i} (\bar{s}_{j,0} - t_n) \middle| F_{i,n}\right) = \frac{\bar{\Gamma}_{n-1}(G)}{\bar{\Lambda}_{n-1}(G)} (\bar{s}_{i,0} - t_{n-1}). \quad (16)$$

where $\bar{\Lambda}_{n-1}(G)$ (resp. $\bar{\Gamma}_{n-1}(G)$) is the variance of the residual informational advantage (resp. covariance between $\sum_{k \neq i} (\bar{s}_{k,0} - t_{n-1})$ and $(\bar{s}_{i,0} - t_{n-1})$) conditional on F_{n-1} . Note that the regression coefficients in Eqs. (15)-(16) are identical for all traders, despite the asymmetry in the variance-covariance matrix of average signals. As a consequence, the trading intensity in Eq. (13) does not depend on each trader's location. This property keeps our model markedly distinct from the two traders "simultaneous information acquisition" model considered by Bernhardt and

Miao (2004), in which signals have different variances (see their Section 3.2 and Proposition 5). A further difference between our analysis and that in Bernhardt and Miao (2004) is that in finding the equilibrium, we take into account *off* equilibrium paths. To this end, we follow the approach in Foster and Viswanathan (1996), and conjecture that the price induced by past suboptimal play shows up both in trader i strategy and in his value function. The necessary and sufficient conditions for an equilibrium hinge upon the mutual consistency between these two conjectures:

Proposition 1. *There exists a unique Bayesian Nash equilibrium in which trading strategies and prices are as in Eqs. (13)-(14).*

The complete equilibrium characterization is detailed in Appendix B. Note that in our model, not only are traders concerned with forecasting the information the remaining traders possess. This learning process is also complicated by the existence of heterogeneous distances among the traders, and the number of information linkages every trader has with his neighbors. Indeed, trading strategies and prices are sensibly affected by the heterogeneous correlation structure arising from the information linkages among traders, as we shall explain in detail in Section 3.

2.2 Value estimates and subjective mispricings

Back, Cao and Willard (2000) show that in continuous time, the equilibrium trading strategies and the expectation of the sum of others' trades are linear in the subjective mispricing perceived by any trader i , i.e. $E(f|F_{i,n}) - p_{n-1}$. This property survives in our model with information linkages, since traders adopt identical strategies. Our model displays the additional features that the forecast of individual trades is linear in the subjective mispricing, and that it *does* depend upon the traders' distance. We now produce the formal argument, which we shall use to develop the interpretation of the model predictions in the next section.

We denote with Δ_n the ratio of the market maker's residual variance of the average signal in the market, $var(\bar{s}|F_n)$, to the residual variance of each trader's average signal, $var(\bar{s}_{i,0}|F_n)$, viz

$$\Delta_n = \frac{var(\bar{s}|F_n)}{var(\bar{s}_{i,0}|F_n)}. \quad (17)$$

Furthermore, let $\sigma_{f,n}^2$ and $\zeta_{f,n}^2$ denote respectively the market maker and trader i residual variance

of the full information asset value after n rounds of trading

$$\sigma_{f,n}^2 = \text{var} [E(f | \mathbf{s}_0) | F_n] \quad (18)$$

$$\varsigma_{f,n}^2 = \text{var} [E(f | \mathbf{s}_0) | F_{i,n+1}]. \quad (19)$$

The equilibrium trading strategies in Eq. (13) can then be expressed as in the following proposition:

Proposition 2. *Any trader i strategy in Eq. (13) is linear in the subjective mispricing $E(f | F_{i,n}) - p_{n-1}$, i.e.*

$$x_{i,n} = \frac{\hat{G}\beta_n}{\Delta_{n-1}\theta} [E(f | F_{i,n}) - p_{n-1}], \quad (20)$$

and

$$\Delta_n = \frac{\sigma_{f,n}^2 - \varsigma_{f,n}^2}{\sigma_{f,n}^2}. \quad (21)$$

Moreover,

$$E\left(\sum_{j \neq i} x_{j,n} \middle| F_{i,n}\right) = (M\Delta_{n-1} - 1) x_{i,n}, \quad (22)$$

and

$$\begin{aligned} E(x_{i+k,n} | F_{i,n}) &= \bar{\rho}_{n-1}(k, G) x_{i,n} \\ &= \underbrace{E\left(\sum_{j \neq i} x_{j,n} \middle| F_{i,n}\right)}_{\equiv \text{Market side direction}} - \underbrace{\left(\sum_{\ell \notin \{i, i+k\}} \bar{\rho}_{n-1}(\ell, G)\right) x_{i,n}}_{\equiv \text{Information linkages direction}}, \end{aligned} \quad (23)$$

where, for a given G , $\bar{\rho}_n(k, G)$ is the correlation, at the n batch, between the average signals available to any two traders who are located k positions apart.

Proposition 2 generalizes, in a discrete time setting, Lemma 6, and other results, in Back, Cao and Willard (2000). As in Back, Cao and Willard, Eq. (20) indicates that any trader buys/sells the asset if he believes the asset is undervalued/overvalued. Moreover, Eq. (21) reveals that Δ_n is a measure of the relative “tightness of beliefs” between the market maker and the traders, i.e. the percentage of the market maker’s residual uncertainty that is accounted for by each trader’s information.

Eqs. (20) and (22) imply that any trader expects to trade in the same direction of his peers and, thus, to be on the same side of the market if and only if the market maker’s residual

uncertainty is sufficiently large, i.e. $\Delta_n > M^{-1}$. This property generates a “rat race.” When $\Delta_n < M^{-1}$, each trader expects to be on the opposite side of the market, which generates a “waiting game.” The next section demonstrates that in our model, the dynamic properties of the “tightness of beliefs,” Δ_n , are severely affected by the number of information linkages every trader has with his peers, $2G$.

Finally, Eq. (23) highlights the role information linkages and traders’ distance play in our model. Consider, first, a market without information linkages, i.e. $G = 0$. In this case, trader i information set collapses to the signal available at his location and, hence, $\bar{\rho}_n(k, 0) = \Omega_n / \Lambda_n$, which is obviously independent of k . By Eq. (23), then, the i -th trader expects any other trader to submit the same order, i.e. $E(x_{i+k,n} | F_{i,n}) = E(x_{i+j,n} | F_{i,n})$, for all k, j , which implies that,

$$E(x_{i+k,n} | F_{i,n}) = \frac{E\left(\sum_{j \neq i} x_{j,n} | F_{i,n}\right)}{M - 1}, \quad \text{for all } k \neq 0.$$

Therefore the i -th trader believes to be on the same side of the market whenever $\Delta_n > M^{-1}$, and to be so with respect to *every single* trader.

In contrast, Eq. (23) reveals that in the presence of information linkages, any trader i can expect to simultaneously trade in the same direction of the market but against some of his peers. Indeed, Eq. (23) implies that when $G > 0$, the i -th trader’s expects different orders from different traders, i.e. $E(x_{i+k,n} | F_{i,n}) \neq E(x_{i+j,n} | F_{i,n})$, for all k, j , $|j| \neq |k|$.

As we explained in Section 1.2, the correlation between the traders’ information endowments decreases with their relative distance [see Eqs. (6)-(7)], which means that each trader agrees more with his neighbors and less with his distant peers. Therefore, if the i -th trader is on the same side of the market and, for example, believes the asset is undervalued, he *has* to trade in the same direction of his close neighbors, although he *might* trade against peers who are sufficiently far apart. In Eq. (23), the term labeled “Market side direction” captures the first effect (trading with the peers), and the term labeled “Information linkages direction” captures the second effect (trading against the peers). In practice, whether two traders are on the same side of the market depends on their relative distance and the number of information linkages, as we shall show in the next section.

3 Predictions

This section analyzes the properties of the equilibrium predicted by the model. In Section 3.1, we set a benchmark, based on real market data, which we use to calibrate the model parameters; in Section 3.2, we look at the model predictions about correlated trading and volume among traders; in Section 3.3, we analyze how information linkages affect the dynamics of volume, liquidity, market efficiency and profits arising from informed trading.

3.1 The empirical benchmark and model calibration

We identify earnings announcements as natural proxies for the end of the trading period in our theoretical model. Accordingly, we retrieve report dates of quarterly corporate earnings for all AMEX and NYSE stocks during the fiscal year 2006, using the COMPUSTAT database. Our initial sample comprises 10077 report dates for 2723 stocks. We look at the behavior of volume and price impacts during two trading weeks before the announcement, as in other empirical investigations of this sort [e.g., Chae (2005)]. This choice implies a value of $N = 10$ trading days, which is the trading period we shall use to produce the model predictions in Sections 3.2 and 3.3. Data on prices and trading volumes for the eleven days n days prior to announcements are obtained from the CRSP database. They restrict our initial COMPUSTAT sample to 9088 report dates for 2611 stocks, once we exclude data with zero volume in at least one trading period.

Let the observed prices and volumes prior to announcement j be denoted as p_{nj} and v_{nj} , for $n = 1, \dots, 10$. We measure the ten price impacts occurring prior to each announcement j as the ratios of the absolute price change divided by trading volume, $\lambda_{nj} = |\Delta p_{nj}|/v_{nj}$. We rely on a two-way sorting procedure to figure out the volume and price impact patterns for a typical announcement in our sample. For each announcement j , we compute the average volume, $\bar{v}_j = 10^{-1} \sum_{n=1}^{10} v_{nj}$, and the average price impact, $\bar{\lambda}_j = 10^{-1} \sum_{n=1}^{10} \lambda_{nj}$, occurring over the trading period. Then, we select the announcements preceded by (i) average volumes \bar{v}_j within the 14.645-th and 85.355-th percentiles, and (ii) average price impacts $\bar{\lambda}_j$ within the very same previous percentiles. This sorting delivers precisely the middle 50% range of our initial universe of firms, along both volume and price impact dimensions.

To appraise the behavior of volume and price impacts for a typical stock in our sample, we take cross sectional averages relating to all firms surviving our two-way sort, and normalize them to their value in the first trading round. The second and third columns of Table 1 (labeled

“volume” and “lambda”) report these statistics for each trading day, along with cross sectional standard errors. Volume and price impacts display a clear pattern: volume increases slowly over the trading period and peaks up in the last trading round; price impacts, instead, decrease over the trading period, overall.

We further dissect our initial COMPUSTAT sample by the average dollar volume occurring prior to each announcement j , $\bar{pv}_j = 10^{-1} \sum_{n=1}^{10} (p_{nj}v_{nj})$. First, we split the sample into three sorts: large, medium and small firms, which correspond to average dollar volumes \bar{pv}_j above the 80-th percentile, between the 40-th and 60-th percentiles, and below the 20-th percentile. For each of these sub-samples, we apply the two-way sorting procedure described earlier, exclude data with zero volume in at least one trading period, and retrieve the volume and price impacts for a typical large, medium and small firm in our sample. We end up with three sub-samples that include 921 large firms, 913 medium firms, and 693 small firms.

Columns 4 through 9 in Table 1 report summary statistics for the average normalized volume and price impacts, along with cross sectional standard errors, on each of the three sub-samples. Volume is U-shaped and increases towards the end of the trading period, in all sub-samples. Furthermore, price impacts decrease with the trading round in each sub-sample, on average, especially at the level of small and large firms.

Our findings are similar to previous results obtained with US data: for example, Krinsky and Lee (1996) report that at a high frequency level, volume increases prior to earnings announcements; more recently, Chae (2005) uncovers a pattern for volume quite similar to ours, using daily data. As for the price impacts, it is well-known that adverse selection costs widen just around an earnings announcement. Our data display a similar feature. For example, at the announcement date, the cross sectional average price impact for large firms is 1.0604, with a standard error of 0.0259. In the remainder, however, we do not include observations for the announcement date, as we wish to make the trading period end exactly before the announcements, in the spirit of the model. Figure 4 displays the average volume and price impacts for large firms, along with cross sectional 95% confidence bands. We use data related to these firms to calibrate two model parameters: the variance of the asset value, σ_f^2 , and the variance of the liquidity trades, σ_u^2 .

Table 1: Cross sectional properties of volume and price impacts

This table displays statistics for daily trading volume and price impacts prior to corporate earnings announcements during the fiscal year 2006. Earnings announcement dates are retrieved from COMPUSTAT. Trading volume and price impacts for the ten trading days preceding the announcements are computed using the CRSP database. The whole sample is obtained through a two-way sort that includes firms in the middle 50% range of our initial collection of firms, along two dimensions: average volume and average price impacts over the trading period leading to announcements. Columns two and three report cross sectional averages of volume and price impacts (“lambda”) for the whole sample (normalized to their value in the first trading round), with cross sectional standard errors in parenthesis. Columns four to nine report the same statistics for large, medium and small firms, sorted by average dollar volume over the trading period, excluding data with zero volume in at least one trading period. Large, medium and small firms are those in the middle 50% range (along volume and price impact dimensions) of firms above the 80-th percentile, between the 40-th and 60-th percentiles, and below the 20-th percentile for average dollar volume. The last two columns report the model predictions about volume and price impacts for $M = 7$, $\sigma_f^2 = 1$, $\sigma_u^2 = 1.25 \cdot 10^{-1}$, $\rho = -0.15$ and $G = 2$.

trading round	Whole sample [4544 firms]		Large [921 firms]		Medium [913 firms]		Small [693 firms]		Model	
	volume	lambda	volume	lambda	volume	lambda	volume	lambda	volume	lambda
1	1	1	1	1	1	1	1	1	0.881	0.989
2	0.996 (0.017)	0.949 (0.021)	1.012 (0.018)	1.024 (0.028)	0.957 (0.015)	0.938 (0.026)	0.963 (0.039)	0.900 (0.048)	0.871	0.882
3	0.967 (0.016)	0.946 (0.020)	0.995 (0.019)	0.986 (0.026)	0.927 (0.015)	0.984 (0.029)	0.904 (0.043)	0.883 (0.046)	0.863	0.801
4	0.972 (0.016)	0.979 (0.022)	0.982 (0.019)	0.966 (0.027)	0.918 (0.015)	0.978 (0.027)	0.880 (0.033)	0.800 (0.039)	0.861	0.741
5	0.997 (0.016)	0.927 (0.019)	1.004 (0.019)	0.985 (0.026)	0.951 (0.017)	0.995 (0.027)	0.858 (0.032)	0.835 (0.045)	0.861	0.694
6	0.997 (0.016)	0.940 (0.021)	1.024 (0.019)	0.998 (0.027)	0.948 (0.014)	0.982 (0.026)	0.880 (0.033)	0.838 (0.043)	0.862	0.659
7	0.993 (0.016)	0.927 (0.019)	1.037 (0.019)	0.975 (0.026)	0.961 (0.015)	1.008 (0.028)	0.892 (0.036)	0.866 (0.053)	0.869	0.636
8	0.990 (0.017)	0.938 (0.021)	1.043 (0.021)	0.987 (0.026)	0.962 (0.015)	0.982 (0.027)	0.900 (0.037)	0.798 (0.043)	0.883	0.625
9	1.017 (0.016)	0.926 (0.020)	1.102 (0.024)	0.929 (0.026)	0.993 (0.016)	0.938 (0.026)	0.858 (0.031)	0.839 (0.043)	0.912	0.632
10	1.159 (0.018)	0.919 (0.021)	1.234 (0.022)	0.863 (0.023)	1.112 (0.019)	0.931 (0.026)	0.979 (0.042)	0.704 (0.032)	0.991	0.661

To calibrate σ_f^2 , we use the COMPUSTAT database and collect yearly earning-per-share (EPS) series for the stocks we previously sorted as large. We are able to retrieve 293 complete time series for the period 1995 to 2005. For each stock j , we measure its year t EPS growth as the first difference in yearly EPS, and compute the annualized standard deviation for each time series of EPS growth. The cross sectional average standard deviation is 0.902, with 95% confidence bands equal to 0.802 and 1.011. We then set $\sigma_f^2 = 1$ in our theoretical model.

Next, we calibrate the variance of the liquidity trades, σ_u^2 , by matching the volume and price impacts predicted by the model to their empirical counterparts for large firms. This step requires solving numerically the model. Therefore, we need to specify how the information is disseminated and how many traders are active in the market. As for the first issue, we take the asset value, f , to equal the sum of all the signals available to the market,

$$f = \sum_{i=1}^M s_{i,0}. \quad (24)$$

Eq. (24) implies that the covariance between any signal and the asset value equals $c_0 = \Lambda_0 + (M - 1)\Omega_0$. Accordingly, the parameter θ in the full information liquidation value [see Eq. (3)] equals the number of traders, M , which we set equal to 7, based on the empirical evidence in Brennan and Subrahmanyam (1995) and Chordia, Huh and Subrahmanyam (2008).² Note, Eq. (24) also implies that with $\sigma_f^2 = 1$, the correlation between any two signals available at two distinct locations equals,

$$\rho = \frac{1}{M(M-1)} \left(\frac{1}{\Lambda_0} - M \right). \quad (25)$$

That is, the unconditional variance of each signal, Λ_0 , is known, once we specify ρ .

We are now left with calibrating ρ , the number of information linkages each trader has with his peers, $2G$, and, σ_u^2 . We solve the model on a grid of (σ_u^2, ρ) , for every $G \in \{0, 1, 2\}$, and compute a statistics summarizing the distance between the model predictions and their empirical counterparts. The statistics is the sum of two mean squared errors (MSE) calculated over the trading period: the MSE between the model implied volume and the cross sectional average volume for large firms (in percentage); and the MSE between the model implied price impacts and the cross sectional average price impacts for large firms (in percentage). We find that

²To check the robustness of our results, we experimented all combinations of $M \in \{7, 51, 101\}$ and $N \in \{10, 20, 40\}$, and obtained results qualitatively very similar to those we report below. The only exception regards the behavior of λ_n for large values of N , which becomes U-shaped for any G and M , as we explain in Section 3.3.

$(\sigma_u^2, \rho, G) = (1.25 \cdot 10^{-1}, -0.15, 2)$ minimizes this statistics. The last two columns in Table 1 report the model predictions arising in this case. The model reproduces the salient features of data: volume is U-shaped and increases toward the end of the trading period; instead, price impacts decrease, overall.

In the next sections, we discuss the model predictions when the signals correlation is negative, with $\rho = -15\%$, and when it is both zero and positive, with $\rho = 10\%$. These correlation values are those that make our calibrated model fall within the empirically relevant range, although as noted, the last two columns in Table 1 minimize the distance between the model and data. Finally, to further our understanding of how information linkages impart on market variables, we analyze the cases in which $G = 0$ (no information linkages), $G = 1$ (information linkages) and $G = 2$ (many information linkages and “double overlap”).

3.2 Correlated trading and volume

Figure 5 depicts the correlation between the orders of each trader with his peers, close and distant,

$$\text{corr}(x_{i,n}, x_{i+k,n} | F_{n-1}) = \bar{\rho}_{n-1}(k, G),$$

where $\bar{\rho}_n(k, G)$ is the correlation between the average signals of any two traders i and $i + k$ (see Proposition 2).

Figure 5 shows that for a given number of information linkages, $\bar{\rho}_n$ increases with the correlation ρ , for all the batch auctions n and the peers’ distance k . This feature merely reflects the higher correlation between the signals available at each trader’s location. Moreover, for given k and G , the trade correlation decreases over time. This pattern is related to the market maker’s learning process. As in Foster and Viswanathan (1996), over time, the price conveys more and more information about the average signal, rather than about the individual signals available to traders. Similarly, in our model, the market maker learns more about *the average of the average signals*, \bar{s} [see Eq. (3)] than the private signals traders have about the asset value, $\bar{s}_{i,0}$ [see Eq. (1)]. This feature of the learning process implies that the market maker’s resolution of the residual uncertainty, $\sigma_{f,n}^2$ in Eq. (18), takes place somewhat faster than the resolution of the traders’ residual uncertainty, $\varsigma_{f,n}^2$ in Eq. (19). By Eq. (21), then, the tightness of beliefs, Δ_n , decreases over time. At some point, Δ_n becomes so small that the market maker sets the price to a value quite close to the traders’ average opinion of the asset value (see Figure 6), in which case the

traders play a waiting game, as explained in Section 2.2. Inevitably, then, the traders expect to be on opposite sides of the market, as Proposition 2 suggests.

Although each trader expects to trade against the market, our model predicts that the correlation of trades varies across the market. In particular, information linkages induce each trader to trade in the same direction as his close neighbors and in the opposite direction to his distant peers. This property, which is formally due to Eq. (23) in Proposition 2, is well-visible from Figure 5. This figure shows that the trade correlation, $\bar{\rho}_n$, increases with G , especially during the first trading rounds. Indeed, the presence of information linkages leads to a higher duration of the rat race, even when the correlation ρ is negative. This is illustrated by Figure 6, which depicts the dynamics of Δ_n for a large spectrum of values of ρ ($\rho = -0.15, 0$ and 0.90). As a result, for close traders, $\bar{\rho}_n$ remains positive for the entire trading period, reflecting the boost in the correlation between the information endowments and the rat race. For distant traders, however, information linkages are too weak to counterbalance the effects related to the market maker’s learning mechanism, by which each trader expects to trade against the market. In both cases, the trade correlation decreases over time, a critical point we shall discuss below.

Finally, note that compared with the benchmark case of absence of information linkages ($G = 0$), it takes fewer batch auctions for the trade correlation between distant traders to become negative. Moreover, for distant traders, the trade correlation falls dramatically in the presence of information linkages. For example, when $\rho = 10\%$, the trade correlation among distant traders at the end of the trading period is only -11% , in the absence of information linkages ($G = 0$), while it reaches a value of -70% in the presence of information linkages ($G = 1$), and a value of -30% in the presence of many information linkages ($G = 2$). Note, the trade correlation is higher for $G = 2$ than for $G = 1$. This property arises because the presence of many information linkages leads to the “double overlap” property described in Section 1.2 (see Figure 3). Such an information overlap induces a quite substantial boost in the correlation between the information endowments of all traders, which then translates into an increased trade correlation, even for distant traders. Still, the trade correlation for distant traders is significantly lower with $G = 2$ than in the absence of information linkages.

As outlined in the Introduction, these cross sectional properties of trading behavior are consistent with previous findings in the empirical literature. For example, Feng and Seasholes (2004) report that in the Chinese stock market, geographically close investors trade similarly and distant investors submit negatively correlated orders. Bae, Yamada and Ito (2006) find quite similar re-

sults in the Japanese stock market: trades initiated by foreign investors are negatively correlated with trades initiated by local investors. Along similar lines, Dorn, Huberman and Sengmueller (2008, p. 902) argue that correlated trading can be related to “the likelihood that the traders are exposed to the same signals (which are interpreted similarly).” The model predictions about correlated trading suggest additional testable patterns for correlated trading. In our model, trade correlation tends to decrease as the liquidation date approaches. This property holds for all traders, close and distant. By way of contrast, consider the predictions about correlated trading induced by herding behavior. Herding relies upon each trader “imitating” the peers’ previous trades, and would lead to a trade correlation increasing over time. Therefore, the model predictions about trade correlation help discriminate empirically between our information linkages mechanism and herding behavior.

How does volume correlate among traders? Figure 7 depicts the correlation among the volumes generated by the single traders (for brevity, the volume correlation), over the trading period.³ The dynamics of volume correlation are related to, albeit distinct from, those of the trade correlation. Absent any information linkages, individual trades are nearly uncorrelated when the correlation ρ is negative or zero (see Figure 5), which translates into a similar property for the volumes generated by the traders.

In the presence of information linkages, we explained that the correlation among individual trades is positive for close neighbors, while it decays to negative values for distant traders (see Figure 5). Moreover, this decay is more pronounced for $G = 2$ than for $G = 1$. For close neighbors, the volume correlation features a pattern quite similar to that of the trade correlation. However, as the distance between traders increases, the volume correlation decreases, overall, and becomes U-shaped: few information linkages ($G = 1$) now lead to stronger volume correlation than more information linkages ($G = 2$), except for the very first trading rounds. This is because for $G = 1$, the increased correlation of individual trades induced by the presence of information linkages does not offset the correlation decay arising from the market maker’s learning process. Individual trades are strongly negatively correlated for most of the trading rounds and volume correlation peaks up. When $G = 2$, the increase in the trade correlation resulting from the information linkages dominates the decay induced from the market maker’s learning process. As

³Informed volume is estimated as the conditional standard deviation of informed trades, as in Admati and Pfleiderer (1988). We compute the correlation between the individual traders’ volume through simulations, as we do not have a closed-form solution for it.

a result, individual trades and volume are weakly correlated for most of the batch auctions.

Existing empirical studies rarely tackle the issue of correlated volumes. The array of predictions produced by our model suggests a number of relations between correlated volume and distance that warrant empirical scrutiny. To summarize, volume correlation decreases with the traders' distance. Moreover, the correlation of volumes for close traders does not display the same pattern as that for distant traders and, additionally, depends on the number of information linkages in the market, as we explained. Finally, volume correlations are quite flat over the trading period, in the absence of information linkages. The last observation, alone, would be sufficient to test for the very presence of information linkages in a given market.

3.3 Volume, liquidity, efficiency and traders' profits

How do information linkages affect total volume, liquidity, efficiency and traders' profits? On the one hand, a trader with many information linkages is able to estimate more precisely the asset value, since he observes more signals on top of the one available at his location. On the other hand, the same linkages lead a given trader to lose part of his monopolistic information power, since other traders observe his signal. This second effect translates into an incentive for every trader to anticipate his peers: trading aggressiveness increases and so does volume. This clearly emerges from Figure 8, which plots the total volume, estimated as the conditional standard deviation of the trades initiated by the M traders, the market maker, and the liquidity traders. A direct consequence of such a trading behavior is that by impounding more information into their orders, traders reduce the market maker's residual variance of the asset value, $\sigma_{f,n}^2$, thus boosting market efficiency, as Figure 9 reveals for a large spectrum of values of ρ ($\rho = -0.15, 0$ and 0.90). This figure also shows that the rate of the price discovery process increases as the number of information linkages gets larger. In other words, information linkages are sources of enhanced market efficiency. Moreover, the improvement in market efficiency is more pronounced when the correlation ρ is low. These findings are directly related to the dynamics of the tightness of beliefs depicted in Figure 6: information linkages make competition among informed traders more fierce. As a result, traders impound more information into their orders, thus reducing the market maker's residual variance and boosting market efficiency.

The price responsiveness to the order flow, λ_n , is displayed in Figure 10. When the correlation ρ is positive, the combined effect of higher aggressiveness and better price discovery is the

reduction of the adverse selection faced by the market maker. As a result, the price impact of the order flow, λ_n , lowers, which makes the market more liquid, overall. When the correlation ρ is negative, the previous conclusions are reversed: although market efficiency still improves, liquidity deteriorates in the presence of information linkages, especially during the first batch auctions. This outcome is the result of two forces. On the one hand, information linkages enhance the traders' estimates of the asset value, as usual. On the other hand, when $\rho < 0$, the tightness of beliefs between the traders and the market maker, Δ_n , is considerably low (see Figure 6), which leads the traders to play a waiting game for the entire trading period (for $G = 1$) or for nearly all the batch auctions (for $G = 2$). Thus, the market maker faces relatively more informed traders postponing their trades for most of the time and, hence, he raises the price responsiveness. Naturally, these same effects explain the dynamic behavior of total volume depicted in Figure 8.

Our analysis shows that information linkages among traders may lead to a variety of empirically plausible patterns for total volume and liquidity occurring prior to a corporate event. As we explained in Section 3.1, the pattern that best fits our sample is that with many information linkages, $G = 2$, and a negative correlation among the signals available at each trader's location. This combination results in increasing volume and decreasing price impacts over the trading period, as in our data. At the same time, the model is able to generate additional predictions that might possibly apply to alternative samples or sorting criteria, as we discussed. Moreover, our model can explain the high frequency behavior of liquidity detected in the empirical literature. For example, Krinsky and Lee (1996) show that with half-hour trading intervals, adverse selection costs increase around the announcement date, a feature our sample displays as well, once we include data for the announcement dates (see Section 3.1). Results not reported here confirm that our model, like others [e.g., Foster and Viswanathan (1996) and Back, Cao and Willard (2000)] does indeed predict a U-shaped pattern for the price impacts, once we increase the frequency of trading from daily to, say, hourly.

The last task of this section is to analyze the effects information linkages produce on the expected profits from informed trading. Figure 11 illustrates two main results. First, information linkages damage the expected profits when the correlation ρ is positive, but enhance them when the correlation is negative. Second, for any fixed G , expected profits are non-monotonic in the correlation ρ .

What are the origins of these findings? Changes in the correlation ρ generate two effects.

First, as ρ increases, the signals available at each location become informationally closer to one other, thus reducing each trader's monopolistic power. Second, as Eq. (25) reveals, an increase in ρ obviously implies a drop in Λ_0 , which makes the traders' estimates of the asset value more precise. In the absence of information linkages, the losses in the monopolistic power dominate over the precision gains, when ρ is high or negative. When, instead, ρ is positive and sufficiently low, the losses in the monopolistic power are more than offset by the precision gains. Figure 11 reveals that these properties are preserved once we introduce information linkages, although the threshold value of ρ below which information linkages enhance the expected profits becomes simply zero.

In summary, the first model prediction in this section is that information linkages boost both market efficiency and aggregate volume. As we explained in the Introduction, the extant empirical literature is supportive of this prediction [e.g., Dorn, Huberman and Sengmueller (2008), Hong, Lim and Stein (2000)], although there does not exist yet a systematic analysis of how such a piece of empirical evidence is related to additional properties of our model. For example, we find that information linkages can affect liquidity in a quite rich manner: they damage it, when the signals at the traders' location are negatively correlated, and they improve it, otherwise. Interestingly, profits from informed trading are affected by the presence of information connections in the exact opposite way: traders' expected profits may now increase or decrease, according to whether the linkages convey negatively or positively correlated pieces of information.

To date, there are no empirical studies aiming at deciphering how aggregate volume, liquidity and trading profits, all relate to the information connectedness among traders. Brennan and Subrahmanyam (1995) find a positive relation between liquidity and analyst coverage (a proxy for information connectedness), and Hau (2001) provides some evidence that information connections can lower traders' profits. As discussed, our model is consistent with these empirical findings, if the information linkages available to traders convey positively correlated information. At the same time, the evidence in these papers obviously relates to different samples and does not tackle all the diverse implications of our model. Our findings in this section, in conjunction with those for the cross sectional properties of informed trading and volume (discussed in Section 3.2), lead instead to a rich and well-defined battery of testable predictions, which can systematically be dealt with in future empirical research.

4 Discussion of related work

The literature has produced prominent instances of rational explanations for heterogenous trading behavior. Brennan and Cao (1997) develop a rational expectations model (not one with strategic agents) in which a domestic asset is also traded by foreign investors. In their model, local investors are able to estimate the asset value with higher precision, compared to foreign investors. Thus, a good piece of public news makes those traders with more precise information (local) trade less aggressively than those with less precise information (foreign), which might lead to negative correlation between trades initiated by distant investors. Our paper offers an alternative mechanism, in which heterogenous trade correlations occur even if the precision of the private signals is the same for all the traders.

DeMarzo, Kaniel and Kremer (2004) consider a model in which agents belong to different communities, compete for local resources, and face capital market imperfections. These imperfections lead to restricted stock market participation, which makes the marginal investors' utility decreasing in the community wealth, provided risk-aversion is sufficiently high. Then, correlated portfolio choices within a community arise because the marginal investors price assets taking into account their "community factors." Our explanation of correlated behavior is also based on the broad idea that investors belong to different communities (geographic, demographic, etc.). At the same time, our model is built upon the informational distance between traders, and does not rely on a consumption-based rationale.

Our paper also relates to the literature on social learning and information networks. (See Goyal (2007) and Jackson (2008) for surveys of the literature on the broader area of the economics of networks.) Early papers in this literature are those of Ellison and Fudenberg (1993, 1995), who study learning environments in which agents base their decisions on their peers past experience. A more recent paper is that by DeMarzo, Vayanos and Zwiebel (2003). The authors study the process of opinion formation in social networks, and propose persuasion bias as a boundedly rational heuristic for processing information. As explained in Section 1.1, a new contribution to the role of communication among strategic competitors is that of Stein (2007), who develops a model in which competitors can find it fruitful to engage in truthful conversations, provided these conversations boost the quality of their initial estimates about an asset payoff. While our contribution shares similarities with all these papers in so far as we consider information networks, our work explicitly analyzes the role played by these networks in the context of financial markets.

In independent work, Ozsoylev (2006) develops a model in which every investor observes the expectations of his neighbors. (See, also, Malinova and Smith (2006) for a related rational

expectation model.) His information network is similar in spirit to our local information linkages mechanism. Indeed, our model can be interpreted as one in which every trader observes the signals of the neighbors with some error, as we explained in Section 1.1. However, our work differs from Ozsoylev’s for two reasons. First, Ozsoylev considers a setup in which the investors do not enjoy market power. The assumption of no-market power allows the author to investigate asymmetric information networks. In our model, traders do enjoy monopolistic market power and, hence, need to forecast the forecasts of others. To simplify this dimensionality issue, we consider a symmetric network. The second difference with Ozsoylev’s model is that ours is dynamic. Ozsoylev’s work and ours therefore complement each other.

5 Conclusion

Why do individual investors exhibit a correlated trading behavior? One explanation put forward in the empirical literature is that proximity affects portfolio choice, which makes informationally close investors trade in the same direction, and informationally distant investors exhibit quite distinct portfolio choices. Although this argument is simple and appealing, it is not clear whether it can be made consistent with rational behavior. For example, it may be argued, the mere observation of the equilibrium price and additional public signals might induce investors to *all* trade in the same direction. Nor are the implications of an heterogeneous trading behavior clear, both empirically and theoretically. For example, what are the implications of correlating trading on market efficiency, liquidity, volume, or profitability from informed trading? Under which conditions do correlated trading and volume arise in a market and, eventually, disappear from it? How differently should we expect informationally distant investors to trade?

This paper contributes to providing answers to these questions. We consider a market in which traders belong to overlapping information networks, connected through what we term information linkages. We show that these linkages exert a quite substantial impact on trading strategies as well as market variables. First, they affect the cross sectional properties of correlated trading: our model predicts that “neighbor” trades are positively correlated and “distant” trades are negatively correlated. While this prediction is consistent with the extant empirical literature, the model also leads to one additional testable implication that calls for further empirical scrutiny: the correlation among trades (close and distant) decreases over the trading period. This property helps discriminate our rationale for correlated trading from others. For example, a theory based

on herding would imply an increasing correlation among close trades, due to each trader imitating the previous moves made by his peers.

The second testable implication of the model relates to the cross section of volumes initiated by informed traders. In the presence of linkages, the correlation of volumes depends on the traders' location as well: it is always positive, and larger for close traders than for distant traders. Moreover, over the trading period, this correlation is decreasing for close traders, and non-monotonic for distant traders. Instead, we show that in the absence of information linkages, the correlation among volumes is quite flat over the trading period, and close to zero. This feature is particularly useful in testing for the existence of information networks among traders.

The third set of implications of our model relates to aggregate variables. We show that information linkages raise volume and price informativeness, and significantly affect liquidity conditions and gains from informed trading. The model predicts that the traders' expected profits increase (or decrease), and liquidity worsens (or improves), according to whether the information linkages convey quite distinct (or positively correlated) signals about the fundamentals. These properties are broadly consistent with previous empirical findings, although their definite empirical validation would need to be performed in conjunction with a systematic assessment of the cross sectional properties our theoretical work suggests for informed trading and volume.

Appendix

A. Preliminary results

Derivation of Eqs. (5)-(8). To derive Eq. (5), we use the definition of the average signal in Eq. (1). A simple computation leaves:

$$\text{var}(\bar{s}_{i,0}) = \frac{\hat{G}\Lambda_0 + 2G\hat{G}\Omega_0}{\hat{G}^2},$$

or equivalently (5). Next, we derive Eqs. (6) and (7). These equations correspond to two cases: (i) $2G \leq (M-1)/2$, and (ii) $2G \geq (M-1)/2$. We study these two cases separately. We introduce the following piece of notation

$$\tilde{\mathbf{s}}_{i,0} = \{s_{i-G,0}, \dots, s_{i,0}, \dots, s_{i+G,0}\},$$

to denote the set of signals each trader i has access to, as a result of the presence of information linkages.

Case (i) ($2G \leq (M-1)/2$). Consider traders i and $j = i+k$, $k \neq 0$. We have: $s_{i+k,0} \notin \tilde{\mathbf{s}}_{i,0}$ for all $|k| > G$. Therefore

$$\bar{\Omega}_0(k, G) = \text{cov}(\bar{s}_{i,0}, \bar{s}_{i+k,0}) = \hat{G}^{-2} \sum_{l=-G}^G \sum_{m=-G}^G \text{cov}(s_{i+l,0}, s_{i+k+m,0}) = \Omega_0,$$

for all $|k| > 2G$, which is the second line in Eq. (6). If, instead, $|k| \leq 2G$, $s_{i+k,0} \in \tilde{\mathbf{s}}_{i,0}$ for all $|k| \leq G$ and $\tilde{\mathbf{s}}_{i+k,0} \cap \tilde{\mathbf{s}}_{i,0} \neq \{\emptyset\}$. In particular, trader i shares $(2G+1-k)$ signals with trader $i+k$. Each of these signals contributes for $(\Lambda_0 + 2G\Omega_0)/(2G+1)^2$ to $\bar{\Omega}_0(k, G)$. Shared signals thus contribute for

$$\frac{\Lambda_0 + 2G\Omega_0}{(2G+1)^2} \cdot (2G+1-k)$$

to $\bar{\Omega}_0(k, G)$. The remaining (not shared) k signals contribute for

$$\frac{(2G+1)\Omega_0}{(2G+1)^2} \cdot k$$

to $\bar{\Omega}_0(k, G)$. Therefore,

$$\bar{\Omega}_0(k, G) = \frac{(\Lambda_0 + 2G\Omega_0)(\hat{G} - k) + \hat{G}\Omega_0 k}{\hat{G}^2}.$$

Grouping terms in the previous expression yields the first line in Eq. (6).

Case (ii) ($2G \geq (M-1)/2$). This case differs from the previous one due to the double overlap discussed in Section 1.2 (see Figure 3). In this case, the number of signals shared by traders i and $i+k$ is:

$$L(k, G) = 2G+1-k+n(k, G), \quad k = 1, \dots, \frac{M-1}{2}. \quad (\text{A1})$$

The term $n(k, G)$ arises because traders located on the trader i 's right hand side semicircle might be sharing signals with traders located between $i+1$ and $i+(M-1)/2$ on the left hand side semicircle (see Figure 3); and obviously the i -th trader shares signals with traders located between $i-1$ and $i-G$ as

well. The double overlap occurs if and only if trader $i + k$ on the left hand side semicircle and trader $i - \ell$ with $\ell \in [1, \frac{M-1}{2}]$ on the right hand side semicircle are such that ℓ and k satisfy:

$$\begin{cases} \frac{M-1}{2} - (\ell - 1) + \frac{M-1}{2} - k \leq G \\ G \geq \ell \geq 1 \\ \frac{M-1}{2} \geq k \geq 1 \end{cases}$$

The first inequality in the previous restrictions requires trader $i - \ell$ to share his signal with trader $i + k$. The second and third constraints restrict trader $i - \ell$ to be on the right hand side semicircle and trader $i + k$ to be on the left hand side semicircle relative to trader i . Thus, for fixed k ($1 \leq k \leq (M-1)/2$), the double overlap occurs if and only if

$$G \geq \ell \geq M - G - k, \quad k = 1, \dots, \frac{M-1}{2},$$

and $\ell \geq 1$. Clearly, $\min_k (M - G - k) = \frac{M-1}{2} - G + 1 \geq 1$. Hence, the constraint that $\ell \geq 1$ is redundant. By the previous inequalities, it follows that:

$$n(k, G) = \max \{G - (M - G - k) + 1, 0\}.$$

By replacing this result into Eq. (A1) leaves:

$$L(k, G) = \begin{cases} 4G + 1 - (M - 1), & \frac{M-1}{2} \geq k \geq 2 \left(\frac{M-1}{2} - G\right) \\ 2G + 1 - k, & 1 \leq k \leq 2 \left(\frac{M-1}{2} - G\right) \end{cases}$$

For all $k \in [1, 2 \left(\frac{M-1}{2} - G\right)]$, $\bar{\Omega}_0(k, G)$ is thus exactly as in case (i) for $k \in [1, 2G]$, and the first line of Eq. (7) follows. For all $k \in [2 \left(\frac{M-1}{2} - G\right), \frac{M-1}{2}]$, tedious but straightforward computations lead to the second line of Eq. (7).

Finally, we demonstrate that Eq. (8) holds true. As usual, we consider the two cases in which $2G \geq (M-1)/2$. If $0 \leq 2G \leq (M-1)/2$, there are $[M - (4G + 1)]$ traders $i + k$ such that $\tilde{\mathbf{s}}_{i+k,0} \cap \tilde{\mathbf{s}}_{i,0} = \{\emptyset\}$. In correspondence of these indexes, $\text{cov}(\bar{s}_{i+k,0}, \bar{s}_{i,0}) = \Omega_0$. Therefore,

$$\bar{\Gamma}_0(G) = 2 \sum_{k=1}^{2G} \bar{\Omega}_0(k, G) + [M - (4G + 1)] \Omega_0.$$

The $2G$ covariances in the summation can be computed through the first line in (6). Eq. (8) follows by the expression of $\bar{\Lambda}_0(G)$ in Eq. (5). Next, consider the case $(M-1)/2 \leq 2G \leq M-1$. We have:

$$\bar{\Gamma}_0(G) = 2 \left\{ \sum_{k=1}^{M-1-2G} \bar{\Omega}_0(k, G) + \left(2G - \frac{M-1}{2}\right) \left[2\bar{\Lambda}_0(G) - \frac{M(\Lambda_0 - \Omega_0)}{\hat{G}^2} - \Omega_0\right] \right\}.$$

By plugging Eqs. (7) and (5) into the previous equation, we find that the expression of $\bar{\Gamma}_0(G)$ is the same as the one obtained in the case $0 \leq 2G \leq (M-1)/2$, and Eq. (8) follows. ■

$$E\left(\sum_{j \neq i} \bar{s}_{j,0} | \mathbf{s}_{i,0}\right) = E\left(\sum_{j \neq i} \bar{s}_{j,0} | \bar{\mathbf{s}}_{i,0}\right) = (M-1) \phi_1 \bar{s}_{i,0}, \quad (26)$$

Derivation of Eq. (26). Consider the projection of $s_{i+k,0}$ onto $\mathbf{s}_{i,0}$ for $|k| \leq G$. We have $s_{i+k,0} \in \mathbf{s}_{i,0}$ for $k = 0, \mp 1, \dots, \mp G$. Hence,

$$E(s_{i+k,0} | \mathbf{s}_{i,0}) = s_{i+k,0}, \quad k = 0, \mp 1, \dots, \mp G. \quad (A1)$$

Next, consider a signal $s_{i+k,0}$ a given trader i does not have access to, i.e. $s_{i+k,0}$, for $k = \mp(G+1), \dots, \mp \frac{M-1}{2}$. Let $\Psi_{0,G} = E(\mathbf{s}_{i,0} \mathbf{s}_{i,0}^\top)$ be the $\hat{G} \times \hat{G}$ variance-covariance matrix of the vector $\mathbf{s}_{i,0}$. $\Psi_{0,G}$ is a $\hat{G} \times \hat{G}$ submatrix extracted from Ψ_0 , and its inverse can be obtained with the same strategy of proof as in Foster and Viswanathan (1996) (p. 1479). Let $K = [(\Lambda_0 - \Omega_0)(\Lambda_0 + 2G\Omega_0)]^{-1}$. Then,

$$\Psi_{0,G}^{-1} = K \cdot \begin{bmatrix} \Lambda_0 + (2G-1)\Omega_0 & -\Omega_0 & \cdots & -\Omega_0 \\ & \Lambda_0 + (2G-1)\Omega_0 & & -\Omega_0 \\ & & \ddots & \vdots \\ & & & \Lambda_0 + (2G-1)\Omega_0 \end{bmatrix}.$$

For $|k| > G$, we have that $\text{cov}(s_{i+k,0}, \mathbf{s}_{i,0}) = \Omega_0 \mathbf{1}_{\hat{G}}$, and by the Projection Theorem,

$$\begin{aligned} E(s_{i+k,0} | \mathbf{s}_{i,0}) &= \Omega_0 \mathbf{1}^\top (\Psi_{0,G})^{-1} \mathbf{s}_{i,0} \\ &= \Omega_0 \frac{\Lambda_0 + (2G-1)\Omega_0 - 2G\Omega_0}{(\Lambda_0 - \Omega_0)(\Lambda_0 + 2G\Omega_0)} \mathbf{1}^\top \mathbf{s}_{i,0} \\ &= \frac{\hat{G}\Omega_0}{\Lambda_0 + 2G\Omega_0} \bar{s}_{i,0}, \quad k = 0, \mp 1, \dots, \mp G. \end{aligned} \quad (A2)$$

Therefore, gathering Eqs. (A1) and (A2),

$$E(s_{i+k,0} | \mathbf{s}_{i,0}) = \begin{cases} s_{i+k,0} & \text{for } k = 0, \mp 1, \dots, \mp G \\ \frac{\hat{G}\Omega_0}{\Lambda_0 + 2G\Omega_0} \bar{s}_{i,0} & \text{for } k = \mp(G+1), \dots, \mp \frac{M-1}{2} \end{cases}$$

We now compute, for a given trader i , his forecast of the sum of the remaining traders' average signals onto $\mathbf{s}_{i,0}$. We obviously have $\sum_{j=1}^M \bar{s}_{j,0} = \sum_{j=1}^M s_{j,0}$. Therefore,

$$\begin{aligned} E\left(\sum_{j \neq i} \bar{s}_{j,0} \middle| \mathbf{s}_{i,0}\right) &= \sum_{j=1}^M E(s_{j,0} | \mathbf{s}_{i,0}) - \bar{s}_{i,0} \\ &= \mathbf{1}_G^\top \mathbf{s}_{i,0} + \sum_{|k| > G} E(s_{i+k,0} | \mathbf{s}_{i,0}) - \bar{s}_{i,0} \\ &= (2G+1) \bar{s}_{i,0} + \frac{(M - (2G+1))\Omega_0}{\Lambda_0 + 2G\Omega_0} (2G+1) \bar{s}_{i,0} - \bar{s}_{i,0} \\ &= \frac{(M-1)(2G+1)\Omega_0 + 2G(\Lambda_0 - \Omega_0)}{\Lambda_0 + 2G\Omega_0} \bar{s}_{i,0}. \end{aligned}$$

Eq. (26) follows by rearranging terms in the last equality, and by the definition of ϕ_1 in the main text. ■

Definitions and notation. We make the following definitions:

- The i -th trader residual informational advantage (relative to the market maker) on (i) the signal available at his location, $s_{i,0}$, and (ii) his average signal, $\bar{s}_{i,0}$, are defined to be:

$$s_{i,n} = s_{i,0} - t_n, \quad \bar{s}_{i,n} = \bar{s}_{i,0} - t_n, \quad (\text{A2})$$

after n rounds of trading.

- The market maker's updates of the residual variances are:

$$\Lambda_n = \text{var}(s_{i,0} | F_n), \quad \Omega_n = \text{cov}(s_{i,0}, s_{j,0} | F_n) \quad (\text{A3})$$

and

$$\bar{\Lambda}_n(G) = \text{var}(\bar{s}_{i,0} | F_n), \quad \bar{\Omega}_n(k, G) = \text{cov}(\bar{s}_{i,0}, \bar{s}_{i+k,0} | F_n), \quad \bar{\Gamma}_n(G) = \sum_{k \neq i} \bar{\Omega}_n(k, G), \quad (\text{A4})$$

where Λ_n and Ω_n (resp. $\bar{\Lambda}_n(G)$ and $\bar{\Omega}_n(k, G)$) are the residual variance and covariance of the signals available at each trader's location (resp. of the average signals available to each trader). Note that by Eq. (5),

$$\bar{\Lambda}_n(G) = \frac{\Lambda_n + 2G\Omega_n}{\hat{G}}. \quad (\text{A5})$$

We shall repeatedly use the results recorded in the following three lemmas, which are easy generalizations of results given in Foster and Viswanathan (1996).

Lemma 1. *The relation between the market maker's updated estimate of the asset value, p_n , and the market maker's updated estimate of trader i average signal, t_n , is given by*

$$p_n = \theta t_n. \quad (\text{A6})$$

Proof of Lemma 1. By Semi-Strong market efficiency, $p_n = E(f | F_{M+1,n})$. By Eq. (3), \bar{s} is a sufficient statistic for $E(f | \mathbf{s}_0)$. Therefore,

$$\begin{aligned} p_n &= E(f | F_{M+1,n}) \\ &= E\{E[E(f | \mathbf{s}_0) | \bar{s}, F_{M+1,n}] | F_{M+1,n}\} \\ &= E[E(\theta \bar{s} | \bar{s}, F_{M+1,n}) | F_{M+1,n}] \\ &= \frac{\theta}{M} E\left(\sum_{i=1}^M \bar{s}_{i,0} \middle| F_{M+1,n}\right) \\ &= \theta t_n, \end{aligned}$$

where the last line follows by Eq. (12). ■

Lemma 2. *Let $\sigma_{f,n}^2$ and $\bar{\Gamma}_n(G)$ be as in Eqs. (18) and (A4). We have,*

$$\sigma_{f,n}^2 = \frac{\theta^2}{M} [\Lambda_n + (M-1)\Omega_n]. \quad (\text{A7})$$

Furthermore, the following recursions hold,

$$\Omega_{n-1} - \Omega_n = \Lambda_{n-1} - \Lambda_n; \quad (\text{A8-a})$$

$$\sigma_{f,n-1}^2 - \sigma_{f,n}^2 = \theta^2 (\Lambda_{n-1} - \Lambda_n); \quad (\text{A8-b})$$

$$\bar{\Lambda}_{n-1}(G) - \bar{\Lambda}_n(G) = \Lambda_{n-1} - \Lambda_n, \quad \text{all } G; \quad (\text{A8-c})$$

$$\bar{\Omega}_{n-1}(k, G) - \bar{\Omega}_n(k, G) = \Lambda_{n-1} - \Lambda_n, \quad \text{all } k, G; \quad (\text{A8-d})$$

$$\bar{\Gamma}_{n-1}(G) - \bar{\Gamma}_n(G) = (M-1)(\Lambda_{n-1} - \Lambda_n), \quad \text{all } G. \quad (\text{A8-e})$$

Proof of Lemma 2. First, we derive Eq. (A7). By Eq. (3), and the Law of Iterated Expectations, $E(\theta\bar{s}|F_{M+1,n}) = E[E(f|\mathbf{s}_0)|F_{M+1,n}] = E(f|F_{M+1,n}) = p_n$. Hence, by the first equation in (A3) and Eq. (A6),

$$\begin{aligned} \sigma_{f,n}^2 &= \text{var}(\theta\bar{s} - p_n | F_{M+1,n}) \\ &= E \left[\left(\frac{\theta}{M} \sum_{i=1}^M \bar{s}_{i,0} - \frac{\theta}{M} \sum_{i=1}^M t_{i,n} \right)^2 \middle| F_{M+1,n} \right] \\ &= \frac{\theta^2}{M^2} E \left[\left(\sum_{i=1}^M \bar{s}_{i,n} \right)^2 \middle| F_{M+1,n} \right] \\ &= \frac{\theta^2}{M^2} E \left[\left(\sum_{i=1}^M s_{i,n} \right)^2 \middle| F_{M+1,n} \right] \\ &= \frac{\theta^2}{M} [\Lambda_n + (M-1)\Omega_n]. \end{aligned}$$

Next, we derive Eqs. (A8-a)-(A8-e). Let

$$\tau_n = \text{cov}(s_{i,n-1}, y_n | F_{M+1,n-1}). \quad (\text{A9})$$

Let $\Psi_n = E \left([s_{1,0} - t_n, \dots, s_{M,0} - t_n]^\top [s_{1,0} - t_n, \dots, s_{M,0} - t_n] \middle| F_{M+1,n} \right)$. By the Projection Theorem,

$$\Psi_n = \Psi_{n-1} - \frac{\tau_n^2}{\text{var}(y_n | F_{M+1,n-1})} \mathbf{1}\mathbf{1}^\top,$$

which gives the recursions:

$$\Lambda_n = \Lambda_{n-1} - \frac{\tau_n^2}{\text{var}(y_n | F_{M+1,n-1})}; \quad (\text{A10})$$

$$\Omega_n = \Omega_{n-1} - \frac{\tau_n^2}{\text{var}(y_n | F_{M+1,n-1})};$$

or equivalently (A8-a). Taking one lag in Eq. (A7) yields:

$$\sigma_{f,n-1}^2 = \frac{\theta^2}{M} [\Lambda_{n-1} + (M-1)\Omega_{n-1}],$$

giving the recursion:

$$\sigma_{f,n-1}^2 - \sigma_{f,n}^2 = \frac{\theta^2}{M} [\Lambda_{n-1} - \Lambda_n + (M-1)(\Omega_{n-1} - \Omega_n)] = \theta^2 (\Lambda_{n-1} - \Lambda_n),$$

where the last equality follows by Eq. (A8-a). Eq. (A8-c) follows by Eqs. (A8-a) and (A5), and by simple computations. We now provide the update for $\bar{\Omega}_n(k, G)$, thus completing the specification of the variance-covariance matrix $\bar{\Psi}_n(G) = E\left([\bar{s}_{1,0} - t_n, \dots, \bar{s}_{M,0} - t_n]^\top [\bar{s}_{1,0} - t_n, \dots, \bar{s}_{M,0} - t_n] \middle| F_{M+1,n}\right)$. By Eq. (A8-a) and the expression of the off-diagonal elements in $\bar{\Psi}_n(G)$ [see Eqs. (6) and (7) evaluated at n],

$$\bar{\Omega}_{n-1}(k, G) - \bar{\Omega}_n(k, G) = \Lambda_{n-1} - \Lambda_n, \quad \text{all } k, G.$$

Finally, the following recursion is readily obtained by Eq. (8),

$$\bar{\Gamma}_{n-1}(G) - \bar{\Gamma}_n(G) = (M-1)(\Lambda_{n-1} - \Lambda_n). \quad \blacksquare$$

Lemma 3. *The market maker learning about individual and average signals evolves according to*

$$t_n = t_{n-1} + \zeta_n y_n, \quad \zeta_n = \frac{\tau_n}{\text{var}(y_n | F_{n-1})}. \quad (\text{A11})$$

The relation between the updating parameters λ_n and ζ_n is given by

$$\lambda_n = \theta \zeta_n. \quad (\text{A12})$$

Finally, the trading strategy of any trader i does not depend on the residual order flow $(y_{i,t} - x_{i,t})_{t=1}^{n-1}$. Precisely, we have:

$$x_{i,n} = \hat{G}\beta_n \left(\bar{s}_{i,0} - \sum_{r=1}^{n-1} \zeta_r y_r \right). \quad (\text{A13})$$

Proof of Lemma 3. By the definition of t_n and $s_{i,n}$,

$$t_n - t_{n-1} = E(s_{i,0} - t_{n-1} | F_{M+1,n}) = E(s_{i,n-1} | F_{M+1,n}) = \zeta_n y_n,$$

where ζ_n is the regression coefficient of $s_{i,n-1}$ on y_n , viz

$$\zeta_n = \frac{\text{cov}(s_{i,n-1}, y_n | F_{M+1,n-1})}{\text{var}(y_n | F_{M+1,n-1})} = \frac{\tau_n}{\text{var}(y_n | F_{M+1,n-1})}, \quad (\text{A14})$$

which proves Eq. (A11). As regards Eq. (A12), we have that, by Eqs. (A11) and (A6), $0 = \theta t_n - \theta t_{n-1} - \theta \zeta_n y_n = p_n - p_{n-1} - \theta \zeta_n y_n$, and Eq. (A12) follows by Eq. (14).

Finally, we show Eq. (A13). By the assumption that trading strategies are linear, as in Eq. (13), and the market maker's recursive update in Eq. (A2) and (A11),

$$x_{i,n} = \hat{G}\beta_n \bar{s}_{i,n-1} = \hat{G}\beta_n (\bar{s}_{i,0} - t_{n-1}) = \hat{G}\beta_n \left(\bar{s}_{i,0} - \sum_{r=1}^{n-1} \zeta_r y_r \right). \quad \blacksquare$$

Derivation of Eqs. (15)-(16). We have:

$$E\left(\sum_{j \neq i} \bar{s}_{j,n-1} \middle| F_{i,n}\right) = E\left(\sum_{j \neq i} \bar{s}_{j,n-1} \middle| \bar{s}_{i,n-1}, F_{n-1}\right) = \frac{\bar{\Gamma}_{n-1}(G)}{\bar{\Lambda}_{n-1}(G)} \bar{s}_{i,n-1}, \quad (\text{A15})$$

which is Eq. (16). Moreover:

$$\begin{aligned} E(f - p_{n-1} \middle| F_{i,n}) &= E[E(f - p_{n-1} \middle| \mathbf{s}_0) \middle| F_{i,n}] \\ &= E(\theta \bar{s} - p_{n-1} \middle| F_{i,n}) \\ &= \frac{\theta}{M} E\left(\bar{s}_{i,n-1} + \sum_{j \neq i} \bar{s}_{j,n-1} \middle| F_{i,n}\right) \\ &= \frac{\theta}{M} \left(1 + \frac{\bar{\Gamma}_{n-1}(G)}{\bar{\Lambda}_{n-1}(G)}\right) \bar{s}_{i,n-1}, \end{aligned}$$

where the first line follows by the Law of Iterated Expectations, the second by Eq. (3), the third by Eq. (A6) in Lemma 1 and the fact that $\theta \bar{s} - p_{n-1} = \frac{\theta}{M} \sum_{i=1}^M \bar{s}_{i,n-1}$, and the fourth by Eq. (A15).

B. Proofs of Propositions 1 and 2

We characterize each trader's behavior off the equilibrium path. We conjecture that:

- The i -th trader deviation, $x'_{i,n}$ say, coincides with the equilibrium strategy in Eq. (13), plus an additional term reflecting the price deviation induced by suboptimal play in the previous $n - 1$ rounds,

$$x'_{i,n} = \hat{G} \beta_n \bar{s}_{i,n-1} + \gamma_n (p_{n-1} - p'_{n-1}), \quad (\text{B1})$$

where p'_{n-1} is the price process that would emerge should the i -th trader have decided to deviate in the previous rounds of trading.

- The price induced by past suboptimal play of the i -th trader shows up in the value function in Eq. (9),

$$W_{i,n} = \alpha_n \bar{s}_{i,n}^2 + \psi_n \bar{s}_{i,n} (p_n - p'_n) + \mu_n (p_n - p'_n)^2 + \delta_n. \quad (\text{B2})$$

B.1 Proof of Proposition 1

We claim that the conditions reported below guarantee mutual consistency between Eqs. (B1) and (B2):

Claim 1. *The necessary and sufficient conditions for a unique Bayesian Nash equilibrium in which trading strategies and prices are as in Eqs. (13)-(14) are as follows:*

- (i) *The trading strategy coefficients β_n and γ_n :*

$$\beta_n = \frac{\theta \lambda_n \sigma_u^2}{\hat{G} M \sigma_{f,n}^2}, \quad (\text{B3-a})$$

$$\gamma_n = \frac{(1 - 2\lambda_n \mu_n) \left[1 - \theta^{-1} \hat{G} (M - 1) \beta_n \lambda_n\right]}{2\lambda_n (1 - \lambda_n \mu_n)}. \quad (\text{B3-b})$$

(ii) The price impact λ_n is the unique real, positive solution to:

$$0 = \frac{\theta(M - \hat{G})(\Lambda_n - \Omega_n)\sigma_u^4}{\hat{G}M^2\sigma_{f,n}^4}\lambda_n^4 + \frac{\sigma_u^2\psi_n\bar{\Lambda}_n(G)}{\sigma_{f,n}^2}\lambda_n^3 - \frac{\theta\sigma_u^2[2\Lambda_n + (M-1)\Omega_n - \frac{2G}{\hat{G}}(\Lambda_n - \Omega_n)]}{M\sigma_{f,n}^2}\lambda_n^2 - \psi_n\bar{\Lambda}_n(G)\lambda_n + \frac{\sigma_{f,n}^2}{\theta} \quad (\text{B4})$$

(iii) The value function coefficients satisfy the recursions:

$$\begin{aligned} \alpha_{n-1} &= \alpha_n \left[1 - \theta^{-1}\hat{G}(1 + (M-1)\phi_n)\beta_n\lambda_n \right]^2 + \hat{G}^2\beta_n[\eta_n - \beta_n\lambda_n(1 + (M-1)\phi_n)] \\ \psi_{n-1} &= \psi_n \left[1 - \lambda_n\gamma_n - \theta^{-1}\hat{G}(M-1)\beta_n\lambda_n \right] \left[1 - \theta^{-1}\hat{G}(1 + (M-1)\phi_n)\beta_n\lambda_n \right] \\ &\quad + \hat{G} \left\{ \gamma_n[\eta_n - \beta_n\lambda_n(1 + (M-1)\phi_n)] - \beta_n\gamma_n\lambda_n + \beta_n \left[1 - \theta^{-1}\hat{G}(M-1)\beta_n\lambda_n \right] \right\} \\ \mu_{n-1} &= \mu_n \left[1 - \lambda_n\gamma_n - \theta^{-1}\hat{G}(M-1)\beta_n\lambda_n \right]^2 + \gamma_n \left[1 - \lambda_n\gamma_n - \theta^{-1}\hat{G}(M-1)\beta_n\lambda_n \right] \\ \delta_{n-1} &= \delta_n + \theta^{-2}\alpha_n\lambda_n^2\sigma_u^2 + \theta^{-2}\hat{G}^2\alpha_n\lambda_n^2\beta_n^2\text{var} \left(\sum_{j \neq i} \bar{s}_{j,n-1} \middle| F_{i,n} \right) \end{aligned} \quad (\text{B5})$$

where $\alpha_N = \psi_N = \mu_N = \delta_N = 0$ and

$$\phi_n = \frac{\bar{\Gamma}_{n-1}(G)}{(M-1)\bar{\Lambda}_{n-1}(G)}, \quad (\text{B6-a})$$

$$\eta_n = \frac{\theta(\bar{\Gamma}_{n-1}(G) + \bar{\Lambda}_{n-1}(G))}{\hat{G}M\bar{\Lambda}_{n-1}(G)}, \quad (\text{B6-b})$$

$$\text{var} \left(\sum_{j \neq i} \bar{s}_{j,n-1} \middle| F_{i,n} \right) = M[\Lambda_{n-1} + (M-1)\Omega_{n-1}] - \left[1 + \phi_n^2(M-1)^2 \right] \bar{\Lambda}_{n-1}(G) - 2\bar{\Gamma}_{n-1}(G). \quad (\text{B7})$$

(iv) The full information residual variance satisfies the recursion:

$$\sigma_{f,n}^2 = \left(1 - \theta^{-1}\hat{G}M\beta_n\lambda_n \right) \sigma_{f,n-1}^2.$$

(v) Finally, the following inequality must hold:

$$\lambda_n(1 - \lambda_n\mu_n) > 0. \quad (\text{B8})$$

Note that Proposition 1 is proven, once we demonstrate that Claim 1 holds true. To prove Claim 1, we proceed in three steps. In the first step, we derive a recursive expression for the price deviation induced by traders' suboptimal play. In the second step, we derive the traders' optimality conditions. In the third step, we compute the market maker updates.

We need the results in Lemmas 4 to 6 below.

Lemma 4. Let y'_n and t'_n be the aggregate order flow and the market maker's update of any trader's average signal, when trader i deviates to $(x'_{i,k})_{k=1}^{n-1}$ during the first $n-1$ auctions. The deviation in trader j residual informational advantage, $\bar{s}'_{j,n} = \bar{s}_{i,0} - t'_n$, satisfies:

$$\bar{s}_{j,n-1} - \bar{s}'_{j,n-1} = \frac{1}{\theta} (p'_{n-1} - p_{n-1}). \quad (\text{B9})$$

Proof. By Eq. (10), $y'_n = \sum_{j \neq i} x_{j,n} + x'_{i,n} + u_n$, and by Eq. (A11) in Lemma 3, $t'_n = \sum_{k=1}^n \zeta_k y'_k$. Therefore, by Eq. (A2),

$$\begin{aligned} \bar{s}_{j,n-1} - \bar{s}'_{j,n-1} &= (\bar{s}_{j,0} - t_{n-1}) - (\bar{s}_{j,0} - t'_{n-1}) \\ &= \sum_{k=1}^{n-1} \zeta_k y'_k - \sum_{k=1}^{n-1} \zeta_k y_k \\ &= \frac{1}{\theta} \left(\sum_{k=1}^{n-1} \lambda_k y'_k - \sum_{k=1}^{n-1} \lambda_k y_k \right) \\ &= \frac{1}{\theta} (p'_{n-1} - p_{n-1}), \end{aligned}$$

where the third line follows by Eq. (A12), and the fourth line holds as Eq. (14) implies that $p_n = \sum_{k=1}^n \lambda_k y_k$. ■

Lemma 5. The conditional moments of the residual informational advantage satisfy:

$$E(\bar{s}_{i,n} | F_{i,n}) = \left[1 - \frac{\hat{G}\beta_n \lambda_n}{\theta} (1 + (M-1)\phi_n) \right] \bar{s}_{i,n-1}, \quad (\text{B10})$$

and

$$\begin{aligned} E(\bar{s}_{i,n}^2 | F_{i,n}) &= \left[1 - \theta^{-1} \hat{G}\beta_n \lambda_n (1 + (M-1)\phi_n) \right]^2 \bar{s}_{i,n-1}^2 \\ &\quad + \theta^{-1} \lambda_n^2 \sigma_u^2 + \theta^{-1} \hat{G}^2 \beta_n^2 \lambda_n^2 \text{var} \left(\sum_{j \neq i} \bar{s}_{j,n-1} \middle| F_{i,n} \right), \end{aligned} \quad (\text{B11})$$

where $\text{var} \left(\sum_{j \neq i} \bar{s}_{j,n-1} \middle| F_{i,n} \right)$ is given in Eq. (B7).

Proof. By Eqs. (A2) and (A11),

$$\bar{s}_{i,n} = \bar{s}_{i,0} - t_n = \bar{s}_{i,n-1} - (t_n - t_{n-1}) = \bar{s}_{i,n-1} - \zeta_n y_n.$$

Substituting for the equilibrium order flow, using Eqs. (13) and (A12), and taking expectations yields:

$$E(\bar{s}_{i,n} | F_{i,n}) = \bar{s}_{i,n-1} - \frac{\hat{G}\beta_n \lambda_n}{\theta} \left[\bar{s}_{i,n-1} + E \left(\sum_{i \neq j} \bar{s}_{i,n-1} \middle| F_{i,n} \right) \right].$$

Eq. (B10) follows by Eq. (16) and the definition of ϕ_n in Eq. (B6-a). As for the proof of Eq. (B11), note that by Eqs. (A2) and (A11),

$$\begin{aligned} E(\bar{s}_{i,n}^2 | F_{i,n}) &= \bar{s}_{i,n-1}^2 + \zeta_n^2 E(y_n^2 | F_{i,n}) - 2\zeta_n \bar{s}_{i,n-1} E(y_n | F_{i,n}) \\ &= \left[1 - \theta^{-1} \hat{G}\beta_n \lambda_n (1 + (M-1)\phi_n) \right]^2 \bar{s}_{i,n-1}^2 + \theta^{-1} \lambda_n^2 \sigma_u^2 + \theta^{-1} \hat{G}^2 \beta_n^2 \lambda_n^2 \text{var} \left(\sum_{j \neq i} \bar{s}_{j,n-1} \middle| F_{i,n} \right), \end{aligned}$$

where the expression for $\text{var}\left(\sum_{j \neq i} \bar{s}_{j,n-1} \middle| F_{i,n}\right)$ is given in Eq. (B7) and follows by simple computations. ■

Lemma 6. *The price impact and the residual variance of the full information asset value are given by:*

$$\lambda_n = \frac{\theta M \hat{G} \beta_n \sigma_{f,n-1}^2}{\left(\hat{G} M \beta_n\right)^2 \sigma_{f,n-1}^2 + \theta^2 \sigma_u^2}. \quad (\text{B12-a})$$

$$\sigma_{f,n}^2 = \frac{\theta^2 \sigma_u^2 \sigma_{f,n-1}^2}{\left(\hat{G} M \beta_n\right)^2 \sigma_{f,n-1}^2 + \theta^2 \sigma_u^2}. \quad (\text{B12-b})$$

Proof. We first use Eq. (13) and rewrite the aggregate order flow in Eq. (10) as

$$y_n = \sum_{i=1}^M \hat{G} \beta_n (\bar{s}_{i,0} - t_{n-1}) + u_n. \quad (\text{B13})$$

Since $\sum_{i=1}^M \bar{s}_{i,0} = \sum_{i=1}^M s_{i,0}$ and $t_{n-1} \in F_{n-1}$, then

$$\text{cov}(s_{i,n-1}, y_n | F_{n-1}) = \hat{G} \beta_n \text{cov}\left(s_{i,0}, \sum_{i=1}^M s_{i,0} | F_{n-1}\right) = \theta^{-2} \hat{G} M \beta_n \sigma_{f,n-1}^2, \quad (\text{B14})$$

where the last line follows from Eq. (A7) in Lemma 2. Moreover, by Eq. (A6) in Lemma 1 and the fact that $\theta \bar{s} - p_{n-1} = \frac{\theta}{M} \sum_{i=1}^M \bar{s}_{i,n-1}$, the aggregate order flow in Eq. (B13) becomes

$$y_n = \theta^{-1} \hat{G} M \beta_n (\theta \bar{s} - p_{n-1}) + u_n, \quad (\text{B15})$$

so that

$$\text{var}(y_n | F_{n-1}) = \theta^{-2} \left(\hat{G} M \beta_n\right)^2 \text{var}(\theta \bar{s} - p_{n-1} | F_{n-1}) + \sigma_u^2 = \theta^{-2} \left(\hat{G} M \beta_n\right)^2 \sigma_{f,n-1}^2 + \sigma_u^2, \quad (\text{B16})$$

where the last line follows from Eq. (3) and the definition of $\sigma_{f,n}^2$ in Eq. (18). Since $\lambda_n = \theta \zeta_n$ by Eq. (A12) in Lemma 3, then λ_n in Eq. (B12-a) obtains from Eqs. (B14) and (B16).

By Eq. (3) and the Law of Iterated Expectations $\text{cov}(f, y_n | F_{M+1,n-1}) = \text{cov}(\theta \bar{s}, y_n | F_{n-1})$, and Eq. (B15) yields

$$\text{cov}(f, y_n | F_{n-1}) = \theta^{-1} \hat{G} M \beta_n \text{var}(\theta \bar{s} - p_{n-1} | F_{n-1}) = \theta^{-1} \hat{G} M \beta_n \sigma_{f,n-1}^2. \quad (\text{B17})$$

Equation (B17) together with the Projection Theorem gives

$$\sigma_{f,n}^2 = \sigma_{f,n-1}^2 - \lambda_n \text{cov}(f, y_n | F_{n-1}) = \sigma_{f,n-1}^2 \left(1 - \theta^{-1} \hat{G} M \beta_n \lambda_n\right), \quad (\text{B18})$$

and Eq. (B12-b) follows replacing λ_n from Eq. (B12-a) in the previous equality. ■

We are now ready to prove Claim 1 in three steps.

Step 1: Price deviation. Using the equilibrium strategies in Eqs. (13)-(14), the price deviation writes as:

$$p_n - p'_n = p_{n-1} - p'_{n-1} + \lambda_n \left[\sum_{j \neq i} \hat{G} \beta_n (\bar{s}_{j,n-1} - \bar{s}'_{j,n-1}) + \hat{G} \beta_n \bar{s}_{i,n-1} - x'_{i,n} \right].$$

Plugging Eq. (B9) in Lemma 4 into the previous equation, we obtain the following recursive equation for the price deviation:

$$p_n - p'_n = (p_{n-1} - p'_{n-1}) \left[1 - \theta^{-1} \hat{G} (M-1) \beta_n \lambda_n \right] + \hat{G} \lambda_n \beta_n \bar{s}_{i,n-1} - \lambda_n x'_{i,n}. \quad (\text{B19})$$

Step 2: Traders' strategies. First, we show that the strategy in Eq. (B1) and the value function in Eq. (B2) are mutually consistent. Any trader i faces the following recursive problem:

$$W_{i,n-1} = \max_{x'_{i,n}} E \left[\left(f - \underbrace{(p'_{n-1} + \lambda_n x'_{i,n} + \lambda_n \sum_{j \neq i} x_{j,n})}_{=p'_n} \right) x'_{i,n} + W_{i,n} \middle| F_{i,n} \right]. \quad (\text{B20})$$

Given the conjectured value function in Eq. (B2), the trading strategies in Eq. (13) and Eq. (B19), the optimality conditions of the previous problem lead to:

$$0 = E(f - p_{n-1} | F_{i,n}) + (p_{n-1} - p'_{n-1}) - \hat{G} \beta_n \lambda_n \sum_{j \neq i} (\bar{s}'_{j,n-1} - \bar{s}_{j,n-1}) \\ - \hat{G} \beta_n \lambda_n E \left(\sum_{j \neq i} \bar{s}_{j,n-1} \middle| F_{i,n} \right) - 2\lambda_n x'_{i,n} - \lambda_n \psi_n E(\bar{s}_{i,n} | F_{i,n}) - 2\lambda_n \mu_n E(p_n - p'_n | F_{i,n}).$$

together with the second order condition in (B8). By replacing Eq. (B9) in Lemma 4, Eq. (B10) in Lemma 5 and Eq. (B19) in the previous equation, and by rearranging terms, we obtain Eq. (B1), where γ_n is as in Eq. (B3-b) and

$$\beta_n = \frac{\eta_n - \hat{G}^{-1} \lambda_n \psi_n}{\lambda_n [1 + (1 - \theta^{-1} \lambda_n \psi_n) (1 + (M-1) \phi_n)]}. \quad (\text{B21})$$

Next, we use Eq. (B1), and find that the expected profit in any single auction is:

$$E[(f - p_n) x'_{i,n} | F_{i,n}] \\ = \hat{G}^2 \beta_n [\eta_n - \beta_n \lambda_n (1 + (M-1) \phi_n)] \bar{s}_{i,n-1}^2 \\ + \gamma_n \left[1 - \lambda_n (\gamma_n + \theta^{-1} \hat{G} (M-1) \beta_n) \right] (p_{n-1} - p'_{n-1})^2 \\ + \left\{ \gamma_n (\eta_n - 2\beta_n \lambda_n) + \beta_n \left[1 - (M-1) \lambda_n (\gamma_n \phi_n + \theta^{-1} \hat{G} \beta_n) \right] \right\} \hat{G} \bar{s}_{i,n-1} (p_{n-1} - p'_{n-1}). \quad (\text{B22})$$

By taking the conditional expectation of the value function in Eq. (B2) leaves:

$$E(W_{i,n} | F_{i,n}) = \alpha_n E(\bar{s}_{i,n}^2 | F_{i,n}) + \psi_n (p_n - p'_n) E(\bar{s}_{i,n} | F_{i,n}) + \mu_n (p_n - p'_n)^2 + \delta_n. \quad (\text{B23})$$

Next, use Eqs. (B10) and (B11) in Lemma 5 together with Eqs. (B7) and (B19) into Eq. (B23). Finally, plug the resulting expression for Eqs. (B22) and (B23) into Eq. (B20) and identify terms to obtain the

recursions for the coefficients α_n, μ_n, ψ_n and δ_n in Eq. (B5). Finally, Eqs. (B6-a), (B6-b) and (B7) follow by a direct computation and Lemma 3.

Step 3: Market maker updates. By combining Eqs. (B12-a) and (B12-b) in Lemma 6 we find an alternative expression for λ_n ,

$$\lambda_n = \frac{\hat{G}M\beta_n \sigma_{f,n}^2}{\theta \sigma_u^2}, \quad (\text{B24})$$

which is Eq. (B3-a). By solving Eq. (B12-b) in Lemma 6 for $\sigma_{f,n-1}^2$ gives

$$\sigma_{f,n-1}^2 = -\frac{\theta^2 \sigma_u^2 \sigma_{f,n}^2}{\left(\hat{G}M\beta_n\right)^2 \sigma_{f,n}^2 - \theta^2 \sigma_u^2}. \quad (\text{B25})$$

By combining Eqs. (B3-a), (B24) and (B25) together with $\Lambda_n = \Lambda_{n-1} - \zeta_n \cdot \text{cov}(s_{i,n-1}, y_n | F_{n-1})$, we find that Λ_n solves,

$$\Lambda_{n-1} - \Lambda_n = -\frac{\sigma_u^4}{\left(\hat{G}M\beta_n\right)^2 \sigma_{f,n}^2 - \theta^2 \sigma_u^2} \lambda_n^2 = -\frac{\lambda_n^2 \sigma_u^2 \sigma_{f,n}^2}{\theta^2 \left(\lambda_n^2 \sigma_u^2 - \sigma_{f,n}^2\right)}, \quad (\text{B26})$$

Also, Eqs. (A8-c) and (A8-e) in Lemma 2 imply that

$$\frac{\bar{\Gamma}_{n-1}(G)}{\bar{\Lambda}_{n-1}(G)} = \frac{\bar{\Gamma}_n(G) + (M-1)(\Lambda_{n-1} - \Lambda_n)}{\bar{\Lambda}_n(G) + (\Lambda_{n-1} - \Lambda_n)}. \quad (\text{B27})$$

By eliminating β_n between Eqs. (B3-a) and (B21), we find that

$$\theta \sigma_u^2 \lambda_n^2 \left[1 + (1 - \theta^{-1} \lambda_n \psi_n) (1 + (M-1) \phi_n)\right] = \left(\hat{G}\eta_n - \lambda_n \psi_n\right) M \sigma_{f,n}^2,$$

and substituting Eqs. (B6-a) and (B6-b) in the previous equation leaves

$$\theta \sigma_u^2 \lambda_n^2 \left[1 + (1 - \theta^{-1} \lambda_n \psi_n) \left(1 + \frac{\bar{\Gamma}_{n-1}(G)}{\bar{\Lambda}_{n-1}(G)}\right)\right] = \left[\frac{\theta}{M} \left(1 + \frac{\bar{\Gamma}_{n-1}(G)}{\bar{\Lambda}_{n-1}(G)}\right) - \lambda_n \psi_n\right] M \sigma_{f,n}^2.$$

The quartic equation $F(\lambda_n) = 0$ in Eq. (B4) is obtained by substituting Eqs. (5) and (8) evaluated at n , Eq. (A7) in Lemma 2 and Eqs. (B26) and (B27) into the previous equation, and by tedious computations.

To show that Eq. (B4) admits a unique positive solution, note that the constant and the coefficient of λ_n^4 are both positive, and that the coefficient of λ_n^2 is negative. (Let $\rho_n = \Omega_n/\Lambda_n$ be the correlation coefficient between individual signals. Since $|\rho_n| \leq 1$, then $\Lambda_n - \Omega_n \geq 0$ and the coefficient for λ_n^4 is non-negative. Moreover, $\Lambda_n + (M-1)\Omega_n > 0$ and, hence, $2\Lambda_n + (M-1)\Omega_n > 0$.) On the other hand, the sign of ψ_n determines the sign of the terms in λ_n^3 and λ_n . However, regardless of whether ψ_n is positive or negative, there are only two sign changes. By Descartes' rule, Eq. (B4) has at most two real positive roots. By Eq. (B18), $\sigma_{f,n}^2 < \sigma_{f,n-1}^2 \Leftrightarrow \theta^{-1} \beta_n \lambda_n \hat{G}M < 1$. By Eq. (B3-a) this restriction becomes $\lambda_n^2 < \sigma_{f,n}^2 \sigma_u^{-2}$ or equivalently $\lambda_n < \sigma_u^{-1} \sigma_{f,n}$. By Eq. (B4), $F(\lambda = 0) = \theta^{-1} \sigma_{f,n}^2 > 0$, $F(\lambda_n = \sigma_u^{-1} \sigma_{f,n}) = -(\theta M)^{-1} \sigma_{f,n}^2 < 0$ and $F(\lambda = +\infty) = +\infty$; hence, there is one and only one positive root between 0 and $\sigma_u^{-1} \sigma_{f,n}$. ■

B.2 Proof of Proposition 2

We need the following preliminary result.

Lemma 7. *Let Δ_n be as in Eq. (17). Then,*

$$\Delta_n = \frac{(\bar{\Gamma}_n(G) + \bar{\Lambda}_n(G))/M}{\bar{\Lambda}_n(G)}. \quad (\text{B28})$$

Proof. Define,

$$\bar{\Gamma}_n(G) = (M-1)\Omega_n + \frac{2G}{\hat{G}}(\Lambda_n - \Omega_n). \quad (\text{B29})$$

Then,

$$\begin{aligned} \text{var}(\bar{s} | F_n) &= \text{var}\left(M^{-1} \sum_{i=1}^M \bar{s}_{i,0} \middle| F_n\right) \\ &= M^{-2} \left[M\bar{\Lambda}_n(G) + M \cdot \text{cov}\left(\sum_{j \neq i} \bar{s}_{j,n}, \bar{s}_{i,n} \middle| F_n\right) \right] \\ &= M^{-2} (M\bar{\Lambda}_n(G) + M\bar{\Gamma}_n(G)) \\ &= (\bar{\Gamma}_n(G) + \bar{\Lambda}_n(G))/M, \end{aligned}$$

where the first line follows by the definition of \bar{s} . Then, Eq (B28) follows by the previous equality and the definition of Δ_n in Eq. (17). ■

We now proceed with the proof of Proposition 2. First, by Lemma 7 and Eq. (B6-b),

$$\Delta_{n-1} = \frac{\hat{G}\eta_n}{\theta}. \quad (\text{B30})$$

By plugging Eq. (15) into Eq. (13) leaves

$$x_{i,n} = \frac{\beta_n}{\eta_n} [E(f | F_{i,n}) - p_{n-1}] = \frac{\hat{G}\beta_n}{\theta\Delta_{n-1}} [E(f | F_{i,n}) - p_{n-1}], \quad (\text{B31})$$

where the last equality holds by Eq. (B30). This is Eq. (20).

Next, we prove Eq. (22). We have,

$$\begin{aligned} E\left(\sum_{j \neq i} x_{j,n} \middle| F_{i,n}\right) &= E\left(\sum_{j \neq i} \hat{G}\beta_n \bar{s}_{j,n-1} \middle| F_{i,n}\right) \\ &= \hat{G}\beta_n (M-1) \phi_n \bar{s}_{i,n-1} \\ &= \hat{G}\beta_n \frac{\bar{\Gamma}_{n-1}(G)}{\bar{\Lambda}_{n-1}(G)} \bar{s}_{i,n-1} \\ &= \frac{\bar{\Gamma}_{n-1}(G)}{\bar{\Lambda}_{n-1}(G)} x_{i,n}, \\ &= (M\Delta_{n-1} - 1) x_{i,n} \\ &= \frac{\hat{G}\beta_n}{\theta\Delta_{n-1}} (M\Delta_{n-1} - 1) [E(f | F_{i,n}) - p_{n-1}], \end{aligned}$$

where the first line holds by the expression for the trading strategy in Eq. (13), the second line follows by the expression for the forecast of the traders' informational advantage in Eq. (16), the third line follows by Eq. (B6-a), the fourth line follows, again, by Eq. (13), the fifth line follows by Eq. (B28) in Lemma 7, and, finally, the sixth line holds by Eq. (B31).

To prove Eq. (21), note that by Eqs. (A5) and (B29),

$$\bar{\Lambda}_n(G) + \bar{\Gamma}_n(G) = \frac{\Lambda_n + 2G\Omega_n}{\hat{G}} + (M-1)\Omega_n + \frac{2G}{\hat{G}}(\Lambda_n - \Omega_n) = \Lambda_n + (M-1)\Omega_n.$$

Hence, we can rewrite Eq. (A7) in Lemma 2 as

$$\sigma_{f,n}^2 = \frac{\theta^2}{M} (\bar{\Lambda}_n(G) + \bar{\Gamma}_n(G)). \quad (\text{B32})$$

In terms of Eq. (B32), Eq. (B7) is,

$$\text{var} \left(\sum_{j \neq i} \bar{s}_{j,n} | F_{i,n+1} \right) = M (\bar{\Lambda}_n(G) + \bar{\Gamma}_n(G)) - \left[1 + (M-1)^2 \phi_{n+1}^2 \right] \bar{\Lambda}_n(G) - 2\bar{\Gamma}_n(G).$$

Next, we substitute ϕ_{n+1} from Eq. (B6-a) in the previous equation, and obtain,

$$\text{var} \left(\sum_{j \neq i} \bar{s}_{j,n} | F_{i,n+1} \right) = M (\bar{\Lambda}_n(G) + \bar{\Gamma}_n(G)) - \frac{\bar{\Lambda}_n(G)^2 + \bar{\Gamma}_n(G)^2}{\bar{\Lambda}_n(G)} - 2\bar{\Gamma}_n(G). \quad (\text{B33})$$

We now compute $\varsigma_{f,n}^2$ defined in Eq. (19),

$$\varsigma_{f,n}^2 = \text{var}(\theta \bar{s} | F_{i,n+1}) = \text{var} \left(\theta \frac{1}{M} \sum_{i=1}^M \bar{s}_{i,0} \Big| F_{i,n+1} \right) = \left(\frac{\theta}{M} \right)^2 \text{var} \left(\sum_{j \neq i} \bar{s}_{j,n} | F_{i,n+1} \right),$$

where we use $\bar{s}_{j,n} = \bar{s}_{j,0} - t_n$ and the fact that $t_n \in F_{i,n+1}$ to get the last equality. Plugging Eq. (B33) into the previous equation yields,

$$\varsigma_{f,n}^2 = \left(\frac{\theta}{M} \right)^2 \left[M (\bar{\Lambda}_n(G) + \bar{\Gamma}_n(G)) - \frac{\bar{\Lambda}_n(G)^2 + \bar{\Gamma}_n(G)^2}{\bar{\Lambda}_n(G)} - 2\bar{\Gamma}_n(G) \right]. \quad (\text{B34})$$

By Eqs. (B32) and (B34), we have

$$\sigma_{f,n}^2 - \varsigma_{f,n}^2 = \left(\frac{\theta}{M} \right)^2 \frac{(\bar{\Lambda}_n(G) + \bar{\Gamma}_n(G))^2}{\bar{\Lambda}_n(G)}.$$

Eq. (21) follows by the previous equality, the expression for $\sigma_{f,n}^2$ in Eq. (B32), and the expression for Δ_n in Eq. (B28) in Lemma 7.

Finally, we prove Eq. (23). For $k \neq 0$, we have,

$$E(x_{i+k,n} | F_{i,n}) = \hat{G} \beta_n E(\bar{s}_{i+k,n-1} | F_{i,n}) = \hat{G} \beta_n \frac{\bar{\Omega}_{n-1}(k, G)}{\bar{\Lambda}_{n-1}(G)} \bar{s}_{i,n-1} = \bar{\rho}_{n-1}(k, G) x_{i,n},$$

which is the first line in Eq. (23). Moreover, we have,

$$M\Delta_{n-1} - 1 = \frac{\bar{\Gamma}_{n-1}(G)}{\bar{\Lambda}_{n-1}(G)} = \bar{\rho}_{n-1}(k, G) + \sum_{\ell \notin \{i, i+k\}} \bar{\rho}_{n-1}(\ell, G),$$

where the first equality follows by Eq. (B28) in Lemma 7. The second line in Eq. (23) follows by the previous equality, Eq. (22) and the first line in Eq. (23). ■

C. Computation of the equilibrium

We solve for the equilibrium using backward induction. By Eq. (A8-a), $\Lambda_n - \Omega_n = \Lambda_0 - \Omega_0$. We fix a terminal value for Λ_N and compute $\Omega_N = \Lambda_N + \Omega_0 - \Lambda_0$. $\sigma_{f,N}^2$ then follows by Eq. (A7). Since $\alpha_N = \psi_N = \mu_N = \delta_N = 0$, we solve for λ_N in Eq. (B4), which yields β_N and γ_N through Eqs. (B3-a)-(B3-b). To compute the value function coefficients as of at time $N-1$, one needs to express Λ_{N-1} and Ω_{N-1} in terms of variables known at time N . Below, we show that:

$$\Lambda_{n-1} = \frac{\theta\Lambda_n - \hat{G}(M-1)\lambda_n\beta_n(\Lambda_n - \Omega_n)}{\theta - \hat{G}M\lambda_n\beta_n}. \quad (\text{C1})$$

Then, Λ_{N-1} is obtained by evaluating Eq. (C1) at $n = N$, and Ω_{N-1} is obtained by the equality $\Omega_{N-1} = \Lambda_{N-1} + \Omega_0 - \Lambda_0$. Finally, we retrieve the regression coefficients ϕ_N and η_N through Eqs. (B6-a)-(B6-b) and the equality $\bar{\Gamma}_{N-1} = \bar{\Gamma}_N + (M-1)(\Lambda_{N-1} - \Lambda_N)$ [see Eq. (A8-e)]. The value function coefficients at $N-1$, then, are uniquely determined by Eq. (B5).

The above procedure is applied at each trading round $n \in [1, N]$, yielding the initial value of Λ_0 implied by the choice of the terminal value of Λ_N . Then, the resulting initial value of Λ_0 is compared to that we posited as the initial parameter, and repeat the procedure for different choices of Λ_N , until we achieve convergence.

Derivation of Eq. (C1). By Eqs. (A10), (A14), the expression for τ_n found above, and Eq. (A12),

$$\Lambda_n = \Lambda_{n-1} - \zeta_n\tau_n = \Lambda_{n-1} - \frac{M\hat{G}\beta_n\lambda_n}{\theta^3}\sigma_{f,n-1}^2. \quad (\text{C2})$$

Taking one lag in Eq. (A7) and substituting the result into Eq. (C2) yields:

$$\Lambda_{n-1} = \Lambda_n + \frac{\hat{G}\beta_n\lambda_n}{\theta}[\Lambda_{n-1} + (M-1)\Omega_{n-1}].$$

Since $\Omega_{n-1} - \Omega_n = \Lambda_{n-1} - \Lambda_n$ [see Eq. (A8-a)],

$$\Omega_{n-1} = \Omega_n + \frac{\hat{G}\beta_n\lambda_n}{\theta}(\Lambda_{n-1} - \Omega_{n-1} + M\Omega_{n-1}) = \frac{\theta\Omega_n + \hat{G}\beta_n\lambda_n(\Lambda_{n-1} - \Omega_{n-1})}{\theta - M\hat{G}\beta_n\lambda_n}.$$

By solving for Ω_{n-1} we find that

$$\Omega_{n-1} = \frac{\theta\Omega_n + \hat{G}\beta_n\lambda_n(\Lambda_{n-1} - \Omega_{n-1})}{\theta - \hat{G}\beta_n\lambda_n M}. \quad (\text{C3})$$

Finally,

$$\Lambda_{n-1} = \Lambda_n + \Omega_{n-1} - \Omega_n = \frac{\theta\Lambda_n - \hat{G}\beta_n\lambda_n[M(\Lambda_n - \Omega_n) - (\Lambda_{n-1} - \Omega_{n-1})]}{\theta - \hat{G}\beta_n\lambda_n M},$$

where we have used Eq. (C3). Eq. (C1) follows by replacing $\Lambda_n - \Omega_n = \Lambda_{n-1} - \Omega_{n-1}$ in the previous equation. ■

Computation of the correlation among volumes. By definition, the correlation among the volume generated by any two traders i and j is,

$$\rho_{V,n}(i,j) = \frac{\text{cov}(|x_{i,n}|, |x_{j,n}|)}{\text{var}(|x_{i,n}|)}, \quad \text{for } i \neq j. \quad (\text{C4})$$

To produce the results in Figure 11, we set $\rho = -0.15$, $\rho = 0.10$ or $\rho = 0.90$. We simulate $S = 5000$ values of the initial signals $(s_{i,0})_{i=1}^M$, and compute the dynamics of the individual trades over the ten batch auctions. Thus, we obtain a sequence of trades $(x_{i,n}^s)_{s=1}^S$ for each trader i and each batch auction n . We estimate $\rho_{V,n}(i,j)$ in Eq. (C4) through the estimator $\hat{\rho}_{S,V,n}(i,j)$ given below,

$$\hat{\rho}_{S,V,n}(i,j) = \frac{\sum_{s=1}^S \left(|x_{i,n}^s| - \frac{1}{S} \sum_{s=1}^S |x_{i,n}^s| \right) \left(|x_{j,n}^s| - \frac{1}{S} \sum_{s=1}^S |x_{j,n}^s| \right)}{\sqrt{\sum_{s=1}^S \left(|x_{i,n}^s| - \frac{1}{S} \sum_{s=1}^S |x_{i,n}^s| \right)^2 \cdot \sum_{s=1}^S \left(|x_{j,n}^s| - \frac{1}{S} \sum_{s=1}^S |x_{j,n}^s| \right)^2}}.$$

References

- Admati, A. R., and P. C. Pfleiderer. 1988. A Theory of Intraday Patterns: Volume and Price Variability. *Review of Financial Studies* 1:3-40.
- Barber, B.M., T. Odean and N. Zhu. 2006. Systematic Noise. Working paper, Graduate School of Management, University of California, Davis.
- Bae, K.-H., T. Yamada, and K. Ito. 2006. How do Individual, Institutional, and Foreign Investors Win and Lose in Equity Trades? Working paper, Queen's School of Business, Queen's University.
- Bernhardt, D., and J. Miao. 2004. Informed Trading When Information Becomes Stale. *Journal of Finance* 59:339-390.
- Bikhchandani, S., and S. Sharma. 2001. Herd Behavior in Financial Markets. *IMF Staff Papers* 47:279-310.
- Brennan, M. J., and H. H. Cao. 1997. International Portfolio Investment Flows. *Journal of Finance* 52:1851-1880.
- Brennan, M. J. and A. Subrahmanyam. 1995. Investment Analysis and Price Formation in Securities Markets. *Journal of Financial Economics* 38:361-381.
- Back, K., H. H. Cao, and G. A. Willard. 2000. Imperfect Competition Among Informed Traders. *Journal of Finance* 55:2117-2155.
- Chae, J. 2005. Trading Volume, Information Asymmetry, and Timing Information. *Journal of Finance* 60:413-442.
- Chordia, T., S-W Huh, and A. Subrahmanyam. 2008. Theory-Based Illiquidity and Asset Pricing. *Review of Financial Studies*, forthcoming.
- Cohen, L., A. Frazzini, and C. Malloy. 2007. The Small World of Investing: Board Connections and Mutual Fund Returns. *Journal of Political Economy*, forthcoming.
- Colla, P., and A. Mele. 2008. Information Linkages and Correlated Trading: Unabridged Version. Working paper, London School of Economics: <http://fmg.lse.ac.uk/~antonio/>.
- Coval, J. D., and T. J. Moskowitz. 1999. Home Bias at Home: Local Equity Preference in Domestic Portfolios. *Journal of Finance* 54:2045-2073.
- DeMarzo, P. M., D. Vayanos, and J. Zwiebel. 2003. Persuasion Bias, Social Influence, and Unidimensional Opinions. *Quarterly Journal of Economics* 118:909-968.
- DeMarzo, P. M., R. Kaniel, and I. Kremer. 2004. Diversification as a Public Good: Community Effects in Portfolio Choice. *Journal of Finance* 59:1677-1715.
- Dorn, D., G. Huberman, and P. Sengmueller. 2008. Correlated Trading and Returns. *Journal of Finance* 63:885-920.
- Ellison, G., and D. Fudenberg. 1993. Rules of Thumb for Social Learning. *Journal of Political Economy* 101:612-643.
- Ellison, G., and D. Fudenberg. 1995. Word-of-Mouth Communication and Social Learning. *Quarterly Journal of Economics* 110:93-125.
- Feng, L., and M. S. Seasholes. 2004. Correlated Trading and Location. *Journal of Finance* 59:2117-2144.

- Foster, F. D., and S. Viswanathan. 1996. Strategic Trading When Agents Forecast the Forecasts of Others. *Journal of Finance* 51:1437-1478.
- Goyal, S. 2007. *Connections*. Princeton: Princeton University Press.
- Grinblatt, M., and M. Keloharju. 2001. How Distance, Language, and Culture Influence Stockholdings and Trades. *Journal of Finance* 56:1053-1073.
- Guiso, L., P. Sapienza, and L. Zingales. 2006. Does Culture Affect Economic Outcomes?" *Journal of Economic Perspectives* 20:23-48.
- Hau, H. 2001. Location Matters: An Examination of Trading Profits. *Journal of Finance* 56:1959-1983.
- Hertz, E. 1998. *The Trading Crowd: An Ethnography of the Shanghai Stock Market*. Cambridge: Cambridge University Press.
- Holden, C. W., and A. Subrahmanyam. 1992. Long-Lived Private Information and Imperfect Competition. *Journal of Finance* 47:247-270.
- Hong, H., J. D. Kubik, and J. C. Stein. 2005. Thy Neighbor's Portfolio: Word-of-Mouth Effects in the Holdings and Trades of Money Managers. *Journal of Finance* 60:2801-2824.
- Hong, H., T. Lim and J. C. Stein. 2000. Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies. *Journal of Finance* 55:265-295.
- Ivković, Z., and S. Weisbenner. 2005. Local Does as Local Is: Information Content of the Geography of Individual Investors' Common Stock Investments. *Journal of Finance* 60:267-306.
- Jackson, M. O. 2008. *Social and Economic Networks*. Princeton: Princeton University Press.
- Kelly, M., and C. O'Grada. 2000. Market Contagion: Evidence from the Panics of 1854 and 1857. *American Economic Review* 90:1110-1124.
- Krinsky, I. and J. Lee. 1996. Earnings Announcements and the Components of the Bid-Ask Spread. *Journal of Finance* 51:1523-1535.
- Kyle, A. S. 1985. Continuous Auctions and Insider Trading. *Econometrica* 53:1335-1355.
- Lakonishok, J., A. Shleifer, and R. W. Vishny. 1992. The impact of institutional trading on stock prices. *Journal of Financial Economics* 32:23-43.
- Malinova, K., and L. Smith. 2006. Informational Diversity and Proximity in Rational Expectations Equilibrium. Working paper, University of Toronto and University of Michigan.
- Ozsoylev, H. 2006. Asset Pricing Implications of Social Networks. Working paper, Oxford, Said Business School.
- Shiller, R. J. 1984. Stock Prices and Social Dynamics. *Brooking Papers on Economic Activity* 2:457-498.
- Shiller, R. J. 2005. *Irrational Exuberance* (2nd edition). Princeton: Princeton University Press.
- Shiller, R., and J. Pound. 1989. Survey Evidence on Diffusion of Interest and Information Among Investors. *Journal of Economic Behavior and Organization* 12:47-66.
- Stein, J. C. 2007. Conversations Among Competitors. *American Economic Review*, forthcoming.
- Watts, D. J., and S. H. Strogatz. 1998. Collective Dynamics of 'Small-World' Networks. *Nature* 393:440-442.

Figures

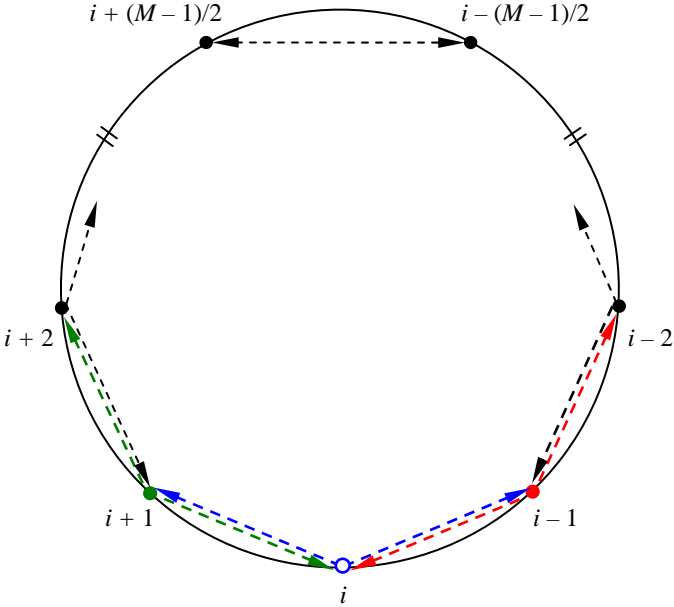


Figure 1: Geographical location of traders

This figure depicts an example of a network of information linkages among M traders who are physically located around a circle. Traders are ordered clockwise: trader i has trader $i+1$ as his clockwise neighbor and trader $i-1$ as his counterclockwise neighbor, so that each trader has $(M-1)/2$ clockwise peers and $(M-1)/2$ counterclockwise peers. In this example, each trader has $2G=2$ information linkages: the signal available at the location of the i -th trader is also observed by one clockwise and one counterclockwise neighbor.

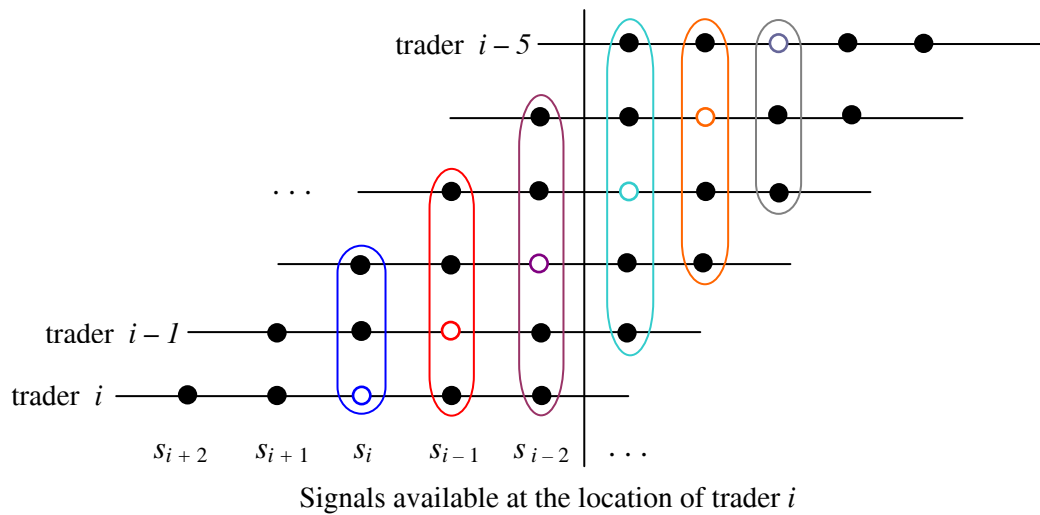


Figure 2: Overlapping information sets

This figure illustrates how signals overlap across the traders, when there are $M \geq 11$ such traders and each of them has $2G=4$ information linkages with his peers. The empty circles denote the signals available at the location of the traders. The filled circles denote the signals every trader observes on top of the signal available at his location. The signals on the left (resp. right) of the empty circles are those that are available at the location of left (resp. right) neighbors.

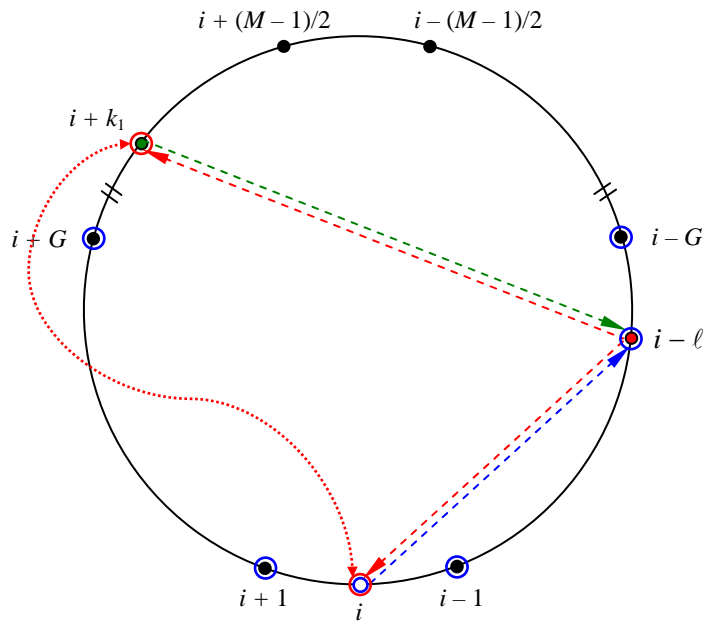


Figure 3: “Small worlds”

This figure illustrates a “double overlap” in the signals available at the location of the traders. This property arises when the number of information linkages among traders is large, i.e. $2G \geq (M-1)/2$. In this figure, the i -th trader does not observe the signal available at the location of the $(i+k_1)$ -th trader, but only the signals available at the location of all traders up to $(i+G)$ and $(i-G)$ – e.g., the signal at the $(i-l)$ -th trader’s location. Yet the $(i+k_1)$ -th trader observes the signal available at the location of the $(i-l)$ -th trader, which makes the i -th and the $(i+k_1)$ -th traders have correlated information endowments.

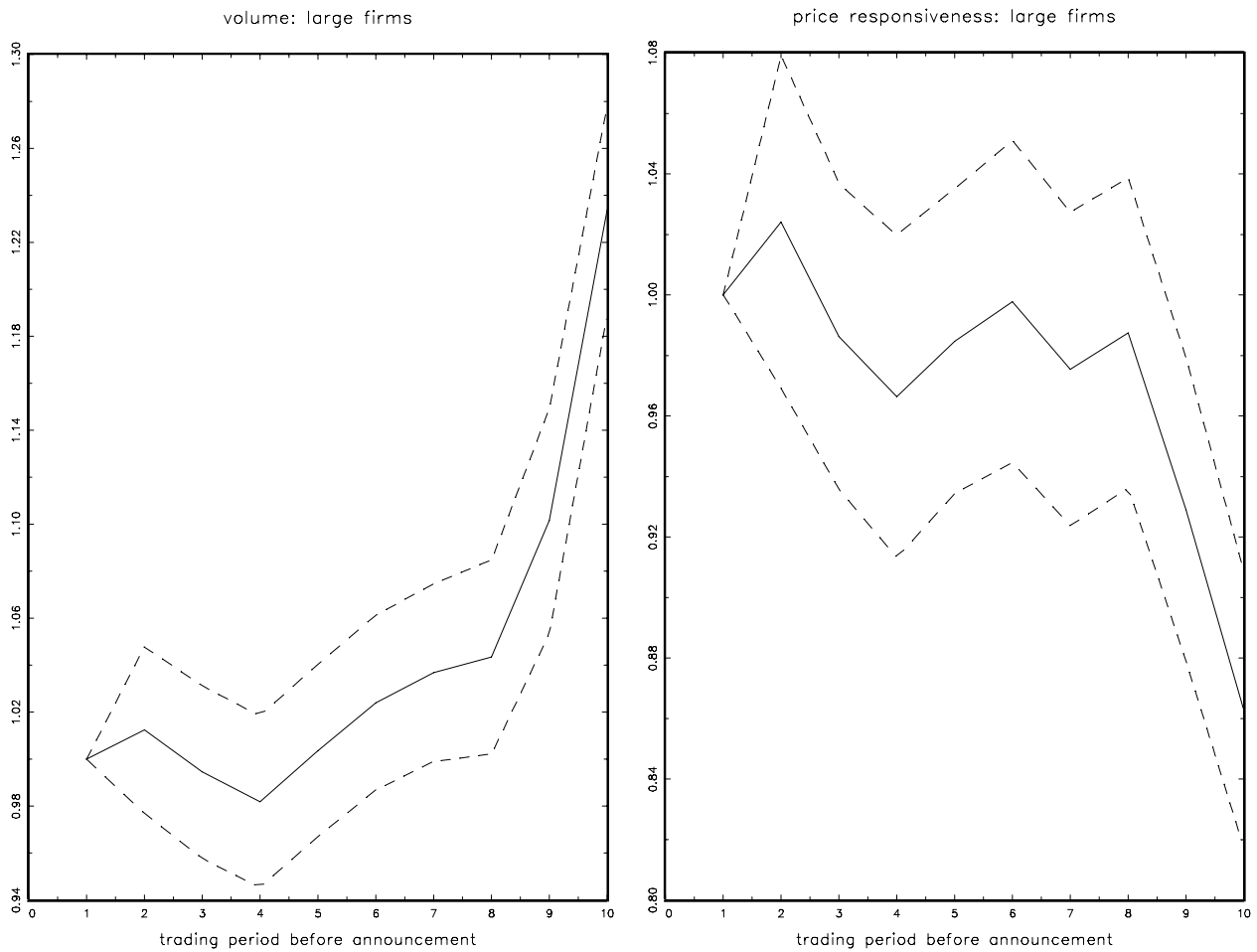


Figure 4: Volume and price impacts for large firms

Cross sectional properties of volume and price impacts for 921 firms during the fiscal year 2006. Firms are sorted by the average dollar volume over the trading period leading to a corporate earnings announcement, and by a two-way sorting procedure (along volume and price impact dimensions) that leaves the middle 50% range of the firms above the 80-th percentile for average dollar volume. The left-hand side panel displays the cross sectional averages (solid line) and 95% confidence bands (dashed lines) for volume up to, and excluding, the announcement date. The right-hand side displays cross sectional averages (solid line) and 95% confidence bands (dashed lines) for the price impacts.

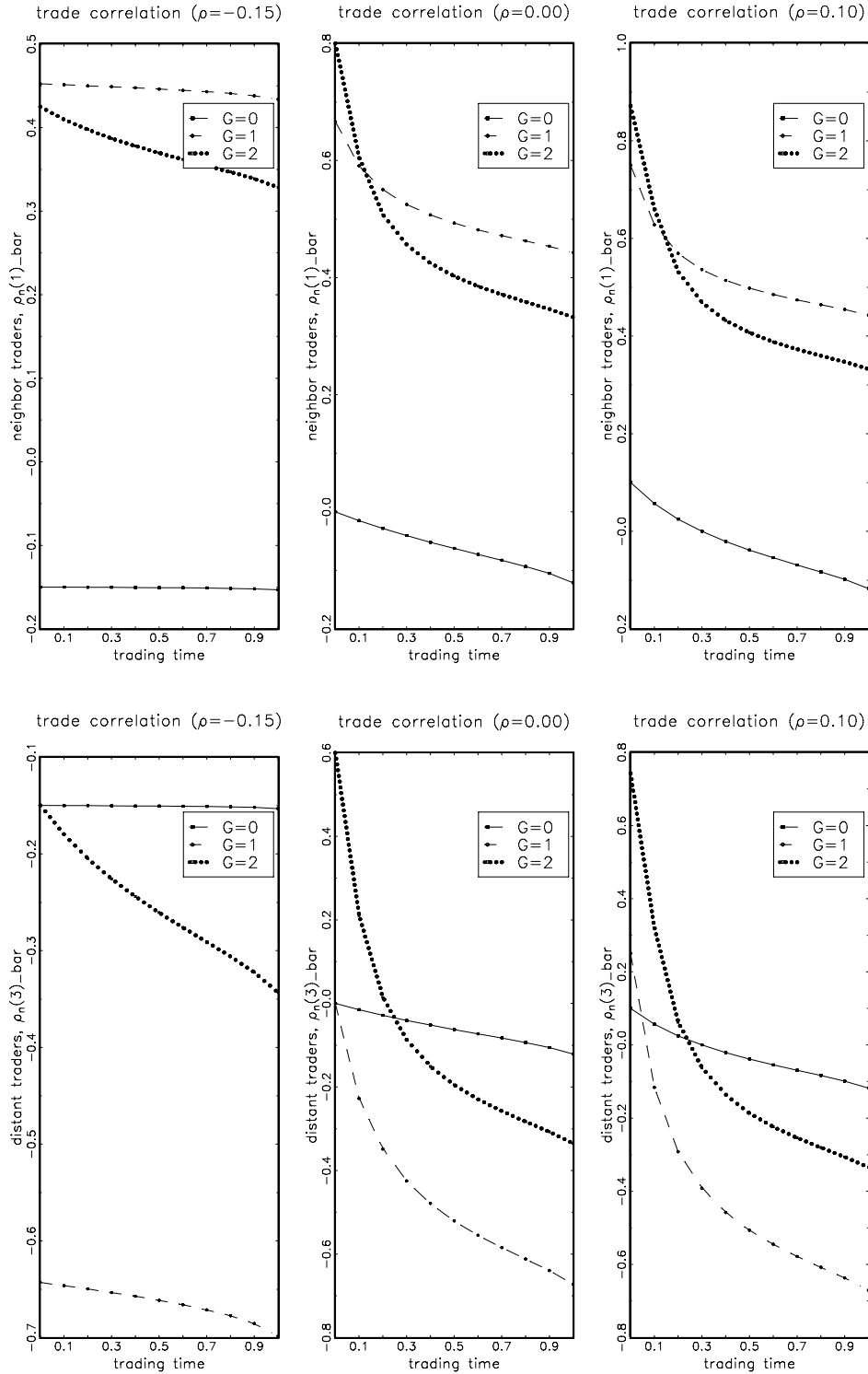


Figure 5: Heterogeneity in the correlation among trades

Correlation among trades in a market with information linkages, seven informed traders, one unit of initial variance of information, one and $\frac{1}{4}$ units of liquidity trader variance across all periods and ten trading rounds. The left-hand side panels depict the dynamics of correlation among trades arising when the initial correlation among the signals available at the traders' information linkages is negative ($\rho = -0.15$). The remaining panels depict the dynamics of correlation among trades when this correlation is zero (middle panels) and positive ($\rho = 0.10$) (right-hand side panels). The top panels depict the correlation dynamics between two close neighbors, i.e. the dynamics of trade correlation between traders i and $i-1$. The bottom panels depict the correlation dynamics between two distant traders, i.e. the dynamics of the correlation among the trade emanating from traders i and $i-3$. Each panel displays the dynamics of correlation arising when each trader has a number of information linkages equal to $2G$, with $G=0,1$ and 2 .

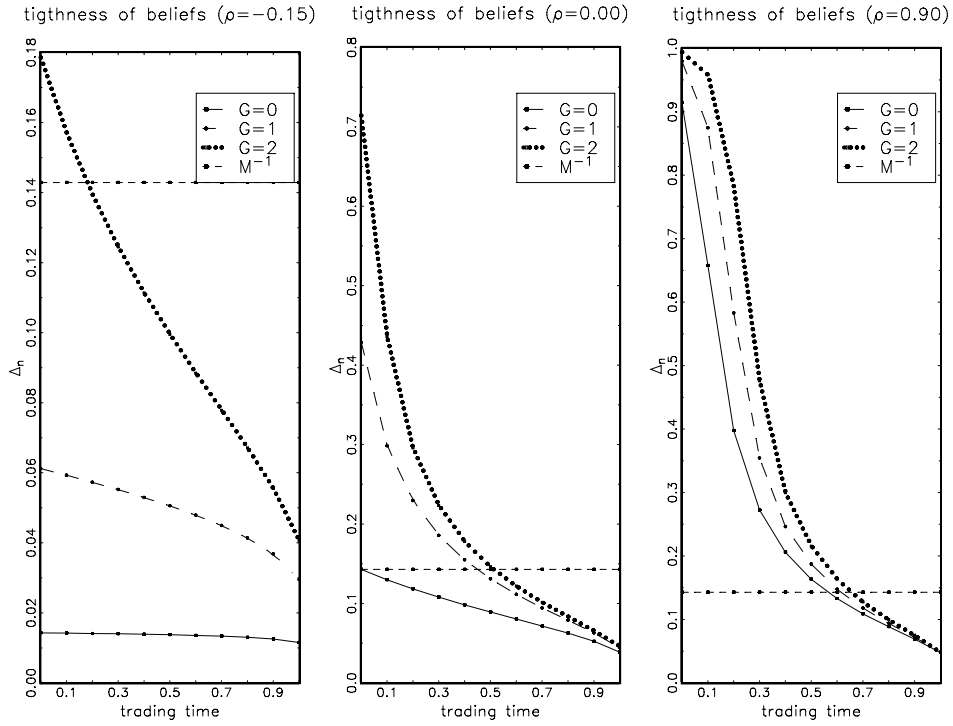


Figure 6: Tightness of beliefs

The dynamics of Δ_n , the tightness of beliefs between the market maker and any trader, in a market with information linkages, seven informed traders, one unit of initial variance of information, one and $\frac{1}{4}$ units of liquidity traders variance across all periods and ten trading rounds. The horizontal line is the inverse of the number of traders, M^{-1} . When $\Delta_n > M^{-1}$, traders engage in a rat race. When $\Delta_n < M^{-1}$, traders play a waiting game. The left-hand side panel depicts the dynamics of Δ_n when the initial correlation among the signals available at the traders' information linkages is negative ($\rho = -0.15$). The remaining panels depict the dynamics of Δ_n when this correlation is zero (middle panel) and positive, with $\rho = 0.90$ (right-hand side panel). Each panel displays the dynamics of Δ_n arising when each trader has a number of information linkages equal to $2G$, with $G = 0, 1$ and 2 .

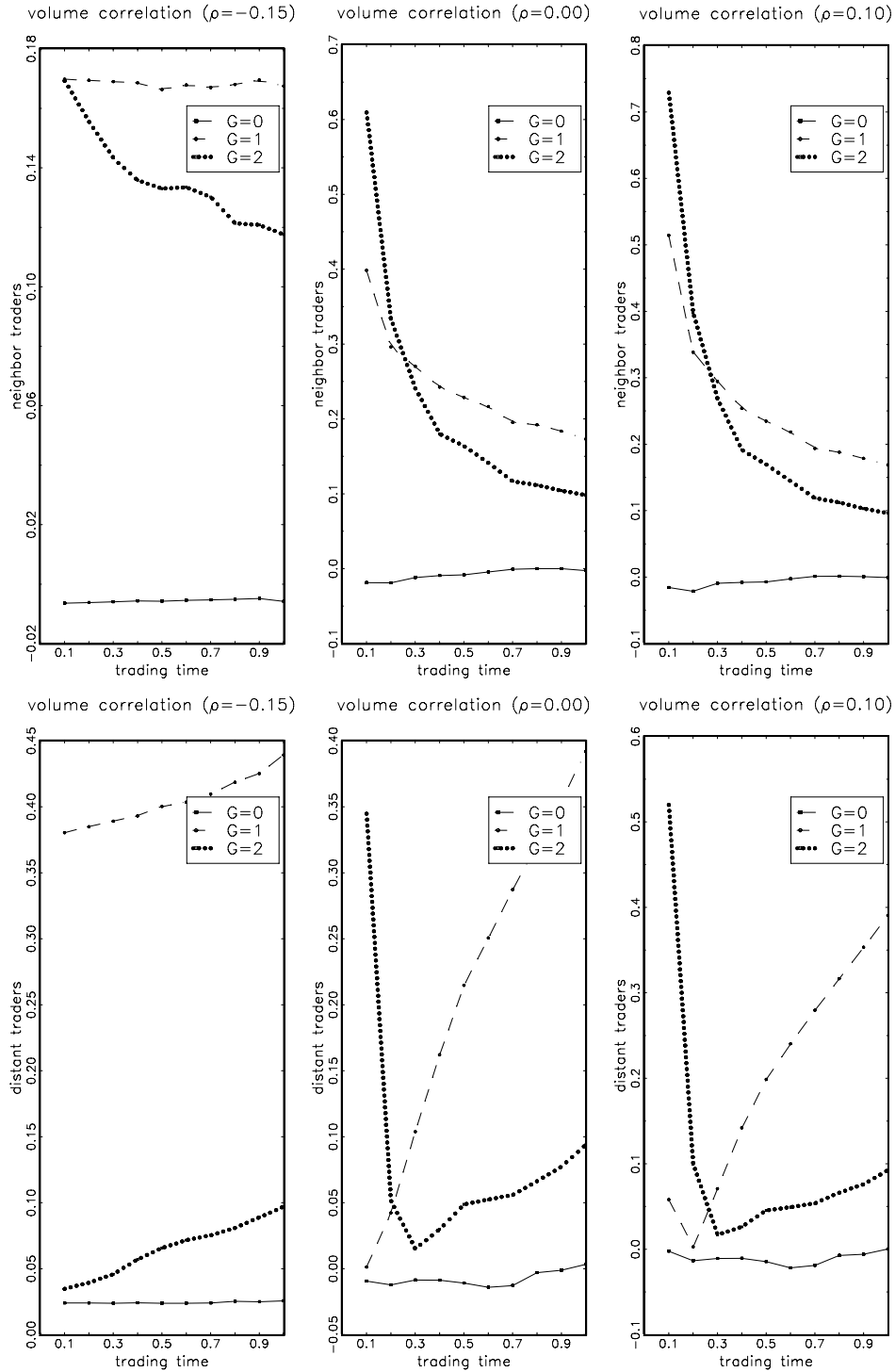


Figure 7: Heterogeneity in the correlation among volumes

Correlation among volumes initiated by informed traders in a market with information linkages, seven informed traders, one unit of initial variance of information, one and $\frac{1}{4}$ units of liquidity trader variance across all periods and ten trading rounds. The left-hand side panels depict the dynamics of correlation among volume arising when the initial correlation among the signals available at the traders' information linkages is negative ($\rho=-0.15$). The remaining panels depict the dynamics of correlation among volume when this correlation is zero (middle panels) and positive ($\rho=0.10$) (right-hand side panels). The top panels depict the correlation dynamics between two close neighbors, i.e. the dynamics of volume correlation between traders i and $i-1$. The bottom panels depict the dynamics of volume correlation between two distant traders, i.e. the dynamics of correlation among the volume generated by traders i and $i-3$. Each panel displays the dynamics of the volume correlation arising when each trader has a number of information linkages equal to $2G$, with $G=0,1$ and 2 .

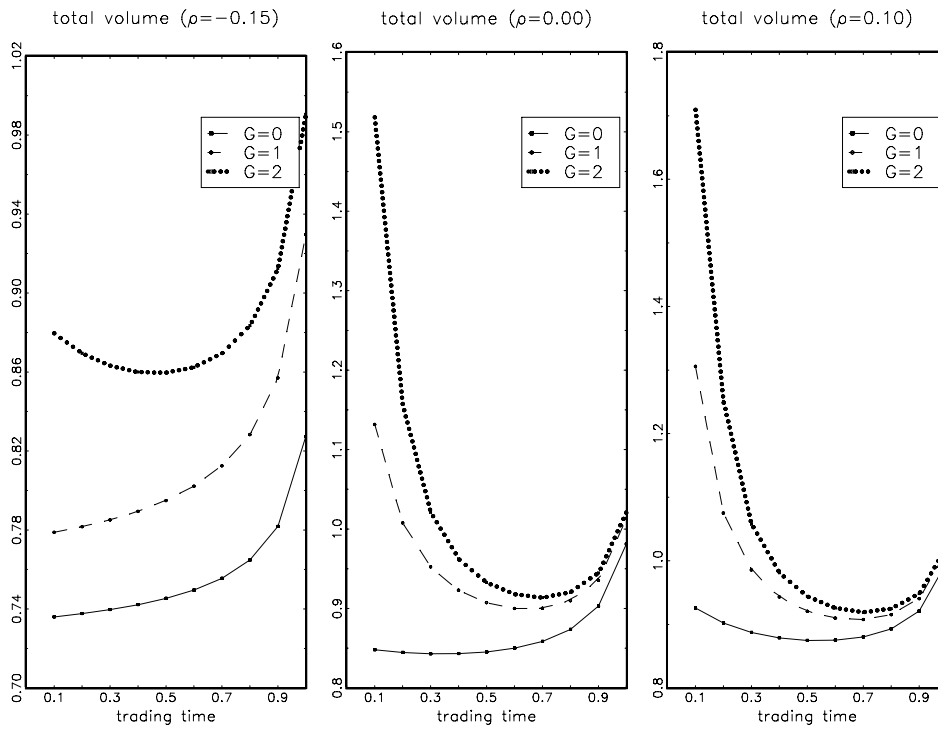


Figure 8: Total volume

The dynamics of total volume in a market with information linkages, seven informed traders, one unit of initial variance of information, one and $\frac{1}{4}$ units of liquidity traders variance across all periods and ten trading rounds. The left-hand side panel depicts the dynamics of volume when the initial correlation among the signals available at the traders' information linkages is negative ($\rho = -0.15$). The remaining panels depict the dynamics of volume when this correlation is zero (middle panel) and positive ($\rho = 0.10$) (right-hand side panel). Each panel displays the dynamics of volume arising when each trader has a number of information linkages equal to $2G$, with $G=0, 1$ and 2 .

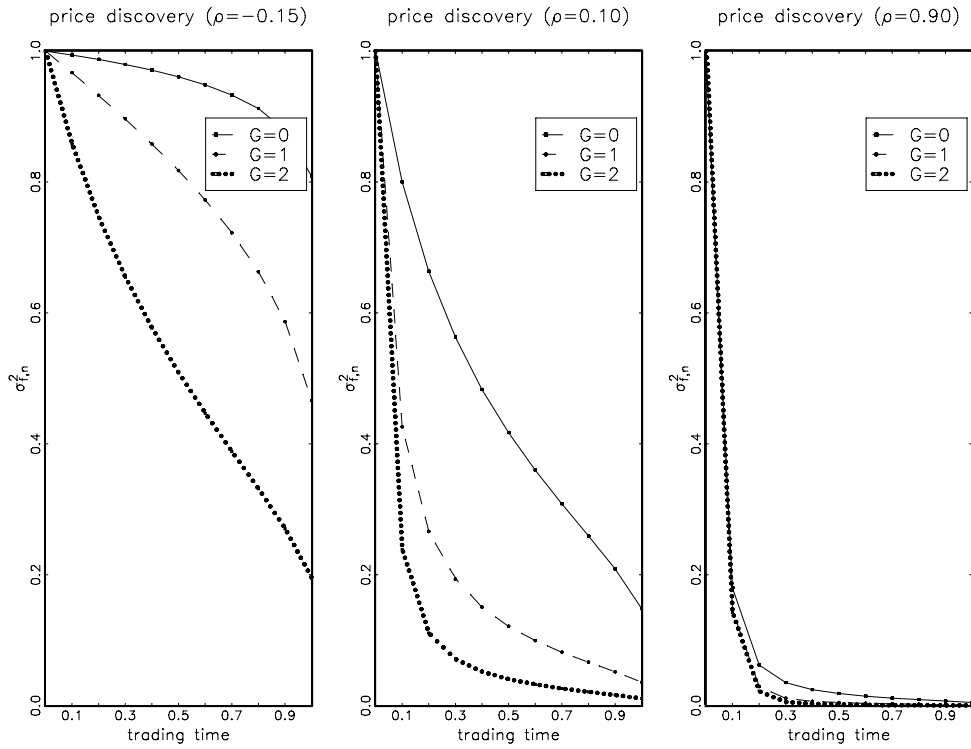


Figure 9: Efficiency

The price-discovery process in a market with information linkages, seven informed traders, one unit of initial variance of information, one and $\frac{1}{4}$ units of liquidity trader variance across all periods and ten trading rounds. The left-hand side panel depicts the price-discovery process arising when the initial correlation among the signals available at the traders' information linkages is negative ($\rho=-0.15$). The remaining panels depict price-discovery process arising when this correlation is positive but low ($\rho=0.10$) (middle panel) and high ($\rho=0.90$) (right-hand side panel). Every panel displays the price-discovery process arising when each trader has a number of information linkages equal to $2G$, with $G=0,1$ and 2 .

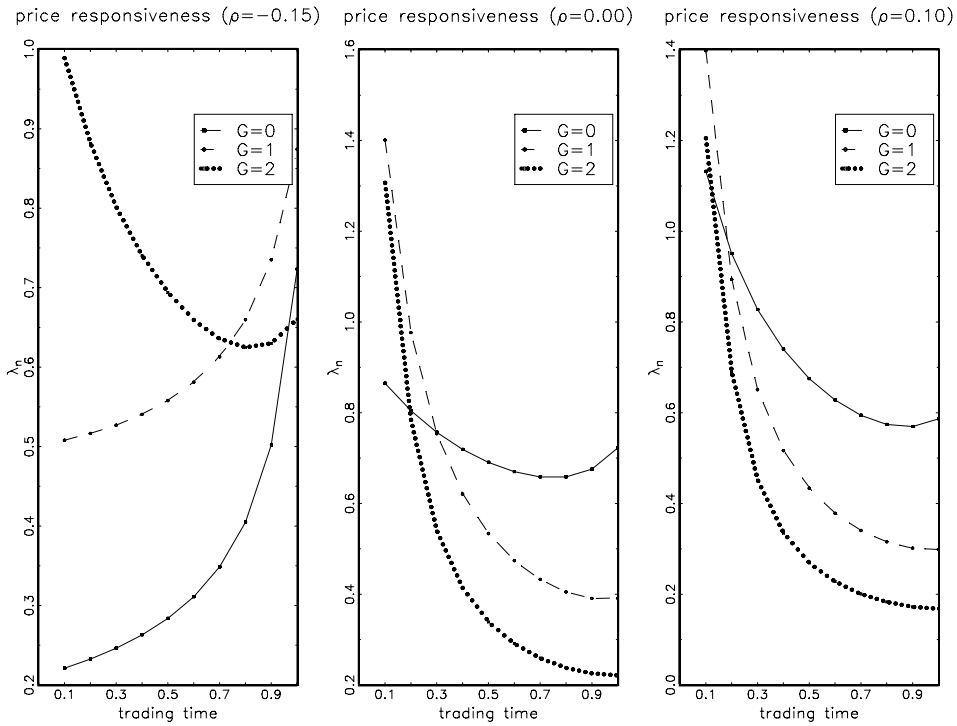


Figure 10: Price impacts

Price-impacts in a market with information linkages, seven informed traders, one unit of initial variance of information, one and $\frac{1}{4}$ units of liquidity trader variance across all periods and ten trading rounds. The left-hand side panel depicts the dynamics of the price impacts arising when the initial correlation among the signals available at the traders' information linkages is negative ($\rho=-0.15$). The remaining panels depict the dynamics of price impacts when this correlation is zero (middle panel) and positive ($\rho=0.10$) (right-hand side panel). Each panel displays the dynamics of price-impacts arising when each trader has a number of information linkages equal to $2G$, with $G=0,1$ and 2 .

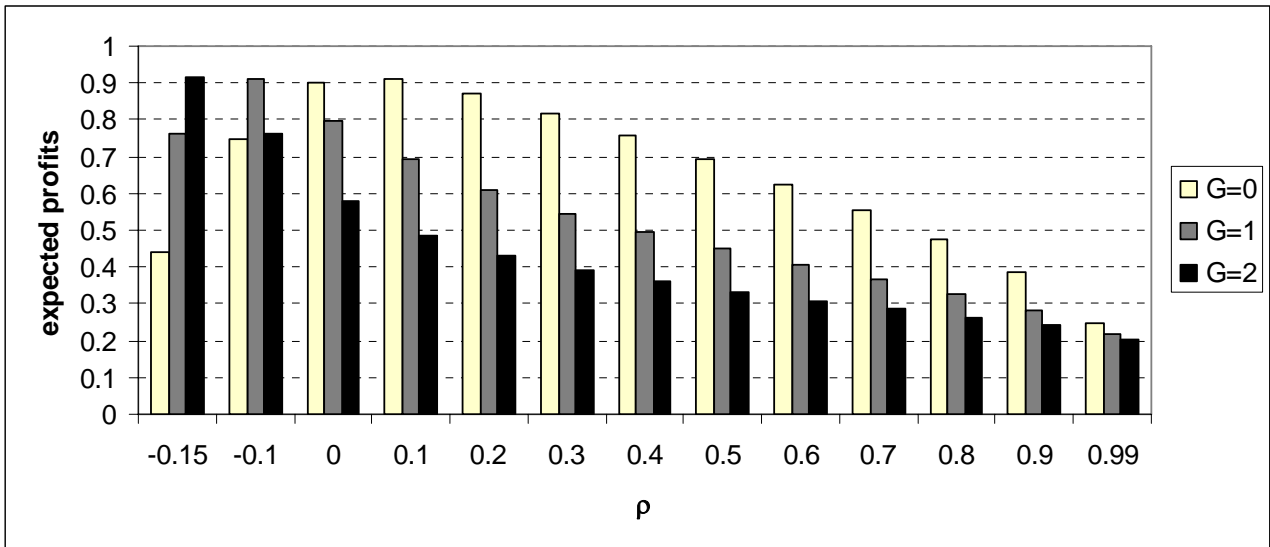


Figure 11: Traders' expected profits

The traders' expected profits in a market with seven informed traders, one unit of initial variance of information, one and $\frac{1}{4}$ units of liquidity trader variance across all periods and ten trading rounds. Displayed are the expected profits arising when each trader has a number of information linkages equal to $2G$, with $G=0,1$ and 2 ; and the initial correlations among individual signals equals $\rho=-0.15, -0.10, 0.01, 0.1, 0.2, \dots, 0.9$, and 0.99 .