Informed Trading and Portfolio Returns[*]

Alex Boulatov[†]
Bauer College of Business
University of Houston

Terrence Hendershott[‡]
Walter A. Haas School of Business
University of California, Berkeley

Dmitry Livdan[§]
Walter A. Haas School of Business
University of California, Berkeley

October 23, 2009

Abstract

We solve a multi-period model of strategic trading with long-lived information in multiple assets with correlated innovations in fundamental values. Market makers in each asset can only condition their price functions on trading in the that asset (but not on trading in the other asset). Using daily non-public data from the New York Stock Exchange we test the model's predictions on the conditional and unconditional lead-lag relations of institutional order flows and returns within portfolios. We find support for the model prediction of positive autocorrelations in portfolio returns as well as the predictions for how informed order flow positively predicts future returns and future informed order flow. As the model predicts we find these relations strengthen for portfolios formed from assets with higher correlation of fundamental values.

1 Introduction

Asset prices following a random walk is the basis for much of theoretical and empirical asset pricing. Therefore, an examination of the predictability of asset returns is the first non-introduction chapter in Campbell, Lo, and MacKinlay (1997). Lo and MacKinley (1988)

show that the autocorrelations of equal-weighted portfolio returns are significantly positive while the autocorrelations of individual asset returns are generally negative. Together these point to positive cross-autocorrelations in asset returns (Lo and MacKinley (1990a)). We study whether or not informed institutional trading is a source of these correlation patterns. We construct a multi-period model of strategic trading with long-lived information in multiple assets with correlated fundamental values. The model's key assumption is that in each trading period prices in each asset are functions of only trading in that asset (and not trading in the other asset). We find support for the model's prediction of positive autocorrelations in daily portfolio returns as well as the predictions for how informed order flow positively predicts future returns and future informed order flow. We also test and find support for the model's comparative static predictions on the fundamental correlation of assets by examining portfolios formed from stocks in the same industry versus portfolios formed from stocks randomly chosen from different industries.

Explanations other than asymmetric information and the slow diffusion of information have been suggested to explain the positive autocorrelation in daily and weekly portfolio returns. Conrad and Kaul (1988) propose time varying risk premia. Lo and MacKinley (1990b) and Boudoukh, Richardson, and Whitelaw (1994) explore asynchronous trading. However, these have proven unsatisfactory, e.g., using several tests based on transaction data Anderson, Eom, Hahn, and Park (2008) find that partial price adjustment is the major source of the portfolio autocorrelations. Lo and MacKinley (1990b), Brennan, Jagadeesh and Swaminathan (1993), Chan (1993), and others suggest that slow adjustment to common information is the source for partial price adjustment; but, Bernhardt and Mahani (2008) show that it is difficult to construct a model in which asymmetric information with strategic trading that leads to positive autocorrelations in portfolio returns.

We propose a simple friction that generates positive autocorrelation in portfolio returns in a multi-period Kyle (1985) style model with strategic trading on long-lived information in multiple assets with correlated innovations in the assets' fundamental values. The key assumption is that in each asset the market makers can only condition their pricing rule on trading in that asset. This friction exists in virtually all markets.[1] This same

---

[1]Allowing orders to be contingent on trading in multiple asset can prevent market clearing as the contingencies on the different orders can be mutually exclusive.

assumption arises from any friction that prevents market makers from observing and perfectly interpreting information from trading in all other assets continuously and instantaneously.[2] Because assets' fundamental values are correlated, order flows in both assets are informative about payoffs in both assets. The informed trader strategically anticipate this by conditioning his orders in each asset on information not directly relevant to that asset's signal (the signal in the other asset).

The model has an analytical solution in the case of two assets traded over two trading rounds when assets have symmetrically distributed fundamental values and liquidity trades are independent across time and assets with the same variance. The informed trader minimizes the informational impact of his trades in both assets by strategically choosing informational trading intensities across time and assets. The informed trader does this across time by curbing the aggressiveness of his trades in both assets in the first period in a manner similar to a two-period version of the original Kyle (1985) model. His cross-asset strategy is quite different. The informed trader sets the sensitivities of his trading demand in either asset to be positive in his information about the asset he trades in and negative in his information about the other asset. This signal jamming strategy reduces the market maker's ability to learn about the value of either asset from both order flows.

The main time-series implications of the model is that individual assets' returns are independent across time while returns on a portfolio of the assets are positively auto-correlated. The individual assets' returns are independent across time because asset-specific order flows are sufficient statistics for prices of their respective assets. In other words, the informed trader does not reuse any information he has traded on in the past since it has been already incorporated into the price. This property of the model also makes current order flow independent of lagged returns.

The intuition behind positive cross-autocorrelation can be illustrated by the following example. Suppose that there are two assets and that in the first trading round the market

[2]In a model without informed trading Chan (1993) uses a similar assumption. In a model with informed trading Chordia, Sarkar, and Subrahmanyam (2008) use a related assumption to examine short-lived information—the informed traders only have one period to trade on their information which is revealed before the informed can trade again. The short-lived information along with the assumption that common factor information is only traded on in the large stocks allows Chordia, Sarkar, and Subrahmanyam to examine time variation in liquidity and information diffusion.

maker in the first asset receives a large positive order flow in his asset but cannot condition on order flow in the other asset. Because the total order flow contains a noisy liquidity component, the market maker adjusts her assets's price only partially upward in response to the large positive order flow. After the transaction takes place, she examines the transaction price of the second asset and uncovers its order flow which contains additional information about her asset. If it is also large and positive, she is more confident about the favorable information and adjusts her asset's price further upward. If it is not, she is less confident about the favorable information and revises the asset's price downward. Therefore, the price change of the first asset in the second trading round is positively correlated with the price change of the second asset in the first trading round. It is worth noting that there is no opportunity for anyone without the informed trader's information to construct a profitable trading strategy based on the positive autocorrelation in portfolio returns. In this sense our results support market efficiency even though there is predictability in portfolio returns.

Except for the zero correlation between informed order flow and lagged returns, the model predicts that portfolios lead-lag autocorrelations between returns and informed order flows increase in the correlation between innovations in the assets' fundamental values. The model also allows for a joint test of return and informed order flow predictability in a conventional vector autoregression setting.

We use daily non-public data from the New York Stock Exchange (NYSE) to test the model's predictions on the correlations of institutional order flows and returns within portfolios and within the asset making up the portfolios. The data set contains seven years from January 1999 to December of 2005 of daily buy and sell volume of executed institutional investor orders for a large cross-section of NYSE stocks.[3] The data set was constructed from the NYSE's Consolidated Equity Audit Trail Data (CAUD) files that contain detailed information on all executed orders. One of the fields associated with each order, Account Type, specifies whether the order comes from an institution. We use institutional order flow (buy volume minus sell volume) normalized by the market cap as a data proxy for the informed order flow in the model. We first test the model's predictions on the market

---

[3]Badrinath, Kale, and Noe (1995) and Sias and Starks (1997) show how the fraction of firms owned by institutions affects portfolio return autocorrelations, but do not examine the relations between institutional flows and returns that is the focus of this paper.

portfolio. We find support for the model's predictions: positive autocorrelations in portfolio returns as well as the predictions for how informed order flow positively predicts future returns and future informed order flow, while returns do not predict informed order flow.

We also test and find support for the model's comparative static predictions on the fundamental correlation of assets by examining portfolios formed from stocks in the same industry versus portfolios formed from stocks randomly chosen from different industries. We show that the lead-lag correlations between returns and informed order flow is higher in industry portfolio than in randomly formed portfolios. We also show that the relevant vector autoregression coefficients for industry portfolios are higher and lower, as predicted, than the same coefficients for random portfolios.

The the paper is organized as follows. The model is presented in Section 2. Our empirical analysis are reported in Section 3. Section 4 concludes.

2 The Model

2.1 Set Up

We consider an economy in which two risky assets are traded in the financial market over two trading rounds.[4] There are three types of risk-neutral agents in the economy: an informed trader, competitive market makers, and a number of liquidity ("noise") traders. At $t = 0$ the informed trader learns the fundamental values of both risky assets, $\mathbf{V} = (V_1, V_2)'$, simultaneously drawn from the joint normal distribution

$$\widetilde{\mathbf{V}} = \begin{pmatrix} \widetilde{V}_1 \\ \widetilde{V}_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sigma_0^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right). \tag{1}$$

The statistical properties of the assets' fundamental values are summarized in Assumption 1.[5]

ASSUMPTION 1: *The fundamental values of the assets, $\widetilde{V}_{1,2}$, are positively correlated, $\rho > 0$*

---

[4]The analysis can be extended to the case of multiple risky assets

[5]This and some other assumptions are made for the sake of clarity and tractability. Under these assumptions the model can be solved analytically. By reducing the degrees of freedom these assumptions also put a much higher hurdle on the model fit to the data.

, *have means normalized to zero, and the same variance,* σ_0^2.

Each asset is handled by competitive market makers. Market makers do not know V_1 and V_2 until the announcement after the second and final trading round. Before any trading takes place, market makers know the unconditional joint distribution of $\widetilde{\mathbf{V}}$ and thus quote $\mathbf{P}(0) = \mathrm{E}[\widetilde{\mathbf{V}}] = \mathbf{0}$ at $t = 0$. The aggregate liquidity demand in asset $i = 1, 2$ at time $t = 1, 2$, $\mathbf{u}(t) = (u_1(t), u_2(t))'$, is normally distributed with zero mean and variance equal to σ_u

$$\mathbf{u}(t) \sim N\left(\mathbf{0}, \sigma_u^2\mathbf{I}\right), \tag{2}$$

where $\mathbf{I}$ is a 2×2 unit matrix. Assumption 2 outlines cross-sectional and intertemporal properties of the liquidity demands. This assumption simplifies the market makers' inference problem, but is not crucial.

ASSUMPTION 2: *The liquidity demands are independent across assets,* $\mathrm{Cov}(u_i(t), u_j(t)) = 0$, *and time,* $\mathrm{Cov}(u_i(1), u_i(2)) = 0$.

Figure 1 illustrates the sequence of events and the information sets of all agents. In the first trading rounds the informed trader takes into account the correlation between assets by conditioning his demand in each asset, $\mathbf{x}(t) = (x_1(t), x_2(t))'$, on the value of the other asset, $V_{j \neq i}$. The informed order flows at $t = 1$ are given by

$$\mathbf{x}(1) = \begin{pmatrix} \beta_{11}(1) & \beta_{12}(1) \\ \beta_{12}(1) & \beta_{22}(1) \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \beta(1)\mathbf{V}, \tag{3}$$

and the corresponding total order flows, $\mathbf{y}(t) = (y_1(t), y_2(t))'$, are given by

$$\mathbf{y}(1) = \beta(1)\mathbf{V} + \mathbf{u}(1). \tag{4}$$

[insert Figure 1]

During the trading rounds market makers observe order flows and use Bayes rule to update their beliefs about V_1 and V_2. Assumption 3 outlines the strategies available to the market makers.

ASSUMPTION 3: *In both trading rounds the market makers condition their pricing functions in an asset on order flow in that asset (and not on order flow in the other asset). After the first trading round, time $t = 1^+$, market makers observe the prices of both assets, infer order flow in both assets, and adjust prices before the second trading round.*

At a high enough frequency, Assumption 3 is consistent with the fact that no market structure allows market makers (or any traders) to condition their prices in one asset on trading in another asset. Even if market makers could condition their pricing function in an asset on order flow in other assets, Assumption 3 is a reduced-form way of capturing any friction precluding the market makers' from instantaneously and fully processing and acting on all information in all securities.[6]

Assumption 3 is crucial for our analysis. If market makers observe and condition prices on order flows in both assets, then, as in Kyle (1985), prices would be fully informationally efficient: information in the current order flow for either asset being orthogonal to the information in the next period order flow for both assets. As a result, order flows and, therefore, returns for individual assets and their equal-weighted portfolio would not be predictable.

When Assumption 3 holds market makers condition prices only on their own order flows their prices are less than fully informationally efficient and the pricing error of the first asset is correlated with the order flow in the second asset. This is because the assets' fundamental values are correlated and, therefore, the order flows they do not observe contain additional information about the asset they make. Further, although each asset price is an unbiased estimate of the true asset value conditional on its own order flow, the sum of asset prices (the price of an equal-weighted portfolio) is not an unbiased estimate of the true aggregate value conditional on both order flows. Consequently, when market makers correct pricing errors using order flow inferred from the price change of the other asset, stock returns will be positively cross-autocorrelated.

In accordance with Assumption 3, competitive market makers observe the order flows in

---

[6]The informed trader acts strategically to take advantage of this inefficiency in order to maximize his expected profits.

the assets they make and set prices at $t = 1$ according to the schedules

$$\mathbf{P}(1) = \begin{pmatrix} \mathrm{E}[\widetilde{V}_1|y_1(1)] \\ \mathrm{E}[\widetilde{V}_2|y_2(1)] \end{pmatrix} = \lambda(1)\mathbf{y}(1). \tag{5}$$

In the spirit of Kyle (1985), the inverse market depth parameter $\lambda_i(1)$ in (5) is a slope coefficient in the linear regression of $\widetilde{V}_i$ on $y_i(1)$:

$$\lambda(1) = \begin{pmatrix} \frac{\mathrm{Cov}(\widetilde{V}_1, y_1(1))}{\mathrm{Var}(y_1(1))} & 0 \\ 0 & \frac{\mathrm{Cov}(\widetilde{V}_2, y_2(1))}{\mathrm{Var}(y_2(1))} \end{pmatrix} = \tag{6}$$

$$= \begin{pmatrix} \frac{\beta_{11}(1) + \rho\beta_{12}(1)}{\beta_{11}^2(1) + 2\rho\beta_{11}(1)\beta_{12}(1) + \beta_{12}^2(1) + \beta_K^2} & 0 \\ 0 & \frac{\beta_{22}(1) + \rho\beta_{21}(1)}{\beta_{22}^2(1) + 2\rho\beta_{22}(1)\beta_{21}(1) + \beta_{21}^2(1) + \beta_K^2} \end{pmatrix},$$

where $\beta_K \equiv \frac{\sigma_u}{\sigma_0}$ refers to the elasticity of informed order flow to private information in Kyle (1985). At the end of the first trading round, $t = 1^+$, after trades take place the market makers observe prices of the assets they do not make. Market makers use posted prices to infer order flows they did not observe at the start of the first trading round and use them to adjust the prices of their assets to the full information level according to

$$\mathbf{P}(1^+) = \mathrm{E}[\widetilde{\mathbf{V}}|y_1(1), y_2(1)] = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12} & \Lambda_{22} \end{pmatrix} \begin{pmatrix} y_1(1) \\ y_2(1) \end{pmatrix} = \mathbf{\Lambda}\mathbf{y}(1). \tag{7}$$

where Λ_{ij} are slope coefficients in the linear regression of $\widetilde{V}_i$ on $y_j(1)$

$$\Lambda_{i1} = \frac{\mathrm{Cov}(\widetilde{V}_i, y_1(1))\mathrm{Var}(y_2(1)) - \mathrm{Cov}(\widetilde{V}_i, y_2(1))\mathrm{Cov}(y_1(1), y_2(1))}{\mathrm{Var}(y_1(1))\mathrm{Var}(y_2(1)) - (\mathrm{Cov}(y_1(1), y_2(1)))^2}, \tag{8}$$

$$\Lambda_{i2} = \frac{\mathrm{Cov}(\widetilde{V}_i, y_2(1))\mathrm{Var}(y_1(1)) - \mathrm{Cov}(\widetilde{V}_i, y_1(1))\mathrm{Cov}(y_1(1), y_2(1))}{\mathrm{Var}(y_1(1))\mathrm{Var}(y_2(1)) - (\mathrm{Cov}(y_1(1), y_2(1)))^2}, \quad i = 1, 2.$$

Evaluating the slope coefficients yields

$$\mathbf{\Lambda} = \mathbf{\Psi}^{-1}\beta(1)', \tag{9}$$

where the covariance matrix of the fundamental values conditional on the order flows,

$\operatorname{Cov}\left(\widetilde{\mathbf{V}}|\,\mathbf{y}\,(1)\right)$, is

$$\boldsymbol{\Psi} = \begin{pmatrix} \psi_{11} & \psi_{12} \\ \psi_{12} & \psi_{22} \end{pmatrix} = \begin{pmatrix} \frac{\beta_K^2}{1-\rho^2} + \beta_1\,(1)\,\beta_1'\,(1) & -\frac{\rho\beta_K^2}{1-\rho^2} + \beta_1\,(1)\,\beta_2'\,(1) \\ -\frac{\rho\beta_K^2}{1-\rho^2} + \beta_1\,(1)\,\beta_2'\,(1) & \frac{\beta_K^2}{1-\rho^2} + \beta_2\,(1)\,\beta_2'\,(1) \end{pmatrix}, \tag{10}$$

where $\beta_i\,(1)$ denotes ith row of the matrix $\beta\,(1)$. The "efficient" prices (7) are quoted to the informed trader at the start of the second trading round.

At the beginning of the second trading round, $t = 2$, the informed trader submits market orders

$$\mathbf{x}\,(2) = \beta\,(2)\left(\mathbf{V} - \mathbf{P}(1^+)\right), \tag{11}$$

and market makers receive the following total order flows

$$\mathbf{y}\,(2) = \beta\,(2)\left(\mathbf{V} - \mathbf{P}(1^+)\right) + \mathbf{u}\,(2)\,. \tag{12}$$

Analogous to the first trading period, market makers can only condition their pricing functions on the order flows in each assets and not on the order flows in the other asset:

$$\mathbf{P}\,(2) = \mathbf{P}(1^+) + \lambda\,(2)\,\mathbf{y}\,(2)\,. \tag{13}$$

where the inverse market depth parameter $\lambda\,(2)$ in (13) is the slope coefficient in the linear regression of : $\widetilde{V}_i - P_1(1^+)$ on $y_i\,(2)$

$$\lambda\,(2) = \begin{pmatrix} \frac{\operatorname{Cov}(\widetilde{V}_1 - P_1(1^+),y_1(2))}{\operatorname{Var}(y_1(2))} & 0 \\ 0 & \frac{\operatorname{Cov}(\widetilde{V}_2 - P_2(1^+),y_2(2))}{\operatorname{Var}(y_2(2))} \end{pmatrix}. \tag{14}$$

At the end of the second trading round $\mathbf{V}$ is revealed to all agents and payoffs are realized.

2.2 Solution and Results

The model is solved by finding a vector of trading strategies, $\mathbf{x}^*(t)$, which maximizes the informed trader's expected profits over both trading rounds

$$\mathbf{x}^*\,(t) = \underset{\mathbf{x}(t)}{\arg\max}\ \mathrm{E}\underbrace{\left[\sum_{t=1}^{2}\sum_{i=1}^{2} x_i\,(t)\,(V_i - P_i(t))\,|\widetilde{\mathbf{V}} = \mathbf{V}\right]}_{\text{Informed trader's total expected payoff}}. \tag{15}$$

9

The following theorem summarizes the solution of the model. The symmetry of the asset values' statistical properties together with the independence of liquidity demand across assets and time allow for the model to be solved up to a system of nonlinear equations.

THEOREM 1: *There exists a linear solution to the informed trader's profit maximization problem* (15) *characterized by the following parameters.*
At time $t = 1$ *the informed trader's strategy,* $\mathbf{x}(1) = \beta(1)\mathbf{V}$, *is characterized by* $\beta_{11}(1) = \beta_{22}(1) \equiv \beta_+$ *and* $\beta_{12}(1) = \beta_{21}(1) \equiv \beta_-$ *which are the solutions to the system of nonlinear equations:*

$$\beta_+ + \beta_- = \beta_K \left(2\beta_K\lambda_1 + \left(2\lambda_2\beta_K(1+\rho) \left(\frac{1}{(1+\rho)} + \left(\frac{\beta_+ + \beta_-}{\beta_K} \right)^2 \right)^2 \right)^{-1} \right)^{-1}, \quad (16)$$

$$\beta_+ - \beta_- = \beta_K \left(2\beta_K\lambda_1 + \left(2\lambda_2\beta_K(1-\rho) \left(\frac{1}{(1-\rho)} + \left(\frac{\beta_+ - \beta_-}{\beta_K} \right)^2 \right)^2 \right)^{-1} \right)^{-1}.$$

The market maker pricing function is $\mathbf{P}(1) = \lambda(1)(\beta(1)\mathbf{V} + \mathbf{u}(1)) = \lambda(1)\mathbf{y}(1)$, *with:*

$$\lambda_1(1) = \lambda_2(1) \equiv \lambda_1 = \frac{\beta_+ + \rho\beta_-}{\beta_+^2 + 2\rho\beta_+\beta_- + \beta_-^2 + \beta_K^2}. \quad (17)$$

At time $t = 1^+$ *the asset prices are updated using both order flows* $\mathbf{P}(1^+) = \mathbf{\Lambda}\mathbf{y}(1)$, *with*

$$\Lambda_{11} = \Lambda_{22} \equiv \Lambda_+ = \frac{\psi_+\beta_+ - \psi_-\beta_-}{\psi_+^2 - \psi_-^2}, \quad (18)$$

$$\Lambda_{12} = \Lambda_{21} \equiv \Lambda_- = \frac{\psi_+\beta_- - \psi_-\beta_+}{\psi_+^2 - \psi_-^2},$$

where $\psi_+ = \frac{\beta_K^2}{1-\rho^2} + \beta_+^2 + \beta_-^2$ *and* $\psi_- = -\frac{\rho\beta_K^2}{1-\rho^2} + 2\beta_+\beta_-$.
At time $t = 2$ *the informed trader's strategy is* $\mathbf{x}(2) = \beta(2)(\mathbf{V} - \mathbf{P}(1^+))$ *and the market maker pricing function is* $\mathbf{P}(2) = \mathbf{P}(1^+) + \lambda(2)(\mathbf{x}(2) + \mathbf{u}(2))$, *with*

$$\beta_{11}(2) = \beta_{22}(2) = \frac{1}{\lambda_2}, \quad (19)$$

$$\beta_{12}(2) = \beta_{21}(2) = 0,$$

$$\lambda_1(2) = \lambda_2(2) \equiv \lambda_2 = \sqrt{\frac{1 - \beta_+\Lambda_+ - \beta_-\Lambda_- - \rho(\beta_+\Lambda_- + \beta_-\Lambda_+)}{4\beta_K^2}}. \quad (20)$$

Proof: All proofs are in the Appendix.

Figure 2 provides a graphical illustration of the parameters that characterize Theorem 1 as a function of the correlation of the fundamental values of the assets (ρ). The symmetry of the assets and the independence of the liquidity demand across both time and assets leads to a symmetric solution. The friction of the market makers not being able to condition their prices on order flow in both assets leads to complex first-period trading strategy by the informed traders characterized by (16). The informed trader's strategy solves the fixed point problem discussed in the proof that translates into the system of nonlinear equations (16). The solution is unique and can be solved for numerically. Figure 2 does this fixing the parameter $\beta_K = 1$ and solving equations (16) for all values of ρ.

[insert Figure 2]

Panel A of Figure 2 graphs the informed trader's first-period trading strategy. The solution is symmetric and is described by the linear trading intensity in each asset as a function of the final payoff in that asset, β_+, and the linear trading intensity of in each asset as a function of the final payoff in the other asset, β_-. Panel A has 4 lines: β_+, β_-, $\beta_+ + \beta_-$, and $\beta_+(\rho = 0)$. The latter two lines are included to illustrate how the solution changes with ρ. When the correlation between assets is zero β_+ is simply the first-period β in a standard 2-period Kyle (1985) model. Because the informed knows the final payoffs with certainty there is no reason to trade on the information in one asset in the other asset and the cross-asset β is zero, $\beta_- = 0$. As the assets become more correlated, ρ increases, the informed trader increases his first-period trading intensity, β_+. This is because the correlation of the assets improves the market makers ability to learn when observing prices of both assets. This makes the second-period prices $\mathbf{P}(1^+)$ more informative and decreases the profitability of the informed's second-period profits. Knowing this the informed increases his first-period trading intensity.

Because the asset correlation improves the market makers' learning from trading in both assets, the informed tries to disguise his first-period trading by negatively trading on the cross-asset information: β_- is negative and decreasing in ρ. This negatively trading on information that is irrelevant to the final value of that asset, but relevant to the market

11

makers' learning, reduces the informed trader's first-period profits.[7] However, jamming the market makers' inference problem increases the informed's second-period profits even more. This leads to the overall trading intensity dropping with the asset correlation, $\beta_- + \beta_+$ decreases in ρ relative to $\beta_+(\rho = 0)$.

The signal jamming affects the market makers' pricing function in period 1, λ_1 from equation (17), and period 2, Λ_+, Λ_-, and λ_2 from equations (18) and (20). The parameters that characterize these are graphed in Panels B and C of Figure 2. The optimal tradeoff between increasing the within asset trading intensity β_+ and decreasing the cross-asset trading intensity β_- is reflected in the market makers updating after the first period trading. The within asset price impact function is given by Λ_+ and the cross-asset price impact Λ_-. At $\rho = 0$ the price impact functions are the standard Kyle (1985) 2-period solution: $\Lambda_- = 0$ and $\Lambda_+ = \lambda_1$. As the correlation increases the informed trader tries to jam the market maker by making β_- more negative. Because the jamming is less than complete the effect of the correlation in signals is greater than the jamming and the maker makers respond by increasing the sensitivity of price to the cross asset order flows at time $t = 1^+$: Λ_- increases in ρ. The signal jamming does reduce the within asset price impact at time $t = 1^+$, Λ_+. Because signal jamming is costly to first-period profits, the informed trader only partially jams the signal and the total price impact at time $t = 1^+$, $\Lambda_+ + \Lambda_-$, increases. The signal jamming also increases first-period depth as seen by the decrease in the $t = 1$ price impact λ_1. The less than full signal jamming also means that more information in revealed before second trading period begins so the market makers use a smaller second period price impact λ_2.

Using the equilibrium given in Theorem 1 we analyze the lead-lag correlations of individual assets' returns and informed order flows. As is standard we use incremental price changes, $\Delta P_i(t) = P_i(t) - P_i(t-1)$, as a proxy for assets' returns. Proposition 1 summarizes the within asset and across asset lead-lag relations. In the data we will

---

[7]While this is the first instance of signal jamming in the strategic trading literature, signal jamming has been studied in other settings. Fudenberg and Tirole (1986) consider the situation in which the information available to an entrant (who is uncertain about future profitability) is distorted by an incumbent through the use of predatory pricing. Stein (1989) considers a signal-jamming model of managerial incentives in a capital market, where managers distort current earnings by borrowing at adverse rates from the future. We are not aware of prior work on signal jamming in a multidimensional signaling setting.

follow the usual convention of examining autocorrelations and cross-autocorrelations. We present covariances here because the expressions are simpler and of the same sign as the corresponding correlations.

PROPOSITION 1: *The lead-lag covariances of assets' returns are given by*

$$\text{Cov}\left(\Delta P_i\left(1\right), \Delta P_i\left(2\right)\right) = 0, \ i = 1, 2, \tag{21}$$

$$\text{Cov}\left(\Delta P_i\left(1\right), \Delta P_{j\neq i}\left(2\right)\right) = \frac{\left(1 - \rho^2\right)\lambda_1(\lambda_1 - \Lambda_+)\left(\psi_+^2 - \psi_-^2\right)}{\rho\left(\beta_+^2 + \beta_-^2\right) + 2\beta_+\beta_-}\sigma_0^2 \geq 0, \ i, j = 1, 2.$$

The lead-lag covariances of informed order flows are given by

$$\text{Cov}\left(x_i\left(1\right), x_i\left(2\right)\right) = \frac{\Lambda_+}{2\lambda_2}\sigma_u^2 > 0, \ i = 1, 2, \tag{22}$$

$$\text{Cov}\left(x_i\left(1\right), x_{j\neq i}\left(2\right)\right) = \frac{\Lambda_-}{2\lambda_2}\sigma_u^2 \geq 0, \ i, j = 1, 2.$$

The lead-lag covariances of informed order flows and price returns are given by

$$\text{Cov}\left(x_i\left(1\right), \Delta P_i\left(2\right)\right) = (2\lambda_1 - \Lambda_+)\frac{\sigma_u^2}{2} > 0, \ i = 1, 2, \tag{23}$$

$$\text{Cov}\left(x_i\left(1\right), \Delta P_{j\neq i}\left(2\right)\right) = -\frac{\Lambda_-}{2}\sigma_u^2 + \frac{1}{\lambda_1}\text{Cov}\left(\Delta P_i\left(1\right), \Delta P_{j\neq i}\left(2\right)\right) \geq 0, \ i, j = 1, 2,$$

and

$$\text{Cov}\left(\Delta P_i\left(1\right), x_j\left(2\right)\right) = 0, \ i, j = 1, 2. \tag{24}$$

The intuition behind the autocovariances and cross-autocovariances is straightforward. Because the informed trader strategically trades so as to not reveal all of his information in the first period the within asset informed order flow positively autocovary, $\text{Cov}(x_i(1), x_i(2)) > 0$. Because the assets final values are positively correlated the informed order flow has positive contemporaneous covariance, $\text{Cov}(x_1(1), x_2(1)) > 0$, and cross autocovariance, $\text{Cov}(x_1(1), x_2(2)) > 0$. Because order flow has positive price impact and the uniformed order flow is not observable to the market makers, the covariance of informed order flows with subsequent price changes has the positive signs as with subsequent informed order flow: $\text{Cov}(x_i(1), \Delta P_j(2)) > 0$.

As is standard in Kyle models price changes in each asset are a martingale:

$\text{Cov}(\Delta P_i(1), \Delta P_i(2)) = 0$. This is because asset-specific order flows are sufficient statistics for prices of their respective assets. In other words, the informed trader does not reuse any information he has traded on in the past because it has been already incorporated into the price. This property of the model also makes current order flows independent of lagged returns: $\text{Cov}(\Delta P_i(1), x_j(2)) = 0$.

The positive cross-autocovariance follows from market makers' cross-asset learning about information that is common across assets. Upon observing the order flow in one asset the market makers fully update that asset's price based on the information in order flow. Upon observing the price in the other asset the market makers can infer the order flow in the other asset. The correlation in the assets fundamental values allows the market makers to use both asset prices/order flows to better filter out the noise trading. If the price change in the second asset is consistent with the price change in the first asset, then the market makers further update prices in same direction. If the price change in the second asset is not consistent with the price change in the first asset, then the market makers revises price in the opposite direction.

Proposition 1's results on the within and cross asset autocovariances provide the intuition necessary to understand the results for a portfolio of the assets. Next we combine both assets into the equal-weighted portfolio

$$\Delta P_p(t) = \frac{1}{2}(\Delta P_1(t) + \Delta P_2(t)), \tag{25}$$
$$x_p(t) = \frac{1}{2}(x_1(t) + x_2(t)),$$

and study its lead-lag covariances in Proposition 2.

PROPOSITION 2: *The equal-weighted portfolio of securities is characterized by the following lead-lag covariances*

$$\text{Cov}\left(\Delta P_p(1), \Delta P_p(2)\right) = \frac{1}{2}\text{Cov}\left(\Delta P_1(1), \Delta P_2(2)\right) \geq 0, \tag{26}$$
$$\text{Cov}\left(x_p(1), x_p(2)\right) = \frac{\Lambda_+ + \Lambda_-}{4\lambda_2}\sigma_u^2 > 0,$$
$$\text{Cov}\left(x_p(1), \Delta P_p(2)\right) = \frac{\sigma_u^2}{2}\left(\lambda_1 - \frac{\Lambda_+ + \Lambda_-}{2}\right) + \frac{1}{\lambda_1}\text{Cov}\left(\Delta P_1(1), \Delta P_2(2)\right) \geq 0,$$
$$\text{Cov}\left(\Delta P_p(1), x_p(2)\right) = 0.$$

The portfolio autocoviances in Proposition 2 are averages of the within and across asset autocovariances in Proposition 1. Because within asset price changes have zero autocovariance and price changes have positive cross autocovariance the portfolio price changes positively autocovary. Informed trading in each individual asset predicts subsequent informed order flow and price changes in both assets so the informed order flow in the portfolio has positive autocovariance with the subsequent informed order flow and price changes in the portfolio. Past price changes in each asset do not covary with subsequent order flow in either asset so consequently the price change of the portfolio has zero covariance with informed trading in the portfolio.

3 Empirical Tests

To test the empirical predictions of the model we need to choose a time horizon corresponding to the trading periods in the model. In our standard Kyle setting trading occurs in batches. Assumption 3 requires that the time period is short enough that the market makers cannot condition their trading on order flow in both assets. However, we also would like to see if the effects of the friction in Assumption 3 are relevant at more than very short horizons. To balance these considerations we focus on daily returns and trading. We expect the effects to be larger at shorter horizons and attenuate at longer horizons. Below we examine correlations and regressions using lags longer than one and the results confirm the intuition that the results are stronger at shorter horizons. To proxy for informed order flow we will use a measure of institutional trading from the NYSE.

3.1 Data

The data set contains seven years of daily buy and sell volume of executed institutional investor orders for all NYSE common domestic stocks that were traded on the NYSE any time between January 1, 1999 and December 31, 2005. The data set was constructed from the NYSE's Consolidated Equity Audit Trail Data (CAUD) files that contain detailed information on all orders that execute on the exchange, both electronic (from the limit order book) and manual (those handled by floor brokers). One of the fields associated with the

buyer and seller of each order, Account Type, specifies whether the order comes from an institutional investor. We use the institutional order flow (buy volume minus sell volume) as a proxy for informed trading. The CAUD data was first provided to academics as part of the TORQ dataset constructed by Joel Hasbrouck. Our CAUD data is complemented by the daily data on returns (from transaction price in CRSP and from closing quotes in TAQ),trading volume (CRSP), and market capitalization (number of shares outstanding times price from CRSP).

3.2 Market-Level Tests

We begin testing the lead-lag portfolio relations between informed order flows and returns given in Proposition 2. To do this we construct an equal-weight market portfolio for returns and informed order flow for all stocks each day. We use institutional order flow (buy volume minus sell volume) measured as a percent of a total market cap as a proxy for informed order flows in the model. Table 1 provides overview statistics for these market portfolios. The mean market return is 5.7 basis points per day and the standard deviation of the market return is almost 1% per day. Institutions' order flow is slightly positive over the sample period. The standard deviation of institutional order flow is a little less than 1 basis point of market capitalization per day. The last two rows of Table 1 provide evidence on the size of institutional trading volume (as opposed to order flow). Institutional trading is 40.95% of NYSE trading volume. Unless such a large fraction of the total trading is due to informed trading the institutional trading proxy likely includes some noise trading as well, e.g., mutual funds facing redemptions. This does not affect the signs of the lead-lag covariances in Proposition 2, but can affect any analysis where both returns and our informed order flow proxy are both included as explanatory variables.

[insert Table 1]

The model has 2 trading periods so the correlation calculations in Proposition 2 are for a single lag. As discussed above it is an open question as to what calendar time period corresponds to the trading periods in the model, we also examine lags up to 4 days. This allows for time-series dependencies in returns and institutional order flow beyond one day.

16

The correlation results are shown in Table 2. Panel A of this table reports the correlation coefficients of equal-weighted market returns, $R_m(t)$, with lagged equal-weighted market returns and market institutional order flow, $x_m(t)$. In agreement with the model's predictions, lagged returns do not predict institutional order flow, but lagged returns do predict returns. Panel B of this table reports correlation coefficients of daily market institutional order flow with lagged equal-weighted market returns and market institutional order flow. With an exception of a two-day lag, correlations between returns and lagged institutional order flow are positive and statistically significant, as the model predicts. Institutional order flow is positively serially correlated, again in an agreement with the model.

[insert Table 2]

In the data, however, cross-autocorrelations between informed order flows and returns can be the manifestation of the high autocorrelations of institutional order flow combined with a high contemporaneous correlation between institutional order flow and returns. Under such a scenario, the lagged institutional order flow is a noisy proxy for the lagged returns and, once we control for lagged returns, the lead-lag effect will disappear. To address this alternative hypothesis we estimate joint autoregression described in Proposition 3.

PROPOSITION 3: *Consider the following vector autoregression (VAR) of the equal-weighted portfolios*

$$\begin{pmatrix} \Delta P_p(t) \\ x_p(t) \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} \Delta P_p(t-1) \\ x_p(t-1) \end{pmatrix} + \begin{pmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \end{pmatrix}. \tag{27}$$

The coefficients b_{ij} are given in the model by

$$b_{11} = \frac{1}{2}\left(\frac{\Lambda_+ + \Lambda_-}{2\lambda_1} - 1 \right) < 0, \tag{28}$$

$$b_{12} = -\left(1 + \frac{\beta_K^2}{(1+\rho)(\beta_+ + \beta_-)^2} \right)\lambda_1 b_{11} + \frac{\beta_K^2}{\lambda_1 \sigma_u^2 (1+\rho)(\beta_+ + \beta_-)^2}\text{Cov}\left(\Delta P_1(1), \Delta P_2(2) \right) > 0,$$

$$b_{21} = -\frac{\Lambda_+ + \Lambda_-}{4\lambda_1 \lambda_2} < 0,$$

$$b_{22} = \left(1 + \frac{\beta_K^2}{(1+\rho)(\beta_+ + \beta_-)^2} \right)\frac{\Lambda_+ + \Lambda_-}{4\lambda_2} > 0.$$

17

Regression models similar to Proposition 3 are used in Brennan et al. (1993) and Chordia and Swaminathan (2000) to study the lead-lag relations between stocks with different levels of analyst coverage and trading volume and by Hou (2007) to study lead-lag relations between large and small stocks within and across industries. The unconditional relations (correlations) in Proposition 2 and the conditional relations (regression coefficients) in Proposition 3 have the same sign when the lagged variable is informed order flow as b_{12} and b_{22} are positive. However the conditional relations between past price changes and subsequent price changes and informed order flow, b_{11} and b_{21}, become negative because of the contemporaneous correlation between the informed order flow and price changes.

In Proposition 2 there is no relation between past prices changes and informed order flow. However, liquidity trading is a component of returns. In the second period the informed trader trades against price changes due to noise. As a result, b_{21} is negative. In addition, Cov $(\Delta P_p(1), \Delta P_p(2))$ is less than Cov $(x_p(1), \Delta P_p(2))$ which makes b_{11} negative.

We estimate equations (27) with one ($K = 1$) and four (Panel B, $K = 4$) lags in the data

$$R_m(t) = a_{0,1} + \sum_{k=1}^{K} a_k R_m(t - k) + \sum_{k=1}^{K} b_k x_m(t - k) + \varepsilon_1(t), \tag{29}$$

$$x_m(t) = a_{0,2} + \sum_{k=1}^{K} c_k R_m(t - k) + \sum_{k=1}^{K} d_k x_m(t - k) + \varepsilon_2(t). \tag{30}$$

Table 3 reports the VAR results with Newey-West corrected t- and F-statistics. In agreement with the model, the coefficients on institutional order flow in the returns equation, the sum of b_k, are positive and statistically significant at the 1% level for one-lag regressions (Panel A, 5.1719, t-statistic $= 4.70$) and at the 5% level for four-lag regressions (Panel B, 7.284, F-statistic $= 3.96$). Thus, lagged institutional order flow contains information about future market returns beyond that contained in lagged market returns. Also in agreement with the model the coefficients on institutional order flow in the institutional order flow equation, the sum of d_k, are positive and statistically significant at the 1% level, 0.0810 with a t-statistic $= 10.83$ for one-lag (Panel A) and 0.1575 with a F-statistic $= 11.83$ for four lags (Panel B). The provide further support our use of institutional order flow as a proxy for informed order flow and also support the model.

The signs of the coefficients on returns in the returns equation, the sum of a_k, are negative in Proposition 3 while positive and statistically significant in the data, at the 1% level for one-lag (Panel A, 0.0810, t-statistic = 10.83) and four-lag (Panel B, 0.1575, F-statistic = 11.83) regressions. The coefficients on returns in the institutional order flow equation, the sum of c_k, is negative in Proposition 3. In the data there is no relation between returns and subsequent informed order flow. These both may be due to institutional order flow being an imperfect proxy of the informed order flow.

3.3 Industry-Level Test: The Impact of Asset Correlation (ρ)

While the market-level analysis is supportive of the model, the correlation in the fundamental values of the assets is a key driver of the model. To further study the model's relation to the data we examine how increasing this correlation (ρ) change the correlation and VAR coefficients in Propositions 2 and 3 and Tables 2 and 3. Panel A of Figure 3 presents the correlation coefficients as a function of ρ. As noted in Proposition 2 the covariance of informed order flow with lagged price changes, $\text{Cov}\left[\Delta P_p\left(1\right), x_p\left(2\right)\right]$, is zero for all values of ρ so the corresponding correlation is omitted. The other correlation coefficients all increase with ρ. These follow from the informed trader's order flow becoming more correlated as the fundamental correlation increases. However, the informed tries to jam the market makers' inference problem which leads to the autocorrelation of informed order flow growing more slowly than the other two correlations.

Similarly, Panel B of Figure 3 graphs the VAR coefficients from Proposition 3 as a function of ρ. As seen in Panel A the autocorrelation of the informed trader's trading increases in ρ, this leads to the coefficients on lagged order flow in both the returns and order flow equations, b_{12} and b_{22}, increasing with ρ. The coefficient on lagged returns in the returns equation, b_{11}, starts off negative at $\rho = 0$ and increases with ρ; b_{11} increases faster than b_{12} and b_{22} because the informed trader increases the signal jamming strategy with ρ, but not fast enough to counteract the market makers' improved cross-asset learning. The coefficient

of lagged returns in the informed order flow equation becomes increasingly negative with ρ. This is because the increased correlation of fundamentals means that the informed trades more aggressively against noise in the portfolio return.

Testing the impact of ρ requires identification of assets that have higher or lower fundamental correlation. We follow Hou (2007) in using industry designation as a proxy for where the fundamental correlation is higher. We use the 12 industry SIC-code-based classifications from Ken French's website. The final of the 12 industries includes stocks that do not fit the first 11 industries and is referred to as "other." The "other" group is by far the largest group (763 stocks in our sample period as compared to the next largest which is "Manufacturing" with 314 stocks). To ensure that all of our tests do not reuse any data we will use this other category to form random (non-industry) portfolios.

To construct portfolios we take stocks in the the first 11 industry portfolios and calculate returns and informed order flows as we did for the market portfolio. For the other category we randomly divide stocks in it into 11 portfolios and calculate returns and informed order flows for each. Thus, we have 22 portfolios where 11 are formed from stocks within the same industry and 11 which are formed from stocks chosen from random industries. To focus on the industry-level effects we remove the market-level effects from the within industry and random portfolios by using the residuals from the regression of returns (order flows) of each portfolio on the market returns (order flows). To avoid creating correlation in the residuals across the industry and random categories portfolios in each group are regressed on the "market" formed from stocks only within the 11 portfolios in the same category.

Similar to Table 2 for the market-level analysis, Table 4 examines the lead-lag correlations coefficients for the 11 portfolios formed within industries and the 11 randomly formed portfolios. The correlation coefficients for each of the 22 portfolios are estimated individually. To allow for statistical inference that properly accounts for possible correlations across portfolios and time we estimate the full covariance matrix for the 22 coefficients via seemingly unrelated regression (SUR) with the Newey-West approach to control for heteroskedasticity and autocorrelation.

[insert Table 4]

Panel A of Table 4 reports the correlation coefficients for the 11 within industry portfolios. Average coefficients across each category of portfolios are reported along with the corresponding F-statistics. As in the market portfolio and in the model returns are significantly positively correlated with lagged returns and lagged informed order flow. Order flow is positively autocorrelated. Order flow is not correlated with the first two lags of returns. These correlations are consistent with the model.

Panel B of Table 4 reports the correlation coefficients for the 11 randomly formed portfolios. For the random portfolios the only autocorrelation that is reliably different from zero is order flow. Given that lagged order flow is not positively correlated with returns, this suggests that institutional order flow that is orthogonal to the systematic institutional order has some persistent component that appears unrelated to information.

Panel C tests the differences in each correlation coefficient between the industry portfolios (Panel A) and random portfolios (Panel B). The differences in average coefficient between the two categories is given along with the F-statistic that the difference does not equal zero. Given that the coefficients in Panel B other than the autocorrelation of order flow show little difference from zero, it is not surprising that the differences in Panel C are similar in magnitude and statistical significance to the within industry correlations in Panel A. The differences in the autocorrelation of order flow is positive and statistically significant. These differences match the model predictions that the correlation coefficients are increasing in ρ shown in Panel A of Figure 3.

Using the same approach as Table 4, Table 5 presents the VAR for returns and order flows for the industry and random portfolios. As in Table 3 the VARs are estimated for both 1 lag and 4 lags. As with the correlation coefficients in Table 4 the VAR coefficients for each of the 22 portfolios are estimated individually. To allow for statistical inference that properly accounts for possible correlations all the VAR coefficients across the 22 portfolios are estimated simultaneously via seemingly unrelated regression (SUR).

[insert Table 5]

Panel A of Table 5 presents the 1-lag and 4-lag VAR for the within industry portfolios. As with the market portfolio the coefficients on lagged order flow are positive and statistically

significant in both the return and order flow equations. The coefficients on lagged returns in the return equation are positive and statistically significant. Unlike the market results, but consistent with the model's predictions, the coefficients on returns in the order flow equation are negative.

Panel B of Table 5 provides the 1-lag and 4-lag VAR for the randomly formed portfolios. The cross-equation coefficients, order flow in the return equations and returns in the order flow equation, do not differ from zero. As with the correlation coefficients the coefficients on lagged order flow are positive in the order flow equations. The coefficients on lagged returns in the returns equation differ from zero, but have opposite signs in the 1-lag and 4-lag specifications.

Panel C tests the differences in the VAR coefficients between the industry portfolios (Panel A) and random portfolios (Panel B). All the differences in coefficients have the signs predicted by the model in Panel B of Figure 3.

4 Conclusion

We solve a multi-period model of strategic trading with long-lived information in multiple assets with correlated innovations in fundamental values. The model assumes that market makers in assets cannot condition their price functions in either asset on trading in the other asset. Using daily non-public data from the New York Stock Exchange we test the model's predictions on the unconditional and conditional relations between our proxy for informed order flow (institutional order flow) and returns within portfolios. We find support for the model's prediction of positive autocorrelations in portfolio returns as well as the predictions for how informed order flow positively predicts future returns and future informed order flow. We also test the model's comparative static predictions for how the relations between informed order flow and returns depend on and the fundamental correlation of assets within a portfolio. We find support for many of the predicted relations in the data.

The predictability of a portfolio of asset returns, but no predictability within the individual assets, follows from the market makers' inability to condition their pricing functions in each asset on trading in all the assets. The correlation in the informed order flow

across assets reveals additional information which the market makers incorporate into price before the next trading opportunity. Therefore, while returns are predictable, there is no opportunity for anyone without the informed trader's information to construct a profitable trading strategy based on the positive autocorrelation in portfolio returns. In this sense our results support market efficiency even though there is predictability in portfolio returns.

Appendix A. Proofs

Proof of Theorem 1. The informed trader's total expected payoff, π_I, is given by

$$\pi_I = \mathrm{E}\left[\sum_{t=1}^{2}\sum_{i=1}^{2} x_i(t)\left(V_i - P_i(t)\right)|\widetilde{\mathbf{V}} = \mathbf{V}\right], \tag{A1}$$

where the expectation is performed with respect to both the liquidity demand and the realizations of the fundamentals[8].

Combining (A1) with (5), (7), and (13), we obtain

$$
\begin{aligned}
\pi_I = {}& \mathrm{E}\left[x_1(1)\left(V_1 - \lambda_1(1)x_1(1)\right)\right] + \tag{A2}\\
& + \mathrm{E}\left[x_1(2)\left(V_1 - \left(\Lambda_{11}y_1(1) + \Lambda_{12}y_2(1)\right)\right) - \lambda_1(2)x_1(2)\right)\right]\\
& + \mathrm{E}\left[x_2(1)\left(V_2 - \lambda_2(1)x_2(1)\right)\right]\\
& + \mathrm{E}\left[x_2(2)\left(V_2 - \left(\Lambda_{21}y_1(1) + \Lambda_{22}y_2(1)\right)\right) - \lambda_2(2)x_2(2)\right)\right].
\end{aligned}
$$

Following Kyle (1985) the optimization is performed in two steps using backward induction. First, we optimize (A2) with respect to $x_1(2)$ and $x_2(2)$, to obtain

$$x_i^*(2) = \underset{x_i(2)}{\arg\max}\; x_i(2)\left(V_i - P_i(1^+) - \lambda_i(2)x_i(2)\right),\quad i = 1, 2, \tag{A3}$$

which after comparing with (11) implies that the optimal matrix $\beta(2)$ is diagonal

$$\beta_{11}(2) = \frac{1}{2\lambda_1(2)}, \qquad \beta_{22}(2) = \frac{1}{2\lambda_2(2)}, \tag{A4}$$
$$\beta_{21}(2) = \beta_{12}(2) = 0.$$

This is because all agents know that $t = 2$ is the last trading round before the fundamentals are revealed. Clearly, this is not the case for the first trading round, and, therefore, we do not expect $\beta(1)$ to be diagonal.

Second, we substitute $x_i^*(2)$ back into (A2) and optimize it with respect to $x_1(1)$ and $x_2(1)$ taking $x_i^*(2)$ as given

$$x_1^*(1) = \underset{x_1(1)}{\arg\max}\; x_1(1)\left(V_1 - \Lambda_{11}x_1^*(2) - \Lambda_{21}x_2^*(2) - \lambda_1(1)x_1(1)\right), \tag{A5}$$
$$x_2^*(1) = \underset{x_2(1)}{\arg\max}\; x_2(1)\left(V_2 - \Lambda_{12}x_1^*(2) - \Lambda_{22}x_2^*(2) - \lambda_2(1)x_2(1)\right),$$

The first order conditions for $x_i^*(1)$ which follow from (A5) can be written in the matrix form as

$$\mathbf{x}^*(1) = \frac{1}{2\lambda(1)}\left(\mathbf{V} - \mathbf{\Lambda}'\frac{1}{2\lambda(2)}\left(\mathbf{V} - \mathbf{\Lambda}\mathbf{x}^*(1)\right)\right), \tag{A6}$$

---

[8]Note that the expectations with respect to $\mathbf{u}(1)$ and $\mathbf{u}(2)$ are "nested" because the liquidity demand in the first trading round is observed in the second trading round and the insider's strategy is conditioned on this.

taking into account that

$$\mathbf{x}^* (2) = \frac{1}{2\lambda (2)} \left(\mathbf{V} - \mathbf{\Lambda}\mathbf{x}^* (1) \right). \tag{A7}$$

Combining (A6) with (3), $\beta (1)$ can be found as a solution of the following fixed-point condition for

$$\beta (1) = \left(\mathbf{I} - \frac{1}{2\lambda (1)}\mathbf{\Lambda}' \frac{1}{2\lambda (2)}\mathbf{\Lambda} \right)^{-1} \frac{1}{2\lambda (1)} \left(\mathbf{I} - \mathbf{\Lambda}' \frac{1}{2\lambda (2)} \right), \tag{A8}$$

where $\mathbf{I}$ stands for the unit matrix.

We proceed with a proof in two steps. First, we guess that solution is symmetric

$$\lambda_1 (1) = \lambda_2 (1) = \lambda_1, \tag{A9}$$
$$\lambda_1 (2) = \lambda_2 (2) = \lambda_2,$$
$$\beta_{11} (1) = \beta_{22} (1) = \beta_+,$$
$$\beta_{12} (1) = \beta_{21} (1) = \beta_-,$$
$$\Lambda_{11} = \Lambda_{22} = \Lambda_+,$$
$$\Lambda_{12} = \Lambda_{21} = \Lambda_-.$$

Second, we will verify that solution (A9) exists and satisfies all the necessary conditions.

In the symmetric case, the fixed point condition (A8) is simplified to

$$\left(\mathbf{I} - \frac{1}{4\lambda_1\lambda_2}\mathbf{\Lambda}^2 \right) \beta (1) = \frac{1}{2\lambda_1} \left(\mathbf{I} - \frac{1}{2\lambda_2}\mathbf{\Lambda} \right), \tag{A10}$$

and effectively represents a system of two equations for β_+ and β_-

$$\left(1 - \frac{\Lambda_+^2 + \Lambda_-^2}{4\lambda_1\lambda_2} \right) \beta_+ - \frac{\Lambda_+\Lambda_-}{2\lambda_1\lambda_2}\beta_- = \frac{1}{2\lambda_1} \left(1 - \frac{\Lambda_+}{2\lambda_2} \right), \tag{A11}$$

$$-\frac{\Lambda_+\Lambda_-}{2\lambda_1\lambda_2}\beta_+ + \left(1 - \frac{\Lambda_+^2 + \Lambda_-^2}{4\lambda_1\lambda_2} \right) \beta_- = -\frac{\Lambda_-}{4\lambda_1\lambda_2}.$$

Next, we introduce new "auxiliary" variables

$$z_\pm \equiv \sqrt{1 \pm \rho}\frac{(\beta_+ \pm \beta_-)}{\beta_K}. \tag{A12}$$

The subscript $\pm$ on a variable is used to represent two equations where in the first (second) equation the relevant variable with subscript $\pm$ is replaced by that variable with subscript $+$ $(-)$. Similarly if $\pm$ is used as an operator it means $+$ in the first equation and $-$ in the second equation. Therefore (A12) represents two equations: $z_+ = \sqrt{1 + \rho}\frac{(\beta_+ + \beta_-)}{\beta_K}$ and $z_- = \sqrt{1 - \rho}\frac{(\beta_+ - \beta_-)}{\beta_K}$. z_+ represents a rescaling of the expected informed trading intensity, $\beta_+ + \beta_-$, and z_- represents a rescaling of the difference between the within and across asset expected informed trading intensity, $\beta_+ - \beta_-$. The properties of $z_\pm$ which we will examine later simplify many of the following proofs. Next we guess that

$$\Lambda_+ \pm \Lambda_- = \frac{1}{\beta_K}\frac{\sqrt{1 \pm \rho}z_\pm}{1 + z_\pm^2}. \tag{A13}$$

25

Substituting both (A12) and (A13) back into the system (A11) yields a system of nonlinear equations for $z_\pm$. Next we verify that our guess for $\Lambda_\pm$ is self-consistent. It can be easily shown that in the case of the symmetric solution matrix $\mathbf{\Psi}$ takes the following form

$$\psi_{11} = \psi_{22} \equiv \psi_+ = \frac{\beta_K^2}{2}\left(\frac{1+z_+^2}{1+\rho} + \frac{1+z_-^2}{1-\rho}\right), \tag{A14}$$

$$\psi_{12} = \psi_{21} \equiv \psi_- = \frac{\beta_K^2}{2}\left(\frac{1+z_+^2}{1+\rho} - \frac{1+z_-^2}{1-\rho}\right),$$

It follows from (9) that

$$\Lambda_+ = \frac{\psi_+\beta_+ - \psi_-\beta_-}{\psi_+^2 - \psi_-^2} = \frac{1}{2\beta_K}\left(\frac{\sqrt{1+\rho}\,z_+}{1+z_+^2} + \frac{\sqrt{1-\rho}\,z_-}{1+z_-^2}\right), \tag{A15}$$

$$\Lambda_- = \frac{\psi_+\beta_- - \psi_-\beta_+}{\psi_+^2 - \psi_-^2} = \frac{1}{2\beta_K}\left(\frac{\sqrt{1+\rho}\,z_+}{1+z_+^2} - \frac{\sqrt{1-\rho}\,z_-}{1+z_-^2}\right).$$

which verifies (A13). Finally, using (A14) we find λ_1

$$\lambda_1 = \frac{\beta_+ + \rho\beta_-}{\beta_+^2 + 2\rho\beta_+\beta_- + \beta_-^2 + \beta_K^2} = \frac{1}{\beta_K}\frac{\sqrt{1+\rho}\,z_+ + \sqrt{1-\rho}\,z_-}{2 + z_+^2 + z_-^2}. \tag{A16}$$

Using equations (14) and (A25) we find λ_2

$$\lambda_2^2 = \frac{1 - \beta_+\Lambda_+ - \beta_-\Lambda_- - \rho\left(\beta_+\Lambda_- + \beta_-\Lambda_+\right)}{4\beta_K^2} = \frac{1}{8\beta_K^2}\left(\frac{1+\rho}{1+z_+^2} + \frac{1-\rho}{1+z_-^2}\right), \tag{A17}$$

which completes the proof.∎

LEMMA 1: $\Lambda_-(\rho)$ *is nonnegative on the interval* $\rho \in [0,1]$.

Proof:

$$\Lambda_-(\rho)\Lambda_+(\rho) = \frac{1}{4\beta_K^2}\left(\frac{(1+\rho)\,z_+^2}{\left(1+z_+^2\right)^2} - \frac{(1-\rho)\,z_-^2}{\left(1+z_-^2\right)^2}\right) = \tag{A18}$$

$$\frac{1}{4\beta_K^2}\frac{z_+^2\left(1+z_-^2\right)^2 - z_-^2\left(1+z_+^2\right)^2 + \rho\left(z_+^2\left(1+z_-^2\right)^2 + z_-^2\left(1+z_+^2\right)^2\right)}{\left(1+z_+^2\right)^2\left(1+z_-^2\right)^2}.$$

We need to show that

$$z_+^2\left(1+z_-^2\right)^2 - z_-^2\left(1+z_+^2\right)^2 = (z_+ - z_-)(1 - z_+z_-)\left[z_+\left(1+z_-^2\right) + z_-\left(1+z_+^2\right)\right] \geq 0, \tag{A19}$$

which follows from the property described below in (A20).

The below Figure graphs the numerical solution for $z_\pm(\rho)$ as functions of ρ. It shows that $z_+(\rho)$ is monotonically increasing on the interval $\rho \in [0,1]$ while $z_-(\rho)$ is monotonically

decreasing on the same interval. Importantly, the auxiliary functions $z_{\pm}(\rho)$ and $z_{+}(\rho)$ satisfy the condition

$$1 > z_{+}(\rho) \geq z_{-}(\rho) \geq 0, \quad \rho \in [0,1], \tag{A20}$$

Plot of Auxiliary Variables $z_{\pm}$:
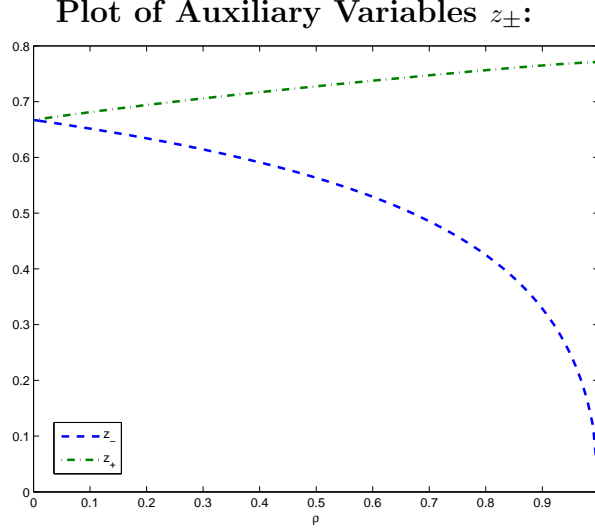
The result of Lemma 1 follows since $\Lambda_{+}(\rho)$ is positive and $\Lambda_{-}(\rho)\Lambda_{+}(\rho) \geq 0$ on $\rho \in [0,1]$. ∎

We will need the following Lemma for our further proofs.

LEMMA 2: *Order flows have the following variances*

$$\mathrm{Var}\,(y_1\,(1)) = \mathrm{Var}\,(y_2\,(1)) = \left(1 + \frac{z_+^2 + z_-^2}{2}\right)\sigma_u^2, \tag{A21}$$

$$\mathrm{Var}\,(y_1\,(2)) = \mathrm{Var}\,(y_2\,(2)) = 2\sigma_u^2,$$

$$\mathrm{Var}\,(x_1\,(1)) = \mathrm{Var}\,(x_2\,(1)) = \left(\frac{z_+^2 + z_-^2}{2}\right)\sigma_u^2,$$

$$\mathrm{Var}\,(x_1\,(2)) = \mathrm{Var}\,(x_2\,(2)) = \sigma_u^2.$$

Total order flows, $y_i(t)$, have the following covariances

$$\mathrm{Cov}(y_1\,(1)\,,y_2\,(1)) = \left(\frac{z_+^2 - z_-^2}{2}\right)\sigma_u^2, \tag{A22}$$

$$\mathrm{Cov}(y_1\,(1)\,,y_1\,(2)) = \mathrm{Cov}(y_1\,(1)\,,y_2\,(2)) = 0,$$

$$\mathrm{Cov}(y_1\,(2)\,,y_2\,(2)) = \mathrm{Cov}(x_1\,(2)\,,x_2\,(2)) = \frac{\sigma_0^2}{8\lambda_2^2}\left(\frac{1+\rho}{1+z_+^2} + \frac{1-\rho}{1+z_-^2}\right).$$

Price changes have the following covariances

$$\text{Var}\left(\Delta P_1\left(1\right)\right) = \text{Var}\left(\Delta P_2\left(1\right)\right) = \lambda_1^2 \text{Var}\left(y_1\left(1\right)\right), \tag{A23}$$

$$\text{Var}\left(\Delta P_1\left(2\right)\right) = \text{Var}\left(\Delta P_2\left(2\right)\right) = 2\lambda_2^2\sigma_u^2 + \left(\Lambda_-^2 - \left(\Lambda_+ - \lambda_1\right)^2\right)\text{Var}\left(y_1\left(1\right)\right),$$

$$\text{Cov}(\Delta P_1\left(2\right), \Delta P_2\left(2\right)) = \lambda_2^2 \text{Cov}(y_1\left(2\right), y_2\left(2\right)) +$$

$$+ \left(\left(\Lambda_+ - \lambda_1\right)^2 - \Lambda_-^2\right)\text{Cov}(y_1\left(1\right), y_2\left(1\right)).$$

Proof: $\text{Var}\left(y_1\left(1\right)\right)$ is equal to $\text{Var}\left(y_2\left(1\right)\right)$ by symmetry and can be calculated directly

$$\text{Var}\left(y_1\left(1\right)\right) = \sigma_0^2\left(\beta_+^2 + \beta_-^2 + 2\rho\beta_+\beta_-\right) + \sigma_u^2 = \tag{A24}$$

$$= \frac{\sigma_0^2\beta_K^2}{4}\left\{\frac{2z_+^2}{1+\rho} + \frac{2z_-^2}{1-\rho} + 2\rho\left(\frac{z_+^2}{1+\rho} - \frac{z_-^2}{1-\rho}\right)\right\} + \sigma_u^2 =$$

$$= \left(1 + \frac{z_+^2 + z_-^2}{2}\right)\sigma_u^2.$$

Because $\text{Var}\left(y_1\left(1\right)\right) = \text{Var}\left(x_1\left(1\right)\right) + \sigma_u^2$ the expression for $\text{Var}\left(x_1\left(1\right)\right)$ follows immediately. We can write $x_1\left(2\right)$ as

$$x_1\left(2\right) = \frac{1}{2\lambda_2}\left\{\left(1 - \beta_+\Lambda_+ - \beta_-\Lambda_-\right)\widetilde{V}_1 - \left(\beta_-\Lambda_+ + \beta_+\Lambda_-\right)\widetilde{V}_2 - \Lambda_+u_1(1) - \Lambda_-u_2(1)\right\}. \tag{A25}$$

Then

$$\text{Var}\left(x_1\left(2\right)\right) = \frac{1}{4\lambda_2^2}\left\{\left[\left(1 - \beta_+\Lambda_+ - \beta_-\Lambda_-\right)^2 + \left(\beta_-\Lambda_+ + \beta_+\Lambda_-\right)^2 - \right.\right. \tag{A26}$$

$$\left.\left. -2\rho\left(1 - \beta_+\Lambda_+ - \beta_-\Lambda_-\right)\left(\beta_-\Lambda_+ + \beta_+\Lambda_-\right)\right]\sigma_0^2 + \left(\Lambda_+^2 + \Lambda_-^2\right)\sigma_u^2\right\}.$$

It can be easily verified that

$$1 - \beta_+\Lambda_+ - \beta_-\Lambda_- = \frac{1}{2}\left(\frac{1}{1+z_-^2} + \frac{1}{1+z_+^2}\right), \tag{A27}$$

$$\beta_-\Lambda_+ + \beta_+\Lambda_- = \frac{1}{2}\left(\frac{1}{1+z_-^2} - \frac{1}{1+z_+^2}\right),$$

$$\left(\Lambda_+^2 + \Lambda_-^2\right)\sigma_u^2 = \left(\frac{1+\rho}{2}\frac{z_+^2}{1+z_+^2} + \frac{1-\rho}{2}\frac{z_-^2}{1+z_-^2}\right)\sigma_0^2.$$

Substituting (A27) into (A26) yields the desired result.

$\text{Cov}(y_1(1), y_2(1))$ can be obtained as follows

$$\text{Cov}(y_1(1), y_2(1)) = \sigma_0^2 \left(\rho \left(\beta_+^2 + \beta_-^2 \right) + 2\beta_+ \beta_- \right) = \tag{A28}$$

$$= \frac{\sigma_0^2 \beta_K^2}{4} \left\{ \frac{2\rho z_+^2}{1+\rho} + \frac{2\rho z_-^2}{1-\rho} + 2 \left(\frac{z_+^2}{1+\rho} - \frac{z_-^2}{1-\rho} \right) \right\} =$$

$$= \left(\frac{z_+^2 - z_-^2}{2} \right) \sigma_u^2.$$

In order to prove that the total order flows are not correlated across time we note that using the symmetry of the solution (specifically that $\text{Var}(y_1(1)) = \text{Var}(y_2(1))$) the equations (8) can be written as

$$\Lambda_+ = \frac{\text{Cov}(\widetilde{V}_1, y_1(1))\text{Var}(y_1(1)) - \text{Cov}(\widetilde{V}_1, y_2(1))\text{Cov}(y_1(1), y_2(1))}{(\text{Var}(y_1(1)))^2 - (\text{Cov}(y_1(1), y_2(1)))^2}, \tag{A29}$$

$$\Lambda_- = \frac{\text{Cov}(\widetilde{V}_1, y_2(1))\text{Var}(y_1(1)) - \text{Cov}(\widetilde{V}_1, y_1(1))\text{Cov}(y_1(1), y_2(1))}{(\text{Var}(y_1(1)))^2 - (\text{Cov}(y_1(1), y_2(1)))^2}.$$

Substituting $y_1(2)$ into the covariance yields

$$\text{Cov}(y_1(1), y_1(2)) = \frac{1}{2\lambda_2} \left\{ \text{Cov}(\widetilde{V}_1, y_1(1)) - \Lambda_+ \text{Var}(y_1(1)) - \Lambda_- \text{Cov}(y_1(1), y_2(1)) \right\}. \tag{A30}$$

We just need to show that $\Lambda_+ \text{Var}(y_1(1)) + \Lambda_- \text{Cov}(y_1(1), y_2(1)) = \text{Cov}(\widetilde{V}_1, y_1(1))$, which directly follows from (A29).

To prove the last identity note that $\Delta P_1(2)$ can be written as

$$\Delta P_1(2) = \lambda_2 y_1(2) + (\Lambda_+ - \lambda_1) y_1(1) + \Lambda_- y_2(1). \tag{A31}$$

Consider now

$$\text{Var}(\Delta P_1(2)) = 2\lambda_2^2 \sigma_u^2 + \left(\Lambda_-^2 + (\Lambda_+ - \lambda_1)^2 \right) \text{Var}(y_1(1)) + 2\Lambda_- (\Lambda_+ - \lambda_1) \text{Cov}(y_1(1), y_2(1)). \tag{A32}$$

It follows from (A30) that

$$\frac{\lambda_1 - \Lambda_+}{\Lambda_-} = \frac{\text{Cov}(y_1(1), y_2(1))}{\text{Var}(y_1(1))} = \frac{z_+^2 - z_-^2}{2 + z_+^2 + z_-^2}, \tag{A33}$$

which upon substitution into (A32) yields the desired result.

$\text{Cov}(y_1(2), y_2(2))$ can be obtained using (A25) and (A27)

$$\text{Cov}(y_1(2), y_2(2)) = \text{Cov}(x_1(2), x_2(2)) = \tag{A34}$$

$$= \frac{\sigma_0^2}{4\lambda_2^2} \left\{ \rho \left((1 - \beta_+ \Lambda_+ - \beta_- \Lambda_-)^2 + (\beta_- \Lambda_+ - \beta_+ \Lambda_-)^2 \right) - \right.$$

$$\left. -2\left(1 - \beta_+ \Lambda_+ - \beta_- \Lambda_-\right)\left(\beta_- \Lambda_+ - \beta_+ \Lambda_-\right) + 2\beta_K^2 \Lambda_+ \Lambda_- \right\} =$$

$$= \frac{\sigma_0^2}{8\lambda_2^2} \left(\frac{(1+\rho) z_+^2}{\left(1 + z_+^2\right)^2} - \frac{(1-\rho) z_-^2}{\left(1 + z_-^2\right)^2} + \frac{1+\rho}{\left(1 + z_+^2\right)^2} - \frac{1-\rho}{\left(1 + z_-^2\right)^2} \right),$$

and the result follows immediately.

Finally we find $\text{Cov}(\Delta P_1(2), \Delta P_2(2))$. Using (A31) we obtain

$$\text{Cov}(\Delta P_1(2), \Delta P_2(2)) = \lambda_2^2 \text{Cov}(y_1(2), y_2(2)) + \tag{A35}$$

$$+ \left(\Lambda_-^2 + (\Lambda_+ - \lambda_1)^2 \right) \text{Cov}(y_1(1), y_2(1)) + 2\Lambda_-(\Lambda_+ - \lambda_1) \text{Var}(y_1(1)).$$

The result follows immediately after applying (A33) in (A35).■

Proof of Proposition 1.

Because all variances are provided by Lemma 2, we only need to calculate all the necessary covariances in order to prove Proposition 1. We will use result of Lemma 2 that total order flows are not correlated across time (second relation in (A22)) to prove (22)

$$\text{Cov}(y_1(1), y_{1,2}(2)) = \text{Cov}(x_1(1) + u_1(1), x_{1,2}(2)) \tag{A36}$$

$$= \text{Cov}(x_1(1), x_{1,2}(2)) - \frac{1}{2\lambda_2} \text{Cov}(u_1(1), P_{1,2}(1^+))$$

$$= \text{Cov}(x_1(1), x_{1,2}(2)) - \frac{1}{2\lambda_2} \text{Cov}(u_1(1), \Lambda_\pm u_1(1)) = 0.$$

It immediately follows that

$$\text{Cov}(x_1(1), x_{1,2}(2)) = \frac{\Lambda_\pm}{2\lambda_2} \sigma_u^2. \tag{A37}$$

The lead-lag covariances of price changes are given by

$$\text{Cov}(\Delta P_1(1), \Delta P_1(2)) = \lambda_1 (\Lambda_+ - \lambda_1) \text{Var}(y_1(1)) + \lambda_1 \Lambda_- \text{Cov}(y_1(1), y_2(1)), \tag{A38}$$

$$\text{Cov}(\Delta P_1(1), \Delta P_2(2)) = \lambda_1 \Lambda_- \text{Var}(y_1(1)) + \lambda_1 (\Lambda_+ - \lambda_1) \text{Cov}(y_1(1), y_2(1)).$$

Combining (A38) and (A33), we obtain

$$\text{Cov}(\Delta P_1(1), \Delta P_1(2)) = 0, \tag{A39}$$

$$\text{Cov}(\Delta P_1(1), \Delta P_2(2)) = \frac{\lambda_1 \Lambda_-}{\text{Var}(y_1(1))} \left(\text{Var}(y_1(1))^2 - \text{Cov}(y_1(1), y_2(1))^2 \right).$$

Taking into account that

$$\text{Var}(y_1(1))^2 - \text{Cov}(y_1(1), y_2(1))^2 = \left(1 + z_+^2\right)\left(1 + z_-^2\right) \sigma_u^4, \tag{A40}$$

we immediately obtain

$$\text{Cov}(\Delta P_1\,(1)\,,\Delta P_2\,(2)) = \frac{2\lambda_1\Lambda_-\left(1+z_+^2\right)\left(1+z_-^2\right)}{2+z_+^2+z_-^2}\sigma_u^2 \geq 0. \tag{A41}$$

(A41) combined with

$$\psi_+^2 - \psi_-^2 = \beta_K^4\frac{\left(1+z_+^2\right)\left(1+z_-^2\right)}{1-\rho^2}, \tag{A42}$$

(A33) and (A28) yields (21).

The lead-lag covariances of informed order flows and returns are given by

$$\begin{aligned}
\text{Cov}(x_1\,(1)\,,\Delta P_1\,(2)) &= \lambda_2\text{Cov}(x_1\,(1)\,,x_1\,(2)) + (\Lambda_+ - \lambda_1)\,\text{Var}\,(x_1\,(1)) + \Lambda_-\text{Cov}(y_1\,(1)\,,y_2\,(1)) \\
&= \frac{\Lambda_+}{2}\sigma_u^2 - (\Lambda_+ - \lambda_1)\,\sigma_u^2 + \underbrace{(\Lambda_+ - \lambda_1)\,\text{Var}\,(y_1\,(1)) + \Lambda_-\text{Cov}(y_1\,(1)\,,y_2\,(1))}_{0} \\
&= \left(\lambda_1 - \frac{\Lambda_+}{2}\right)\sigma_u^2. \tag{A43}
\end{aligned}$$

$$\begin{aligned}
\text{Cov}\,(x_1\,(1)\,,\Delta P_2\,(2)) &= \lambda_2\text{Cov}(x_1\,(1)\,,x_2\,(2)) + \Lambda_-\text{Var}\,(x_1\,(1)) + (\Lambda_+ - \lambda_1)\,\text{Cov}\,(y_1\,(1)\,,y_2\,(1)) \\
&= \frac{\Lambda_-}{2}\sigma_u^2 - \Lambda_-\sigma_u^2 + \Lambda_-\text{Var}\,(y_1\,(1)) + (\Lambda_+ - \lambda_1)\,\text{Cov}\,(y_1\,(1)\,,y_2\,(1)) \\
&= -\frac{\Lambda_-}{2}\sigma_u^2 + \frac{1}{\lambda_1}\text{Cov}(\Delta P_1\,(1)\,,\Delta P_2\,(2)). \tag{A44}
\end{aligned}$$

which in combination with (A33) yields

$$\text{Cov}(x_1\,(1)\,,\Delta P_1\,(2)) = \Lambda_+\sigma_u^2\left(\frac{1}{2} + \frac{\Lambda_-\left(z_+^2 - z_-^2\right)}{\Lambda_+\left(2+z_+^2+z_-^2\right)}\right) > 0, \tag{A45}$$

$$\text{Cov}(x_1\,(1)\,,\Delta P_2\,(2)) = \Lambda_-\sigma_u^2\left(\frac{1}{2} + \frac{(z_+ + z_-)^2}{2+z_+^2+z_-^2}\right) \geq 0. \tag{A46}$$

Combining the covariances with results of Lemma 2 completes the proof.∎

Proof of Proposition 2. Result for $\text{Cov}\,(\Delta P_p\,(1)\,,\Delta P_p\,(2))$ and $\text{Cov}\,(\Delta P_p\,(1)\,,x_p\,(2))$ follow immediately from Proposition 1. Next we consider

$$\text{Cov}(x_p\,(1)\,,x_p\,(2)) = \frac{1}{2}\left(\text{Cov}(x_1\,(1)\,,x_1\,(2)) + \text{Cov}(x_1\,(1)\,,x_2\,(2))\right), \tag{A47}$$

and the result follows when we use (A37) in (A47). Finally, consider

$$\text{Cov}(x_p\,(1)\,,\Delta P_p\,(2)) = \frac{1}{2}\left[\text{Cov}\,(x_1\,(1)\,,\Delta P_1\,(2)) + \text{Cov}\,(x_1\,(1)\,,\Delta P_2\,(2))\right], \tag{A48}$$

and the result follows immediately.∎

We need the following Lemma to prove Proposition 3.

LEMMA 3: *Portfolio informed order flows have the following variances*

$$\text{Var}\left(x_p\left(1\right)\right) = \frac{z_+^2}{2}\sigma_u^2, \tag{A49}$$

$$\text{Var}\left(x_p\left(2\right)\right) = \frac{1}{2}\left(\sigma_u^2 + \text{Cov}(x_1\left(2\right), x_2\left(2\right))\right),$$

Portfolio price changes have the following variances

$$\text{Var}\left(\Delta P_p\left(1\right)\right) = \frac{\lambda_1^2\left(1+z_+^2\right)}{2}\sigma_u^2, \tag{A50}$$

$$\text{Var}\left(\Delta P_p\left(2\right)\right) = \lambda_2^2\left(\sigma_u^2 + \frac{1}{2}\text{Cov}(y_1\left(2\right), y_2\left(2\right))\right) + \frac{2\Lambda_-^2\left(1+z_+^2\right)\left(1+z_-^2\right)^2}{\left(2+z_+^2+z_-^2\right)^2}.$$

Portfolio price changes and informed order flows have the following covariance

$$\text{Cov}(x_p\left(1\right), \Delta P_p\left(1\right)) = \lambda_1 \text{Var}\left(x_p\left(1\right)\right). \tag{A51}$$

Proof: The variance of $x_p\left(1\right)$ can be calculated as follows

$$\text{Var}\left(x_p\left(1\right)\right) = \frac{1}{2}\left(\text{Var}\left(x_1\left(1\right)\right) + \text{Cov}(x_1\left(1\right), x_2\left(1\right))\right) \tag{A52}$$

$$= \frac{\sigma_u^2}{2}\left(\frac{z_+^2+z_-^2}{2} + \frac{z_+^2-z_-^2}{2}\right).$$

$\text{Var}\left(x_p\left(2\right)\right)$ is equal to

$$\text{Var}\left(x_p\left(2\right)\right) = \frac{1}{2}\left(\underbrace{\text{Var}\left(x_1\left(2\right)\right)}_{\sigma_u^2} + \text{Cov}(x_1\left(2\right), x_2\left(2\right))\right), \tag{A53}$$

and the result follows. The other proofs are straightforward except for $\text{Var}\left(\Delta P_p\left(2\right)\right)$, which we prove next.

$$\text{Var}\left(\Delta P_p\left(2\right)\right) = \frac{1}{2}\left(\text{Var}(\Delta P_1\left(2\right)) + \text{Cov}(\Delta P_1\left(2\right), \Delta P_2\left(2\right))\right) = \tag{A54}$$

$$\lambda_2^2\left(\sigma_u^2 + \frac{1}{2}\text{Cov}(y_1\left(2\right), y_2\left(2\right))\right) + \Lambda_-^2\left(\frac{\lambda_1-\Lambda_+}{\Lambda_-}-1\right)^2\left(\text{Var}\left(y_1\left(1\right)\right) + \text{Cov}(y_1\left(1\right), y_2\left(1\right))\right).$$

The result follows after substituting (A33) into (A32).∎

Proof of Proposition 3: The VAR coefficients b_{11} and b_{12} are found from

$$\begin{pmatrix} b_{11} \\ b_{12} \end{pmatrix} = \underbrace{\begin{pmatrix} \text{Var}\left(\Delta P_p\left(1\right)\right) & \text{Cov}(x_p\left(1\right), \Delta P_p\left(1\right)) \\ \text{Cov}(x_p\left(1\right), \Delta P_p\left(1\right)) & \text{Var}\left(x_p\left(1\right)\right) \end{pmatrix}^{-1}}_{A} \begin{pmatrix} \text{Cov}(\Delta P_p\left(1\right), \Delta P_p\left(2\right)) \\ \text{Cov}(x_p\left(1\right), \Delta P_p\left(2\right)) \end{pmatrix}. \tag{A55}$$

We need to calculate

$$\det(A) = \lambda_1^2 \sigma_u^2 \mathrm{Var}\left(x_p\left(1\right)\right). \tag{A56}$$

Next we find b_{11}

$$b_{11} = \frac{1}{\det(A)} \left(\mathrm{Cov}(\Delta P_p\left(1\right), \Delta P_p\left(2\right))\mathrm{Var}\left(x_p\left(1\right)\right) - \mathrm{Cov}(x_p\left(1\right), \Delta P_p\left(2\right))\mathrm{Cov}(x_p\left(1\right), \Delta P_p\left(1\right))\right) = \tag{A57}$$

$$= \frac{\mathrm{Cov}(\Delta P_p\left(1\right), \Delta P_p\left(2\right)) - \lambda_1 \mathrm{Cov}(x_p\left(1\right), \Delta P_p\left(2\right))}{\lambda_1^2 \sigma_u^2} = \frac{1}{2}\left(\frac{\Lambda_+ + \Lambda_-}{2\lambda_1} - 1\right).$$

Using relation (A33), which can be rewritten as

$$\lambda_1 - \frac{\Lambda_+}{2} = \frac{\Lambda_+}{2} + \frac{\Lambda_-\left(z_+^2 - z_-^2\right)}{2 + z_+^2 + z_-^2},$$

in (A57) we obtain

$$b_{11} = -\frac{1}{2\lambda_1}\left(\frac{\Lambda_+ - \Lambda_-}{2} + \frac{\Lambda_-\left(z_+^2 - z_-^2\right)}{2 + z_+^2 + z_-^2}\right) < 0.$$

b_{12} can be found analogously

$$b_{12} = \frac{1}{\det(A)}\left(\mathrm{Cov}(x_p\left(1\right), \Delta P_p\left(2\right))\mathrm{Var}\left(\Delta P_p\left(1\right)\right) - \mathrm{Cov}(\Delta P_p\left(1\right), \Delta P_p\left(2\right))\mathrm{Cov}(x_p\left(1\right), \Delta P_p\left(1\right))\right) =$$

$$\tag{A58}$$

$$= \frac{1 + z_+^2}{z_+^2}\left(\frac{\Lambda_+ + \Lambda_-}{4} + \frac{\Lambda_-\left(z_+^2 - z_-^2\right)}{2\left(2 + z_+^2 + z_-^2\right)} + \frac{\Lambda_-\left(z_+ + z_-\right)^2}{2 + z_+^2 + z_-^2}\right) - \frac{\Lambda_-\left(1 + z_+^2\right)\left(1 + z_-^2\right)}{2 + z_+^2 + z_-^2} =$$

$$= \left(1 + \frac{1}{z_+^2}\right)\left(\frac{\Lambda_+ + \Lambda_-}{4} + \frac{\Lambda_- z_+ z_-\left(1 - z_+ z_-\right)}{2 + z_+^2 + z_-^2}\right) > 0.$$

Alternatively, it can be rewritten as

$$b_{12} = \frac{1}{\sigma_u^2}\left[\left(1 + \frac{1}{z_+^2}\right)\mathrm{Cov}(x_p\left(1\right), \Delta P_p\left(2\right)) - \frac{1}{\lambda_1}\mathrm{Cov}(\Delta P_p\left(1\right), \Delta P_p\left(2\right))\right] =$$

$$= \frac{1}{2}\left(1 + \frac{1}{z_+^2}\right)\left(\lambda_1 - \frac{\Lambda_+ + \Lambda_-}{2}\right) + \frac{1}{\lambda_1 \sigma_u^2 z_+^2}\mathrm{Cov}\left(\Delta P_1\left(1\right), \Delta P_2\left(2\right)\right).$$

We can now find b_{22} and b_{21} which are equal to

$$\begin{pmatrix} b_{21} \\ b_{22} \end{pmatrix} = \begin{pmatrix} \mathrm{Var}\left(\Delta P_p\left(1\right)\right) & \mathrm{Cov}(x_p\left(1\right), \Delta P_p\left(1\right)) \\ \mathrm{Cov}(x_p\left(1\right), \Delta P_p\left(1\right)) & \mathrm{Var}\left(x_p\left(1\right)\right) \end{pmatrix} \begin{pmatrix} \mathrm{Cov}(\Delta P_p\left(1\right), x_p\left(2\right)) \\ \mathrm{Cov}(x_p\left(1\right), x_p\left(2\right)) \end{pmatrix}. \tag{A59}$$

Therefore,

$$b_{21} = \frac{1}{\det(A)} \left(\underbrace{\mathrm{Cov}(\Delta P_p\left(1\right), x_p\left(2\right))\mathrm{Var}\left(x_p\left(1\right)\right)}_{0} - \mathrm{Cov}(x_p\left(1\right), x_p\left(2\right))\mathrm{Cov}(x_p\left(1\right), \Delta P_p\left(1\right)) \right) = \quad (A60)$$

$$= -\frac{\Lambda_+ + \Lambda_-}{4\lambda_1\lambda_2}.$$

Finally,

$$b_{22} = \frac{1}{\det(A)} \left(\mathrm{Cov}(x_p\left(1\right), x_p\left(2\right))\mathrm{Var}\left(\Delta P_p\left(1\right)\right) - \underbrace{\mathrm{Cov}(\Delta P_p\left(1\right), x_p\left(2\right))\mathrm{Cov}(x_p\left(1\right), \Delta P_p\left(1\right))}_{0} \right) = \quad (A61)$$

$$= \frac{1}{4} \left(1 + \frac{1}{z_+^2} \right) \frac{\Lambda_+ + \Lambda_-}{\lambda_2}.$$

The result follows after using the definition of z_+ (equation (A12)) in (A61).∎

References

Campbell, J.Y., A.W. Lo, A.C. MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press, Princeton, NJ.

Anderson, R.M., K.S. Eom, S.B. Hahn, and Jong-Ho Park, 2008, Stock return autocorrelation is not spurious, U.C. Berkeley Working Paper.

Badrinath, S.G., J.R. Kale, T.H. Noe, 1995, Of shepherds, sheep, and the crossautocorrelations in equity returns, *Review of Financial Studies* 8(2): 401-430.

Bernhardt, D. and R. Mahani, 2007, Asymmetric Information and Stock Return Cross-Autocorrelations, *Economics Letters*, 14-22.

Boudoukh, J., M. Richardson, R. Whitelaw, 1994, A tale of three schools: Insights on autocorrelation of short-horizon returns, *Review of Financial Studies* 7(3): 539-573.

Brennan, M.J., N. Jagadeesh, B. Swaminathan, 1993, Investment Analysis and The Adjustment of Stock Prices to Common Information, *Review of Financial Studies* 6: 799-824.

Chan, K., 1993. Imperfect information and cross-autocorrelation among stock prices, *Journal of Finance*, 1211-1230.

Chordia, T., B. Swaminathan, 2000, Trading volume and cross-autocorrelations in stock returns, *Journal of Finance* 55: 913-936.

Conrad, J., G. Kaul, 1988, Time-variation in expected returns, *Journal of Business* 61: 409-425.

Fudenberg, D., J. Tirole, 1986, A signal-jamming theory of predation, *Rand Journal of Economics* 17(3), 366-376.

Hou, K., 2007, Industry information diffusion and the lead-lag effect in stock returns, *Review of Financial Studies* 20(4): 1113-1138.

Kyle, Albert S., 1985, Continuous auctions and informed trading, *Econometrica* 53, 1315-1335.

Lo, A., A.C. MacKinlay, 1988, Stock market prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies* 3(2): 175-205.

Lo, A., A.C. MacKinlay, 1990a, When are contrarian profits due to stock market overreaction?, *Review of Financial Studies* 3(2): 175-205.

Lo, A., A.C. MacKinlay, 1990b, An econometric analysis of nonsynchronous trading, *Journal of Econometrics*, 45: 181-211.

Sias, R.W., L.T. Starks, 1997, Return autocorrelation and institutional investors, *Journal of Financial Economics* 46: 103-131.

Stein J., 1989, Efficient capital markets, inefficient firms: A model of myopic corporate behavior, *Quarterly Journal of Economics* 104(4): 655-669.

Table 1: **Summary Statistics**

This table reports summary statistics of equal-weighted market returns, $R_m(t)$, market institutional order flows (buy volume minus sell volume), $x_m(t)$, as well as ratios of total market institutional orders to market cap and trading volume. Market portfolio is formed on any give day by equally weighing returns for all stocks with available data that day from January 1999 to December 2005. Institutional order flow is measured as a percent of a total market cap. Newey-West t-statistics are reported in the parentheses.

	Mean	Std. dev.	Min.	Max.
$R_m(t)$	5.7524×10^{-4}	9.7248×10^{-3}	-0.0521	0.0517
$x_m(t)$	2.0824×10^{-5}	1.0079×10^{-4}	-3.9424×10^{-4}	5.5299×10^{-5}
$\frac{buy(t)+sell(t)}{2*mktcap(t)}$	2.5120×10^{-3}	6.8296×10^{-4}	5.6341×10^{-4}	4.7912×10^{-3}
$\frac{buy(t)+sell(t)}{2*vol(t)}$	0.4095	0.0491	0.2704	0.5985

Table 2: **Lead-Lag Correlation Coefficients for the Market Portfolio**

Panel A of this table reports correlation coefficients of equal-weighted market returns, $R_m(t)$, with lagged equal-weighted market returns and market institutional order flows, $x_m(t)$. Panel B of this table reports correlation coefficients of daily net market institutional order flow with lagged equal-weighted market returns and market institutional volume. The correlations are reported up to four lags. All variables are daily. Market portfolios are formed each day by equally weighing returns/institutional order flows for all stocks with available data that day from January 1999 to December 2005. Institutional order flow is measured as a percent of a total market cap. Newey-West t-statistics are reported in the parentheses. ***, **, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

	$R_m(t-1)$	$R_m(t-2)$	$R_m(t-3)$	$R_m(t-4)$
	Panel A: Return			
$R_m(t)$	0.0673	0.0143	0.0458	0.0261
	(2.14**)	(0.42)	(1.34)	(0.76)
$x_m(t)$	−0.0273	−0.0615	−0.0516	−0.0359
	(−0.01)	(−0.01)	(−0.01)	(−0.01)
	Panel B: Informed Order Flow			
	$x_m(t-1)$	$x_m(t-2)$	$x_m(t-3)$	$x_m(t-4)$
$R_m(t)$	0.0292	−0.0015	0.0115	0.0219
	(27.99***)	(−1.45)	(11.83***)	(21.89***)
$x_m(t)$	0.3071	0.2488	0.2324	0.2059
	(10.65***)	(9.08***)	(9.19***)	(7.55***)

Table 3: **Vector Autoregressions for the Market Portfolio**

This table reports results of jointly estimating the following one-lag (Panel A, $K = 1$) and four-lag (Panel B, $K = 4$) vector autoregressions, using daily equal-weighted market returns, $R_m(t)$, and institutional order flow, $x_m(t)$, from January 1999 to December 2005

$$R_m(t) = a_{0,1} + \sum_{k=1}^{K} a_k R_m(t-k) + \sum_{k=1}^{K} b_k x_m(t-k) + \varepsilon_1(t),$$

$$x_m(t) = a_{0,2} + \sum_{k=1}^{K} c_k R_m(t-k) + \sum_{k=1}^{K} d_k x_m(t-k) + \varepsilon_2(t).$$

$R_m(t-1 : t-k)$, $k = 1$ or 4, reports $\sum_{k=1}^{K} a_k$ or $\sum_{k=1}^{K} c_k$, depending on the left-hand side variable. Similarly, $x_m(t-1 : t-k)$, $k = 1$ or 4, reports $\sum_{k=1}^{K} b_k$ or $\sum_{k=1}^{K} d_k$, depending on the left-hand side variable. All variables are daily. Market portfolios are formed each day by equally weighing returns/institutional order flow for all stocks with available data that day. Newey-West t-statistics are reported in the parentheses. *Italics* indicate the F-statistics for the hypothesis that the sum of the coefficients equals zero. ***, **, and * denote statistical significance at the 1, 5, and 10% levels, respectively. Institutional order flow is measured as a percent of a total market cap.

	Panel A: One-Lag	
LHS	$R_m(t-1)$	$x_m(t-1)$
$R_m(t)$	0.0810	5.1719
	(10.83***)	(4.70***)
$x_m(t)$	0.0003	0.2623
	(1.12)	(119.52***)
	Panel B: Four-Lag	
LHS	$R_m(t-1 : t-4)$	$x_m(t-1 : t-4)$
$R_m(t)$	0.1576	7.2841
	*11.83***	*3.96**
$x_m(t)$	−0.0001	0.5070
	0.04	*198.46***

Table 4: **Lead-Lag Correlation Coefficients for Within-Industry and Random Portfolios**

This table compares lead-lag correlations between portfolios of assets with high and low correlation in fundamental values, ρ. We start with 12 Fama and French industry portfolios using daily data from January 1999 to December 2005. To construct portfolios we take stocks in the the first 11 industry portfolios and calculate returns and informed order flows as we did for the market portfolio. For the other category we randomly divide stocks in it into 11 portfolios and calculate returns and informed order flows for each. Thus, we have 22 portfolios where 11 are formed from stocks within the same industry and 11 which are formed from stocks chosen from random industries. To focus on the industry-level effects we remove the market-level effects from the within industry and random portfolios by using the residuals from the regression of returns (order flows) of each portfolio on the market returns (order flows). To avoid creating correlation in the residuals across the industry and random categories portfolios in each group are regressed on the "market" formed from stocks only within the 11 portfolios in the same category. Panel A reports average correlation coefficients of returns, $R_i(t)$, and institutional order flow, $x_i(t)$, with their lagged counterparts for high ρ portfolios. Panel B reports average correlation coefficients of returns, $R_i(t)$, and institutional order flow, $x_i(t)$, with their lagged counterparts for low ρ portfolios. The correlations are reported up to four lags in both cases. Panel C reports the differences between the average correlation coefficients from Panels A and B. *Italics* indicate the F-statistics for the hypothesis that the sum of the correlation coefficients equals zero (Panels A and B) and that the difference is equal to zero (Panel C). Institutional order flow is measured as a fraction of a total market cap. Coefficients are estimated at the same time using seemingly unrelated regressions; ***, **, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

	$R_i(t-1)$	$R_i(t-2)$	$R_i(t-3)$	$R_i(t-4)$	$x_i(t-1)$	$x_i(t-2)$	$x_i(t-3)$	$x_i(t-4)$
			Panel A.	Within-Industry (High ρ)				
$R_i(t)$	0.1075	0.0186	0.0175	-0.0002	0.0350	0.0017	0.0150	0.0074
	226.32***	*6.72****	*5.89****	*0.00*	*25.44****	*0.00*	*3.87***	*0.89*
$x_i(t)$	-0.0089	0.0026	0.0245	0.0161	0.2581	0.1700	0.1477	0.1268
	0.81	*0.16*	*6.04****	*3.71***	*1380.70****	*575.96****	*431.00****	*315.95***
			Panel B.	Random (Low ρ)				
$R_i(t)$	0.0021	-0.0116	-0.0147	-0.0193	-0.0011	-0.0053	-0.0109	-0.0090
	0.09	*2.62**	*4.16****	*7.18****	*0.01*	*0.50*	*2.39**	*1.56*
$x_i(t)$	0.0228	-0.0002	0.0051	0.0202	0.2314	0.1222	0.1012	0.0851
	10.29***	*0.00*	*0.49*	*7.28****	*1094.55****	*292.94****	*200.90****	*140.85****
			Panel C.	High Minus Low				
$R_i(t)$	0.1053	0.0303	0.0321	0.0191	0.0361	0.0070	0.0259	0.0164
	107.97***	*8.87****	*9.98****	*3.51***	*21.43****	*0.10*	*5.91****	*1.89*
$x_i(t)$	-0.0317	0.0028	0.0194	-0.0041	0.0267	0.0478	0.0465	0.0418
	10.63***	*0.04*	*0.26*	*2.26**	*7.33****	*22.68****	*21.33****	*17.10****

Table 5: **Vector Autoregressions for Within-Industry and Random Portfolios**

This table reports results of jointly estimating the following one-lag (Panel A, $K = 1$) and four-lag (Panel B, $K = 4$) vector autoregressions. We start with 12 Fama and French industry portfolios using daily data from January 1999 to December 2005 which we split into two groups of 11 portfolios each. For within industry portfolios (high ρ) we use the first 11 industries—everything but the "other" group. For the random portfolios (low ρ) we randomly divide stocks in the "other" group into 11 equal-sized portfolios. Because assets in this case are industry specific, they have a common industry component and thus proxy for assets with high ρ. We then remove a common market component from each portfolio by regressing its return (institutional order flow) on the market return (institutional order flow) and using the residuals. For each portfolio we calculate the daily equal-weighted returns, $R_i(t)$, and institutional order flow, $x_i(t)$. We estimate the following VAR for each portfolio:

$$R_i(t) = a_{0,i,1} + \sum_{k=1}^{K} a_{i,k} R_i(t-k) + \sum_{k=1}^{K} b_{i,k} x_i(t-k) + \varepsilon_{i,1}(t),$$

$$x_i(t) = a_{0,i,2} + \sum_{k=1}^{K} c_{i,k} R_i(t-k) + \sum_{k=1}^{K} d_{i,k} x_i(t-k) + \varepsilon_{i,2}(t).$$

The results in panel A are for portfolios of firms within the same industry. $R_i(t-1:t-k)$, $k = 1$ or 4, reports the average $\sum_{k=1}^{K} a_{i,k}$ or $\sum_{k=1}^{K} c_{i,k}$, depending on the left-hand side variable. Similarly, $x_i(t-1:t-k)$, $k = 1$ or 4, reports the average $\sum_{k=1}^{K} b_{i,k}$ or $\sum_{k=1}^{K} d_{i,k}$, depending on the left-hand side variable. *Italics* indicate the F-statistics for the hypothesis that the average of the coefficients equals zero (Panels A and B) and that the difference is equal to zero (Panel C). Institutional order flow is measured as a fraction of a total market cap. Coefficients are estimated at the same time using seemingly unrelated regressions; ***, **, and * denote statistical significance at the 1, 5, and 10% levels, respectively.

	One-Lag		Four-Lags	
	$R_i(t-1)$	$x_i(t-1)$	$R_i(t-1:t-4)$	$x_i(t-1:t-4)$
Panel A. Within-Industry (High ρ)				
$R_i(t)$	0.1060	0.8280	0.1159	1.0094
	*211.99***	*4.24***	*75.75***	*2.58**
$x_i(t)$	−0.0011	0.2701	−0.0012	0.4387
	*55.15***	*1460.26***	*18.71***	*1609.73***
Panel B. Random (Low ρ)				
$R_i(t)$	0.0134	−0.0819	−0.0304	−0.3477
	*3.45***	*0.21*	*4.44***	*1.40*
$x_i(t)$	−0.0000	0.2315	0.0002	0.3592
	0.01	*108.40***	*0.16*	*999.00***
Panel C. High Minus Low				
$R_i(t)$	0.0926	0.9099	0.1463	1.3571
	*81.40***	*4.26***	*55.56***	*3.83***
$x_i(t)$	−0.011	0.0386	−0.0014	0.0795
	*10.98***	*15.00***	*4.87***	*25.42***

Figure 1: **Timeline of the model**

This figure shows the timeline of the model.

INFORMATION

- MM_1 knows dist. of $V_{1,2}$
- MM_2 knows dist. of $V_{1,2}$

- MM_1 observes $y_1(1)$
- MM_2 observes $y_2(1)$

- MM_1 observes $y_2(1)$
- MM_2 observes $y_1(1)$

- MM_1 observes $y_1(2)$
- MM_2 observes $y_2(2)$

Time $\longrightarrow$ 0   1 1^+ 2

- I learns V_1 and V_2
- $MM_{1,2}$ quotes $P_{1,2}(0) = 0$

- I submits $x_{1,2}(1)$
- $u_{1,2}(1)$ arrives
- $MM_{1,2}$ quotes $P_{1,2}(1)$
- Trades take place

- MM_1 adjusts quote to $P_1(1^+)$
- MM_2 adjusts quote to $P_2(1^+)$

- I submits $x_{1,2}(2)$
- $u_{1,2}(2)$ arrives
- $MM_{1,2}$ quotes $P_{1,2}(2)$
- Trades take place
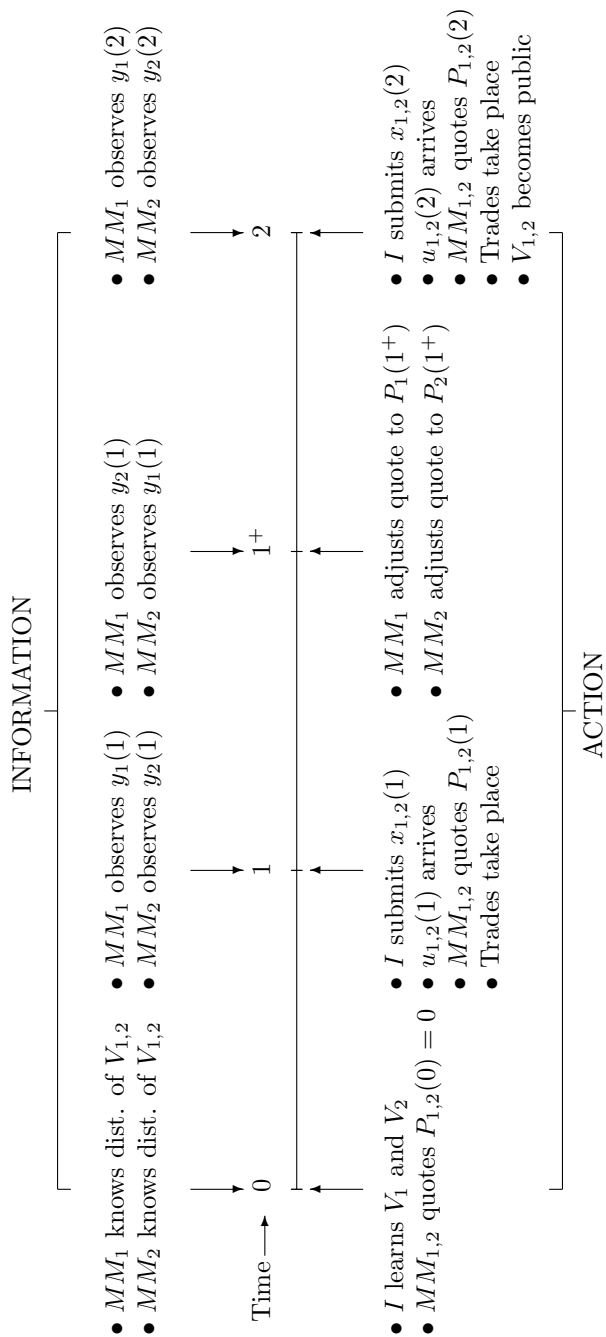- $V_{1,2}$ becomes public

ACTION

Figure 2: **Graphical Illustration of Theorem 1**

We plot the parameters of the solution defined in Theorem 1 as functions of ρ with $\beta_K = 1$. Panel A shows elasticities of the order flow to information in the first trading round, $\beta_{\pm}$, given by (16). Panel B shows $\Lambda_{\pm}$ given by (18). Panel C shows inverse market depth parameter for trading rounds one, λ_1, given by (17), and two, λ_2, given by (20).

Panel A: $\beta_{\pm}(\rho)$

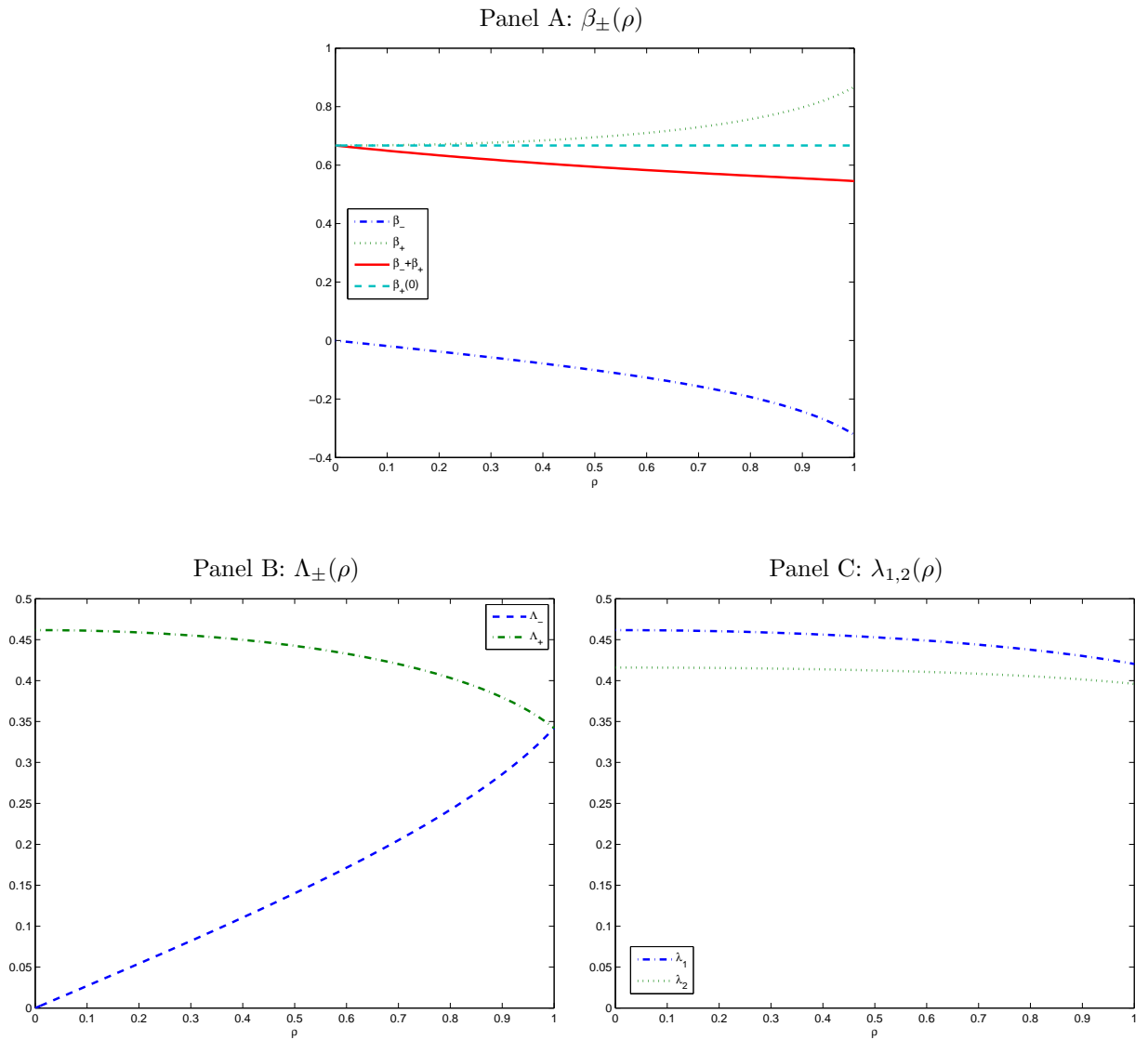Panel B: $\Lambda_{\pm}(\rho)$

Panel C: $\lambda_{1,2}(\rho)$

Figure 3: **Model Correlations Coefficients and VAR Coefficients as a Function of Asset Correlation**

This figure compares comparative statics results with respect to ρ for equal-weighted portfolios of two stocks implied by the model. Panel A plots various lead-lag correlation coefficients implied by the model as functions of ρ. Panel B plots vector autoregression coefficients from the following regression

$$\begin{pmatrix} \Delta P_p(t) \\ x_p(t) \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} \Delta P_p(t-1) \\ x_p(t-1) \end{pmatrix} + \begin{pmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \end{pmatrix}$$

as functions of ρ in the model with $\beta_K = 1$.

Panel A: Correlation Coefficients

Panel B: VAR Coefficients

44