THE JOURNAL OF FINANCE • VOL. LI, NO. 4 • SEPTEMBER 1996

Strategic Trading When Agents Forecast the Forecasts of Others

F. DOUGLAS FOSTER and S. VISWANATHAN*

ABSTRACT

We analyze a multi-period model of trading with differentially informed traders, liquidity traders, and a market maker. Each informed trader's initial information is a noisy estimate of the long-term value of the asset, and the different signals received by informed traders can have a variety of correlation structures. With this setup, informed traders not only compete with each other for trading profits, they also learn about other traders' signals from the observed order flow. Our work suggests that the initial correlation among the informed traders' signals has a significant effect on the informed traders' profits and the informativeness of prices.

MANY FINANCIAL AND COMMODITY markets can be characterized by a number of informed traders, each with different information. As an example, in agricultural commodity markets differing assessments of crop quality, weather conditions, and demand for products means investors will have many, distinct views about the future value of a harvest. Depending on the crop and the growing environment, differences in the opinions of traders can be significant and persistent. Such heterogeneity of private information is not limited to agricultural markets. In financial markets, information may be related to corporate acquisitions, Federal Reserve policy, patent filings, mineral finds, marketing policy, et cetera. Investors' beliefs about these values and their subsequent implications for corporate value can be varied. Because of the richness and diversity of information in actual markets, one would like to model speculative trading with private information where informed traders have disparate information, and where this disparate information takes on a variety of correlation structures. Further, one would like the model to be

* College of Business Administration, The University of Iowa, The Fuqua School of Business, Duke University, respectively. This article was first presented in October 1993 at the Fourth Annual Accounting and Finance Conference at the Washington University at St. Louis. We owe a special debt to Pete Kyle for discussions that led to this article and to Kerry Back, René Stulz (the editor), and an anonymous referee for detailed comments that have considerably improved our article. Also, we are grateful to Mark Bagnoli, Utpal Bhattacharya, Peter Carr, Bernard Dumas, Phil Dybvig, Larry Glosten, Kevin McCardle, Duane Seppi, Avanidhar Subrahmanyam, Sheridan Titman, Xavier Vives, seminar participants at Duke University, the 1994 EFA meetings at Brussels, the Fourth Annual Accounting and Finance Conference at Washington University at St. Louis, Hong Kong University of Science and Technology, The University of Iowa, The University of Missouri, University of Notre Dame, and the 1994 WFA meetings at Santa Fe for their suggestions and comments. Viswanathan thanks the Business Associates Fund, the Isle Maligne Fund, and the Futures and Options Research Center (FORCE) of the Fuqua School of Business for financial support.

dynamic to address the endogenous evolution of beliefs, and to relate this evolution of beliefs to trading volume and the informativeness of prices.

Modeling markets with heterogeneous information among the traders becomes complex very quickly because traders infer the value of an asset from not only their own private information, but also using any information revealed by other traders through trading. For example, an informed trader may increase her estimate of the value of a security if she notices unusually large purchases of the asset that are unlikely to come from liquidity (uninformed) motivated trades. Any model that incorporates differential information in this setting must also address the inferences that traders make once trading occurs.

This richer, and more realistic, view of a market raises a number of interesting and fundamental economic questions. How is information incorporated into prices in a market with differentially informed traders? How quickly do prices reflect the information known jointly by all informed traders? How do informed traders learn about each others' information? How might inferences of the informed traders affect market maker behavior? What happens to trading volume and the volatility of prices? What happens to informed traders' profits, and hence the resources they will expend to become informed? How does competition among informed traders change when no information is to be released publicly for many periods?

While the role of information in price formation has long been a central concern of finance researchers, there have been few models of the complexities that arise with speculative trading when traders have heterogeneous information. In particular, extensions of the Kyle (1985) monopolistic trader model assume that all the informed traders have identical information and thus are unable to address the issues that arise with heterogeneously informed traders (see Michener and Tighe (1991), Holden and Subrahmanyam (1992), and Foster and Viswanathan (1993)).[1] These articles show that with identical information, informed traders compete very aggressively, and most of the information is impounded in prices within a few trading periods (we refer to this intense competition as a "rat race"). These papers indicate a dichotomy between strategic trading with a monopolistically informed trader (analyzed in Kyle (1985)) where information is released slowly over time, and strategic trading with oligopolistically informed traders where information is released almost instantly.

However, it is not clear how critical the identical information assumption is for these results. It may be that if informed traders have different signals, competition would not be so aggressive in an oligopolistic setting. If this is the case, a better understanding of the economics of competition with disparate information is needed. The goal of this article is to develop a model to analyze these questions. Our primary focus is on the inference problem facing the market maker and informed traders when informed traders have different initial information about the value of an asset.

[1] See also Dutta and Madhavan (1993) for a related identical information model.

Our analysis is broken into two major sections. In the first, we develop a generalization of Kyle's (1985) monopolist-informed trader model, where we assume that there are a number of informed traders, each with a unique information endowment. In this market, informed traders learn from the order flow, and use the order flow and their private information (which includes a history of their own trades) to forecast the forecasts of other informed traders. Normally, discrete time models where traders forecast the forecasts of others about the liquidation value of the asset are plagued by the problem of an increasing state history with time. This is due to a regress in expectations as follows: traders forecast the forecasts of other traders; forecast the forecasts that other traders make of their forecasts; and so on (see Townsend (1983a, 1983b), Marcet and Sargent (1989a, 1989b), and Singleton (1987) for more on these issues). To avoid this problem, we use an approach that (in the context of our model) yields a sufficient statistic for the trading history and avoids this regress in expectations, both when all traders play their equilibrium strategies and when one trader deviates from his equilibrium strategy given the equilibrium strategies of others. Because of this, our work is similar to recent articles by He and Wang (1995) and Vayanos (1993). However, He and Wang (1995) consider a competitive model with a continuum of traders and, as a consequence, ignore strategic trading. Cao (1994) solves a model similar to our own.

The technique we employ to solve the inference problem in this article can be used in a variety of other settings in economics and finance. First, it has potential applications in models of industrial organization, where firms learn about the unknown cost structure of the industry (see Vives (1988)) either in markets with homogeneous products (Cournot competition) or differentiated products (Bertrand competition). Second, the approach of this article can be directly used to model a dynamic limit order model (thus extending Kyle (1989)). Finally, this paper has potential applications in the rate of convergence to rational expectations literature that has traditionally used a continuum, as opposed to a finite number, of agents (see Vives (1993) for an example of the continuum of agents approach).

In the second section of the article, we use the model to solve numerically a number of examples. These examples provide important insights about how the information structure influences the dynamic competition among informed traders. The results in this setting are quite different from those in prior work that has assumed informed traders have identical information. As noted earlier, in models of competition with identical information, as the number of trading periods is increased (holding the liquidity trading constant), there is more intense trading in early periods, and the total profits to informed traders fall. With heterogeneous information, we find that with a small initial positive correlation between the private signals of the informed traders, increasing the number of trading periods (while holding the total liquidity trading constant) leads to greater profits to informed traders. In fact, total informed trader profits may be close to the maximum profit of a monopolist-informed trader.

We find that trading outcomes depend critically on the initial correlation of informed traders' private signals. With heterogeneous information, each trader has some degree of monopoly power, because part of his information is known only to him. This reduces the degree of competition between traders, which provides an incentive to trade less aggressively. In addition, the correlation between the signals of the informed traders falls deterministically as more trading occurs. After a number of trading periods, this yields a negative correlation between traders' private information (more precisely between the residual information known by each informed trader relative to the information known by the market maker). This is akin to a difference of opinion among the informed traders about the value of the asset.[2] This difference of opinion arises because the market maker is learning more accurately about the average signal and less accurately about any individual signal; hence, the remaining private information becomes more disparate through time. Knowing this will happen, every informed trader has an incentive to make smaller trades, while hoping that other informed traders make large trades. We refer to this standoff characteristic of the market as the "waiting game." The "waiting game" cannot occur with identically informed traders, but appears to be an intriguing feature of markets with differing information.

Through a number of numerical examples, we find that if the initial correlation between the signals is not too high, the competitive ("rat race") incentive of positive correlation occurs in very early periods, followed by a period with negatively correlated signals and a tendency of traders to then play the "waiting game." This means that there is less competition and greater profits to the informed traders than would be found with highly correlated private information. We also use numeric solutions to the model to explore the implications of this phenomenon for trading costs, volatility, and the profits for informed traders.

Our work has a number of important empirical implications. First, it suggests that the cross-sectional correlation of information with analysts is an important determinant of trading patterns. While Brennan, Jegadeesh, and Swaminathan (1993) test the relation between price informativeness and the number of analysts, further explorations (using analyst data) of the richer implications of our model seem worthwhile. Second, if new information arrives daily (as suggested by the work of Cho (1995)), then different cross-sectional correlations of information can potentially explain why there is a U-shaped pattern in bid-ask spreads in the stock market and why this U-shaped pattern may vary across stocks. Also, differences in cross-sectional correlation may explain why the options market exhibits declining bid-ask spreads over the day, while the stock market exhibits a U-shaped pattern (see Chan, Chung, and Johnson (1995) for this evidence). Third, the model adds to the new literature that seeks to explain the volume-volatility relationship using structural models of speculative trading (see Easley, Kiefer, and O'Hara (1993) and Foster and Viswanathan (1995)). Considering dynamic trading with heteroge-

[2] This is in contrast to other models with heterogeneous beliefs (Harris and Raviv (1993) and Wang (1993)), where beliefs are exogenously specified to be different and continue to differ over time.

neous information may result in a better structural model to explain the volume-volatility relation.

The article is organized as follows. Section I presents the model. Section II discusses our equilibrium concept, and Section III shows how we solve the inference problem in forecasting the forecasts of others when all traders follow their equilibrium strategies. Section IV resolves the dimensionality issue when one trader deviates from his optimal strategy, keeping the equilibrium strategies of other traders fixed. In addition, this section provides necessary and sufficient conditions for an equilibrium. Section V presents a partial existence result and an algorithm to compute the equilibrium. Section VI discusses some numerical examples, and Section VII concludes.

I. The Model

In this section of the article, we outline a model based on Kyle (1985). We first list our assumptions and detail the basic features of the model. After introducing the market participants, we describe the information structure used and derive some basic measures for the amount of information stocks available to informed traders.

Consider a market with risk-neutral traders; a market maker, M informed traders, and a number of liquidity traders. These traders buy and sell a single security over N periods. At the beginning of the first period, informed traders receive a private signal that constitutes their initial endowment of information. Based on this initial information, informed traders place orders with the market maker. Each period the market maker sees the total order flow from the informed and liquidity traders, and not knowing which orders come from which traders, announces an adjustment to the current share price at which all orders are filled. The price at which the market maker fills orders is determined by a competitive process. No additional information is revealed, so informed traders trade based on the information that they receive in the initial period and what they learn from watching the order flow. At the end of the last period, the liquidation value of the asset, v, is announced, and holders of the asset are paid.

More specifically, the i^{th} informed trader receives a signal s_{i0} at the start of trading. The signal vector for all informed traders, $(s_{10}, s_{20}, \ldots, s_{M0})$, is drawn from a multivariate normal distribution with mean zero and a variance-covariance matrix Ψ_0. The form of this and all other distributions is known to all market participants. We assume that the matrix Ψ_0 is such that the variance of each signal s_{i0} is Λ_0 and the covariance between any two signals s_{i0} and s_{j0} is Ω_0 (some additional restrictions are needed for Ω_0 and Λ_0 to ensure that the variance-covariance matrix Ψ_0 is positive definite—see equation (11)). The true value of the asset v is normally distributed with mean p_0 and variance σ_v^2. The covariance vector between the signal vector $[s_{10}, \ldots, s_{M0}]$ and the true value of the asset v is defined to be Δ_0, where each element of this vector is c_0. The assumption of symmetry, i.e., that all traders have the same initial variance of information Λ_0, same cross covariance between signals, Ω_0, and the same covariance between signals and true value, c_0, is critical for the analysis that follows.

These assumptions imply:

$$E[v - p_0 | s_{10}, \ldots, s_{M0}] = \Delta_0'[\Psi_0]^{-1} \begin{pmatrix} s_{10} \\ \vdots \\ s_{M0} \end{pmatrix} \tag{1}$$

The symmetrical distribution assumption on informed traders means that the vector $\Delta_0'[\Psi_0]^{-1}$ has identical elements $\kappa = c_0(\Lambda_0 - \Omega_0)/(\Lambda_0[\Lambda_0 + (M-2)\Omega_0] - (M-1)\Omega_0^2)$ and (see details of derivation in Appendix):

$$E[v - p_0 | s_{10}, \ldots, s_{M0}] = \kappa \sum_{i=1}^{M} s_{i0} = \kappa M \left(\frac{1}{M} \sum_{i=1}^{M} s_{i0} \right) = \theta \hat{v} \tag{2}$$

where $\theta = \kappa M$ and $\hat{v} = (1/M)\Sigma_{i=1}^{M} s_{i0}$. Thus the average of the signals, $\hat{v}$, is a sufficient statistic for the information known to all informed traders about the liquidation value of the asset. This means that the market maker and informed traders need be concerned only about inferring the average of the signals from the order flow. It is this fact that allows us to simplify the dynamic inferences of these traders.

In addition to the informed traders there are liquidity traders, whose trade in period n is u_n, the realization of a normally distributed random variable with mean zero and variance σ_u^2. We assume that the liquidity trading is independent of all other random variables; hence, nothing can be learned about the liquidation value of the asset from the liquidity traders' orders.

The last market participant is a market maker who observes the total order flow and sets the price equal to the conditional expected value of the asset, based on the history of orders received up to and including that trading time. If we define the i^{th} informed trader's order at time n to be x_{in}, then the market maker observes the total order flow, $y_n = \Sigma_{i=1}^{M} x_{in} + u_n$ and sets the price at time n so that:

$$p_n = E[v | y_1, \ldots, y_n]. \tag{3}$$

After n periods of trading, the market maker has observed order flows $(y_1, \ldots, y_n)$ and updates his estimate of the signal of the i^{th} trader, s_{i0}, to

$$t_{in} = E[s_{i0} | y_1, \ldots, y_n]. \tag{4}$$

The market maker's update of his estimate of the value of the asset, v, is related to the updating of the signals, s_{i0}, $i = 1, \ldots, M$ according to

$$p_n = E[v | y_1, \ldots, y_n] \doteq \frac{\theta}{M} \sum_{i=1}^{M} t_{in} \tag{5}$$

The last expression in equation (5) follows from the law of iterated expectations and the definition of $\hat{v}$ (a detailed derivation is in the Appendix). Intuitively, the expression follows from the fact that the average of the signals, $\hat{v}$, is a sufficient statistic for the information available with all the informed traders. Hence, the market maker's updating of the true value of the asset is linearly related to the market maker's updating of the vector of signals that is the private information in the model.

The i^{th} informed trader, after each round of trading, observes the order flow y_n and knows his own trade x_{in}. Thus, the informed trader observes

$$z_{in} = y_n - x_{in} \tag{6}$$

and makes a better inference than the market maker about the signals of other traders. Here, z_{in} is the order flow that comes from all the traders excluding the i^{th} trader. This means that the informed trader's information after $n - 1$ periods of trading contains the order flow up to that period $(y_1, \ldots, y_{n-1})$, his original signal s_{i0}, and the net of orders he did not submit $(z_{i1}, \ldots, z_{in-1})$. Based on this information, the i^{th} informed trader forecasts the true value of the stock, v, and the trading decisions of the other traders in the next period, and then decides on his own trade, x_{in}.

Finally, we define the following variables to measure the remaining information at the end of period n (in this context the stock of information available to traders is measured as a variance).

$$\begin{aligned}
\Sigma_n &= \mathrm{Var}(\theta\hat{v}|y_1, \ldots, y_n) = \mathrm{Var}(\theta\hat{v} - p_n) \\
\Lambda_n &= \mathrm{Var}(s_{i0}|y_1, \ldots, y_n) = \mathrm{Var}(s_{i0} - t_{in}) = \mathrm{Var}(s_{in}) \qquad (7) \\
\Omega_n &= \mathrm{Cov}(s_{i0}, s_{j0}|y_1, \ldots, y_n) = \mathrm{Cov}(s_{i0} - t_{in}, s_{j0} - t_{jn}) = \mathrm{Cov}(s_{in}, s_{jn})
\end{aligned}$$

where $s_{in} = s_{i0} - t_{in}$. Here, Σ_n is the conditional variance (given the market maker's information after n periods of trading) of the conditional expectation of the liquidation value given all available information, $\mathrm{E}[v|s_{i0}, \ldots, s_{iM}] = \theta\hat{v}$. Λ_n is the conditional variance of trader i's signal, s_{i0} (given the market maker's information after n periods of trading). Finally, Ω_n is the conditional covariance between any two signals, s_{i0} and s_{j0} (given the market maker's information after n periods of trading).

These variance-covariance measures are related as follows. First note the following relation between Σ_n, Λ_n and Ω_n (see Appendix for derivation),

$$\Sigma_n = \frac{\theta^2}{M}[\Lambda_n + (M-1)\Omega_n]. \tag{8}$$

Also, the symmetric distribution assumption for the informed traders' signals means the covariance between each individual signal and the equilibrium order flow is identical. This imposes the following restriction on the updating process (see Appendix for details):

$$\begin{aligned}\Lambda_{n-1} - \Lambda_n &= \Omega_{n-1} - \Omega_n \\ \Sigma_{n-1} - \Sigma_n &= \theta^2[\Lambda_{n-1} - \Lambda_n]\end{aligned} \tag{9}$$

Hence the evolution of Σ_n, Λ_n, and Ω_n are intimately related. Equation (9) implies

$$\Lambda_n - \Omega_n = \chi, \quad \text{independent of } n. \tag{10}$$

Thus the difference between variances and covariances of the updated signals is independent of the trading period.

The market maker learns about the average signal, $\hat{v} = (1/M)\Sigma_{j=1}^{M} s_{i0}$ over time. After N rounds of trading, there is still a positive value for the variance of the market maker's conditional expectation of the liquidation value of the asset (with a finite number of trading periods the average signal cannot be known perfectly). Thus, $\mathrm{Var}(\hat{v}|y_1, \ldots, y_N) > 0$.[3] But

$$\begin{aligned}\mathrm{Var}(\hat{v}|y_1, \ldots, y_N) &= \frac{1}{M}[\Lambda_N + (M-1)\Omega_N] && \text{using Equation (8)} \\ &= \Lambda_N - \frac{M-1}{M}\chi. && \text{using Equation (10)}\end{aligned} \tag{11}$$

Hence, the conditional variance is greater than zero if and only if $\Lambda_N > [(M-1)/M]\chi$. This also means that $\Omega_N > -(1/M)\chi$.[4] This calculation also implies that the conditional correlation between individual signals (conditional on the order flow) must eventually be negative if the information in the average signal $\hat{v}$ is to be known accurately by the market maker. It is interesting to note that the market maker never knows accurately any individual signal, even after an infinite number of trading periods. However, this is irrelevant, as the average signal is a sufficient statistic (for the purpose of predicting v) for all of the information known by the informed traders.

The fact that the conditional correlation between the signals of the informed traders (conditional on the order flow up that point in time) must eventually be negative has important consequences for informed traders' strategies. We will return to this issue when we consider numerical solutions to the model.

[3] The proof of this fact is as follows. If $\mathrm{Var}(\hat{v}|y_1, \ldots, y_N) = 0$, then there is a period $k < N$ such that $\Sigma_k > 0$, $\Sigma_{k+1} = 0$. This is the last period in which the informed trader trades. But with $\Sigma_k > 0$ and a one-period model, we know that $\Sigma_{k+1} > 0$ from Admati and Pfleiderer (1988), which is a contradiction.

[4] $\Omega_0 > -(1/M)\chi$ ensures that the variance covariance matrix Ψ_0 is positive definite.

II. The Equilibrium Concept

In this section we introduce and discuss the equilibrium notion used in the article. To begin, we define conditions to be satisfied for a Bayesian Nash equilibrium for this model. Then we restrict our search to linear Markov equilibria and conjecture equilibrium strategies for the market maker and informed traders.

We follow Kyle (1985) in our definition of an equilibrium. As discussed in equation (6), the market knows his own order x_{in} and the net order flow y_n. Hence

$$z_{in} = y_n - x_{in}. \tag{12}$$

is the order flow not submitted by the i^{th} informed trader and is used by trader i to learn from the order flow about the liquidation value of the asset.

Just before period n, the i^{th} informed trader's private information consists of his own signal, s_{i0}, plus what he alone learns from the orders up to that point in time, $(z_{i1}, \ldots, z_{in-1})$. In addition, all traders know the past net trades, $(y_1, \ldots, y_{n-1})$. Let

$$x_{in} = X_{in}(s_{i0}, z_{i1}, \ldots, z_{in-1}, y_1, \ldots, y_{n-1}) \tag{13}$$

represent the optimal strategy of the i^{th} informed trader (where the choice of i is arbitrary).

Let

$$p_n = P_n(y_1, \ldots, y_n) \tag{14}$$

represent the optimal strategy of the market maker.

Let $X_i = \langle X_{i1}, \ldots, X_{iN}\rangle$ (for each i) and $P = \langle P_1, \ldots, P_N\rangle$ represent the vector of strategy functions. Define the profit that accrues to the i^{th} informed trader from period n on (given the above strategy functions for traders $i = 1, \ldots, M$ and the market maker) as:

$$\pi_n(X_1, \ldots, X_i, \ldots, X_M, P) = \sum_{k=n}^{N} (v - p_k)x_{ik} \tag{15}$$

A Bayesian Nash equilibrium of the trading game is a $M + 1$ vector of strategies $(X_1, \ldots, X_M, P)$ such that (we follow Kyle (1985) closely here):

1. For any $i = 1, \ldots, M$ and all $n = 1, \ldots, N$ we have for $X_i' = \langle X_{i1}', \ldots, X_{in-1}', X_{in}', \ldots, X_{iN}'\rangle$, we have

$$\begin{aligned} &\mathrm{E}[\pi_n(X_1, \ldots, X_i, \ldots, X_M, P)|s_{i0}, z_{i1}, \ldots, z_{in-1}, y_1, \ldots, y_{n-1}] \\ &\geq \mathrm{E}[\pi_n(X_1, \ldots, X_i', \ldots, X_M, P)|s_{i0}, z_{i1}, \ldots, z_{in-1}, y_1, \ldots, y_{n-1}], \end{aligned} \tag{16}$$

i.e., the optimal strategy function for informed trader i is best no matter which past strategies (or strategy functions) i may have played.

2. For all $n = 1, \ldots, N$, we have

$$p_n = \mathrm{E}[v | y_1, \ldots, y_n], \tag{17}$$

i.e., the market maker sets prices equal to the conditional expected value given the order flow.

This equilibrium concept is the Bayesian Nash equilibrium and is based on a dynamic programming argument. Note that the strategy of the i^{th} informed trader in the n^{th} round is required to be the optimal strategy, not only when trader i plays his optimal strategy in the first $n - 1$ periods, but also when he plays any arbitrary strategy in the first $n - 1$ periods. However, there are no off equilibrium observations of order flow by the other informed traders in the model (even when trader i deviates from his optimal strategy) as every possible order flow is observed. Consequently, we do not have to concern ourselves with the issue of how to assign off equilibrium beliefs.

Given the model structure described above, we are interested in linear Markov equilibria. That is, we search for an equilibrium where the demands of the informed traders, market maker learning about the signal vector, and market maker learning about the true value of the asset take the form:[5]

$$\begin{aligned} x_{in} &= \beta_n s_{in-1} \\ t_{in} &= t_{in-1} + \zeta_n y_n \\ p_n &= p_{n-1} + \lambda_n y_n \end{aligned} \tag{18}$$

While the definition of t_{in} is not a part of the general requirement of an equilibrium, it is crucial for the linear equilibrium we focus on. This is because an essential part of the linear equilibrium is an incorporation of the linear forecasts of the forecasts of other traders.

The parameters ζ_n and λ_n are related in the following way. First, $t_{in} = \mathrm{E}[s_{i0} | y_1, \ldots, y_n]$. Thus

$$\mathrm{E}[s_{i0} | y_1, \ldots, y_n] - t_{in-1} = \mathrm{E}[s_{in-1} | y_1, \ldots, y_n] = \mathrm{E}[s_{in-1} | y_n] = \zeta_n y_n \tag{19}$$

and

$$\mathrm{E}\left[\frac{1}{M} \sum_{i=1}^{M} s_{in-1} \middle| y_n \right] = \zeta_n y_n. \tag{20}$$

[5] This is not the complete specification of the strategy function, as it is only the strategy that is played in equilibrium. At this stage, we do not describe the optimal strategy given prior nonoptimal play by trader i. We discuss the optimal strategy given prior nonoptimal play in Section IV.

But

$$\hat{v} - \frac{1}{\theta} p_{n-1} = \sum_{i=1}^{M} s_{in-1} \tag{21}$$

and hence

$$\mathrm{E}[v - p_{n-1}|y_1, \ldots, y_n] = \mathrm{E}[\theta\hat{v} - p_{n-1}|y_n] = \theta\zeta_n y_n. \tag{22}$$

Therefore

$$\lambda_n = \theta\zeta_n; \tag{23}$$

hence the market maker's updating of the true value of the asset is given by his updating about an individual signal and by the parameter θ.

III. The Dimensionality Issue

In this section we show how the dimensionality issue is resolved (i.e., we avoid the problem of increasing state history with time). We are interested in the case of linear strategies for informed traders and learning by the market maker as given by equation (18).

Consider trader i who is interested in forecasting the true value of the asset that is not predicted by the market after $n - 1$ periods of trading, using his information $(s_{i0}, y_1, y_{n-1}, x_{i1}, \ldots, x_{in-1})$. By equation (18), for $r = 1, \ldots, n - 1$:

$$x_{ir} = \beta_r s_{ir-1} = \beta_r(s_{i0} - t_{ir-1}) = \beta_r\left(s_{i0} - \sum_{s=1}^{r-1} \zeta_s y_s\right) \tag{24}$$

Hence $(x_{i1}, \ldots, x_{in-1})$ is redundant and the meaningful history for trader i is just $(s_{i0}, y_1, \ldots, y_{n-1})$. It is important to note that this only holds when trader i has played his conjectured optimal strategy in past periods. In developing this result, we exploited the fact that past optimal strategies are functions of the private signal s_{i0}, and the order flow up to that point. Hence, any additional information (apart from the initial private signal, s_{i0}) that the informed trader has relative to the market maker is redundant.[6]

The redundancy of $(x_{j1}, \ldots, x_{jn-1})$ in equilibrium for trader j has a key implication on how trader i attempts to manipulate the beliefs of trader j. Because the equilibrium strategy of trader j only depends on his signal and the observed market order flow (and not his own past order $(x_{j1}, \ldots, x_{jn-1})$), trader

[6] It is interesting to compare this result with the hierarchical information structure studied by Foster and Viswanathan (1994). With their asymmetric information structure, any additional information that the lesser informed trader learns from the order flow (relative to market makers) is not redundant.

i can only manipulate trader j's beliefs about the true value of the asset via the publicly observed order flow. This simplifies the strategic gaming in the model, because trader i does not have to worry about trader j's ability to learn more than the market maker.

Thus trader i predicts $v - p_{n-1}$ as follows:

$$\begin{aligned} &E[v - p_{n-1}|s_{i0}, y_1, \ldots, y_{n-1}, x_{i1}, \ldots, x_{in-1}] \\ &= E[v - p_{n-1}|s_{i0}, y_1, \ldots, y_{n-1}] = E[v - p_{n-1}|s_{in-1}, y_1, \ldots, y_{n-1}] \\ &\qquad = E[v - p_{n-1}|s_{in-1}] = \eta_n s_{in-1} \end{aligned} \tag{25}$$

so s_{in-1} is a sufficient statistic for trader i to predict the value of the asset (using information not known to the market maker) after $n - 1$ trading periods. Hence we refer to s_{in-1} as the forecast of trader i after n trading periods. In deriving equation (25), we first use the fact, shown in equation (24), that in equilibrium $(x_{i1}, \ldots, x_{in-1})$ is redundant given $(s_{i0}, y_1, \ldots, y_{n-1})$. Then we exploit the result that when we project s_{i0} on $(y_1, \ldots, y_{n-1})$ to obtain s_{in-1}, both $v - p_{n-1}$ and s_{in-1} are independent of $(y_1, \ldots, y_{n-1})$.

Trader i is also interested in predicting the forecasts of other informed traders, because other traders use these forecasts to submit orders, which in turn determine the market price. Because the j^{th} trader submits an order of the form $x_j = \beta_j s_{jn-1}$, trader i needs to predict s_{jn-1}. However,

$$\begin{aligned} &E[s_{jn-1}|s_{i0}, y_1, \ldots, y_{n-1}] \\ &\qquad = E[s_{jn-1}|s_{in-1}, y_1, \ldots, y_{n-1}] = E[s_{jn-1}|s_{in-1}] = \phi_n s_{in-1} \end{aligned} \tag{26}$$

so trader i's forecast of other traders' forecasts is also a linear function of s_{in-1}. By induction, the n^{th} order forecast of the forecast of others is also a linear function of s_{in-1}. Therefore s_{in-1} is a sufficient statistic, and there is no history-dependent hierarchy of forecasts.

The above discussion shows how the dimensionality issue is resolved when all traders submit their optimal orders (as given by expression (18)). However, because informed traders act strategically, we must consider deviations from the optimal strategy by any one trader (keeping the behavior of other traders fixed). If trader i submits an arbitrary order sequence, $(x_{i1}, \ldots, x_{in-1})$, which is different from the equilibrium orders given by equation (18), the sufficient statistics that we have computed need not be relevant. In particular s_{in-1} is not orthogonal to $(y_1, \ldots, y_{n-1})$ because trader i has not played his optimal strategy in the first $n - 1$ rounds of trading.

In the next section, we resolve the dimensionality issue for the case where one trader deviates from his optimal strategy (keeping the strategies of other traders fixed) and find the necessary and sufficient conditions for equilibrium.

IV. Necessary and Sufficient Conditions For Equilibrium

Suppose that all traders other than trader i play their conjectured equilibrium strategy and the market maker updates beliefs using the linear rules described above. Now consider what will happen if trader i has submitted arbitrary orders in the first $n-1$ periods, $(x_{i1}, \ldots, x_{in-1})$. To solve the model in this setting, we first need to construct the following statistics based on trader i's information from the first $n-1$ periods, $(s_{i0}, z_{i1}, \ldots, z_{in-1}, \ldots, x_{i-1}, \ldots, x_{in-1})$. These statistics correspond to the outcomes that would have occurred had trader i used the equilibrium strategy instead of the arbitrary strategy $(x_{i1}, \ldots, x_{in-1})$ in the first $n-1$ periods.

For each informed trader, starting with $\hat{s}^i_{j0} = s_{j0}$, we recursively define:

$$\hat{y}^i_n = \sum_{j=1}^{M} \beta_n \hat{s}^i_{jn-1} + u_n \qquad \hat{p}^i_n = p_0 + \sum_{k=1}^{n} \lambda_k \hat{y}^i_k$$
$$\hat{t}^i_{jn} = \sum_{k=1}^{n} \zeta_k \hat{y}^i_k \qquad \hat{s}^i_{jn} = s_{j0} - \sum_{k=1}^{n} \zeta_k \hat{y}^i_k \tag{27}$$

where $\hat{y}^i_n$ is the order flow that would have occurred in the n^{th} round of trading if trader i had followed the equilibrium strategy $(\beta\hat{s}^i_{i0}, \ldots, \beta\hat{s}^i_{in-1})$ in the first n periods of trading. Similarly, after n rounds of trading, $\hat{p}^i_n$ is the price that prevails in the n^{th} round of trading, $\hat{t}^i_{jn}$ is the market maker's conditional expected value of informed agent i's information, and $\hat{s}^i_{jn}$ is the information that informed agent j has that is not known to the market maker, if trader i had followed the equilibrium strategy $(\beta\hat{s}^i_{i0}, \ldots, \beta\hat{s}^i_{in-1})$ in the first n periods of trading.[7]

A key observation that we make (and prove) is that the order $(\hat{y}^i_1, \ldots, \hat{y}^i_{n-1})$ is in the information set of trader i, $(s_{i0}, z_{i1}, \ldots, z_{in-1}, x_{i1}, \ldots, x_{in-1})$ (we discuss this next). Hence, so are the variables $(\hat{s}^i_{ik}, \hat{p}^i_k, \hat{t}^i_{jk}), j = 1, \ldots, M, k = 1, \ldots, n-1$.[8] Then by our construction, $\hat{s}^i_{in-1}$ is orthogonal to $(\hat{y}^i_1, \ldots \hat{y}^i_{n-1})$, a fact we use when we consider the value function of the i^{th} informed agent.

Intuitively, $(\hat{y}^i_1, \ldots, \hat{y}^i_{n-1})$ is in the information set of the trader i because trader i knows the strategy that he would have followed in equilibrium, and hence the change in the expectations of the other agents (other informed traders and market makers) that would occur had he played the equilibrium strategy. Consequently, trader i can compute the amount by which trades of

[7] It is important to note that because trader i played nonequilibrium strategies in the past, s_{in-1} is not orthogonal to $y_1, \ldots, y_{n-1}$. The constructed variables $(\hat{s}^i_{in}, \hat{p}^i_n, \hat{t}^i_{jn})$ correspond to outcomes that would have occurred if equilibrium play occurs and are consistent with our discussion in the previous section. One can think of the informed traders keeping two sets of books: the first set corresponds to what happens along the equilibrium path; and the second corresponds to actual outcomes given the nonequilibrium play by the specified informed trader in earlier periods.

[8] However, the variables $\hat{s}^i_{jk}, j \neq i, k = 1, \ldots, n-1$ are not in the information set, as they depend on the unknown information that other informed agents have.

the other informed traders would change if he had played this equilibrium strategy. Therefore, the order flow that would have occurred in equilibrium had he played his equilibrium strategy is part of his information set after $n - 1$ rounds of trading. The proof of this is in the Appendix.

With the variables $(\hat{s}^i_{in-1}, \hat{p}^i_{n-1}, \hat{t}^i_{jn-1})$ in the information set of the informed trader, we provide variables that are sufficient statistics for prediction, taking account of the fact that he has not played his optimal strategy in the past. The sufficient statistics are $\hat{s}^i_{in-1}$ (the forecast along the equilibrium path) and $\hat{p}^i_{n-1} - p_{n-1}$ (the deviation from equilibrium prices induced by past suboptimal play). If trader i had followed his equilibrium strategy, the forecast along the equilibrium path, $\hat{s}^i_{in-1}$, would have been a sufficient statistic. However, one needs to account for past suboptimal play. The price deviation from equilibrium caused by past suboptimal play is the additional variable that is needed to summarize the history observed by trader i. Hence, we show (see Appendix),

$$\begin{aligned} \mathrm{E}[v - p_{n-1}|s_{i0}, y_1, \ldots, y_{n-1}, x_{i1}, \ldots, x_{in-1}] \\ = \mathrm{E}[v - \hat{p}^i_{n-1}|\hat{s}^i_{in-1}] + (\hat{p}^i_{n-1} - p_{n-1}). \quad (28) \end{aligned}$$

These statistics are also sufficient to forecast the forecasts of other traders.

With the above statistics to summarize the information available to trader i after $n - 1$ periods of trading, we conjecture the value function of trader i after stage $n - 1$ to be:

$$\begin{aligned} V_i(\hat{s}^i_{in-1}, \hat{p}^i_{n-1} - p_{n-1}) = \alpha_{n-1}(\hat{s}^i_{in-1})^2 + \psi_{n-1}\hat{s}^i_{in-1}(\hat{p}^i_{n-1} - p_{n-1}) \\ + \mu_{n-1}(\hat{p}^i_{n-1} - p_{n-1})^2 + \delta_{n-1}. \quad (29) \end{aligned}$$

Given past optimal play by trader i, we have $\hat{p}^i_{n-1} = p_{n-1}$, so that only the first and fourth terms on the right hand side of equation (29) remain. However, to find the optimal strategy in equilibrium, we have to consider strategies that are optimal, given past nonoptimal play by this trader. To do this, we have to consider the second and third terms on the right-hand side of expression (29).

We now conjecture the optimal strategy of a trader who has played an arbitrary, suboptimal strategy $x_{i1}, \ldots, x_{in-1}$ (the conjecture is confirmed in equilibrium). Given his past suboptimal play, the future strategy will not be the conjectured optimal strategy in equation (18) (which can only occur if the trader had played optimally in the first $n - 1$ periods). However, his conjectured optimal strategy from trading period n and beyond (given suboptimal play in periods 1 to $n - 1$) is of the form:[9]

$$x_{ik} = \beta_k\hat{s}^i_{ik-1} + \gamma_k(\hat{p}^i_{k-1} - p_{k-1}) \quad \text{for } k = n \text{ to } N. \quad (30)$$

[9] The definition of equilibrium requires us to consider the optimal strategy of trader i for all information set possibilities, not just outcomes given by playing the conjectured equilibrium strategy in the first $n - 1$ periods of trading.

Thus the optimal strategy from this period and beyond is the same as the optimal strategy that would have occurred given past optimal play, plus a second term that depends on the difference between the price that would have occurred in this period had trader i followed the optimal strategy in the past and the actual price this period.

The models of Kyle (1985), Foster and Viswanathan (1993), Michener and Tighe (1991), and Holden and Subrahmanyam (1992) can be considered in this context as follows.[10] In these models, all informed traders have identical information, v (in Kyle (1985) there is the additional assumption of a monopolist informed trader). Thus $\hat{s}^i_{in-1} = v - \hat{p}^i_{n-1}$. With $\beta_n = \gamma_n$, we obtain $x_{in} = \beta_n(v - p_{n-1})$ as is the case in these models. Similarly, setting $\mu_n = \alpha_n$ and $\psi_n = 2\alpha_n$ and using $\hat{s}^i_{in} = v - \hat{p}^i_n$, the value function after n periods becomes $\alpha_n(v - p_n)^2 + \delta_n$, as is the case in these models.[11] Hence the identical information analyses carried out in these models is embedded in our more general approach.

With this background, we state the necessary and sufficient conditions for a linear recursive Markov equilibrium.

Proposition 1: *The necessary and sufficient conditions for a recursive linear Markov equilibrium are:*

$$\beta_n = \frac{\eta_n - \lambda_n\psi_n}{\lambda_n[1+(1+\phi_n(M-1))(1-(\lambda_n\Psi_n/\theta))]}$$

$$\gamma_n = \frac{(1-2\mu_n\lambda_n)(1-(\lambda_n\beta_n(M-1)/\theta))}{2\lambda_n(1-\mu_n\lambda_n)}$$

$$\lambda_n = \frac{\beta_n M\Sigma_n}{\theta\sigma_u^2}$$

plus the value function variables

$$\alpha_{n-1} = (\eta_n - \lambda_n\beta_n[1 + (M-1)\phi_n])\beta_n + \alpha_n\left[1 - \frac{\lambda_n\beta_n}{\theta}[1 + \phi_n(M-1)]\right]^2$$

$$\psi_{n-1} = (\eta_n - \lambda_n\beta_n[1 + (M-1)\phi_n])\gamma_n - \lambda_n\gamma_n\beta_n + \beta_n\left(1 - \frac{\lambda_n\beta_n(M-1)}{\theta}\right)$$
$$+ \psi_n\left(1 - \frac{\lambda_n\beta_n}{\theta}[1 + (M-1)\phi_n]\right)\left(1 - \frac{\lambda_n\beta_n(M-1)}{\theta} - \lambda_n\gamma_n\right)$$

[10] See Back (1992) for a generalization of Kyle's (1985) model in continuous time.

[11] This can be verified using Proposition 1 below, setting $\eta_n = \phi_n = 1$, $\theta = 1$ (the parameters η_n and ϕ_n are defined in Proposition 1) and working backwards from the terminal trading round. Thus, while strategies in these articles do not explicitly account for the fact that informed traders may have played suboptimal strategies in previous periods, these strategies are correct. This is because past strategies are only relevant to the extent they affect the deviation of the true value from the current price.

$$\mu_{n-1} = -\lambda_n\gamma_n^2 + \gamma_n\left(1 - \frac{\lambda_n\beta_n(M-1)}{\theta}\right) + \mu_n\left(1 - \frac{\lambda_n\beta_n(M-1)}{\theta} - \lambda_n\gamma_n\right)^2$$

$$\delta_{n-1} = \delta_n + \frac{\alpha_n\lambda_n^2}{\theta^2}\sigma_u^2 + \alpha_n\frac{\lambda_n^2\beta_n^2}{\theta^2}[(M-1)\ \mathrm{Var}\ (s_{jn-1}|s_{i0}, y_1, \ldots, y_{n-1})]$$
$$+ \alpha_n\frac{\lambda_n^2\beta_n^2}{\theta^2}[[(M-2)(M-1)]\ \mathrm{Cov}\ (s_{jn-1}, s_{kn-1}|s_{i0}, y_1, \ldots, y_{n-1})]$$

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

$$\phi_n = \Omega_{n-1}/\Lambda_{n-1}$$

$$\eta_n = \frac{\theta}{M}[1 + (M-1)\phi_n]$$

$$\mathrm{Var}\ (s_{jn-1}|s_{i0}, y_1, \ldots, y_{n-1}) = \frac{\Lambda_{n-1}^2 - \Omega_{n-1}^2}{\Lambda_{n-1}}$$

$$\mathrm{Cov}\ (s_{jn-1}, s_{kn-1}|s_{i0}, y_1, \ldots, y_{n-1}) = \frac{\Omega_{n-1}(\Lambda_{n-1} - \Omega_{n-1})}{\Lambda_{n-1}}$$

In addition, the second order conditions must hold:

$$\lambda_n(1 - \mu_n\lambda_n) > 0$$

Also, the following recursive equations on the information variances must hold:

$$\Sigma_n = \left(1 - \frac{M\lambda_n\beta_n}{\theta}\right)\Sigma_{n-1}$$

$$\Lambda_n = \Lambda_{n-1} - \frac{M}{\theta^2}\frac{\lambda_n\beta_n}{\theta}\Sigma_{n-1}$$

$$\Omega_n = \Omega_{n-1} - \frac{M}{\theta^2}\frac{\lambda_n\beta_n}{\theta}\Sigma_{n-1}$$

As we have already discussed, these value functions and trading strategies differ considerably from those obtained in Kyle (1985), Foster and Viswanathan (1993), Holden and Subrahmanyam (1992), and Michener and Tighe (1991). The identical information assumption used in these articles implies that informed traders know each other's information precisely and simplifies the value function. (In Kyle (1985) there is only one informed trader, which further simplifies the value function). In contrast, our model involves learning by the informed traders about the information that other informed traders have, so we have to consider explicitly how past suboptimal behavior affects value functions and trading strategies.

V. Computation of the Equilibrium

In this section, we discuss how one can compute numerically the recursions needed to solve the model. First we note that the initial distribution of signals is completely determined by the parameters Λ_0 (the initial variance of each informed trader's signal), Ω_0 (the initial covariance between any two informed traders' signals), and θ (the regression coefficient of the true value v on the average signal $\hat{v}$). Because $\Lambda_n - \Omega_n = \chi$ (see Equation (10)), we need only compute Λ_n in any trading period. Ω_n immediately follows as

$$\Omega_n = \Lambda_n - \chi. \tag{31}$$

We solve the model using a backward induction approach. First, we fix a terminal value of Λ_N such that $\Lambda_N > [(M-1)/M]\chi$ (remember that this is the lower bound on Λ_N given in Equation (11)). Immediately, $\Omega_N = \Lambda_N - \chi$. The value function parameters in the terminal period, α_N, ψ_N, μ_N, and δ_N are all zero.

Now consider any arbitrary trading period n, where we have solved the model backward using the chosen terminal value Λ_N. Thus we know Λ_n, α_n, ψ_n, μ_n, and δ_n. Also, we know θ and χ (both are given parameters), so we know Ω_n and Σ_n. Given ϕ_n, we could use the two equations defining β_n and λ_n to solve for these two variables. However, $\phi_n = \Omega_{n-1}/\Lambda_{n-1}$, and consequently it depends on variables that are unknown in the backward induction algorithm. We can express ϕ_n in terms of χ, Ω_n, and Λ_n as follows. Notice that

$$\begin{aligned}\Lambda_n &= \Lambda_{n-1} - \frac{\lambda_n\beta_n}{\theta}\frac{M}{\theta^2}\sum{}_{n-1} \\ &= \Lambda_{n-1} - \frac{\lambda_n\beta_n}{\theta}[\Lambda_{n-1} + (M-1)\Omega_{n-1}] \\ &= \Lambda_{n-1} - \frac{\lambda_n\beta_n}{\theta}[M\Lambda_{n-1} - (M-1)\chi] \end{aligned} \tag{32}$$

which yields

$$\Lambda_{n-1} = \frac{\Lambda_n - (\lambda_n\beta_n/\theta)(M-1)\chi}{1 - M\lambda_n\beta_n/\theta}. \tag{33}$$

Similarly,

$$\Omega_{n-1} = \frac{\Omega_n + (\lambda_n\beta_n/\theta)\chi}{1-(M\lambda_n\beta_n/\theta)}. \tag{34}$$

Thus

$$\phi_n = \frac{\Omega_n + (\lambda_n\beta_n/\theta)\chi}{\Lambda_n - (\lambda_n\beta_n/\theta)(M-1)\chi}. \tag{35}$$

This representation of ϕ_n in terms of Ω_n, Λ_n, λ_n, and β_n means that the first and third equations in the necessary and sufficient conditions can be solved to yield λ_n and β_n. Substituting for β_n we obtain the following fourth order equation for λ_n:

$$(\lambda_n)^4\frac{\theta\sigma_u^4(M-1)\chi}{M^2\sum_n^2} + (\lambda_n)^3\frac{\sigma_u^2\psi_n\Lambda_n}{\sum_n} - (\lambda_n)^2\frac{\theta\sigma_u^2}{M\sum_n}[(M+1)\Lambda_n - (M-1)\chi].$$

$$- \lambda_n\Lambda_n\psi_n + \frac{\theta}{M}[M\Lambda_n - (M-1)\chi] = 0. \qquad (36)$$

One can verify that (irrespective of whether ψ_n is negative or positive), there are only two sign changes, so there are at most two real roots by Descartes' rule of signs. Also, the left-hand side of equation (36) is strictly greater than zero at $\lambda_n = 0$, negative at $\lambda_n = \sqrt{\Sigma_n}/\sigma_u$, and becomes infinite as λ_n goes to infinity. Thus there is only one real root between 0 and $\sqrt{\Sigma_n}/\sigma_u$, which is the only solution to the necessary conditions.[12] If this solution satisfies the second order conditions, it is sufficient.[13]

Once we compute λ_n, the remaining parameters are easily obtained. The third equation of Proposition 1 yields β_n, and the second equation of the Proposition gives γ_n. Through substitution of these parameters we can compute α_{n-1}, ψ_{n-1}, μ_{n-1}, δ_{n-1}, and Λ_{n-1} which are used as inputs to the prior period.

The solution technique yields an initial value of Λ_0 given the terminal value of Λ_N, keeping χ and θ fixed. If the resulting initial value of Λ_0 is too high relative to value we seek, we reduce Λ_N, and vice versa. We continue to adjust Λ_N in this manner, until we find the terminal value of Λ_N that yields the desired initial Λ_0.

VI. Numerical Results

To illustrate the model, we compute the equilibrium parameter values for a variety of settings. We are interested in seeing how competition among the informed traders is affected by the correlation structure between their initial information endowments, how the correlation structure changes through time, and how lengthening the number of trading periods affects price formation.

The specific parameterization that we choose is as follows. The M informed traders see M different pieces of information, and the truth is the sum of these signals. Hence, by definition the informed traders together know the truth, $v =$

[12] The restriction that $\lambda_n < \sqrt{\Sigma_n}/\sigma_u$ follows from the requirement that Σ_n must be decreasing over time.

[13] This argument guarantees uniqueness in the class of recursive linear Markov equilibria (if an equilibrium exists in this class). Verification of the second order condition is required to ensure existence. In our numerical work, this second order condition is always satisfied.

$\Sigma_{i=1}^{M} s_{i0}$.[14] This parameterization allows us to fix the total information available to the informed traders while altering the correlation between their signals. If $\Omega_0 = \Lambda_0$, all of the signals are perfectly correlated and the truth is just M times each individual signal (this corresponds to all the informed traders having the same information). If $\Omega_0 = 0$, then the signals, s_{i0}, received by the informed traders are independent, and informed traders have the maximal amount of information not known to other traders. Here $\Lambda_0 = \sigma_v^2/M$. By varying Ω_0 between $-(1/M)\chi$ and Λ_0, we can model arbitrary correlations between the signals, while the overall information available to informed traders is kept fixed.

To begin, we consider a four period ($N = 4$) example with three informed traders ($M = 3$) using the parameterization discussed above ($\theta = 3$). Further, we set $\sigma_v^2 = 1$ and $\Sigma_0 = 1$ and consider four different correlation structures (very highly correlated information, a low positive correlation, zero correlation, and low negative correlation). By using these different correlation structures with this parameterization, we are fixing the overall information available to informed traders as a group, and considering different distributions of this information across traders. We set the values of χ to be 0.000004, 0.2, 0.333333, and 1.0 for the four separate cases. Given the other parameter values, these imply Λ_0 values of 0.111114, 0.244445, 0.333333, and 0.777778, respectively. This means that the Ω_0 becomes 0.11111, 0.044444, 0.00, and -0.22222, respectively. The initial correlations are 0.999964, 0.181819, 0.00, and -0.28571, respectively.[15]

In the remainder of this section, we present numerical solutions to the model for different correlation structures. We vary the number of trading periods to understand the effect of more frequent trading. We plot endogenous parameters such as the variance of the true value of the asset, informed traders' profits, market depth, and the conditional correlation between the informed traders' signals.

Figure 1, Panel *A* shows the evolution of the variance of the true value given the order flow for the first n periods, Σ_n. Initially we consider the case of four trading periods ($N = 4$). The lesser the initial correlation in the signals received by the informed traders, the higher is the terminal variance of the true value. This means that less information is being revealed through trading when the signals have lower initial correlation.

Figure 1, Panel *B* describes the total expected profits to all informed trades for each of the correlation structures. The profits to the informed traders are lowest with identical information and are much higher in the other three cases. Surprisingly, the profits are highest in the case with positive but not perfect correlation; the expected profit in this case is 1.45232 versus 1.44285 in the

[14] In fact, this is the most general correlation structure up to a scale parameter θ. This follows from equation (2). In our specific parameterization, the parameter θ in equation (2) is always equal to M. This follows from the fact that $\nu = \Sigma_{i=1}^{M} s_{i0} = M\hat{v}$.

[15] For the first of these cases, the correlation is not exactly one, but is extremely close to one. In our discussion we will refer to this case as informed traders having identical information, even though this is not precisely true.

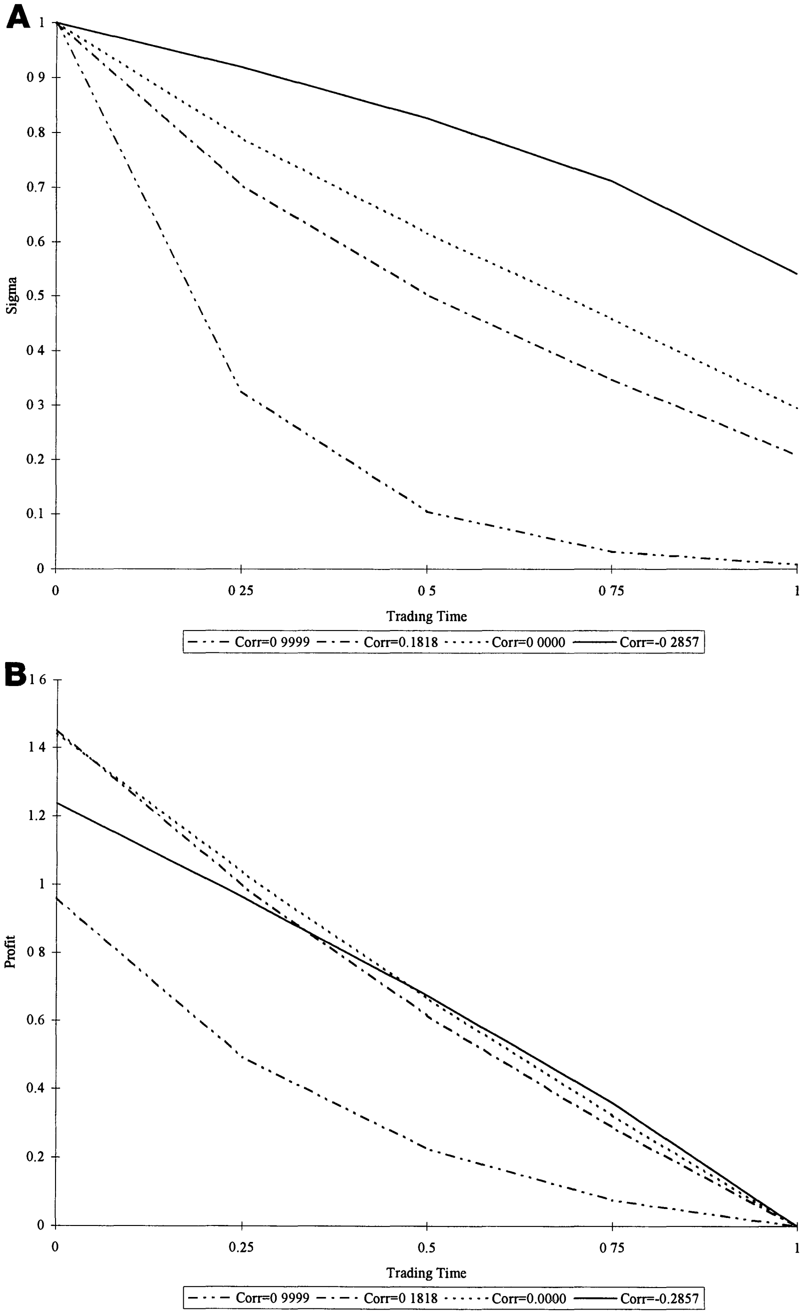


Figure 1. Information Decay, Profit, Market Depth, and Conditional Correlation Dynamics with 4 Periods. Numeric solutions to the model with three informed traders, one unit of initial variance of information, and four units of liquidity trader variance across all periods. Four different initial correlation structures are examined, high (0.9999), low (0.1818), independent (0.0000), and negative (−0.2857). Panel *A*, Information decay; Panel *B*, profit; Panel *C*, market depth; Panel *D*, conditional correlation.

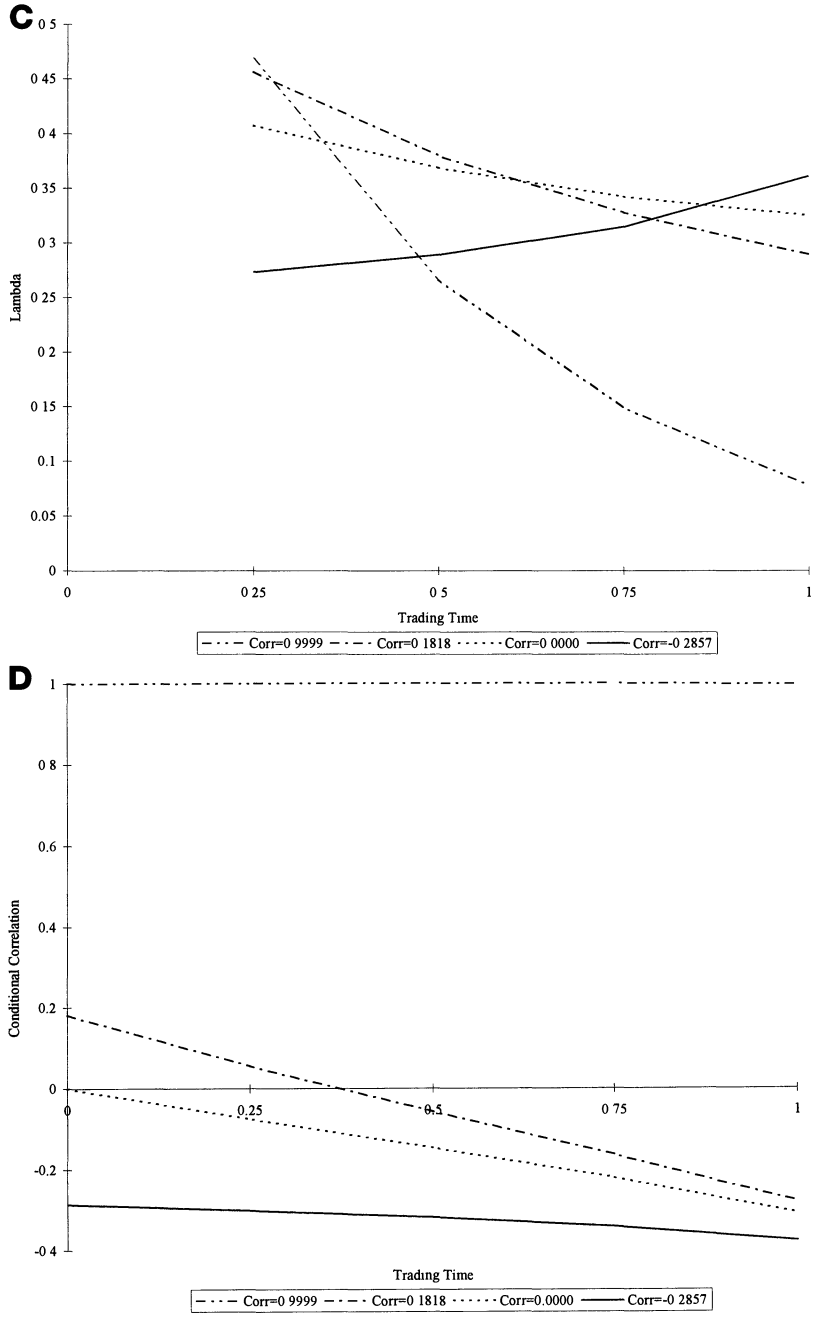


Fig. 1—*continued*

case of initial zero correlation. The expected profit with negative correlation is lower than in the cases of zero and positive correlation.

Figure 1, Panel *C* plots the market maker's sensitivity to the order flow, λ_n, for each period for all correlation structures. The identical information (perfect positive correlation) case has the highest market maker sensitivity in the first period, but as more information is released through trading, it falls in later periods. This is consistent with the intense trading in the first period that occurs with perfect positive correlation. The second highest trading costs occur with positive but not perfect correlation, the third highest with zero correlation, and the lowest with negative correlation. When we look to the terminal period, the order of the market maker's sensitivity to order flow is reversed; the highest sensitivity is found for the case of negative initial correlation, the next with zero initial correlation, then positive but not perfect initial correlation, and finally the lowest sensitivity is found for the case of positive and perfect initial correlation between signals. This occurs because with negative correlation, the least information is released in the first three trading periods, yielding the highest variance of information (Σ_n) at the start of trading in period 4 (see the discussion of Figure 1, Panel *A*).

It is interesting to note that in the negative correlation case, λ_n increases each period and is at its highest in the final period, when the maximum amount of information is released through trading (for this correlation level). This is in contrast with the other three correlation structures, where λ_n decreases with each trading period. This suggests that the initial negative correlation between the informed traders' signals leads to behavior different from that observed with positively correlated signals.[16]

Figure 1, Panel *D* shows the conditional correlation between the informed traders' signals. When the initial correlation is very high, it remains so with only four trading periods. With lower initial correlations the conditional correlation becomes negative even when there are only four trading periods. Also note that the change in the correlation is highest for the case of positive, low initial correlation between the informed traders' signals. From Panel *B* we also know that this is the case with the highest ex ante profits.

The results discussed in the above figures illustrate two effects. First, when traders have heterogeneous information, the competitive pressure from the trades of others is reduced. Each trader has some monopoly power, as part of his information is unique. Thus each trader faces less competition and has less incentive to trade aggressively. Second, with a large enough number of trading periods, the conditional correlation between the informed traders' signals (conditional on the observed order flow) eventually becomes negative (this

[16] While this is peripheral to the main focus of our paper, a substantial empirical literature has documented that trading costs increase after earnings announcements. This has been viewed as contradictory to the results that have been previously obtained in the literature, where λ_n decreases through time. Bushman, Dutta, Hughes, and Indejejikian (1993) provide an excellent discussion of the empirical and theoretical literature and present a model with discretionary liquidity trading where trading costs increase over time. In our model, negatively correlated information yields a similar increase in trading costs over time.

requires more periods the higher the initial correlation). This is because the market maker is learning more about the average of the informed traders' signals than he is about each individual trader's signal. Because the market maker will never learn precisely each informed trader's signal, the conditional correlation between the signals becomes negative over time. This negative correlation can be viewed as an endogenous difference of opinion.

Whenever informed traders' beliefs are negatively correlated (conditional on the public information), each informed trader thinks the market lacks resiliency as follows. With negatively correlated information, each informed trader views the other informed traders as collectively pulling the stock price in the wrong direction. Hence, each informed trader perceives that without his trades, there is no mechanism to push the price to its true value, and all informed traders have an incentive to make smaller trades in the current period, with the hope of making larger trades in the future. We refer to this effect as the "waiting game."

As an example, if an informed trader believes that an asset is undervalued, large trades from other informed traders will lower the price even further, thereby amplifying the mispricing and creating greater future trading profits. This being the case, no informed trader wants to place large orders in the initial trading periods when their signals are negatively correlated. Consistent with this, one needs the liquidity in the market to be lower in latter periods (the market maker's sensitivity to trades, λ_n, must be highest in the terminal rounds of trading) when informed traders have negatively correlated initial signals. Otherwise, informed traders have no incentive to trade in early periods. This effect is borne out in Panel *C* of Figure 1 where the market maker's sensitivity to trades, λ_n, is highest in the terminal period of trading with negatively correlated signals.

These numerical solutions suggest that competition is less intense when there is imperfect correlation among the informed traders signals, as opposed to the case where they have identical information. We now consider what happens if we add more trading periods, while keeping the liquidity trading variance fixed. Consider increasing the number of trading periods between two fixed times, 0 and 1, while overall liquidity trading variance is kept fixed (thus the liquidity trading variance in any given trading period falls). This will allow us to understand how more frequent trading affects order submission and whether our conjectures about the role of the initial correlation between the informed traders' signals in competition is correct.

In Figure 2, Panels *A* and *B*, we show the evolution of variance, Σ_n, for the small but positive correlation case (correlation = 0.181819) and for the perfect correlation case (correlation = 0.999964), respectively. Here $\Sigma_0 = 1$ and $\theta = 3$. As the number of trading periods is varied (between 4, 20, 80, and 800), the total liquidity variance over all trading periods $N\sigma_v^2$ is kept constant at 4. We find that there is a marked contrast in the information decay when we change the correlation structure. With perfect correlation, competition is very intense, and most of variance is released very quickly (the terminal variance is 0.00000154 for the four trading period case). As more trading periods are added, information is released more quickly, and the variance of the liquida-

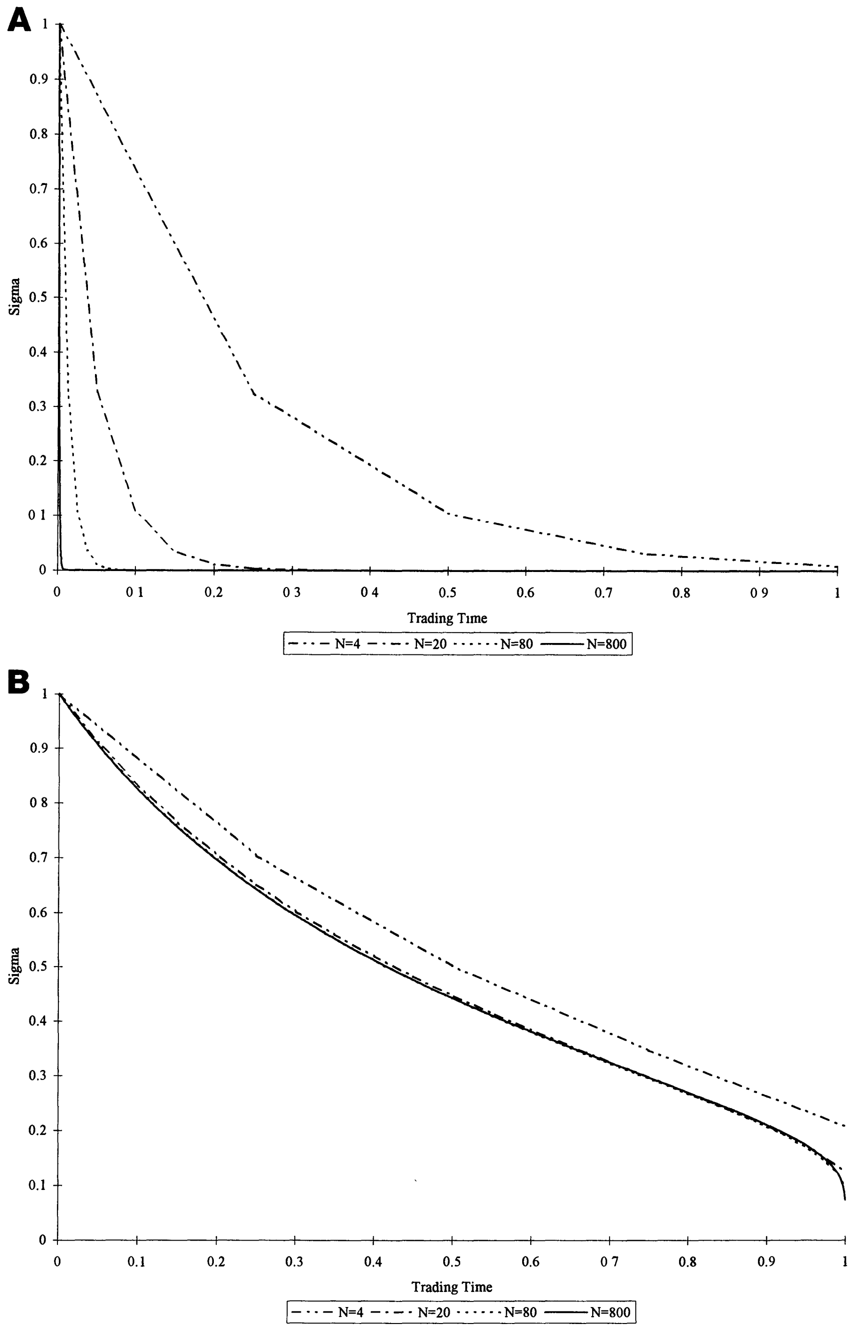


Figure 2. Information Decay with Changing Number of Trading Periods. Numeric solutions to the model with three informed traders, one unit of initial variance of information, and four units of liquidity trader variance across all periods (total liquidity variance is kept constant as new periods are added). Model is solved for 4, 20, 80, and 800 periods. High initial correlation case starts with a correlation of 0.9999. Low initial correlation case starts with a correlation of 0.1818. Panel *A*, high initial correlation; Panel *B*, low initial correlation.

tion value drops dramatically. With small but positive initial correlation, the decline in the variance is much less dramatic. Even after 800 periods, the terminal variance is 0.074128. It appears that competition is much less intense with the lower initial correlation in signals. In fact, the decay in the variance of the liquidation value is much slower than would be the case for a monopolist informed trader. A monopolist informed trader with the same initial variance would trade so that terminal variance after 800 periods would be 0.00125.

In Figure 3, Panels *A* and *B*, we show how the total expected profits made by all informed traders varies by trading periods for the two cases of perfect positive correlation and when the initial correlation is small but positive. From Kyle (1985) we know that a monopolist makes the maximum expected profit with continuous trading of $(\Sigma_0)^{0.5}\sigma_u$. This is 2 in our example. With small but positive correlation in the informed traders' initial information, the expected trading profit increases with the number of trading periods (holding the total liquidity variance constant). With four trading periods, the total expected profit is 1.45232, and with 800 trading periods the total expected informed trader profit is 1.86325. These values are in sharp contrast to the perfect correlation case, where profits fall dramatically as the number of trading periods increases. In fact, with small but positive initial correlation and 800 periods of trading, the three informed traders together make 93 percent of the maximum expected profit of a monopolist informed trader.

The importance of the correlation structure of the informed traders' signals is further confirmed when we consider how the market maker's sensitivity to order flow changes as more trading periods are added. In Figure 4, Panel *A*, we show that the market maker's sensitivity to orders, λ_n, is initially very high and falls quickly with initially perfectly correlated information. This drop is accelerated as more trading periods are added. In contrast, Figure 4, Panel *B* shows the changes in the market maker's sensitivity to order flow when the initial correlation between signals is small but positive. Here changes in market depth follow a distinct U-shaped pattern, and this U-shape becomes more pronounced as trading periods are added. The market maker's sensitivity to trades is high in initial periods and then declines during the middle periods, and increases dramatically in the latter trading periods. The key reason for this difference in market depth is the negative correlation between the informed traders signals that eventually results through repeated trading. From our analysis of the negative correlation case, we know that there is an incentive for the informed traders to delay trades to future periods and play a "waiting game." As the number of trading periods is increased, the part of the informed traders signals not known to the market maker becomes negatively correlated (so long as the initial correlation is not perfect, i.e., $\chi > 0$), which means that informed traders will, after some time, choose to defer their trades to the later periods.[17]

[17] From comparing Panel *B* in Figure 4 with Panel *B* in Figure 5, it is clear that initially when the correlation is negative, the market sensitivity to the trades (λ_n) continues to decline. It is only when the correlation becomes sufficiently negative that the market maker's sensitivity to trades starts to increase and finally jumps dramatically in the final rounds of trading.

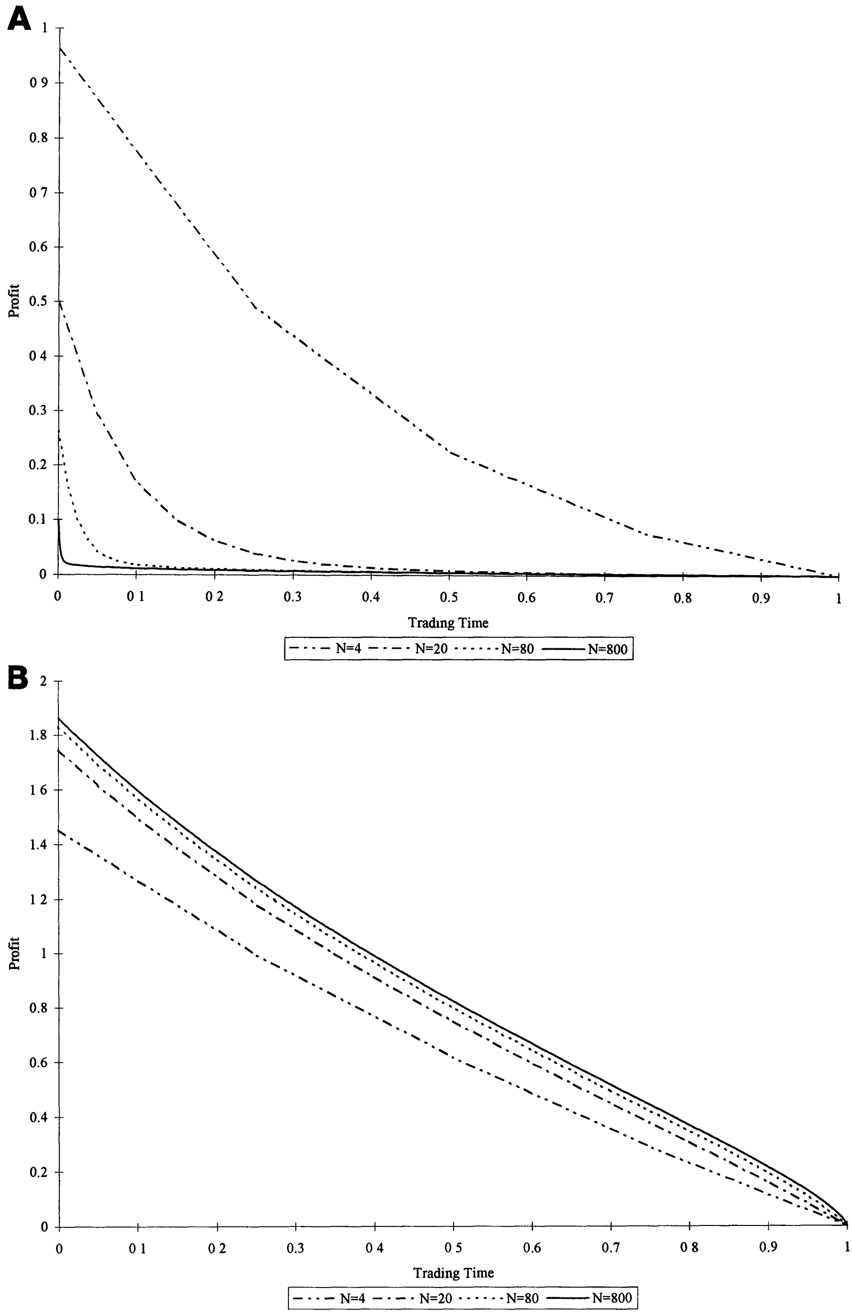


Figure 3. Expected Profits with Changing Number of Trading Periods. Numeric solutions to the model with three informed traders, one unit of initial variance of information, and four units of liquidity trader variance across all periods (total liquidity variance is kept constant as new periods are added). Model is solved for 4, 20, 80, and 800 periods. High initial correlation case starts with a correlation of 0.9999. Low initial correlation case starts with a correlation of 0.1818. Panel *A*, high initial correlation; Panel *B*, low initial correlation.

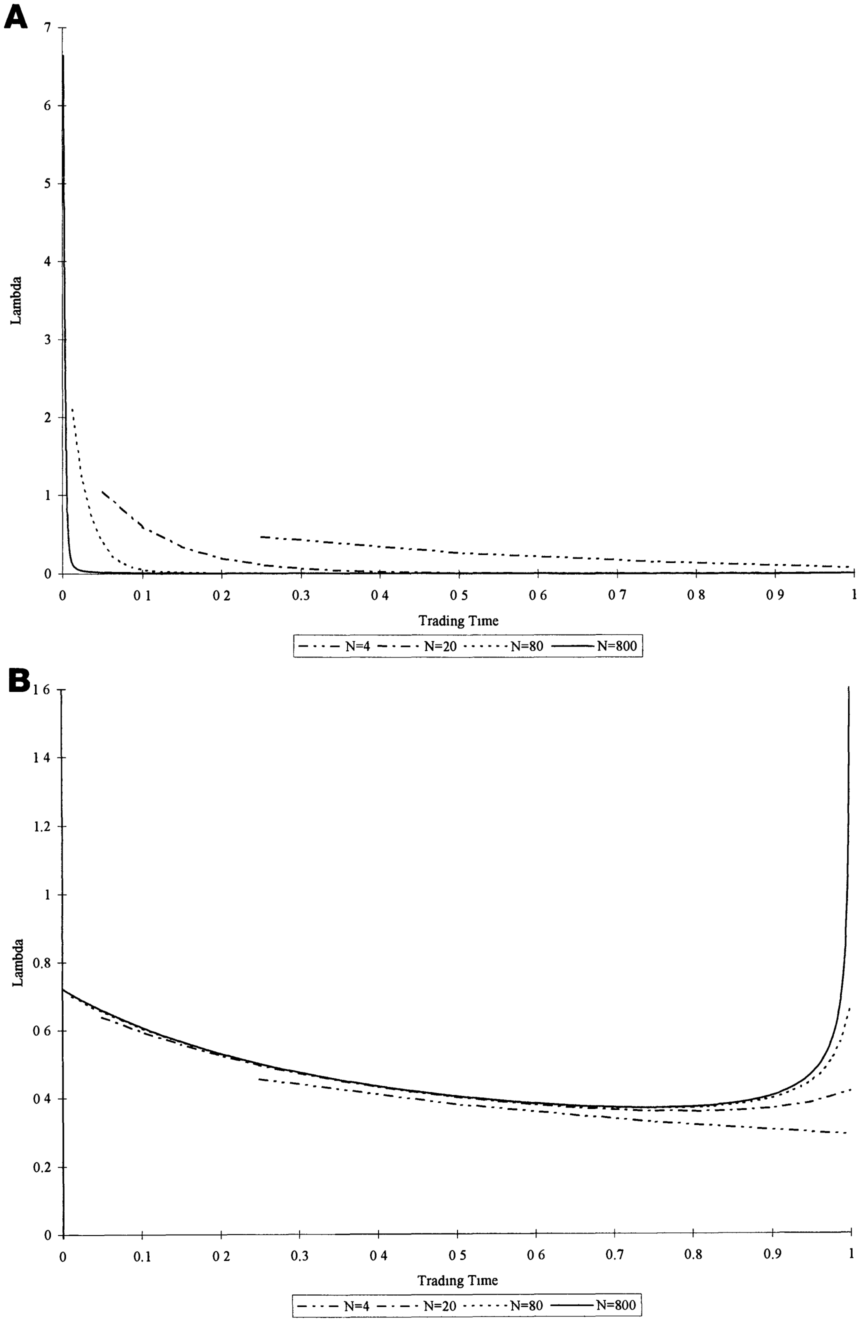


Figure 4. Market Liquidity with Changing Number of Trading Periods. Numeric solutions to the model with three informed traders, one unit of initial variance of information, and four units of liquidity trader variance across all periods (total liquidity variance is kept constant as new periods are added). Model is solved for 4, 20, 80, and 800 periods. High initial correlation case starts with a correlation of 0.9999. Low initial correlation case starts with a correlation of 0.1818. Panel *A*, high initial correlation; Panel *B*, low initial correlation.

In Figure 5, we plot the conditional correlation and confirm that as more periods are added, the correlation becomes negative earlier in the trading process (where periods are considered as a percentage of the total available time). More importantly, the stock of remaining information (as measured as variance) at the time when the correlation becomes negative is higher when more trading periods are added. In Panel *A*, we show the case of initially perfectly correlated signals and note that with only 4 periods, there is very little change in the correlation structure. With 800 periods, however, the correlation becomes negative about period 400. However, the extremely intense trading in the early periods implies that the stock of remaining information at this point in time is 0.0044, which is very close to zero. In Panel *B* of Figure 5, we show the conditional correlation between the informed traders' signals for the case of an initial small, positive correlation. Now, the negative correlation phase starts in period 235, when the stock of remaining information is 0.6007 (as opposed to the 0.50 stock of remaining information in the 4 period case, where the correlation becomes negative in period 3). This high stock of information unknown to the market maker (which increases with the number of trading periods) implies that the negative correlation phase is much more important with initial small, positive correlation. Also, the increase in the initial stock of information (at the point when the correlation turns negative) with the number of trading periods seems to be a fundamental aspect of the increase in profits as the number of trading periods increases.

These numerical results suggest that dynamic competition with heterogeneously informed traders can result in very different competitive outcomes than the case of identically informed traders. In particular, the intense competition that exists with identically informed traders does not occur when the correlation is not perfect. Also, an increase in the number of trading periods does not lead to a decrease in profit and even more intense competition. We find that for our examples with initially small but positive information, an increase in the number of trading periods leads to more cooperation among the informed traders, yielding larger expected trading profits. It appears that these results are driven by the fact that the perfect correlation case leads only to a "rat race," while imperfect correlation eventually leads to a "waiting game."

To investigate the results further, we repeat our analysis in Figure 1 for the 800 trading period case. Figure 6, Panel *A* shows the evolution of the variance of the liquidation value for the four different initial correlations over the 800 trading periods. The interesting aspect of Figure 6, Panel *A* is that the perfect positive correlation case shows a large decay in variance in the first instant and very little subsequent decay; the decay pattern is convex and decreasing. In contrast, the initial negative correlation case shows smaller decays in variance initially and a large decay at the end of trading; the decay pattern is concave and decreasing. Additionally, the initial negative covariance case has a positive terminal variance of information, i.e., some information is not revealed in the market price after 800 periods of trading. The initial small but positive correlation and the initial zero correlation case show an intermediate

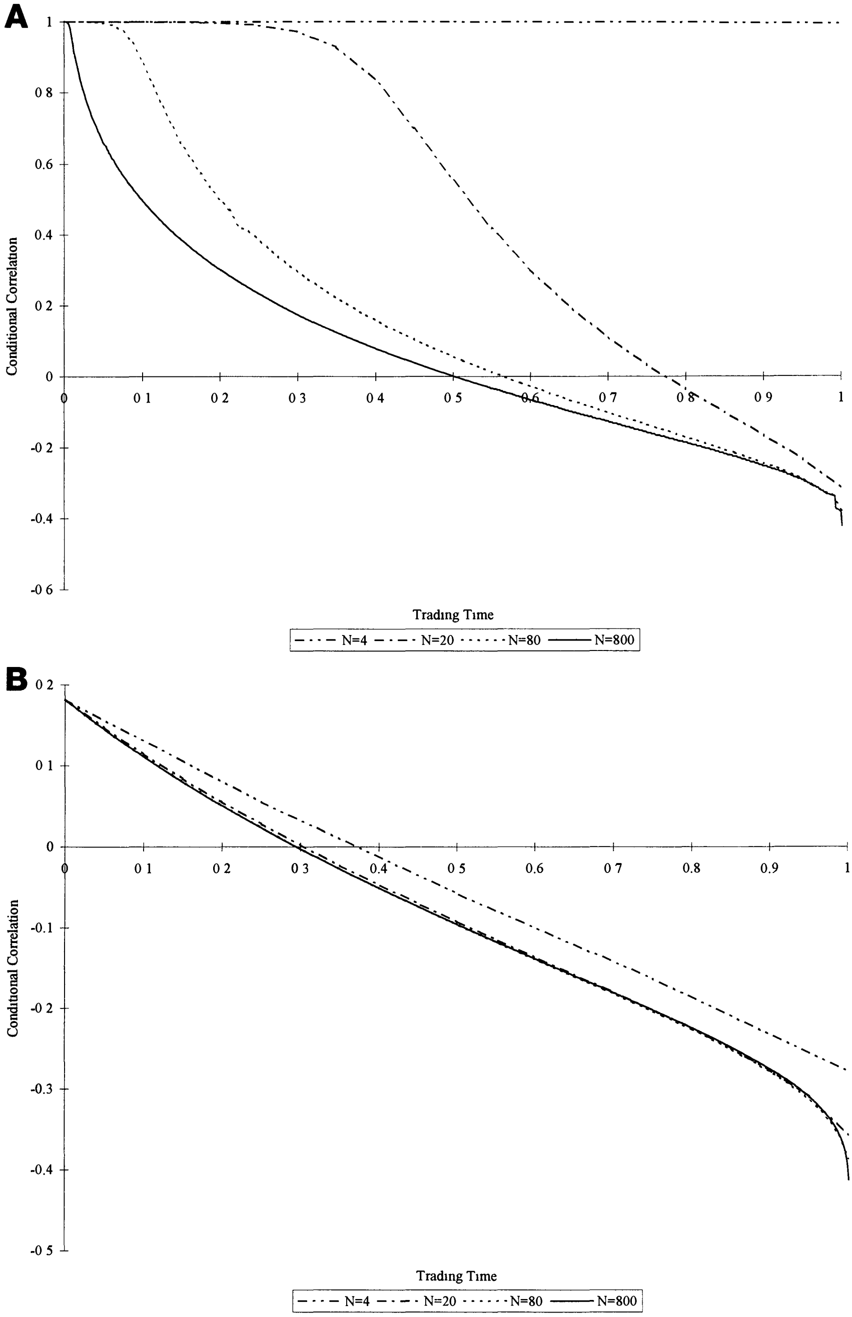
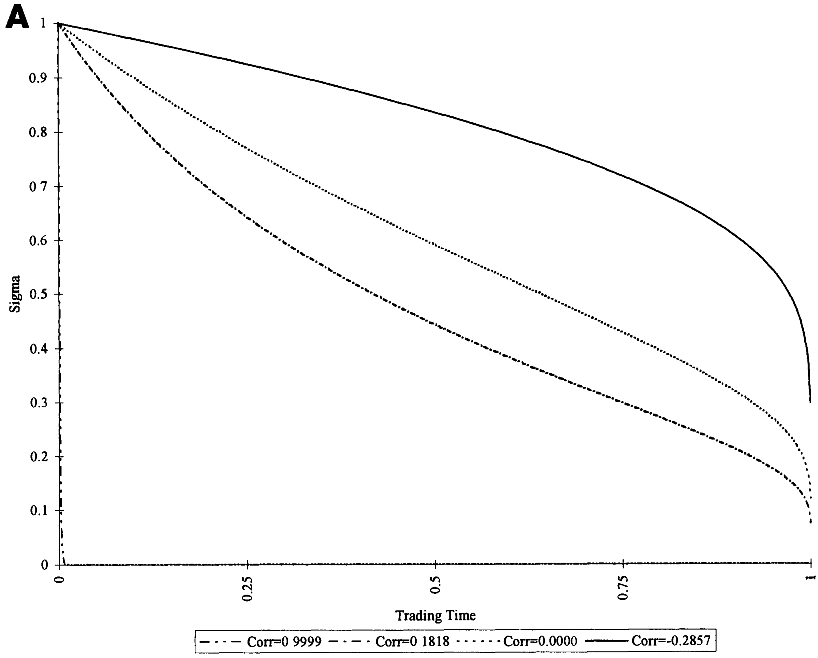


Figure 5. Conditional Correlation with Changing Number of Trading Periods. Numeric solutions to the model with three informed traders, one unit of initial variance of information, and four units of liquidity trader variance across all periods (total liquidity variance is kept constant as new periods are added). Model is solved for 4, 20, 80, and 800 periods. High initial correlation case starts with a correlation of 0.9999. Low initial correlation case starts with a correlation of 0.1818. Panel *A*, high initial correlation; Panel *B*, low initial correlation.



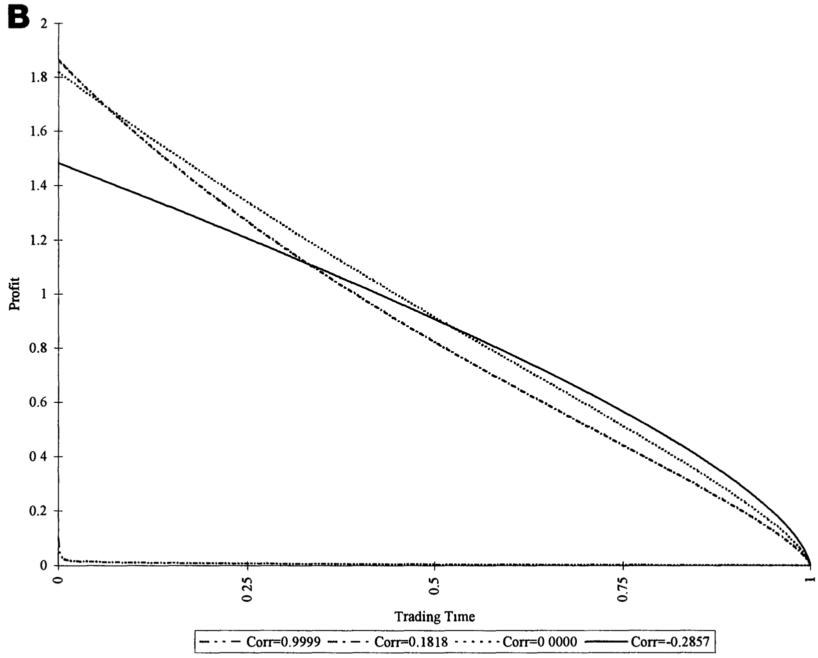
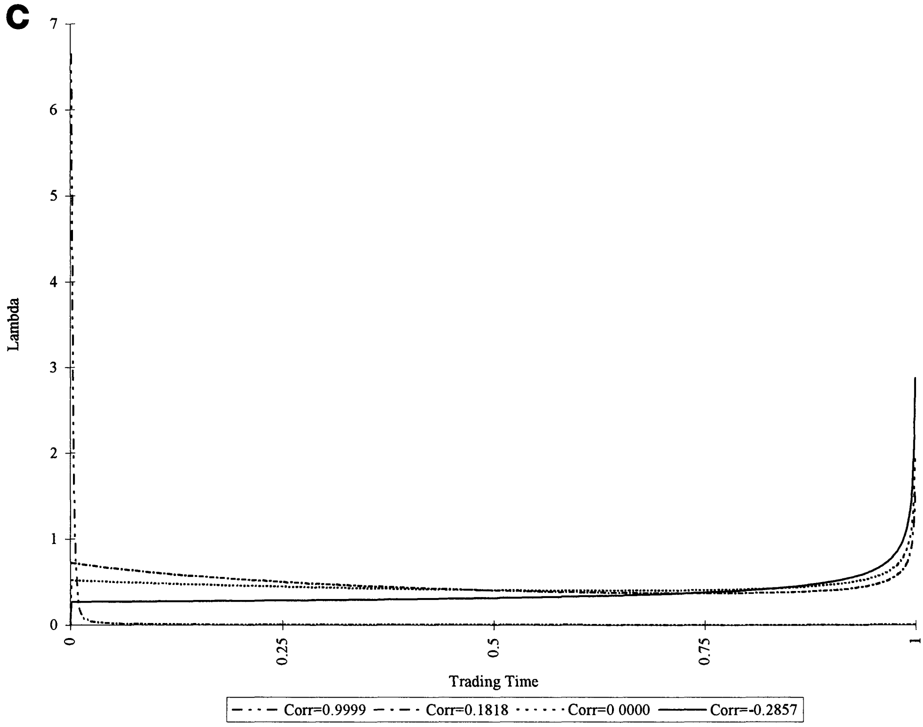


Figure 6. Information Decay, Profit, Market Depth, and Conditional Correlation Dynamics with 800 Periods. Numeric solutions to the model with three informed traders, one unit of initial variance of information, and four units of liquidity trader variance across all periods. Four different initial correlation structures are examined, high (0.9999), low (0.1818), independent (0.0000), and negative (−0.2857). Panel *A*, information decay; Panel *B*, profit; Panel *C*, market depth; Panel *D*, conditional correlation.



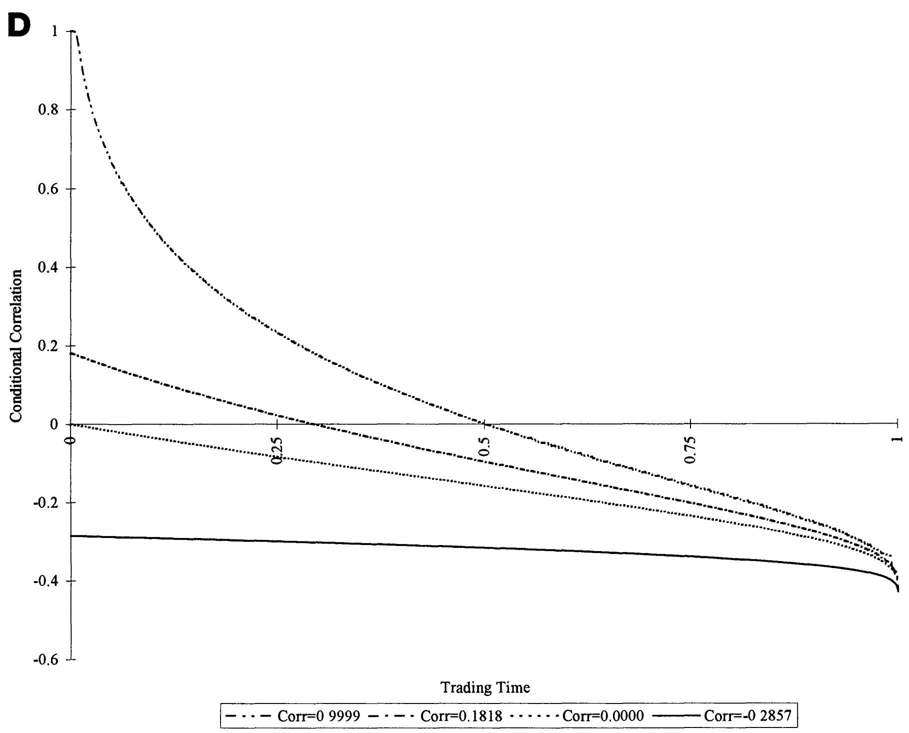


Fig. 6—*continued*

decay pattern that is initially convex and then concave and show a smaller positive terminal variance than the initial negative correlation case.

Figure 6, Panel *B* shows the evolution of total expected profits for the four initial correlations over 800 periods. Again, the small but positive correlation has the highest profits. The reason for this effect is interesting. While initial negative correlation leads to a "waiting game," it also results in information being unused at end of trading (there is a positive terminal variance), which causes some diminution of profits. The initially small, but positive, correlation leads to an initial "rat race" phase followed by the "waiting game" phase and thus results in less information left unused (not impounded into the price) at the end of trading.[18] This leads to greater profits for all the informed traders.

Figure 6, Panel *C* shows the market maker's sensitivity to trades, λ_n, over the 800 periods for the four cases. As discussed previously, the perfect correlation case leads to a rapid drop in the market maker's sensitivity to trades. In contrast, two of the other three cases (the initial small but positive correlation and the initial zero correlation) show a reduction in λ_n followed by an increase toward the end, while the initially negative correlation case shows an increase in λ_n at the end. The jump in the market maker's sensitivity to trades at the end of trading is highest for the initial negative correlation case; then the initial zero correlation case, the initial small but positive correlation case and the initial perfect correlation case.[19]

Finally, in Figure 6, Panel *D*, we show the conditional correlation of the informed trades' signals for the 800 period case, using a variety of initial correlation values. In all four cases, we notice that the conditional correlation is eventually negative. As previously discussed, the conditional correlation becomes negative very quickly with nearly perfectly correlated information. However, the stock of information at this point is so small that it is of little benefit to informed traders. In contrast, with a initial small positive correlation or zero correlation, the stock of information at the point when the correlation is negative is much larger. Finally, with initially negatively correlated information, there is less incentive to trade in early periods and hence little learning by the market maker. Therefore, the change in the correlation structure is slow, except in the terminal rounds of trading.

We believe these numerical results demonstrate that the nature of competition with heterogeneous information depends critically on the correlations of the initial signals and the number of trading periods. When the initial correlation is high, there is an incentive to compete aggressively (the "rat-race"

[18] Back, Cao, and Willard (1995) have recently solved a similar model in continuous time. They find that in continuous time, there is no unused information at the end of trading. Hence, the information that is unused at the terminal time in our approach is traded away in the last instant. However, trading at the last instant is not profitable. This means that our intuitive description works in continuous time if "unused information" is replaced by "information traded away at the last instant."

[19] Because we actually use a nearly perfect positive correlation case, there is a small jump in λ_n at the end of trading. However, it is very small relative to the already insignificant value of λ_n toward the end, and hence we do not discuss it.

effect). However, with low initial correlation, there is also an incentive to postpone trades (the "waiting-game" effect). This effect allows greater implicit cooperation among informed traders and thus greater informed trader profits.

VII. Conclusions

In this article, dynamic trading with heterogeneous information is analyzed. In the context of a specific dynamic trading game, an approach to resolve the dimensionality issue that arises in strategic dynamic models with heterogeneous trading is presented. From this we derive the necessary and sufficient conditions for a linear Markov recursive equilibrium. Using these conditions, the nature of competition with heterogeneous information is analyzed with some numerical examples.

We show that competition with imperfectly correlated information is different from competition with perfectly correlated information. In particular, competition is much less intense because of the existence of a second incentive in equilibrium, which we call the "waiting game." This is in addition to the usual incentive that leads to aggressive competition in the identical information case, which we call the "rat race." Consequently, increasing the number of trading periods while holding the total liquidity variance constant leads to greater profits and less intense competition between the informed traders when the initial correlation is not too high.

Appendix

The Inverse of the Matrix Ψ_0 and the Parameter γ

Let the diagonal elements of $[\Psi_0]^{-1}$ be a and the off diagonal elements all be b. Then the entries of $\Delta_0'[\Psi_0]^{-1}$ are all equal and identical to $c_0[a + (M - 1)b]$. Hence $\kappa = c_0[a + (M - 1)b]$. The values of a and b are given by the linear equations:

$$a\Lambda_0 + (M - 1)b\Omega_0 = 1 \qquad a\Omega_0 + b[\Lambda_0 + (M - 2)\Omega_0] = 0 \tag{A1}$$

which yield the solutions

$$a = \frac{[\Lambda_0 + (M - 2)\Omega_0]}{\Lambda_0[\Lambda_0 + (M - 2)\Omega_0] - (M - 1)\Omega_0^2}$$

$$b = -\frac{\Omega_0}{\Lambda_0[\Lambda_0 + (M - 2)\Omega_0] - (M - 1)\Omega_0^2} \tag{A2}$$

$$\kappa = c_0\frac{\Lambda_0 - \Omega_0}{\Lambda_0[\Lambda_0 + (M - 2)\Omega_0] - (M - 1)\Omega_0^2}. \qquad \text{Q.E.D.}$$

Derivation of Equation (5)

$$
\begin{aligned}
p_n &= \mathrm{E}[v|y_1, \ldots, y_n] \qquad \text{(A3)}\\
&= \mathrm{E}[E[v|\hat{v}, y_1, \ldots, y_n]|y_1, \ldots, y_n]\\
&= \mathrm{E}[E[v|\hat{v}]|y_1, \ldots, y_n]\\
&= \mathrm{E}[\theta\hat{v}|y_1, \ldots, y_n]\\
&= \theta\mathrm{E}[\hat{v}|y_1, \ldots, y_n]\\
&= \frac{\theta}{M}\sum_{i=1}^{M}\mathrm{E}[s_{i0}|y_1, \ldots, y_n]\\
&= \frac{\theta}{M}\sum_{i=1}^{M} t_{in}.
\end{aligned}
$$

Q.E.D.

Derivation of Relations Between Σ_n, Λ_n, and Ω_n

Σ_n can be written as,

$$
\Sigma_n = \mathrm{E}[(\theta\hat{v} - p_n)^2] = \mathrm{E}\left[\left(\frac{\theta}{M}\sum_{i=1}^{M} s_{i0} - \frac{\theta}{M}\sum_{i=1}^{M} t_{in}\right)^2\right] \qquad \text{(A4)}
$$

$$
\begin{aligned}
&= \left(\frac{\theta}{M}\right)^2 \mathrm{E}\left[\left(\sum_{i=1}^{M} s_{in}\right)^2\right]\\
&= \left(\frac{\theta}{M}\right)^2 [M\Lambda_n + M(M-1)\Omega_n]\\
&= \frac{\theta^2}{M}[\Lambda_n + (M-1)\Omega_n].
\end{aligned}
$$

Next let $c_n = \mathrm{Cov}(s_{in-1}, y_n)$. Then the updating for the variance-covariance matrix Ψ_n becomes:

$$
\Psi_n = \Psi_{n-1} - (\mathrm{Cov}(s_{1n-1}, y_n) \cdots \mathrm{Cov}(s_{Mn-1}, y_n))\frac{1}{\mathrm{Var}(y_n)}\begin{pmatrix}\mathrm{Cov}(s_{1n-1}, y_n)\\ \vdots \\ \mathrm{Cov}(s_{Mn-1}, y_n)\end{pmatrix}. \qquad \text{(A5)}
$$

Because $\mathrm{Cov}(s_{in-1}, y_n) = c_n$ for all i, it follows that

$$\Lambda_{n-1} - \Lambda_n = \Omega_{n-1} - \Omega_n \qquad \textstyle\sum_{n-1} - \sum_n = \theta^2[\Lambda_{n-1} - \Lambda_n]. \tag{A6}$$

Q.E.D.

Proof of Sufficient Statistics Claims

We define $\hat{y}^i_{n-1}$ recursively as follows (and also show that it belongs to the information set of trader i after $n - 1$ trading periods):

$$\hat{y}^i_1 = \sum_{j\neq i} \beta_1 \hat{s}^i_{j0} + \beta_1 \hat{s}^i_{i0} + u_1 = \sum_{j\neq i} \beta_1 s_{j0} + \beta_1 s_{i0} + u_1 = z_{i1} + \beta_1 s_{i0} \tag{7}$$

which belongs to the information set at time 1. Next:

$$\begin{aligned}
\hat{y}^i_2 &= \sum_{j\neq i} \beta_1 \hat{s}^i_{j1} + \beta_1 \hat{s}^i_{i1} + u_1 \\
&= \sum_{j\neq i} \beta_1 (s_{j0} - \hat{t}^i_{j1}) + \beta_1 (s_{i0} - \hat{t}^i_{i1}) + u_1 \\
&= \sum_{j\neq i} \beta_1 (s_{j0} - t_{j1}) + \sum_{j\neq i} (t_{j1} - \hat{t}^i_{j1}) + \beta_1 (s_{i0} - \hat{t}^i_{i1}) + u_1 \\
&= z_{i2} + \sum_{j\neq i} (t_{j1} - \hat{t}^i_{j1}) + \beta_1 (s_{i0} - \hat{t}^i_{i1}).
\end{aligned} \tag{A8}$$

Because $\hat{t}^i_{j1} = \zeta_1 \hat{y}^i_1$ and trader i knows $\hat{y}^i_1$, this means that trader i also knows $\hat{t}^i_{j1}$ for arbitrary j (which includes $j = i$). Also, t_{j1} is public information and trader i knows z_{i2} and s_{i0}. Thus $\hat{y}^i_2$ is in trader i's information set after two periods of trading.

Next we show that $\hat{y}^i_3, \ldots, \hat{y}^i_{n-1}$ are in the information set using an induction argument. Assume that $\hat{y}^i_1, \ldots, \hat{y}^i_{r-1}$ is in the information set after $r - 1$ periods of trading. This means:

$$\begin{aligned}
\hat{y}^i_r &= \sum_{j\neq i} \beta_r \hat{s}^i_{jr-1} + \beta_r \hat{s}^i_{ir-1} + u_r \\
&= \sum_{j\neq i} \beta_r (s_{j0} - \hat{t}^i_{jr-1}) + \beta_r (s_{i0} - \hat{t}^i_{ir-1}) + u_r \\
&= \sum_{j\neq i} (s_{j0} - t_{jr-1}) + \beta_r (s_{i0} - \hat{t}^i_{ir-1}) + u_r + \sum_{j\neq i} \beta_r (t_{jr-1} - \hat{t}^i_{jr-1}) \\
&= z_{ir} + \sum_{j\neq i} \beta_r (t_{jr-1} - \hat{t}^i_{jr-1}) + \beta_r (s_{i0} - \hat{t}^i_{ir-1})
\end{aligned} \tag{A9}$$

Because $\hat{t}^i_{jr-1} = \sum_{s=1}^{r-1} \zeta_s \hat{y}^i_s$ is known, it follows that $\hat{y}^i_r$ is known after r periods of trading.

Our claim then is that the variables $\hat{s}^i_{in-1}$ (the forecast along the equilibrium path) and $\hat{p}^i_{n-1} - p_{n-1}$ (the deviation from equilibrium prices induced by past suboptimal play) are sufficient statistics for prediction given past deviations by trader i. We prove this as follows:

$$\begin{aligned} &\mathrm{E}[v - p_{n-1}|s_{i0}, y_1, \ldots, y_{n-1}, x_{i1}, \ldots, x_{in-1}] \\ &\mathrm{E}[v - \hat{p}^i_{n-1} + \hat{p}^i_{n-1} - p_{n-1}|s_{i0}, y_1, \ldots, y_{n-1}, x_{i1}, \ldots, x_{in-1}] \\ &= \mathrm{E}[v - \hat{p}^i_{n-1}|s_{i0}, y_1, \ldots, y_{n-1}, x_{i1}, \ldots, x_{in-1}] + (\hat{p}^i_{n-1} - p_{n-1}) \end{aligned} \tag{A10}$$

Because the second term in the last line of equation (A10) is in the information set, we need only consider the first term in the last line of equation (A10). Now:

$$\begin{aligned} &\mathrm{E}[v - \hat{p}^i_{n-1}|s_{i0}, y_1, \ldots, y_{n-1}, x_{i1}, \ldots, x_{in-1}] \\ &= \mathrm{E}[v - \hat{p}^i_{n-1}|s_{i0}, \hat{y}^i_1, \ldots, \hat{y}^i_{n-1}, x_{i1}, \ldots, x_{in-1}] \end{aligned} \tag{A11}$$

where we use the fact that we can construct $(s_{i0}, y_1, \ldots, y_{n-1}, x_{i1}, \ldots, x_{in-1})$ from $(s_{i0}, \hat{y}^i_1, \ldots, y^i_{n-1}, x_{i1}, \ldots, x_{in-1})$ (the reverse has already been shown to be true).

Hence

$$\begin{aligned} &\mathrm{E}[v - \hat{p}^i_{n-1}|s_{i0}, y_1, \ldots, y_{n-1}, x_{i1}, \ldots, x_{in-1}] \\ &= \mathrm{E}[v - \hat{p}^i_{n-1}|s_{i0}, \hat{y}^i_1, \ldots, \hat{y}^i_{n-1}, x_{i1}, \ldots, x_{in-1}] \\ &= \mathrm{E}[v - \hat{p}^i_{n-1}|\hat{s}^i_{in-1}, \hat{y}^i_1, \ldots, \hat{y}^i_{n-1}, x_{i1}, \ldots, x_{in-1}] \\ &= \mathrm{E}[v - \hat{p}^i_{n-1}|\hat{s}^i_{in-1}, x_{i1}, \ldots, x_{in-1}] \\ &= \mathrm{E}[v - \hat{p}^i_{n-1}|\hat{s}^i_{in-1}] \end{aligned} \tag{A12}$$

The fact that $x_{i1}, \ldots, x_{in-1}$ are arbitrary numbers unrelated to equilibrium play is used in the last step in equation (A12).

Proof of Proposition 1:

In the n^{th} round of trading, the i^{th} informed trader maximizes

$$\begin{aligned} &\mathrm{E}[(v - p_n)x_{in}|s_{i0}, z_{i1}, \ldots, z_{in-1}, y_1, \ldots, y_{n-1}] \\ &\quad + \mathrm{E}[\alpha_n(\hat{s}^i_{in})^2 + \psi_n\hat{s}^i_{in}(\hat{p}^i_n - p_n) \\ &\quad + \mu_n(\hat{p}^i_n - p_n)^2|s_{i0}, z_{i1}, \ldots, z_{in-1}, y_1, \ldots, y_n] + \delta_n \end{aligned} \tag{A13}$$

Because

$$p_n = p_{n-1} + \lambda_n y_n = p_{n-1} + \lambda_n\left(\sum_{j\neq i} \beta_n s_{jn-1} + x_{in} + u_n\right) \tag{A14}$$

we obtain the first order condition for maximization with respect to x_{in} as:

$$\begin{aligned} &\mathrm{E}\left[\left(v - p_{n-1} - \lambda_n \sum_{j \neq i} \beta_n s_{jn-1}\right) \middle| s_{i0}, z_{i1}, \ldots, z_{in-1}, x_{i1}, \ldots, x_{in-1}\right] \\ &\quad - 2\lambda_n x_{in} \\ &\quad + \mathrm{E}[\psi_n \hat{s}^i_{in}(-\lambda_n) + 2\mu_n(\hat{p}^i_n - p_n)(-\lambda_n) | s_{i0}, z_{i1}, \ldots, z_{in-1}, x_{i1}, \ldots, x_{in-1}] = 0 \end{aligned} \tag{A15}$$

and the second order condition is

$$-2\lambda_n + 2\mu_n \lambda_n^2 < 0. \tag{A16}$$

We expand the first order condition to obtain:

$$\begin{aligned} &\mathrm{E}\left[\left(v - \hat{p}^i_{n-1} - \lambda_n \sum_{j \neq i} \beta_n \hat{s}^i_{jn-1}\right) \middle| s_{i0}, z_{i1}, \ldots, z_{in-1}, x_{i1}, \ldots, x_{in-1}\right] \\ &\quad + (\hat{p}^i_{n-1} - p_{n-1}) + \lambda_n \sum_{j \neq i} \beta_n(\hat{s}^i_{jn-1} - s_{jn-1}) - 2\lambda_n x_{in} \\ &\quad - \lambda_n \psi_n \mathrm{E}[\hat{s}^i_{in} | s_{i0}, z_{i1}, \ldots, z_{in-1}, x_{i1}, \ldots, x_{in-1}] \\ &\quad - 2\mu_n \lambda_n \mathrm{E}[(\hat{p}^i_n - p_n) | s_{i0}, z_{i1}, \ldots, z_{in-1}, x_{i1}, \ldots, x_{in-1}] = 0 \end{aligned} \tag{A17}$$

where

$$\hat{p}^i_{n-1} - p_{n-1} = \sum_{r=1}^{n-1} \lambda_r(\hat{y}^i_r - y_r) \tag{A18}$$

is in the information set of informed agent i. Similarly

$$\begin{aligned} \hat{s}^i_{jn-1} - s_{jn-1} &= t_{jn-1} - \hat{t}^i_{jn-1} \\ &= \sum_{r=1}^{n-1} \zeta_r y_r - \sum_{r=1}^{n-1} \zeta_r \hat{y}^i_r \\ &= -\frac{1}{\theta}\left[\sum_{r=1}^{n-1} \lambda_r \hat{y}^i_r - \sum_{r=1}^{n-1} \lambda_r y_r\right] \\ &= -\frac{1}{\theta}(\hat{p}^i_{n-1} - p_{n-1}) \end{aligned} \tag{A19}$$

is also in the information set of informed agent i. We also note that

$$\begin{aligned} \mathrm{E}[v - \hat{p}^i_{n-1} | s_{i0}, z_{i1}, \ldots, z_{in-1}, x_{i1}, \ldots, x_{in-1}] &= \mathrm{E}[v - \hat{p}^i_{n-1} | \hat{s}^i_{in-1}] \\ &= \eta_n \hat{s}^i_{in-1} \end{aligned} \tag{A20}$$

and

$$\mathrm{E}[\hat{s}^i_{jn-1} | s_{i0}, z_{i1}, \ldots, z_{in-1}, x_{i1}, \ldots, x_{in-1}] = \mathrm{E}[\hat{s}^i_{jn-1} | \hat{s}^i_{in-1}] = \phi_n \hat{s}^i_{in-1} \tag{A21}$$

Finally, one can show that

$$\begin{aligned} \hat{p}^i_n - p_n &= (\hat{p}^i_{n-1} - p_{n-1}) + \lambda_n(\hat{y}^i_n - y_n) \\ &= \left[1 - \frac{\lambda_n \beta_n (M-1)}{\theta}\right](\hat{p}^i_{n-1} - p_{n-1}) + \lambda_n \beta_n \hat{s}^i_{in-1} - \lambda_n x_{in} \end{aligned} \tag{A22}$$

Using all of these relations we can show that

$$x_{in} = \beta_n \hat{s}^i_{in-1} + \gamma_n(\hat{p}^i_{n-1} - p_{n-1}) \tag{A23}$$

where

$$\begin{aligned} \beta_n &= \frac{\eta_n - \lambda_n \psi_n}{\lambda_n[1 + (1 + \phi_n(M-1))(1 - ((\lambda_n \psi_n)/\theta))]} \\ \gamma_n &= \frac{(1 - 2\mu_n \lambda_n)(1 - (\lambda_n \beta_n (M-1)/\theta))}{2\lambda_n(1 - \mu_n \lambda_n)} \end{aligned} \tag{A24}$$

Substituting back the optimal choice of x_{in} and simplifying we obtain (more details are available from the authors on request):

$$\alpha_{n-1} = (\eta_n - \lambda_n \beta_n [1 + (M-1)\phi_n])\beta_n + \alpha_n \left[1 - \frac{\lambda_n \beta_n}{\theta}[1 + (M-1)\phi_n]\right]^2$$

$$\begin{aligned} \psi_{n-1} &= (\eta_n - \lambda_n \beta_n [1 + (M-1)\phi_n])\gamma_n - \lambda_n \gamma_n \beta_n + \beta_n \left(1 - \frac{\lambda_n \beta_n (M-1)}{\theta}\right) \\ &\quad + \psi_n \left(1 - \frac{\lambda_n \beta_n}{\theta}[1 + (M-1)\phi_n]\right)\left(1 - \frac{\lambda_n \beta_n (M-1)}{\theta} - \lambda_n \gamma_n\right) \end{aligned} \tag{A25}$$

$$\mu_{n-1} = -\lambda_n \gamma_n^2 + \gamma_n \left(1 - \frac{\lambda_n \beta_n (M-1)}{\theta}\right) + \mu_n \left(1 - \frac{\lambda_n \beta_n (M-1)}{\theta} - \lambda_n \gamma_n\right)^2$$

$$\begin{aligned} \delta_{n-1} &= \delta_n + \frac{\alpha_n \lambda_n^2}{\theta^2}\sigma_u^2 + \alpha_n \frac{\lambda_n^2 \beta_n^2}{\theta^2}[(M-1)\ \mathrm{Var}\ (s_{jn-1} | s_{i0}, y_1, \ldots, y_{n-1})] \\ &\quad + \alpha_n \frac{\lambda_n^2 \beta_n^2}{\theta^2}[[(M-2)(M-1)]\ \mathrm{Cov}\ (s_{jn-1}, s_{kn-1} | s_{i0}, y_1, \ldots, y_{n-1})] \end{aligned}$$

To compute ϕ_n, note that ϕ_n is the regression coefficient of s_{jn-1} on s_{in-1} and is:

$$\phi_n = \frac{\mathrm{Cov}(s_{in-1}, s_{jn-1})}{\mathrm{Var}(s_{in-1})} = \frac{\Omega_{n-1}}{\Lambda_{n-1}} \tag{A26}$$

Next, η_n is the regression of coefficient of $v - p_{n-1}$ on s_{in-1} and is computed as follows:

$$\begin{aligned}
\eta_n &= \frac{\mathrm{Cov}(v-p_{n-1}, s_{in-1})}{\mathrm{Var}(s_{in-1})} \\
&= \frac{(\theta/M)\ \sum_{j=1}^{M} \mathrm{Cov}(s_{jn-1}, s_{in-1})}{\mathrm{Var}(s_{in-1})} \\
&= \frac{\theta}{M}\frac{(M-1)\Omega_{n-1}+\Lambda_{n-1}}{\Lambda_{n-1}} \\
&= \frac{\theta}{M}[1 + (M-1)\phi_n]
\end{aligned} \tag{A27}$$

Next we obtain:

$$\begin{aligned}
&\mathrm{Var}(s_{jn-1}|s_{i0}, y_1, \ldots, y_{n-1}) \\
&\quad = \mathrm{E}[(s_{jn-1} - \phi_n s_{in-1})^2] = \mathrm{E}[(s_{jn-1})^2] - \phi_n^2\mathrm{E}[(s_{in-1})^2] \\
&\quad = \mathrm{Var}(s_{j0}|y_1, \ldots, y_{n-1}) - \phi_n^2\mathrm{Var}(s_{i0}|y_1, \ldots, y_{n-1}) \\
&\quad = [1 - \phi_n^2]\Lambda_{n-1} = \frac{\Lambda_{n-1}^2 - \Omega_{n-1}^2}{\Lambda_{n-1}}
\end{aligned} \tag{A28}$$

Similarly,

$$\begin{aligned}
&\mathrm{Cov}(s_{jn-1}, s_{kn-1}|s_{i0}, y_1, \ldots, y_{n-1}) \\
&= \mathrm{Cov}(s_{jn-1} - \phi_n s_{in-1}, s_{kn-1} - \phi_n s_{in-1}) \\
&= \mathrm{Cov}(s_{jn-1}, s_{kn-1}) - \phi_n\ \mathrm{Cov}(s_{jn-1}, s_{in-1}) \\
&\quad - \phi_n\mathrm{Cov}(s_{in-1}, s_{kn-1}) + \phi_n^2\mathrm{Cov}(s_{in-1}, s_{in-1}) \\
&= \Omega_{n-1} - \phi_n\Omega_{n-1} - \phi_n\Omega_{n-1} + \phi_n^2\Lambda_{n-1} \\
&= \frac{\Omega_{n-1}}{\Lambda_{n-1}}(\Lambda_{n-1} - \Omega_{n-1})
\end{aligned} \tag{A29}$$

To compute λ_n we note that

$$y_n = \sum_{j=1}^{M} \beta_n s_{in-1} + u_n$$

$$= \beta_n M \left[\frac{1}{M} \sum_{j=1}^{M} (s_{i0} - t_{in-1}) \right] + u_n \tag{A30}$$

$$= \frac{\beta_n M}{\theta} (\theta \hat{v} - p_{n-1}) + u_n$$

We then obtain from the projection theorem,

$$\lambda_n = \frac{(\beta_n M/\theta) \sum_{n-1}}{(\beta M/\theta)^2 \sum_{n-1} + \sigma_u^2} \tag{A31}$$

and

$$\sum_n = \sum_{n-1} - \lambda_n^2 \operatorname{Var}(y_n)$$

$$= \sum_{n-1} - \lambda_n \operatorname{Cov}(\theta \hat{v} - p_{n-1}, y_n) \tag{A32}$$

$$= \frac{\sum_{n-1} \sigma_u^2}{(\beta_n M/\theta) \sum_{n-1} + \sigma_u^2}$$

Using equations (A31) and (A32), we obtain that

$$\lambda_n = \frac{\beta_n M \sum_n}{\theta \sigma_u^2} \tag{A33}$$

Finally,

$$\sum_n = \sum_{n-1} - \lambda_n \operatorname{Cov}(\theta \hat{v} - p_{n-1}, y_n)$$

$$= \sum_n - \lambda_n \frac{\beta_n M}{\theta} \sum_{n-1} \tag{A34}$$

$$= \left(1 - \frac{M \lambda_n \beta_n}{\theta}\right) \sum_{n-1}$$

Also note that

$$\Lambda_n = \Lambda_{n-1} - \zeta_n \beta_n [\Lambda_{n-1} + (M-1)\Omega_{n-1}] = \Lambda_{n-1} - \frac{\lambda_n \beta_n}{\theta} \frac{M}{\theta^2} \sum_{n-1} \tag{A35}$$

and similarly

$$\Omega_n = \Omega_{n-1} - \frac{\lambda_n \beta_n}{\theta} \frac{M}{\theta^2} \Sigma_{n-1}. \tag{A36}$$

Q.E.D.

REFERENCES

Admati, A., and P. Pfleiderer, 1988, A theory of intraday patterns: Volume and price variability, *The Review of Financial Studies* 1, 3–40.

Back, K., 1992, Insider trading in continuous time, *The Review of Financial Studies* 5, 387–409.

Back, K., H. H. Cao, and G. A. Willard, 1995, Imperfect competition among informed traders, Unpublished working paper, Washington University, St. Louis.

Brennan, M., N. Jegadeesh, and B. Swaminathan, 1993, Investment analysis and the adjustment of stock prices to common information, *Review of Financial Studies* 6, 799–824.

Bushman, R., S. Dutta, J. Hughes, and R. Indejejikian, 1993, Strategic behavior of differentially informed traders around public announcements, Unpublished working paper, University of Minnesota.

Cao, H., 1994, Imperfect competition in securities markets with diversely informed traders, Unpublished working paper, UCLA.

Chan, K., Y. P. Chung, and H. Johnson, 1995, The intraday behavior of bid-ask spreads for NYSE stocks and CBOE options, *Journal of Finance and Quantitative Analysis* 30, 329–346.

Cho, J. W., 1995, State space representation and estimation of market microstructure models, Unpublished working paper, Georgia Institute of Technology.

Dutta, P., and A. Madhavan, 1993, Price continuity rules and insider trading, Unpublished working paper, University of Pennsylvania.

Easley, D., N. Kiefer, and M. O'Hara, 1993, One day in the life of a very common stock, Unpublished working paper, Cornell University.

Foster, F. D., and S. Viswanathan, 1993, The effect of public information and competition on trading volume and price volatility, *The Review of Financial Studies* 6, 23–56.

Foster, F. D., and S. Viswanathan, 1994, Strategic trading with asymmetrically informed traders and long-lived information, *Journal of Finance and Quantitative Analysis* 29, 499–518.

Foster, F. D., and S. Viswanathan, 1995, Can speculative trading explain the volume-volatility relation?, *Journal of Business Economics and Statistics* 13, 379–396.

Harris, M., and A. Raviv, 1993, Differences of opinion make a horse race, *Review of Financial Studies* 6, 473–506.

He, H., and J. Wang, 1995, Differential information and dynamic behavior of stock volume, *Review of Financial Studies*, 8, 919–972.

Holden, C. W., and A. Subrahmanyam, 1992, Long-lived private information and imperfect competition, *The Journal of Finance* 47, 247–270.

Kyle, A., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315–1335.

Kyle, A., 1989, Informed speculation with imperfect competition, *Review of Economic Studies* 56, 317–356.

Marcet, A., and T. J. Sargent, 1989a, Convergence of least squares learning in environments with hidden state variables and private information, *Journal of Political Economy* 97, 1306–1322.

Marcet, A., and T. J. Sargent, 1989b, Convergence of least squares learning in environments with self referential linear stochastic models, *Journal of Economic Theory* 48, 337–368.

Michener, R., and C. Tighe, 1991, A Cournot model of insider trading, Unpublished working paper, University of Virginia, Charlottesville.

Singleton, K., 1987, Asset prices in a time-series model with disparately informed, competitive traders, in William Barnett and Kenneth Singleton, Eds.: *New Approaches to Monetary Economics* (Cambridge University Press, Cambridge, MA).

Townsend, R. M., 1983a, Forecasting the forecasts of others, *Journal of Political Economy* 91, 546–88.

Townsend, R. M., 1983b, Equilibrium theory with learning and disparate expectations: Some issues and methods, in Roman Frydman and Edmund S. Phelps, Eds.: *Individual Forecasting and Aggregate Outcomes: "Rational Expectations" Examined* (Cambridge University Press, Cambridge).

Vayanos, D., 1993, A dynamic model of an imperfectly competitive bid-ask market, Unpublished working paper, MIT.

Vives, X., 1988, Aggregation of information in large Cournot markets, *Econometrica* 56, 851–876.

Vives, X., 1993, How fast do rational agents learn?, *Review of Economic Studies* 60, 329–347.

Wang, A., 1993, Strategic trading with heterogeneous prior beliefs, Unpublished working paper, University of North Carolina, Chapel Hill.