WP 29_09

Yi Xue
Simon Fraser University, Canada

Ramazan Gençay
Simon Fraser University, Canada
and
The Rimini Centre for Economic Analysis, Italy

HIERARCHICAL INFORMATION AND THE RATE OF INFORMATION DIFFUSION

Hierarchical Information and the Rate of Information Diffusion

Yi Xue[*]   Ramazan Gençay[†]

December 2008

Abstract

The rate of information diffusion and consequently price discovery, is conditional upon not only the design of the market microstructure, but also the informational structure. This paper presents a market microstructure model showing that an increasing number of information hierarchies among informed competitive traders leads to a slower information diffusion rate and informational inefficiency. The model illustrates that informed traders may prefer trading with each other rather than with noise traders in the presence of the information hierarchies. Furthermore, we show that momentum can be generated from the predictable patterns of noise traders, which are assumed to be a function of past prices.

Key Words: Information hierarchies, Information diffusion rate, Momentum

JEL No: G10, G11, D43, D82

[*]Corresponding author: Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, V5A 1S6, Canada, Email: `yxue@sfu.ca`.

[†]Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, V5A 1S6, Canada and Rimini Center for Economic Analysis, Rimini, Italy, Email: `gencay@sfu.ca`. Ramazan Gençay gratefully acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada and the Social Sciences and Humanities Research Council of Canada.

The rate of information diffusion, and, consequently, price discovery, is conditional on not only the design of the market microstructure, but also the informational structure. The existing finance literature demonstrates that by polarizing traders into informed traders and noise traders, price discovery can be very slow (see Kyle (1985)) or very fast (see Foster and Viswanathan (1993)) depending on the market microstructure.[1] The goal of this paper is to understand the impact of the informational structure on the degree of information diffusion inefficiency. We show for instance, that even in a market where there are many informed traders, the rate of information diffusion can be very slow if the information is distributed hierarchically.

Specifically, we consider a discrete time, infinite trading horizon model in which traders trade a single asset with a probability of liquidation in every period. The traders differ in two ways. First, traders are in different information hierarchies such that traders in a higher information hierarchy embed the information of those in lower hierarchies. Such a hierarchy leads to a mechanism where informed traders in a higher hierarchy may prefer trading against informed traders in lower hierarchies. In order to trade against informed traders in lower hierarchies, a trader needs to maintain the informational advantage by preventing the spread of the information incorporated into the price (at a fast rate). Thus, this informational mechanism can lead to inefficiency in information diffusion. This is more likely to happen when the number of information hierarchies increases. Second, traders in the lowest information hierarchy, i.e., noise traders, do not receive signals about fundamentals and trade on price direction. Thus, our setting includes the traditional stylized setting of the dichotomy of informed traders and noise traders as a special case. When the informed traders trade with the noise traders, they take into account the fact that noise traders trade on price direction. This can lead to the generation of the momentum.

The informativeness of price, or the rate of the information diffusion, is a function of the layers of hierarchies among informed traders for the following reasons. First, the expectation error of the value of the underlying asset originating from the signal extraction of the partially informed

---

[1]Kyle (1985) shows that when a monopolistically informed trader strategically trades with noise traders, the monopolist will prevent the information from being released, which in turn leads to a slower rate of information diffusion. On the other hand, information is released almost instantly when informed traders possessing identical information compete very aggressively, as analyzed in Foster and Viswanathan (1993).

traders makes the price less informative. The accuracy of the expectation formed by the partially informed traders decreases as the number of the layers of the hierarchies increases. This decrease in accuracy is partly due to the fact that it is increasingly hard for the partially informed traders to infer whether the movement in price is due to a change in fundamental value of the underlying asset or the liquidity shock brought by noise traders.[2] Second, the prevention of information disclosure by informed traders in higher information hierarchies makes the price less informative. When the riskiness of trading with other informed traders relative to the riskiness of trading with noise traders decreases, the informed traders will prevent information disclosure in order to make a profit from other informed traders. We show that informed traders in higher information hierarchies are more likely to be profitable if they trade against other informed traders when the price deviates sufficiently far away from the fundamental value of the underlying asset and the number of information hierarchy layers increases. As a result, the information diffusion is slower when the number of information hierarchy layers among informed traders increases.

In addition, rather than assuming that the behavior of noise traders is independently and identically distributed across time, as in the standard literature, we argue that the behavioral pattern of noise traders consists of predictable (based on price direction) and an unpredictable (idiosyncratic liquidity shock) components. In turn, our model generates a number of interesting and testable implications that are absent from existing models of rational noisy equilibrium. For instance, the predictable pattern of noise traders' behavior resulting in persistence structure in prices, regardless of the distributional properties of the fundamentals. In addition, interaction between informed traders and noise traders can lead to various market liquidity levels. Perhaps the most novel feature of our model is that the trend-following behavior pattern of noise traders can generate momentum in returns, i.e., a positive autocorrelation in returns. One explanation comes from the self-fulfilling type argument that when noise traders believe there is momentum, they will follow the trend; while when noise traders follow the trend, momentum can be generated.

Overall, the main contributions of this paper are as follows. First, this paper proposes a frame-

---

[2]For instance, a positive innovation in the value of the underlying asset is the observational equivalent to a shortage of liquidity supply provided by noise traders from the perspective of the partially informed traders.

work to study the impact of hierarchical information and the layers of hierarchy on the speed of price discovery. We show that the speed of price discovery and the informativeness of price decrease when the number of the information hierarchy layers increases. Second, the predictable behavior pattern of noise traders has a significant impact on the persistence structure of prices and returns. The trend-following behavior pattern facilitates the formation of momentum in returns. Third, bounded rationality of noise traders (the unpredictable component) may generate profits for them when informed traders believe it is too risky to trade with them and choose to trade with other informed traders, thus providing justification for the existence of noise traders.

The rest of the paper is organized as follows. In Section II, a case of two information hierarchies is discussed where only one informed trader and noise traders are presented in the market. This model serves as a benchmark to motivate the extensions that follow and demonstrates that the predictable pattern of noise traders results in the persistence in prices and facilitates the formation of persistence in returns. In addition, the trend-following behavior of noise traders can generate momentum in returns. In section III, an information hierarchy with three levels, two types of informed traders and noise traders, is studied. The information diffusion is slower compared to the benchmark case. In section IV, a general case of N information hierarchies is investigated. The information diffusion speed decreases as the number of information hierarchy levels increases. In section V, Monte Carlo simulations are carried out to demonstrate the impact of the predictable pattern of noise traders and of the number of hierarchical information levels. Finally, we conclude.

I.   Case of Two Information Hierarchies

We begin by providing a simple version of our model without modeling the interactions between informed traders at different hierarchical information levels. This special case helps develop the intuition for how the interaction between a competitive informed trader and noise traders affects the equilibrium properties and the formation of return predictability, in particular, the formation of momentum. The benchmark case is also interesting in its own right as it provides a framework to study how the behavior of noise traders affects the equilibrium properties that can not be captured

3

by existing noisy rational expectation literatures. In Section III, we extend this version to allow for interactions between informed traders by explicitly modeling the information hierarchy among informed traders.[3]

A. Financial Assets

Consider two traded assets. One is a riskless asset with a fixed rate of return of $1 + r$. The other asset is composed of shares or claims on a hypothetical firm, which pays no dividends but has a chance of being liquidated every period. The probability of being liquidated in period $t + 1$, conditional on the firm's surviving until period t, is assumed to be constant λ. When liquidation happens, the firm pays the shareholders a liquidation value V_t, which is assumed to follow a stochastic process. Throughout the paper, we assume that V_t is Gaussian. V_t can be independently and identically distributed (iid) or serially correlated. The normality assumption is for the purpose of simplicity as it permits the existence of a linear equilibrium. The market price for the risky asset is P_t.

B. Game Structure

Two groups of investors, one informed trader and continuum of noise traders, trade the asset on every trading date. At the beginning of each trading date, noise traders supply a certain amount of the shares of the underlying asset to the market. The supply of noise traders is stochastic, which captures the bounded rationality of noise traders. The informed trader observes the current market price and submits his demand. The price can be adjusted upward or downward, and the informed trader will adjust his optimal demand for the number of shares of the underlying asset accordingly. The market clears when the demand from the informed trader equals to the supply from noise traders.

---

[3]We borrow the basic setting from Makarov and Rytchkov (2007).

C.   Noise Traders

Traditionally, the behavior of liquidity suppliers, i.e., noise traders has been assumed in the literature to be independently and identically distributed (*iid*) over time. With the purpose of investigating the effects of the interactions between noise and informed traders on the equilibrium outcomes, we model the behavior of noise traders in a way consistent with Harris (2003). Harris (2003) describes the noise traders as types of traders who trade based on their beliefs concerning the price change direction rather than the fundamentals of the underlying asset. According to this description, a typical example of noise trader is a technical trader. Technical traders trade based on "pattern recognition" type techniques which aid in the formation of the beliefs concerning the direction of prices.[4] Therefore, we assume that there is a predictable components of the aggregate supply of noise traders (technical traders) that should be a function of past prices. In addition, there is an unpredictable component which captures the difference in opinions among noise traders or technical traders. Hence, we model the aggregate supply of the shares of the underlying asset from noise traders throughout the paper Q_t^s as $\beta P_{t-1} + \Theta_t$, where P_t is the price of the trading asset. βP_{t-1} captures the predictable component of the aggregate supply of shares from noise traders. Θ_t is the *iid* Gaussian with mean 0 and variance σ_Θ^2.[5] The random component is for the purpose of preventing prices from being fully revealing, in the spirit of Grossman and Stiglitz (1980).[6]

β is an aggregate measure of noise traders' supply of shares of the underlying asset. The sign of β indicates the aggregate response of the noise traders to the price change direction. For instance, if $\beta < 0$, noise traders will sell the traded asset when the unit price is negative (price is decreasing). It seems that the noise traders behave as if they follow the trend of the price. We label noise traders with negative β *"trend followers"*. In contrast, if $\beta > 0$, noise traders will sell the trading asset when the unit change of price of the traded asset is positive (price is increasing), i.e., noise traders

---

[4]There are two different types of technical analysis: subjective and objective analysis. Subjective analysis captures the fact that different traders may come up with different conclusions based on the same information set due to subjective judgments or priors. Even in objective analysis, different traders may have different confidence levels or tolerance levels which lead to different trading decisions.

[5]The mean is not necessarily 0. It could be a positive number with a large magnitude, which would guarantee that noise traders always supply a positive number of shares. The zero mean assumption could be interpreted as demeaned supply.

[6]The predictable component is specified for simplicity, although it is easy to extend our analysis to allow for many price lags without changing our main results.

act against the trend of the price, and we label the noise traders with positive β *"contrarians"*. The magnitude of β measures the sensitivity of noise traders to price direction changes. The larger the magnitude of β, the more sensitive noise traders are to the changes in price direction. That is, if the magnitude of β is large, with a small change of price, noise traders adjust their holdings of the underlying asset to a large extent. The magnitude of β can be interpreted as the aggregate number of noise traders present in the market as well. β could be time-varying. Throughout this paper, however, we do not intend to model the evolution of noise traders and assume β to be a time invariant parameter. We will elaborate more on the relationship between β and equilibrium properties later.

D. Informed Traders

In this benchmark model, we only consider a single informed investor. In each trading period t, the investor receives a signal S_t about the fundamental value of the underlying asset V_t:

$$S_t = V_t + b_s \eta_t \tag{1}$$

where $\eta_t \sim iidN(0,1)$. b_s is the standard deviation of the signal and $1/b_s^2$ is the precision of the signal. The signal is more informative when b_s decreases. The investor is assumed to have a mean-variance preference, i.e., $E(W_t) - \frac{1}{2}\delta Var(W_t)$, where W_t is his wealth level at t and δ is the preference parameter. When $\delta = 0$, the traders are risk neutral, and when δ increases, the traders become more risk averse. In order to obtain a closed-form solution, we use myopic preference to abstract away from dynamic hedging considerations.[7] Let Q_{t+1} be excess return, i.e.,

$$Q_{t+1} = \lambda V_{t+1} + (1-\lambda)P_{t+1} - (1+r)P_t \tag{2}$$

where $\lambda V_{t+1} + (1-\lambda)P_{t+1}$ is the expected return from the holding of one share of the trading asset and $(1+r)P_t$ is the opportunity cost of holding. Therefore, the per period utility maximization

---

[7]Myopic preference is a simplifying assumption, and its main purpose is to obtain an analytically tractable solution.

problem for the informed trader is

$$\max_{X_t} E[Q_{t+1}|F_t]X_t - \frac{1}{2}\delta Var[Q_{t+1}|F_t]X_t^2$$

where F_t is the information set available for informed traders that contains the current price, the history of past prices, and all the received signals, that is, $F_t = \{P_t, P_{t-1}, \ldots, P_0, S_t, S_{t-1}, \ldots, S_0\}$, and X_t is the holding of the informed investor at trading period t. For the expositional purposes, we use $E_t[.] = E[.|F_t]$ throughout the paper. Hence, the optimal demand of the shares of the underlying asset from informed traders at t is

$$X_t = \frac{E_t[Q_{t+1}]}{\delta Var_t[Q_{t+1}]} \tag{3}$$

Let $\omega_t = \frac{1}{\delta Var_t[Q_{t+1}]}$ so that $X_t = \omega_t E_t[Q_{t+1}]$. Thus, the informed trader adjusts his holding of the risky asset proportional to the expected return from the holdings. The risk averse coefficient affects the proportion of the informed trader's investment in the risky asset. When $\delta = 0$, the optimal holding X_t is not well-defined for $E_t[Q_{t+1}] \neq 0$, and X_t could be any amount for $E_t[Q_{t+1}] = 0$. Namely, if the informed trader is risk-neutral, he would like to borrow an infinite amount of money to invest in the risky asset if its expected return is greater than zero. He will spend any portion of his wealth on the risky asset if its expected return is equal to zero because the risky asset is indifferent from the riskless asset in this case. Similarly, when δ increases, the trader becomes more risk averse; he will invest less and less of his wealth into the risky asset. In particular, if $\delta = \infty$, he will not invest any of his wealth into risky asset.

E. Equilibrium

We focus on stationary and linear expectation equilibrium, where $\text{Var}_t[Q_{t+1}]$ is constant such that $\omega_t = \frac{1}{\delta \text{Var}_t[Q_{t+1}]} = \omega$ is time invariant and there is no bubble in prices. Hence $X_t = \omega E_t[Q_{t+1}]$.

Market clearing implies $X_t = \beta P_{t-1} + \Theta_t$. Hence,

$$P_t = -\frac{\beta P_{t-1}}{\omega(1+r)} - \frac{\Theta_t}{\omega(1+r)} + \frac{1}{1+r} E_t[\lambda V_{t+1} + (1-\lambda)P_{t+1}] \tag{4}$$

Defining $P_t^* = P_t - \rho P_{t-1}$, we can rewrite Equation 4 as

$$P_t^* = -\frac{\Theta_t}{\omega(1+r)} + \frac{1}{1+r} E_t[\lambda V_{t+1} + (1-\lambda)P_{t+1}^*]$$

where ρ solves:

$$(1-\lambda)\rho^2 + (1+r)\rho + \beta = 0. \tag{5}$$

The detailed derivation is given in Appendix. Iterating Equation 5 forward and invoking the no-bubble constraint, i.e., $lim_{j\to\infty}(\frac{1-\lambda}{1+r})^j P_{t+j} = 0$, we have:

$$
\begin{aligned}
P_t^* &= -\frac{\Theta_t}{\omega(1+r)} + \sum_{s=0}^{\infty} (\frac{1-\lambda}{1+r})^s E_t[\lambda V_{t+s+1} - \frac{(1-\lambda)\Theta_{t+s+1}}{\omega(1+r)}] \\
&= -\frac{\Theta_t}{\omega(1+r)} + \sum_{s=0}^{\infty} (\frac{1-\lambda}{1+r})^s E_t[\lambda V_{t+s+1}]
\end{aligned}
$$

Hence,

$$P_t = \rho P_{t-1} - \frac{\Theta_t}{\omega(1+r)} + \sum_{s=0}^{\infty} (\frac{1-\lambda}{1+r})^s E_t[\lambda V_{t+s+1}] \tag{6}$$

Equation 6 shows that the equilibrium price at time t is the sum of three terms. The first term incorporates the predictable pattern of noise traders' behavior. As shown in Equation 5, the serial correlation structure of prices (ρ) is a function of the predictable pattern of noise traders' behavior (β). Intuitively, because the informed trader understands that the predictable pattern of the aggregate supply from the noise traders is a function of past prices, he can forecast the mean of the aggregate supply from the noise traders. Thus, the trader can form a better forecast of the price in the next period by taking into account this information, which will help increase his profit. As a result, the price in the next period is correlated with the past prices, which generates the

8

serial correlation. As demonstrated in a later section, this serial correlation in prices brought by the predictable pattern in noise traders' behavior generates a serial correlation in returns as well. Therefore, the momentum could be generated. Notice that this result is independent of the assumption of the fundamental value of the underlying asset. We will elaborate more on this in later sections. The second term compensates for the risks for informed traders, which originate from the non-predictable component of aggregate noise traders' behavior. The third term is the expected payoff of asset holdings.

F.   Equilibrium Properties

Stationary Equilibrium

We then study the conditions for the equilibrium price to be stationary. We rewrite Equation 5 as

$$P_t - \rho P_{t-1} = -\frac{\Theta_t}{\omega(1+r)} + \sum_{s=0}^{\infty} (\frac{1-\lambda}{1+r})^s E_t[\lambda V_{t+s+1}] \tag{7}$$

Notice that in Equation 7, the right hand side (RHS) is assumed to be covariance stationary and has finite variance given the stationary assumption of V_t. Therefore, the price is covariance stationary when $|\rho| < 1$. Formally,

Proposition 1 *The relationship between the stationarity of equilibrium and β can be summarized as follows:*

1. *When noise traders are trend followers ($\beta < 0$), the maximum number of stationary equilibria is 1. Formally, there exist a stationary equilibrium if and only if the probability of liquidation is not sufficiently large ($\lambda < \beta + r + 2$).*

2. *When noise traders are contrarians and not sensitive to price direction change ($\frac{1+r}{2} > \beta > 0$) and the probability of liquidation is sufficiently large ($\lambda > \beta - r$), there is a unique stationary equilibrium. However, if the probability of liquidation is not sufficiently large, ($\lambda < \beta - r$), there exists two stationary equilibria.*

9

3. *When noise traders are contrarians and they are sensitive to price direction change ($\beta > \frac{1+r}{2}$) and if the probability of liquidation is sufficiently small ($\lambda < \frac{1+r}{2}$), there exist two stationary equilibria. Additionally, if $\lambda > \beta - r$, there is a unique equilibrium. Otherwise, if $\frac{1+r}{2} < \lambda < \beta - r$, there is no stationary equilibrium.*

Proposition 1 suggests that the predictable component β and the probability of liquidation λ play important roles in determining of the stationarity of the equilibrium price process. As shown in Proposition 1, the equilibrium can be stationary only if the liquidation probability is not large. This result is consistent with the observation that when the probability of being liquidated in the following period is very large, an informed trader is reluctant to adjust the optimal holding in response to the short run profitable opportunity originating from the trading pattern of noise traders. When the probability of liquidation is small, the riskiness of exploiting the short run profitable opportunity from noise traders is relatively low. In that case, the informed trader may be more willing to adjust his holding accordingly. Therefore, the market will be cleared because the demand from the informed trader will match the supply from noise traders. Therefore, a stationary equilibrium can exist only if the probability of liquidation is sufficiently low.

Momentum

Momentum is defined as the rate of acceleration of a security's price or volume in technical analysis terms. Once a momentum trader sees an acceleration in a stock's price, earnings, or revenues, the trader will often take a long or short position in the stock with the hope that its momentum will continue in either an upward or downward direction. This strategy relies more on short-term movements in price rather than on the fundamental value of companies. Jegadeesh and Titman (1993) show that the momentum trading strategy can generate abnormal profit. Since its discovery, momentum has been one of the most resilient anomalies challenging the market efficiency hypothesis. It is well known that any theory seeking to explain momentum should be able to generate positive serial correlations in returns. In this section, we investigate the ability of our model to generate momentum, and particularly the positive autocorrelation in returns.

We consider two specifications of V_t. To start, we assume V_t are *iid* with $E_t[V_{t+s}] = \mu$. Hence, Equation 6 becomes

$$P_t = \rho P_{t-1} - \frac{\Theta_t}{\omega(1+r)} + \frac{\lambda\mu}{r+\lambda} \tag{8}$$

Equation 8 demonstrates that if $|\rho| < 1$, P_t follows an autoregressive (AR) (1) process and $\text{Corr}[P_t, P_{t-s}] = \rho^s$ follows directly. Defining return as the difference between the price levels, i.e., $r_t = P_t - P_{t-1}$, we have

$$r_t = \rho r_{t-1} - \frac{\Theta_t - \Theta_{t-1}}{\omega(1+r)} \tag{9}$$

Equation 9 shows that given $|\rho| < 1$, returns follow an autoregressive moving average model (ARMA)(1,1) process with mean 0. It can further be shown that $\text{Corr}[r_t, r_{t-1}] = \rho^2 + \rho - 1$. It follows directly that there can be a positive serial correlation of returns at the first lag. Formally, if $\rho < \frac{-\sqrt{5}-1}{2}$ or $\rho > \frac{\sqrt{5}-1}{2}$, the returns display a positive autocorrelation at first lag, i.e., momentum is generated.[8]

The result of positive serial correlations in returns provides a rational explanation of momentum. From Proposition 1, if we restrict ourselves to stationary equilibrium, the only possible scenario resulting in positive serial correlations in returns is when ρ is positive which corresponds to $\beta < 0$ and $\lambda < \beta + r + 2$ case. This suggests that it is only possible in our model to generate a momentum anomaly when noise traders behave like trend followers. This is consistent with a self-fulfilling explanation. The reason why noise traders behaves like trend-followers is because the momentum anomaly exists, and when noise traders behaves like trend-followers, the momentum anomaly can be generated.

Next, we consider V_t to be a stationary AR(1) process. That is, $V_t = aV_{t-1} + b_v\epsilon_t$, where ϵ_t are *iid* standard normals, b_v is the inverse of the square root of the precision of the innovations, and $a < 1$. When b_v increases, the precision of the signal decreases; that is, the V_t becomes noisier.

---

[8]In order to have a positive ρ which is consistent with the observed price autocorrelation coefficient, we choose $\rho > \frac{\sqrt{5}-1}{2}$. Combined with the annual risk free rate $r = 0.03$ and liquidation probability $\lambda = 10\%$, we need to have $\beta < -1$.

Then $E_t[V_{t+s}] = a^s V_t$. Hence

$$P_t = \rho P_{t-1} - \frac{\Theta_t}{\omega(1+r)} + \frac{\lambda a V_t}{1+r-(1-\lambda)a} \tag{10}$$

Defining return as the difference in price, i.e., $r_t = P_t - P_{t-1}$, we have:

$$r_t = \rho r_{t-1} - \frac{\Delta\Theta_t}{\omega(1+r)} + \frac{\lambda a}{1+r-a(1-\lambda)}\Delta V_t, \tag{11}$$

where $\Delta\Theta_t = \Theta_t - \Theta_{t-1}$ and $\Delta V_t = V_t - V_{t-1}$. It can be shown that the unconditional covariance of returns is

$$\text{Cov}[r_t, r_{t-1}] = \rho\text{Var}[r_{t-1}] + \frac{\sigma_\Theta^2}{\omega^2(1+r)^2} + \frac{\lambda^2 a^2}{(1+r-a(1-\lambda))^2}\text{Var}[V_{t-1}] \tag{12}$$

We are still able to generate positive serial correlations in returns when $\rho > 0$. From Proposition 1, when noise traders are trend-followers, the ρ can be positive and momentum can exist. In addition, even if ρ is negative, it is still possible for momentum to exist as long as the right hand side (RHS) of Equation 12 is positive. In other words, if the fundamental value of the underlying asset is serially correlated, the momentum could exist regardless of the behavior pattern of the noise traders.

What if There is no Pattern in Noise Traders' Behavior?

To illustrate of the benefits of the assumption that there is a predictable component in noise traders' behavior, i.e. βP_{t-1}, we investigate the case where there is no predictable pattern in noise traders' behavior, i.e., $\beta = 0$. That is, the liquidity supply from noise traders is Θ_t, where Θ_t is *iid* normal. This is consistent with the standard assumption about noise traders' behavior in the literature (for instance, Kyle (1985)). The equilibrium price can be shown to be

$$P_t = -\frac{\Theta_t}{\omega(1+r)} + \sum_{s=0}^{\infty}(\frac{1-\lambda}{1+r})^s E_t[\lambda V_{t+s+1}] \tag{13}$$

This shows that the statistical properties of prices are fully determined by the statistical properties of the fundamental value of the underlying asset in the absence of the predictable pattern in noise traders' behavior. Hence, the extra gains from the assumption of a predictable pattern in noise traders' behavior can be summarized as follows:

1. The assumption of a predictable pattern, i.e., $\beta \neq 0$, results in persistence in price regardless of the statistical assumption of fundamentals. On the contrary, if there is no predictable pattern, i.e., $\beta = 0$, the persistence structure in price depends on the statistical assumption of fundamentals. With the presence of the predictable pattern in noise traders' behavior, price is persistent even when V_t is iid. The persistence in prices is partly due to the fact that the informed trader adjusts his optimal holding based on the expectation of the predictable pattern in the noise traders' behavior. This predictable pattern in noise traders' behavior preserves the correlation structure of prices across time. The persistence in prices is also due to the fact that the noise traders utilize the information concerning past prices to adjust their position, which determines the price in the current period. This feedback mechanism can also help the formation of the persistence structure in prices. On the contrary, if there is no predictable pattern in noise traders' behavior, the equilibrium price is entirely determined by the distributional assumption of fundamentals, i.e., V_t. To see this, from Equation 13 we can demonstrate that if V_t is iid, P_t is iid. If V_t is an AR(1) process, then P_t is an AR(1) process.

2. The assumption of a predictable pattern, i.e., $\beta \neq 0$, generates momentum (the positive serial correlation in returns), while if there is no predictable pattern, i.e., $\beta = 0$, momentum cannot be generated in our model. When there is no predictable pattern in noise traders' behavior, generally, asymmetric information cannot generate momentum alone. Intuitively, rational traders require higher compensation for holding a larger amount of the risky asset. This leads to a positive relationship between returns in the next period and the supply from noise traders. In the meantime, there is a negative relationship between price and the supply from noise traders. Recall that return in this period is the difference between prices. If supply

is assumed to be *iid*, realized returns are negatively correlated.[9] On the contrary, when there is a predictable pattern in noise traders' behavior, as indicated earlier, momentum or positive serial correlation in returns can be generated in our model. Notice that there are two opposing effects that generate momentum. One is the negative slope in the demand curve of informed traders and the other one is trend-following behavior of noise traders. We show that under some parameters values, the noise traders' trend following behavior may dominate the downward demand curve effects, which generates a positive serial correlation in returns and provides an explanation for the existence of momentum.

II.   Case of Three Information Hierarchies

We now extend the simple model of the previous section to allow for the interaction of informed traders at different hierarchical information levels. We consider two types of informed traders, corresponding to two informational hierarchical levels. To obtain a closed form solution, we further simplify the setting. Assume there are two factors that jointly determine the fundamental values of the underlying asset. That is, V_t is a function of the two factors V_t^1 and V_t^2. For simplicity, we assume that the function is linear, i.e., $V_t = V_t^1 + V_t^2$. In addition, we assume that there is no noise in the signals. Namely, a type 1 trader (fully informed trader) receives two signals per period, $S_{1,t} = V_t^1$ and $S_{2,t} = V_t^2$. A Type 2 trader (partially informed trader) receives only one signal $S_{2,t} = V_t^2$. Furthermore, we assume that V_t^1 and V_t^2 are AR(1) processes, i.e., $V_t^1 = aV_{t-1}^1 + b_V \epsilon_t^1$ and $V_t^2 = aV_{t-1}^2 + b_V \epsilon_t^2$.

The information set for the type 2 trader is $F_{2,t} = \{P_t, P_{t-1}, \ldots, V_t^2, V_{t-1}^2, \ldots\}$, and the information set for the type 1 trader is $F_{1,t} = \{P_t, P_{t-1}, \ldots, V_t^1, V_{t-1}^2 1, \ldots, V_t^2, V_{t-1}^2, \ldots\}$. By construction, $F_{2,t} \subset F_{1,t}$, and it captures the idea of a hierarchical information structure. It is well known from the existing literature on forecasting the forecasts of others, the infinite regress problem can be avoided with a hierarchical information structure. Intuitively, a fully informed trader knows everything a partially informed trader knows. Therefore, a fully informed trader knows exactly the expectation

---

[9] $\text{Cov}[Q_t, r_t] \rightarrow \text{Cov}[\Theta_t, P_t - P_{t-1}] = \text{Cov}[\Theta_t, P_t] < 0$.

formed by a partially informed trader on the signal received by the fully informed trader. Then the infinite regress problem collapses.[10] Formally, the three information hierarchies' equilibrium is characterized as:

Proposition 2 *If a type 1 trade observes V_t^1 and V_t^2 and a type 2 trader only observes V_t^2, the equilibrium price is*

$$P_t = \rho P_{t-1} + P_V V_t + P_\Theta \Theta_t + P_\Delta(\widehat{V}_t^1 - V_t^1),$$

where

$$P_\Theta = \frac{\rho}{\beta}$$

$$P_V = -\frac{a\lambda}{\frac{\beta}{\rho\Omega} + a(1-\lambda)}$$

$$P_\Delta = -\frac{a\omega_2(\lambda + (1-\lambda)\frac{a\lambda}{-\frac{\beta}{\rho\Omega}-a(1+\lambda)})}{\frac{\beta}{\rho} + a\omega_1(1-\lambda)c}$$

where $\widehat{V}_t^1 = E[V_t^1|F_{2,t}]$, $\Omega = \omega_1 + \omega_2$, $\omega_1 = Var[Q_{t+1}|F_{1,t}]$, $\omega_2 = Var[Q_{t+1}|F_{2,t}]$, and where ρ solves

$$(1-\lambda)\Omega\rho^2 - (1+r)\Omega\rho - \beta = 0$$

and c solves

$$\frac{\rho^2}{\beta^2}\sigma_\Theta^2(1-c)(1-a^2c) - cb_V^2(\frac{a\lambda}{\frac{\beta}{\rho\Omega} + a(1+\lambda)} + P_\Delta)^2 = 0$$

The expectation errors follow an AR(1) process:

$$\widehat{V}_t^1 - V_t^1 = ac(\widehat{V}_{t-1}^1 - V_{t-1}^1) - b_V c\epsilon_t^1 + kP_\Theta\Theta_t$$

---

[10]It is well known that when a hierarchical information structure exists, the fully informed traders can infer the exact expectations of the partially informed traders. Then signal extraction problem between these traders can be characterized by a finite number of the state variables that include the expectations of the partially informed traders. If there is no hierarchical information structure, the signal extraction problem needs to be characterized by an infinite number of the states variables that include the infinite iteration of expectation among traders, for instance, trader A's expectation of trader B's expectation, trader B's expectation of trader A's expectation of trader B's expectation, and so on.

where $k = -\frac{\frac{1-c}{a\lambda}}{\frac{B}{\rho\Omega}+a(1+\lambda)}+P_\Delta$.

Proposition 2 shows that the equilibrium price in the case of three information hierarchies consists of four terms, instead of three terms as in the case of two information hierarchies . The extra term is $P_\Delta(\widehat{V}_t^1 - V_t^1)$, which captures the expectation error of type 2 traders in guessing the signal received by type 1 traders. That is, the forecasting behavior of type 2 traders adds noise to the equilibrium price. Further, as shown in Proposition 2, the expectation error follows an AR(1) process and is thus persistent.

Intuitively, the persistent structure in prices leads to a slower information diffusion rate. If there is no persistent structure in prices, the innovation in fundamentals can be incorporated immediately. But when prices are persistent, the innovation in fundamentals can have long lasting effects so that the prices can not immediately be adjusted to reflect the fundamental value of the underlying asset. The persistent structure is due to the combined effects of persistent structure through the predictable pattern in noise traders and the persistent expectation errors formed by partially informed traders. Thus, we label the former effect the "beta effect" and the latter the "hierarchical effect". We will continue to investigate the roles of these two effects on the information diffusion rate in the next section.

A. Information Diffusion Rate

As mentioned earlier, the persistent structure in prices leads to a slower information diffusion rate. The slower information diffusion rate comes from two sources: the "beta effect" and the "hierarchical effect".

Beta effect

To study the role of the predictable pattern of noise traders' behavior in generating a slower information diffusion rate, we abstract from the hierarchical information setting. That is, we consider the two information hierarchical levels case. Without loss of generality, we restrict ourselves to the AR specification of fundamental evolution, i.e., $V_t = aV_{t-1} + b_v\epsilon_t$, where ϵ_t are *iid* standard normals,

16

b_v is the inverse of the square root of the precision of the innovations, and $a < 1$. As b_v increases, the V_t becomes noisier. Then $E_t[V_{t+s}] = a^s V_t$. Hence

$$P_t = \rho P_{t-t} - \frac{\Theta_t}{\omega(1+r)} + \frac{\lambda a V_t}{1+r-(1-\lambda)a} \tag{14}$$

When there is no predictable pattern in noise traders' behavior, i.e., $\beta = 0$, the equilibrium price is

$$P_t = -\frac{\Theta_t}{\omega(1+r)} + \frac{\lambda a V_t}{1+r-(1-\lambda)a} \tag{15}$$

To demonstrate a slower information diffusion rate, we show that the impulse response in prices to an innovation in liquidity shocks brought by noise traders, i.e., in Θ_t, is larger. Intuitively, when there is an innovation in fundamentals, if the impulse response of prices is larger, it will take a longer time for prices to "settle down". That is, the information diffusion rate is slower.

Lemma 3 *The information diffusion rate is slower if there is a predictable pattern in noise traders' behavior.*

Proof: If there is no predictable pattern in noise traders' behavior, i.e., $\beta = 0$, the impulse response of prices to an innovation in liquidity shocks brought by noise traders, i.e., Θ_t, is IR_t^1 such that

$$
\begin{aligned}
IR_t^1 &= -\frac{1}{(1+r)\Omega} \\
IR_t^1 &= 0
\end{aligned}
\tag{16}
$$

for $t = 0, 1, \ldots$. Notice that $IR_t \geq 0$ for all t. Meanwhile, if there is a predictable pattern in noise traders' behavior, i.e., $\beta \neq 0$, the impulse response of prices to an innovation in the fundamental value of the underlying asset, i.e., V_t, is IR_t^2 such that

$$
\begin{aligned}
IR_0^2 &= -\frac{1}{(1+r)\Omega} \\
IR_t^2 &= \rho IR_{t-1}^2
\end{aligned}
\tag{17}
$$

for $t = 0, 1, \ldots$. It is easy to see that $IR_t^2 - IR_t^1 = \rho IR_{t-1}^2 > 0$ for all $t \geq 0$. That is, the impulse response in prices to an innovation in liquidity shocks is greater when $\beta \neq 0$ than when $\beta = 0$. $IR_t^2 > IR_t^1$ for all t implies that the difference between the current price and equilibrium price is smaller when $\beta = 0$. Therefore, it takes less time for price to converge in the case where noise traders do not have predictable pattern, i.e., $\beta = 0$. In other words, the information diffusion rate is slower if there is a predictable pattern in noise traders' behavior.

Q.E.D

Lemma 3 shows that the existence of a predictable pattern in noise traders' behavior leads to a slower information diffusion rate. The slower information diffusion rate is due to the fact that informed traders make trading decisions taking into account the predictable pattern in noise traders' behavior that generates the persistent structure in prices.

In addition, we investigate the deviation of prices from the fundamental value of the underlying asset due to the existence of the predictable pattern in noise traders' behavior. If the deviation is different than zero persistently, then the information is incorporated into prices at a slower rate.

Formally, if there is no predictable pattern in noise traders' behavior, i.e., $\beta = 0$, there is no deviation of prices from the fundamental value of the underlying asset because of the complete market. That is, when $\beta = 0$, the price in every period fully reflects the fundamental value of the underlying asset. Meanwhile, if there is a predictable pattern in noise traders' behavior, i.e., $\beta \neq 0$, as shown in Equation 14, the price in every period deviates from the fundamental value of the underlying asset. The deviation is

$$D_t = \rho P_{t-1} \tag{18}$$

for $t = 0, 1, \ldots$. It is easy to see that $D_t > 0$ for all t. That is, the predictable pattern in noise traders' behavior leads to deviation of prices from the fundamental value of the underlying asset. The information about the fundamental value of the underlying asset, i.e., innovation in V_t, is incorporated into the price at a longer horizon. It also takes prices a longer time to adjust to account for the innovation in fundamental value of the underlying asset.

18

Hierarchical Effect

This section study the role of hierarchical information levels in generating a slower information diffusion rate. To isolate the beta effect brought by the predictable pattern in noise traders' behavior, we impose $\beta = 0$ in this part.

Proposition 4 *The information diffusion rate is slower in the case of three information hierarchies than in the case of two information hierarchies.*

Proof: Formally, the three information hierarchies' equilibrium when $\beta = 0$ is characterized as: If a type 1 trader observes V_t^1 and V_t^2 and a type 2 trader only observes V_t^2, the equilibrium price is

$$P_t = P_V V_t + P_\Theta \Theta_t + P_\Delta (\widehat{V}_t^1 - V_t^1),$$

where

$$P'_\Theta = -\frac{1}{\Omega(1+r)}$$

$$P'_V = \frac{a\lambda}{1+r-(1-\lambda)a}$$

$$P'_\Delta = \frac{\omega_2(1+r)P'_V}{\Omega(1+r) - \omega_1 ac}$$

where $\widehat{V}_t^1 = E[V_t^1 | F_{2,t}]$, $\Omega = \omega_1 + \omega_2$, $\omega_1 = \mathrm{Var}[Q_{t+1}|F_{1,t}]$, $\omega_2 = \mathrm{Var}[Q_{t+1}|F_{2,t}]$, where c solves

$$\frac{\sigma_\Theta^2}{b_V^2} \frac{(1-a\frac{1-\lambda}{1+r})^2}{\Omega^2 a^2 \lambda^2}(1-c)(1-a^2 c)(1+\frac{\omega_1}{\omega_2}(1-ac\frac{1-\lambda}{1+r}))^2 - c(\frac{\omega_1}{\omega_2}(1-ac\frac{1-\lambda}{1+r}))^2 = 0$$

The expectation errors follow an AR(1) process:

$$\widehat{V}_t^1 - V_t^1 = ac(\widehat{V}_{t-1}^1 - V_{t-1}^1) - b_V c\epsilon_t^1 + kP'_\Theta \Theta_t$$

where $k = (1-c)/(\frac{a\lambda}{1+r-a(1-\lambda)}\frac{\frac{\omega_1}{\omega_2}(1-ac\frac{1-\lambda}{1+r})}{1+\frac{\omega_1}{\omega_2}(1-ac\frac{1-\lambda}{1+r})})$. It is easy to see that $P'_V > 0$. In addition, $P'_V > P'_\Delta > 0$. To see that,

$$P'_\Delta = \frac{\omega_2(1+r)P'_V}{\Omega(1+r) - \omega_1 ac}$$

19

$$= \frac{1}{\frac{\omega_1}{\omega_2}(1 - ac\frac{1-\lambda}{1+r}) + 1} P_V' \tag{19}$$

Notice that $\frac{\omega_1}{\omega_2}(1 - ac\frac{1-\lambda}{1+r}) + 1 > 1$, so that $P_V' > P_\Delta' > 0$. The impulse response of prices to an innovation in liquidity shocks of the underlying asset IR_t^3 is

$$IR_0^3 = (kP_\Delta + 1)P_\Theta'$$
$$IR_t^3 = a^t c^t k P_\Theta' \quad for \quad t > 0 \tag{20}$$

Notice that because shocks in the supply by noise traders, i.e., the innovations in Θ_t, affect the price persistently because that they are incorporated into the persistent expectation errors formed by the partially informed trader. It takes time for the partially informed traders to learn that liquidity shocks are irrelevant to the fundamental value of the underlying asset. Thus, it takes prices a longer time to adjust to the shocks in supply by noise traders and the information diffusion rate is slower.

<div align="right">Q.E.D</div>

We further investigate the deviation of prices from the fundamental value of the underlying asset due to the existence of the predictable pattern in noise traders' behavior. If the deviation is consistently non-zero, then the information is incorporated into prices at a slower rate.

Formally, if there is no predictable pattern in noise traders' behavior, i.e., $\beta = 0$, there is no deviation of prices from the fundamental value of the underlying asset because of the complete market. That is, when $\beta = 0$, the price in every period fully reflects the fundamental value of the underlying asset. Meanwhile, with three hierarchical information levels, the price in every period deviates from the fundamental value of the underlying asset. The deviation is

$$D_t = P_\Delta'(\widehat{V}_t^1 - V_t^1) \tag{21}$$

for $t = 0, 1, \ldots$. It is easy to see that $D_t > 0$ for all t. That is, the case of three information hierarchies leads to the deviation of prices from the fundamental value of the underlying asset. The information about the fundamental value of the underlying asset, i.e., innovation in V_t, is incorporated into the

price at a longer horizon. Therefore, it also takes prices a longer time to adjust to the innovation in fundamental value of the underlying asset.

B. Stationary Equilibrium

Similar to the case of two information hierarchies, there are two possible values of ρ for the equilibrium. Solving the roots explicitly, we have

$$\rho_{1,2} = \frac{(1+r) \pm \sqrt{(1+r)^2 + 4(1-\lambda)\frac{\beta}{\Omega}}}{2(1-\lambda)} \tag{22}$$

From Proposition 2,we have $P_t - \rho P_{t-1} = +P_V V_t + \frac{\rho}{\beta}\Theta_t + P_\Delta(\widehat{V}_t^1 - V_t^1)$. The right hand side (RHS) of the equation is assumed to be stationary and has finite variance given the stationary assumption of V_t. Therefore, the price is stationary when $|\rho| < 1$. Formally,

Proposition 5 *The relationship between the stationarity of equilibrium and β can be summarized as follows:*

1. *When noise traders are contrarians, the maximum number of stationary equilibria is 1. Formally, there exist a unique stationary equilibrium if and only if the probability of liquidation is not sufficiently large ($\lambda < 2 + r - \frac{\beta}{\Omega}$).*

2. *When noise traders are trend followers and they are not too sensitive to price direction changes ($-\frac{\Omega(1+r)}{2} < \beta < 0$), and if the probability of liquidation is not sufficiently large ($\lambda < -r - \frac{\beta}{\Omega}$), two stationary equilibria exist. Otherwise, if liquidation is sufficiently large ($\lambda > -r - \frac{\beta}{\Omega}$), there exists a unique stationary equilibrium.*

3. *When noise traders are contrarians and they are sensitive to price direction changes ($\beta < -\frac{\Omega(1+r)}{2}$), and the probability of liquidation is sufficiently large ($-r - \frac{\beta}{\Omega} < \lambda$, there is a unique stationary equilibrium. If the probability of liquidation is sufficiently small ($\lambda < \frac{1-r}{2}$), there are two stationary equilibria. Otherwise, i.e., $\frac{1-r}{2} < \lambda < -r - \frac{\beta}{\Omega}$, there is no stationary equilibrium.*

21

Proposition 5 suggests that in addition to the number of equilibria, the predictable pattern of noise traders' behavior can affect the stationarity of the equilibrium price process as well. It is interesting to note that the impact of the predictable pattern of noise traders' behavior affects the stationarity of the equilibrium price process in a different way. For instance, in the case of two information hierarchies, when noise traders are trend-followers, there can be two stationary equilibria, while at most one equilibrium can exist in the case of three information hierarchies when noise traders are trend-followers.

C. Persistent Prices

Proposition 2 allows us to calculate the correlation structure of prices explicitly. It can be shown that in stationary equilibrium, the variance of price is

$$\text{Var}[P_t] = \frac{1}{1-\rho^2} \frac{a^2\lambda^2}{(\frac{\beta}{\rho\Omega} + a(1+\lambda))^2} \text{Var}[V_t] + \frac{\rho^2}{\beta^2}\sigma_\Theta^2 + P_\Delta^2 \frac{1}{1-\rho^2 c^2}(b_V^2 c^2 + \frac{k^2\sigma_\Theta^2}{\omega^2(1+r)^2}).$$

This demonstrates that there are two sources of noise in the price. One is the random (unpredictable) behavior of noise traders, and the other is from the expectation errors of partially informed trader. Notice that if the noise traders' behavior becomes noisier, i.e., σ_Θ increases, the variance of expectation errors of the partially informed trader also increases. Intuitively, partially informed trader will try to distinguish the effects of noise traders' behavior and of the signals of the fully informed trader on prices. Because noise traders' behavior does not reveal any information on fundamental value of the underlying asset, it is not helpful in forming expectations of the fundamental value of the underlying asset. If the behavior of noise traders becomes noisier, it is harder for the partially informed traders to extract useful information on fundamentals. The partially informed trader will make greater errors in forming expectations of fundamental value of the underlying asset, which also leads to a noisier market price.

Furthermore, the correlation structure of prices can be recursively represented as:

$$\text{Cov}[P_t, P_{t-1}] = \rho\text{Var}[P_{t-1}] + \frac{1}{1-\rho^2}(\frac{a^2\lambda^2}{(\frac{\beta}{\rho\Omega} + a(1+\lambda))^2})\text{Cov}[V_t, V_{t-1}]$$

22

$$+P_\Delta^2 \text{Cov}[\widehat{V}_t^1 - V_t^1, \widehat{V}_{t-1}^1 - V_{t-1}^1]$$

$$= \rho \text{Var}[P_t] + \frac{a}{1-\rho^2}\left(\frac{a^2\lambda^2}{(\frac{\beta}{\rho^\Omega} + a(1+\lambda))^2}\right) + P_\Delta^2 ac$$

$$\text{Cov}[P_t, P_{t-s}] = \rho \text{Cov}[P_t, P_{t-s+1}] + \frac{a^s}{1-\rho^2}\left(\frac{a^2\lambda^2}{(\frac{\beta}{\rho^\Omega} + a(1+\lambda))^2}\right) + P_\Delta^2 (ac)^s$$

This shows that the prices are more persistent compared to the case of two information hierarchies because of the additional correlation structure brought about by the expectation errors, i.e., $P_\Delta^2(ac)^2$.

D. Return Predictability

The return is defined by the difference in prices, i.e., $r_t = P_t - P_{t-1}$. From Proposition 2, return is

$$r_t = \rho r_{t-1} - \frac{a\lambda}{\frac{\beta}{\rho^\Omega} + a(1+\lambda)}\Delta V_t + \frac{\rho}{\beta}\Delta\Theta_t + P_\Delta\Delta(\widehat{V}_t^1 - V_t^1), \tag{23}$$

where $\Delta V_t = V_t - V_{t-1}$, $\Delta\Theta_t = \Theta_t - \Theta_{t-1}$ and $\Delta(\widehat{V}_t^1 - V_t) = (\widehat{V}_t^1 - V_t^1) - (\widehat{V}_{t-1}^1 - V_{t-1}^1)$.

Notice that the statistical structure of returns is determined by the statistical properties of ΔV_t, $\Delta\Theta_t$, and $\Delta(\widehat{V}_t^1 - V_t)$, which correspond to evolution of the fundamental value of underlying asset, noise traders' behavior, and expectation errors. Notice that ΔV_t follows an ARMA(1,1) process, $\Delta(\widehat{V}_t^1 - V_t)$ follows an ARMA(1,1) process and $\Delta\Theta_t$ follows a MA(1) process. Formally,

$$\Delta V_t = V_t - V_{t-1} = a(V_{t-1} - V_{t-2}) + b_V(\epsilon_t - \epsilon_{t-1}) = a\Delta V_{t-1} + b_V(\epsilon_t - \epsilon_{t-1})$$

$$\Delta(\widehat{V}_t^1 - V_t) = ac(\Delta(\widehat{V}_t^1 - V_t)) - (b_V c\epsilon_t^1 - k\frac{\rho}{\beta}\Theta_t - b_V c\epsilon_{t-1}^1 + k\frac{\rho}{\beta}\Theta_{t-1}) \tag{24}$$

Similar to the analysis of prices, the correlation structure of returns is

$$\text{Cov}[r_t, r_{t-1}] = \rho\text{Var}[r_{t-1}] + \frac{a^2\lambda^2}{(\frac{\beta}{\rho^\Omega} + a(1+\lambda))^2}\text{Cov}[\Delta V_t, \Delta V_{t-1}] - \frac{\rho^2}{\beta^2}\sigma_\Theta^2$$

$$+ P_\Delta^2\text{Cov}[\Delta(\widehat{V}_t^1 - V_t), \Delta(\widehat{V}_{t-1}^1 - V_{t-1})]$$

$$\text{Cov}[r_t, r_{t-s}] = \rho\text{Cov}[r_t, r_{t-s+1}] + \frac{a^2\lambda^2}{(\frac{\beta}{\rho^\Omega} + a(1+\lambda))^2}\text{Cov}[\Delta V_t, \Delta V_{t-s}] - \frac{\rho^2}{\beta^2}\sigma_\Theta^2$$

$$+P_\Delta^2 \text{Cov}[\Delta(\widehat{V}_t^1 - V_t), \Delta(\widehat{V}_{t-s}^1 - V_{t-s})] \tag{25}$$

This shows that there is a serial correlation structure in returns. Moreover, it is possible to generate momentum, i.e., a positive correlation in returns, under certain parameter values.

In this section, we investigate the impact of interaction between informed traders at different information hierarchies, namely, the fully informed trader and partially informed trader, on equilibrium properties. We show that compared to the case of two information hierarchies, the market price contains an extra term, which is the expectation error originating from signal extraction by the partially informed trader and adding noise to the market price. This leads to a slower information diffusion rate. The slower information diffusion rate is due to the combined effects of the "beta" effect (the persistence structure brought by the predictable pattern in noise traders' behavior) and the "hierarchical effect" (the extra noise brought by the signal extraction behavior of the partially informed trader). Therefore, the increasing number of information hierarchies decreases the information diffusion rate. That is, when there is an innovation in the fundamental value of the underlying asset, it takes a longer time for the market price to incorporate the innovation and reflects the fundamental value of the underlying asset. Furthermore, prices become more persistent. The extra persistence is also due to the persistence in the expectation errors made by the partially informed trader, which are due to the persistence of the belief updating. In addition, momentum in returns can be generated as well. Two sources contribute to the formation of momentum. First, as in the case of two information hierarchies, the predictable pattern of noise traders' behavior can aid in the formation of momentum via the self-fulfilling mechanism. Second, the autocorrelations in expectation errors can contribute to the momentum as well.

III.   Case of N Information Hierarchies

We extend the Case of three information hierarchies to allow for N information hierarchies to study the impact of increasing the number of information hierarchies. Intuitively, with increasing number of information hierarchies, the fully informed traders may want to trade with the traders

in lower information hierarchies. This is because it is possible that the riskiness of trading with the partially informed traders is lower than trading with noise traders because the riskiness originating from the unpredictable pattern of their behavior may be relatively larger. Formally, we assume there are N factors determining the fundamental value of the underlying asset in a linear fashion, i.e., $V_t = \sum_{n=1}^{N-1} V_t^n$. In addition, we assume that there is no noise in signals for all types of traders, and the type i informed trader observes $V_t^i, \ldots, V_t^{N-1}$. Assume V_t is an AR(1) process, i.e., $V_t^n = aV_{t-1}^n + b_V \epsilon_t^n$ for $n = 1, 2, \ldots, N-1$. The information set for a type i trader is $F_{i,t} = \{P_t, P_{t-1}, \ldots, V_t^i, V_{t-1}^i, \ldots, V_t^{N-1}, \ldots\}$. By construction, $F_{i,t} \subset F_{j,t}$ if $i > j$.

Proposition 6 *If a type i informed trader observes $V_t^i, \ldots, V_t^{N-1}$, for $i = 1, 2, \ldots, N-1$. The equilibrium price is*

$$P_t = \rho P_{t-1} + P_V V_t + P_\Theta \Theta_t + P_{\Delta^1}(\widehat{V}_{2,t}^1 - V_t^1) + \sum_{s=3}^{N-1} P_{\Delta^{s-1}} \left(\sum_{n=1}^{s-1} \widehat{V}_{s,t}^n - \sum_{n=1}^{s-1} V_t^n \right),$$

where

$$P_\Theta = \frac{\rho}{\beta}$$

$$P_V = \frac{a\lambda}{-\frac{\beta}{\rho\Omega} - a(1-\lambda)}$$

$$P_{\Delta^{N-2}} = \frac{a\omega_{N-1}(\lambda + (1-\lambda)P_V)}{-\frac{\beta}{\rho} - a(1-\lambda) \sum_{i=1}^{N-2} \omega_i}$$

for $1 < s < N-2$,

$$P_{\Delta^s} = \frac{a \sum_{i=s}^{N-1} \omega_i(\lambda + (1-\lambda)P_V) - a\omega_{s+1} \sum_{i=s+1}^{N-2} P_{\Delta^i}}{-\frac{\beta}{\rho} - a(1-\lambda) \sum_{i=1}^{s-1} \omega_i}$$

for $s = 1$,

$$P_{\Delta^1} = \frac{a\omega_2(\lambda + (1-\lambda)P_V) - ac\omega_2 \sum_{i=2}^{N-2} P_{\Delta^i}}{-\frac{\beta}{\rho} - a\omega_1(1-\lambda)c}$$

25

where $\widehat{V}_{s,t}^j = E[V_t^j|F_{s,t}]$ for $j = 1, 2, \ldots, s-1$; $\Omega = \sum_{n=1}^{N-1} \omega_n$; $\omega_s = Var[Q_{t+1}|F_{s,t}]$, ρ solves

$$(1-\lambda)\Omega\rho^2 - (1+r)\Omega\rho - \beta = 0$$

and c solves

$$\frac{\rho^2}{\beta^2}\sigma_\Theta^2(1-c)(1-a^2c) - cb_V^2(P_V - \sum_{s=1}^{N-2} P_{\Delta^s})^2 = 0$$

The expectation errors from type 2 traders follow an AR(1) process:

$$\widehat{V}_{2,t}^1 - V_t^1 = ac(\widehat{V}_{2,t-1}^1 - V_{t-1}^1) - b_V c\epsilon_t^1 + kP_\Theta\Theta_t$$

where $k = \dfrac{1-c}{\frac{\beta}{\rho\Omega} - a(1+\lambda)} - \sum_{s=1}^{N-2} P_{\Delta^s}$.

The expectation errors from type s traders follow an AR(1) process, for $s > 2$:

$$\sum_{n=1}^{s-1} \widehat{V}_{s,t}^n - \sum_{n=1}^{s-1} V_t^n = a(\sum_{n=1}^{s-1} \widehat{V}_{s,t-1}^n - \sum_{n=1}^{s-1} V_{t-1}^n) - b_V \sum_{i=1}^{s-1} \epsilon_t^i.$$

Remarks:

1. *Observational equivalence of informed traders in higher information hierarchies.* Notice that, the noise traders' behavior only affects the type 2 trader (the partially informed trader in second highest information hierarchy) directly. Partially informed traders in other information hierarchies need to distinguish more than two sources of randomness, and they only have the market price as an identifying instrument. One interesting result in this setting is that the forecasting of the signals owned by the trader in a higher hierarchy collapses. Formally, $E[E[V_t^s|F_{j,t}]|F_{i,t}] = E[V_t^s|F_{i,t}]$, which follows directly from the law of iterated expectation. It simply states that the traders' best guess in lower hierarchies concerning the signal received by the traders in higher hierarchies is their own expectation of the fundamental value of the underlying asset. This suggests that traders in lower information hierarchies simply cannot

distinguish the traders in higher information hierarchies, implying that all traders in higher information hierarchies are observationally equivalent in the perspective of a trader in a lower hierarchy.

2. *Traders in higher information hierarchies may want to trade with informed traders in lower information hierarchies.* Intuitively, the fully informed trader knows the optimal trading decisions of any partially informed traders in lower hierarchies. Therefore, there is only one type of risk if the fully informed trader wants to trade with the partially informed traders, which originates from the probability of being liquidated. Meanwhile, in addition to the liquidation risk, the fully informed trader faces extra risk when he wants to trade with noise traders, which is due to unpredictable part of noise traders' behavior. In other words, it may be more profitable for the fully informed trader to trade with the partially informed traders rather noise traders. It can be shown that under certain parameter values, the fully informed trader will trade against the partially informed traders rather than noise traders. We will elaborate more on it in a later section.

3. *The rationale for the presence of noise traders.* Following the analysis above, when traders with an informational advantage choose not to trade with noise traders, noise traders could make a profit. The conventional wisdom that noise traders could not make a profit in the long run dictates that informed traders have the informational advantages and do at least well as noise traders. However, in our model, informed traders may prefer trading among themselves rather than trade against noise traders. This may lead to profits for noise traders even in the long run. Thus, the presence of noise traders on the market is justified.

A. Information Diffusion Rate

Similar to the case of three information hierarchies, the persistent structure in prices leads to a slower information diffusion rate. The slower information diffusion rate is due to two sources: the "beta effect" and the "hierarchical effect", as in the case of three information hierarchies. The "beta effect" works similar to that in the case of three informational hierarchies and we restrict ourselves

to the study of the "hierarchical effect". In Particular, we want to study the relationship between information diffusion rate and the number of information hierarchies.

Hierarchical Effect

This section studies the roles of the number of information hierarchies in generating a slower information diffusion rate. To isolate the beta effect brought by the predictable pattern in noise traders' behavior, we impose $\beta = 0$ in this section.

Proposition 7 *The information diffusion rate slows as the number of information hierarchies increases.*

Proof: Formally, the N information hierarchies equilibrium when $\beta = 0$ is characterized as: If a type i informed trader observes $V_t^i, \ldots, V_t^{N-1}$, for $i = 1, 2, \ldots, N-1$. The equilibrium price is

$$P_t = P_V'' V_t + P_\Theta'' \Theta_t + P_{\Delta^1}'' (\widehat{V}_{2,t}^1 - V_t^1) + \sum_{s=3}^{N-1} P_{\Delta^{s-1}}'' \left(\sum_{n=1}^{s-1} \widehat{V}_{s,t}^n - \sum_{n=1}^{s-1} V_t^n \right),$$

where

$$P_\Theta'' = -\frac{1}{\Omega(1+r)}$$

$$P_V'' = \frac{a\lambda}{1 + r - (1 - \lambda)a}$$

$$P_{\Delta^{N-2}}'' = \frac{a\omega_{N-1}(\lambda + (1 - \lambda)P_V'')}{(1+r)\Omega - a(1-\lambda)\sum_{i=1}^{N-2}\omega_i}$$

for $1 < s < N - 2$,

$$P_{\Delta^s}'' = \frac{a\sum_{i=s}^{N-1}\omega_i(\lambda + (1 - \lambda)P_V) - a\omega_{s+1}\sum_{i=s+1}^{N-2}P_{\Delta^i}}{(1+r)\Omega - a(1-\lambda)\sum_{i=1}^{s-1}\omega_i}$$

28

for $s = 1$,

$$P''_{\Delta^1} = \frac{a\omega_2(\lambda + (1-\lambda)P''_V) - ac\omega_2 \sum\limits_{i=2}^{N-2} P''_{\Delta^i}}{(1+r)\Omega - a\omega_1(1-\lambda)c}$$

where $\widehat{V}^j_{s,t} = E[V^j_t | F_{s,t}]$ for $j = 1, 2, \ldots, s-1$; $\Omega = \sum\limits_{n=1}^{N-1} \omega_n$; $\omega_s = \text{Var}[Q_{t+1}|F_{s,t}]$, where c solves

$$\frac{\sigma^2_\Theta}{b^2_V} \frac{(1 - a\frac{1-\lambda}{1+r})^2}{\Omega^2 a^2 \lambda^2}(1-c)(1-a^2c)(1+\frac{\omega_1}{\omega_2}(1 - ac\frac{1-\lambda}{1+r}))^2 - c(\frac{\omega_1}{\omega_2}(1 - ac\frac{1-\lambda}{1+r}))^2 = 0$$

The expectation errors from type 2 traders follow AR(1) process:

$$\widehat{V}^1_{2,t} - V^1_t = ac(\widehat{V}^1_{2,t-1} - V^1_{t-1}) - b_V c\epsilon^1_t + kP''_\Theta \Theta_t$$

where $k = (1-c)/(\frac{a\lambda}{1+r-a(1-\lambda)} \frac{\frac{\omega_1}{\omega_2}(1-ac\frac{1-\lambda}{1+r})}{1+\frac{\omega_1}{\omega_2}(1-ac\frac{1-\lambda}{1+r})})$.

The expectation errors from type s traders follow AR(1) process, for $s > 2$:

$$\sum_{n=1}^{s-1} \widehat{V}^n_{s,t} - \sum_{n=1}^{s-1} V^n_t = a(\sum_{n=1}^{s-1} \widehat{V}^n_{s,t-1} - \sum_{n=1}^{s-1} V^n_{t-1}) - b_V \sum_{i=1}^{s-1} \epsilon^i_t.$$

It is easy to see that $P''_V > 0$.

The impulse response of prices to an innovation in liquidity shocks of the underlying asset IR^4_t, is

$$IR^4_0 = (kP''_{\Delta^1} + 1)P''_\Theta$$

$$IR^4_t = a^t c^t k P''_\Theta \quad for \ \ t > 0 \tag{26}$$

Notice that liquidity shocks, i.e., the innovations in Θ_t, affect the price persistently because that the shocks in supply by noise traders are incorporated into persistent expectation errors formed by the partially informed traders. It takes time for the partially informed traders to learn that the changes in prices are due to shocks in supply and are irrelevant to the fundamental value of the underlying asset. Thus, it takes prices a longer time to adjust to account for the shocks in supply of noise

traders and the information diffusion rate is slower.

<div align="right">Q.E.D</div>

Further, we investigate the deviation of prices from the fundamental value of the underlying asset due to the existence of the predictable pattern in noise traders' behavior. If the deviation is consistently non-zero, the information is incorporated into prices at a slower rate.

Formally, if there is no predictable pattern in noise traders' behavior, i.e., $\beta = 0$, there is no deviation of prices from the fundamental value of the underlying asset because of the complete market. That is, when $\beta = 0$, the price in every period fully reflects the fundamental value of the underlying asset. Meanwhile, with N hierarchical information levels, the price in every period deviates from the fundamental value of the underlying asset. The deviation is

$$D_t = P_{\Delta^1}(\widehat{V}_{2,t}^1 - V_t^1) + \sum_{s=3}^{N-1} P_{\Delta^{s-1}}\left(\sum_{n=1}^{s-1} \widehat{V}_{s,t}^n - \sum_{n=1}^{s-1} V_t^n\right) \tag{27}$$

for $t = 0, 1, \ldots$. It is easy to see that $D_t > 0$ for all t. That is, the case of N information hierarchies leads to deviation of prices from the fundamental value of the underlying asset. That is, the information about the fundamental value of the underlying asset, i.e., innovation in V_t is incorporated into the price at a longer horizon. In addition, as the number of information hierarchies increases, the deviation could be even larger because of the existence of the expectation errors formed by all partially informed traders.

B.   Trading Among Informed Traders

As mentioned earlier, the traders in higher information hierarchies may want to trade against the traders in lower information hierarchies. To illustrate, let us consider the case of three information hierarchies. As shown in Appendix A, the demand from the fully informed trader is

$$\omega_1(((1-\lambda)\rho - (1+r))P_t + (\lambda + (1-\lambda)P_V)aV_t + (1-\lambda)P_\Delta E[(\widehat{V}_{t+1}^1 - V_{t+1}^1)|F_{1,t}] \tag{28}$$

while the demand from the partially informed trader is

$$\omega_2(((1-\lambda)\rho - (1+r))P_t + (\lambda + (1-\lambda)P_V)a(\widehat{V}_t^1 + V_t^2)) \tag{29}$$

where $\widehat{V}_t^1 = a(1 - k(P_V - P_\Delta))\widehat{V}_{t-1}^1 + k(P_V - P_\Delta)V_t^1 + kP_\Theta\Theta_t$.

Consider a scenario where there is a supply shock and no change in the fundamental value of the underlying asset, such that noise traders choose to supply more, i.e., $\Delta\Theta_t > 0$. Although there is no change in the fundamental value of the underlying asset, both types of informed traders will adjust their optimal demands accordingly. Notice that partially informed traders will adjust their expectations of the fundamental value of the asset downward, due to the fact that they cannot distinguish the decrease in fundamental value of the underlying asset from a positive supply shock. Hence, the quantity of adjustment for partially informed traders is $(\lambda + (1-\lambda)P_V)akP_\Theta\Delta\Theta_t$ and for fully informed traders is $(1-\lambda)P_\Delta E[(\widehat{V}_{t+1}^1 - V_{t+1}^1)|F_{1,t}]$.

As indicated earlier in Proposition 2, the expectation error is an AR(1) process and thus persistent. Therefore, from the perspective of the fully informed trader, the expectation error should have the same signs for time t and time $t+1$.

If $P_\Theta < 0$ and thus $(\lambda + (1-\lambda)P_V)akP_\Theta < 0$, partially informed traders will choose to decrease their optimal demand by mistakenly believing that the fundamental value of the asset decreases. This is simply because they cannot distinguish the decrease in fundamental value in the underlying asset from the a positive supply shock. On the other hand, when $P_\Delta < 0$ and thus $(1-\lambda)P_\Delta < 0$, fully informed traders will trade based on the expectation errors from partially informed traders. Notice that partially informed traders will adjust their expectation of the fundamental value of the underlying asset downward, i.e., $\widehat{V}_t^1 - V_t^1 < 0$. From the persistence of the expectation error, fully informed traders will believe $E[(\widehat{V}_{t+1}^1 - V_{t+1}^1)|F_{1,t}] < 0$. Thus, they will increase the holding of the underlying asset in order to make profit from the mistakes of the partially informed traders. Therefore, fully informed traders trade against partially informed traders by trading in opposite directions. In summary, we show that when $P_\Theta < 0$ and $P_\Delta < 0$, fully informed traders trade against partially informed traders in order to profit from the mistakes of the partially informed

31

traders.

Proposition 8 *If $\beta < 0$, $\rho > 0$ and $0 < k < 1$, trading amongst informed traders can happen. In addition, in this case, $P_V < 0$.*

$\beta < 0$ implies that noise traders are trend followers and $\rho > 0$ implies that the price has a positive autocorrelation coefficient which matches the empirical observation of the financial time series of price. According to the Bucy-Kalman filter formula, k is the weight used in the belief updating. A positive k suggests that new information and current belief both receive attention. As a result, the partially informed traders adjust their belief in a wrong way when there is a positive shock in supply.

A negative P_V suggests that when there is a positive innovation in fundamental value of the underlying asset, the price may decrease in response. This is inefficient because such trading amongst informed traders can prolong the procedure of information diffusion and thus the rate of information diffusion is slower.

Proposition 8 characterizes one set of possible parameter values which can generate trading amongst informed traders. That is, partially informed traders mistakenly adjust their beliefs in a wrong way when they cannot distinguish the positive shock in supply from a decrease in fundamental value of the underlying asset. If fully informed traders know that noise traders are trend followers and partially informed traders make a mistake in adjusting their beliefs, they will trade against the partially informed traders to make a profit. This profitable opportunity originates from the mistakes made by the partially informed traders. This is inefficient, because if that is the case, it leads to a negative response of current market price to a positive innovation in the fundamental value of the underlying asset. This distortion of price response is due to the informational arbitrage amongst informed traders.

IV. Numerical Analysis

In this section, we present a numerical study of the model. Using this numerical study, we can show the statistical properties of equilibrium, namely, the price structure, and the return predictability. Second, we want to investigate the "beta effect" and the "hierarchical effect" and their implications.

The parameters used in the simulation study are as follows. The probability of liquidation λ is 5%, which corresponds to an expected lifespan of 20 periods. The risk-free rate r is 1%, which corresponds to the annual yield on a treasury bill. The parameters that describe the fundamental value of the underlying asset are chosen to match the volatility of returns at 12% and we set $a = 0.85$ and $b_V = 0.3$. We vary the above parameters and find that the different sets of parameters do not change the main results qualitatively.

A. Impact of Noise Traders

First, we study the effects of noise traders' behavior. We mainly study the behavior of noise traders in the case of two information hierarchies and the case of three information hierarchies. For the plots, to match the prices' empirical behavior, we want to have a positive AR(1) coefficient that is close to 1. We choose $\beta = -2$ in the case of two information hierarchies and $\beta = -20$ in the case of three information hierarchies. The Figure 1 shows a typical set of plots of price series, return series, and persistent structure which is captured by autocorrelations of prices and returns. It shows that the prices are persistent, and there is a first order positive serial correlation with the magnitude 0.8 in returns. There is a positive serial correlation in returns which suggests the existence of momentum. Similar patterns can also be observed in the plot of the case of three information hierarchies, although the first order serial correlation in returns drops and returns become much less persistent, which is a better approximation of the empirical behavior of returns. This decrease in persistence structure in returns may be due to the existence of interactions between the informed traders which diminishes the effect of the persistent structure brought by the predictable pattern in noise traders' behavior. Thus, when there are more information hierarchies, the magnitude of the

momentum effect is smaller, as shown in Table II. The magnitude of the first order autocorrelation of returns decreases when the number of information hierarchies increases. Furthermore, Table II indicates that the persistence structure in returns decreases when the number of information hierarchies increases.

Table I shows that the simulation results for "beta effect" in the case of three information hierarchies. First column reports AR coefficient of price series, ρ. It demonstrates that when the magnitude of β increases, the AR coefficient ρ decreases. Intuitively, when the size of noise trading on a financial market increases, the behavior of the market price is driven by the noise traders' behavior. The forecasting behavior of informed traders which contribute to the persistence of price in a diminishing manner. The second column reports the first lag serial correlation in returns which also suggests that the magnitude of "momentum effect" decreases as the magnitude of β increases.

There is an ARCH effect in returns which suggests the volatility clustering of returns. Intuitively, because there are potentially two equilibrium with two ρs, the switching from equilibrium can induce the volatility clustering of returns. The persistence structure of squared returns decreases as the magnitude of β increases as evidenced by the decrease in the reported sum of the autocorrelation coefficients of first 10 lags of squared returns.

Define $\Delta_t = \widehat{V}_t^1 - V_t^1$. Table I reports the effects of the predictable pattern of the noise traders behavior. β become more negative implying that the noise traders trade more and more aggressively as trend-followers.

The negative correlation between Δ_t and V_t is interesting. It implies that when there is a positive shock in the fundamental value of the underlying asset, the expectation is biased downward. Therefore, the partially informed trader takes a fundamental risk. When noise traders trade more and more aggressively, the partially informed trader takes more and more fundamental risk. The seventh column shows that given the noise traders are trend-followers, they are right about price direction. It is may be due to that the fully informed trader may find out that it is not that profitable to trade with noise traders and he chooses to trade with the partially informed traders using his own informational advantage. Notice that in our setting, it may be less risky to trade

34

with the partially informed traders than with noise traders. Although there is predictable pattern in noise traders' behavior, it becomes increasingly risky when noise traders trade aggressively. Hence, the fully informed trader instead chooses to exploit his informational advantage over the partially informed traders.

The negative correlation between expectation errors and supply shocks is also consistent with our intuition. Intuitively, if there is a positive shock in liquidity supply, the price will go down. The partially informed trader cannot distinguish whether the drop in price is due to a negative shock in fundamentals or a positive shock in supply. Hence, he adjusts his expectation downward.

The deviation of price from the fundamentals is increasing in β. When the magnitude of β decreases, i.e, traders trade less aggressively, ρ increases. When ρ approaches 1, there is a sharp increase in the deviation corresponding to a 98% increase in mean of price. It seems that as noise traders trade less aggressively, the price becomes more persistent and the a bubble, i.e., the deviation of market price from the fundamental value of the underlying asset, is generated.

B.   Impacts of the Number of Hierarchies

We choose $\beta = -20$ for all values of N. Table II reports the effects of the number of information hierarchies. First thing to note is that the corresponding AR coefficients of price monotonically increase when N increases, implying that price becomes more persistent. Intuitively, if there are more information hierarchies, the expectation errors from the signal extraction behavior of all partially informed traders make the price more persistent. The weight of the past price in determining the current price increases as more partially informed traders try to do signal extraction. Thus, the current price is more correlated with the past price, i.e., the AR coefficient for price increases.

The autocorrelation coefficient of return decreases as the number of the information hierarchies increases as we indicated earlier. To be consistent with the empirical finding, the number of information hierarchies should be greater than 9, which generates a 0.1 to 0.2 autocorrelation coefficient. ARCH effect exists in the returns as well. And the persistence structure of squared returns increases as the number of information hierarchies increase.

As shown in the last column of Table II, the slower information diffusion rate is captured by the number of time periods it takes for the price to converge back to the long run equilibrium price. To do this, we start with a positive supply shock with a magnitude 0.1 and set all V_ts constant. Then we examine the difference between the current price and the fundamental value of the underlying asset. If the difference is sufficiently small, the price converges to the fundamental value.[11] Then we report the number of periods for the price to converge. We can see clearly that as the number of information hierarchies increases, it takes a longer time for price to converge back to the long run fundamental value of the underlying asset. This provides simulation evidence of one of the main results of the paper: As the number of information hierarchies increases, the information diffusion is slower. A similar experiment is studied for the shock in the fundamental value of the underlying asset, and a similar pattern is found.

V.   Conclusion

In this paper, we develop a discrete time, infinite time period model to understand the relationship between hierarchical information and price discovery. The partially informed traders trade a stock based on their signals and expectation of the signals received by fully informed traders. The expectation errors from the signal extraction behavior of the partially informed traders is preserved in the market price, which prevents the information from being released at the same rate as if there was no hierarchical information among traders and the information was released instantaneously. As the number of informational hierarchies increases, it becomes harder for the partially informed traders to distinguish between the shock in fundamental value of the underlying asset and liquidity shock brought about by noise traders. The slower price discovery process of the financial market can also be partly due to the fact that fully informed traders may prefer to trade with partially informed traders. In order to make a profit from the partially informed traders, the fully informed trader should prevent the information disclosure and takes the opposite position of the partially informed traders if they make large expectation errors. Therefore, this informational arbitrage leads

_____
[11]In the simulation, we use 10^{-4} as a threshold level.

36

to a slower information diffusion rate.

In addition, noise traders' behavior is modeled with a predictable behavior pattern. This framework yields a number of interesting findings. For instance, the predictable pattern of noise traders brings the persistence structure in prices regardless of the statistical properties of the fundamentals of the underlying asset. In addition, the predictable pattern of noise traders' behavior can aid in the formation of the momentum, i.e., the positive autocorrelation in returns.

One potentially interesting avenue for future work is to endogenize the evolution of the predictable pattern of noise traders' behavior. Doing so would allow us to determine the relationship between the information diffusion and liquidity of the market. For instance, this extra uncertainty of the predictable pattern in noise traders' behavior may affect the interactions among informed traders and further affect the information diffusion speed. Meanwhile, the evolution of the liquidity supply of noise traders will affect the market depth. This study may provide insights into the correlation of market depth and information diffusion speed.

β	ρ	$\mathrm{Corr}[r_t, r_{t-1}]$	ARCH	$\sum \left|ACF(r_t^2)\right|$	$\mathrm{Corr}[V_t, \Delta_t]$	$\mathrm{Corr}[\Theta_t, \Delta_t]$	Bubble
-11	0.995	0.753	0	1.879	-0.460	-0.508	0.670
-15	0.960	0.736	0	1.092	-0.541	-0.331	0.025
-20	0.922	0.723	0	0.964	-0.586	-0.213	0.010
-25	0.880	0.697	0	0.867	-0.607	-0.130	0.006
-30	0.833	0.657	0	0.775	-0.618	-0.068	0.003
-35	0.776	0.585	0	0.587	-0.624	-0.018	-0.001
-40	0.701	0.536	0	0.640	-0.626	-0.028	0.013
-45	0.594	0.488	0	0.506	-0.630	-0.069	0.003

Table I.: EFFECTS OF THE PREDICTABLE PATTERN OF THE NOISE TRADERS' BEHAVIOR. The first column reports the corresponding AR coefficient of the price series, ρ. The second column reports the first lag serial correlation in returns. The third column reports the P-value of the ARCH test with the H_0: No ARCH effects in returns for the first 10 lags. The fourth column reports the sum of the magnitudes of the autocorrelation function (ACF) coefficients of the squared returns for the first 10 lags. The fifth column reports the correlation between fundamental value and the expectation errors of the partially informed trader. The sixth column reports the correlation between liquidity shocks and expectation errors from the partial informed trader. The seventh column reports the deviation of price from the fundamental value discounted using a risk free rate and corrected for liquidity risk. The fundamental value of the asset is defined as $\frac{\lambda a V_t}{1+r-(1-\lambda)a}$. The equilibrium price without liquidity shocks is used for comparison with the purpose of isolating the effects of β. Formally, $P_t^{comparison} = \rho P_{t-1} - \frac{a\lambda}{\frac{\beta}{\rho\Omega}+a(1-\lambda)} V_t + P_\Delta(\widehat{V}_t^1 - V_t^1)$.

N	ρ	$\mathrm{Corr}[r_t, r_{t-1}]$	ARCH	$\sum \left|ACF(r_t^2)\right|$	Adjustment Periods
5	0.859	0.692	0	0.797	7
6	0.895	0.490	0	0.855	11
7	0.920	0.301	0	1.282	18
9	0.939	0.208	0	1.578	29
10	0.952	0.154	0	1.848	32
14	0.987	0.113	0	2.978	36
15	0.992	0.109	0	2.156	39

Table II.: EFFECTS OF THE INCREASING NUMBER OF HIERARCHICAL INFORMATION LEVELS. $\beta = -20$ for all the cases. The first column reports the corresponding AR coefficient of the price series, ρ. The second column reports the first lag serial correlation in returns. The third column reports the P-value of the ARCH test with the H_0: No ARCH effects in returns for the first 10 lags. The fourth column reports the sum of the magnitudes of the autocorrelation function (ACF) coefficients of the squared returns for the first 10 lags. The fifth column reports the number of time periods for the price to converge back to the stationary equilibrium price level given a positive shock in supply of noise traders.
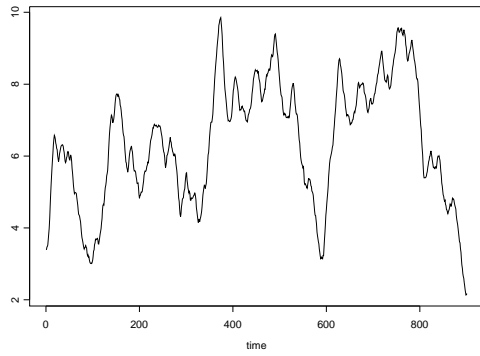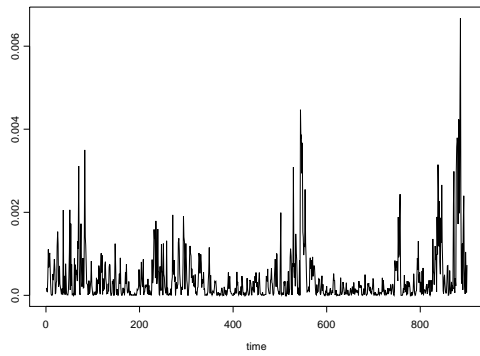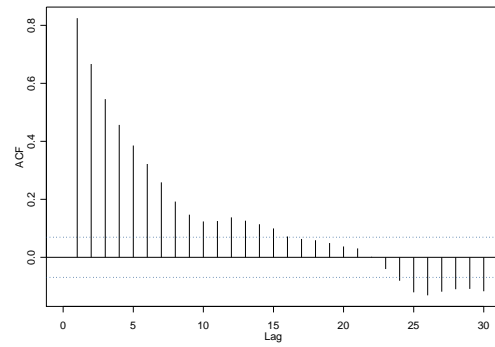
(a) Simulated prices
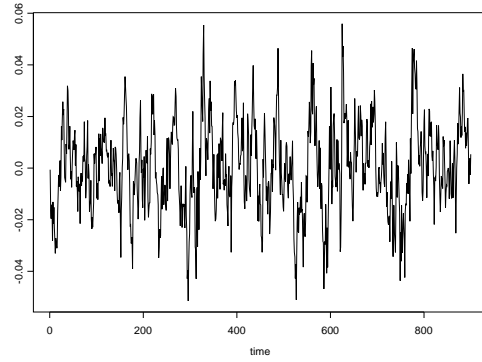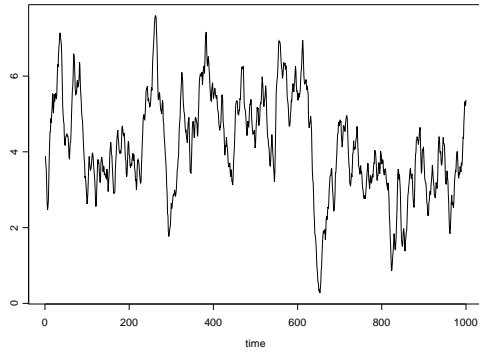
(b) Simulated returns

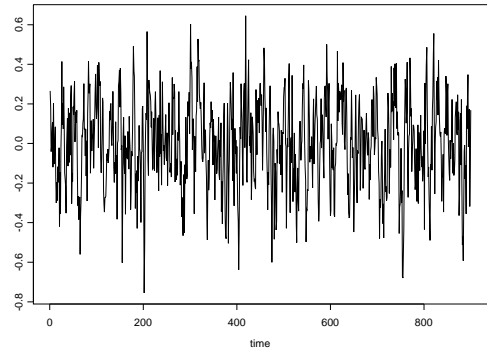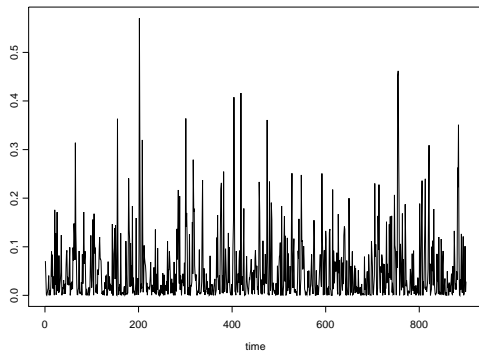(c) Simulated variance of returns

(d) ACF of returns

Figure 1: SIMULATION RESULTS IN TWO INFORMATION HIERARCHIES CASE. (a) Time series of simulated prices. (b) Time series of simulated returns. (c) Time series of simulated variance of returns. (d) Average autocorrelations of returns across 100 simulations.
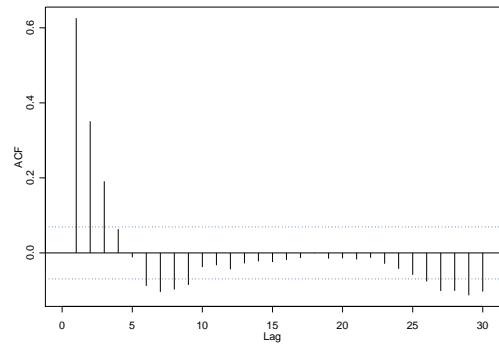
(a) Simulated prices

(b) Simulated returns

(c) Simulated variance of returns

(d) ACF of returns

Figure 2: SIMULATION RESULTS IN THREE INFORMATION HIERARCHIES CASE. (a) Time series of simulated prices. (b) Time series of simulated returns. (c) Time series of simulated variance of returns. (d) Average autocorrelations of returns across 100 simulations.

Appendix 2.A: Proofs:

Proof to Proposition 1

It is straight forward to check that the proposition holds. Then we leave out for readers' exercise.

<div align="right">Q.E.D</div>

Proof to Proposition 2

If hierarchical information structure is assumed, the infinite regress problem collapses. The iterated expectations is reduced to $\widehat{V}_t^1 = E[V_t^1 | F_{2,t}]$. We restrict ourself to the linear rational expectation equilibrium, where the price is a linear function takes the form of

$$P_t = \rho P_{t-1} + P_V V_t + P_\Theta \Theta_t + P_\Delta (\widehat{V}_t^1 - V_t^1) \tag{30}$$

Given Equation 30, the demand for type 1 trader, the fully informed trader and the demand of type 2 trader, the partially informed trader are

$$X_t^1 = \omega_1 E[Q_{t+1} | F_{1,t}]$$

$$X_t^2 = \omega_2 E[Q_{t+1} | F_{2,t}] \tag{31}$$

Using Equation 30, we have

$$
\begin{aligned}
Q_{t+1} &= \lambda V_{t+1} + (1-\lambda) P_{t+1} - (1+r) P_t \\
&= \lambda V_{t+1} + (1-\lambda)(\rho P_t + P_V V_{t+1} + P_\Theta \Theta_{t+1} + P_\Delta (\widehat{V}_{t+1}^1 - V_{t+1}^1)) - (1+r) P_t \\
&= (1-\lambda)\rho P_t + (\lambda + (1-\lambda)P_V)(aV_t + b_V \epsilon_t) + (1-\lambda)P_\Theta \Theta_{t+1} \\
&\quad + (1-\lambda)P_\Delta (\widehat{V}_{t+1}^1 - V_{t+1}^1) - (1+r) P_t
\end{aligned} \tag{32}
$$

Hence, we have

$$
\begin{aligned}
E[Q_{t+1} | F_{2,t}] &= ((1-\lambda)\rho - (1+r)) P_t + [\lambda + (1-\lambda)P_V] a [\widehat{V}_t^1 + V_t^2] + (1-\lambda)P_\Delta (\widehat{V}_{t+1}^1 - \widehat{V}_{t+1}^1) \\
&= ((1-\lambda)\rho - (1+r)) P_t + (\lambda + (1-\lambda)P_V) a (\widehat{V}_t^1 + V_t^2) \\
&= ((1-\lambda)\rho - (1+r)) P_t + (\lambda + (1-\lambda)P_V) a V_t + (\lambda + (1-\lambda)P_V) a (\widehat{V}_t^1 - V_t^1) \\
E[Q_{t+1} | F_{1,t}] &= ((1-\lambda)\rho - (1+r)) P_t + (\lambda + (1-\lambda)P_V) a V_t + (1-\lambda)P_\Delta E[(\widehat{V}_{t+1}^1 - V_{t+1}^1) | F_{1,t}]
\end{aligned} \tag{33}
$$

Therefore, aggregate market demand is

$$
\begin{aligned}
X_t^1 + X_t^2 &= \Omega((1-\lambda)\rho - (1+r))P_t + (\lambda + (1-\lambda)P_V)aV_t) \\
&\quad + (\omega_1(1-\lambda)P_\Delta E[\widehat{V}_{t+1}^1 - V_{t+1}^1|F_{1,t}] + \omega_2(\lambda + (1-\lambda)P_V)a(\widehat{V}_t^1 - V_t^1)
\end{aligned}
\tag{34}
$$

Market clearing condition implies that

$$
\begin{aligned}
\beta P_{t-1} + \Theta_t &= \Omega((1-\lambda)\rho - (1+r))P_t + (\lambda + (1-\lambda)P_V)aV_t) \\
&\quad + (\omega_1(1-\lambda)P_\Delta E[\widehat{V}_{t+1}^1 - V_{t+1}^1|F_{1,t}] + \omega_2(\lambda + (1-\lambda)P_V)a(\widehat{V}_t^1 - V_t^1)
\end{aligned}
\tag{35}
$$

Matching coefficients with Equation 30, we have

$$
\begin{aligned}
\rho &= \frac{\beta}{\Omega((1-\lambda)\rho - (1+r))} \\
P_V &= -\frac{a\Omega(\lambda + (1-\lambda)P_V}{\Omega((1-\lambda)\rho - (1+r))} \\
P_\Theta &= \frac{1}{\Omega((1-\lambda)\rho - (1+r))}
\end{aligned}
\tag{36}
$$

Hence, we can solve all parameters except P_Δ and they are

$$
\begin{aligned}
0 &= (1-\lambda)\Omega\rho^2 - (1+r)\Omega\rho - \beta \\
P_V &= \frac{a\lambda}{-\frac{\beta}{\rho\Omega} - a(1-\lambda)} \\
P_\Theta &= \frac{\rho}{\beta}
\end{aligned}
\tag{37}
$$

Now, we need to calculate P_Δ. In order to do that, we need to model the filtering expectation problem for type 2 trader explicitly. What type 2 trader can effectively observe is $(P_V - P_\Delta)V_t^1 + P_\Theta\Theta_t$. He need to forecast V_t^1. It is essentially a filter problem. We set up the system as

$$
\begin{aligned}
V_t^1 &= aV_{t-1}^1 + b_V\epsilon_t^1 \\
y_t &= (P_V - P_\Delta)V_t^1 + P_\Theta\Theta_t
\end{aligned}
$$

42

Where ϵ_t^1 is iid normal with mean 0 and variance 1 and Θ_t is iid normal with mean 0 and variance σ_Θ^2. Applying Kalman-Bucy filter[12], we have

$$\widehat{V}_t^1 = a(1 - k(P_V - P_\Delta))\widehat{V}_{t-1}^1 + k(P_V - P_\Delta)V_t^1 + kP_\Theta\Theta_t \tag{38}$$

where k solves

$$P_\Theta^2\sigma_\Theta^2 a^2(P_V - P\Delta)k^2 + (P_\Theta^2\sigma_\Theta^2(1 - a^2) + b_V^2(P_V - P_\Delta)^2)k - b_V^2(P_V - P_\Delta) = 0$$

Let $c = 1 - k(P_V - P_\Delta)$, we can rewrite Equation 38 into

$$\widehat{V}_t^1 - V_t^1 = ac(\widehat{V}_{t-1}^1 - V_{t-1}^1) - b_V c\epsilon_t^1 + kP_\Theta\Theta_t \tag{39}$$

If we impose Equation 39 on the market clearing condition, we have

$$P_\Delta \quad = \quad \frac{\omega_1 ac(1 - \lambda)P_\Delta + \omega_2 a(\lambda + (1 - \lambda)P_V)}{\Omega((1 - \lambda)\rho - (1 + r)} \tag{40}$$

Hence, we can solve for P_Δ and c simultaneously

$$P_\Delta \quad = \quad \frac{a\omega_2(\lambda + (1 - \lambda)\frac{a\lambda}{-\frac{\beta}{\rho\Omega} - a(1+\lambda)})}{-\frac{\beta}{\rho} - a\omega_1(1 - \lambda)c}, \tag{41}$$

where c solves

$$\frac{\rho^2}{\beta^2}\sigma_\Theta^2(1 - c)(1 - a^2 c) - cb_V^2(P_V - P_\Delta)^2 = 0$$

Q.E.D

Proof to Proposition 6

We only prove it when $N = 4$. The proof for $N > 4$ is similar. First, we have two partially informed traders, type 2 trader who knows V_t^2, V_t^3, and type 3 trader who only knows V_t^3. Notice that there is only one instrument for them to filter the useful information for them, i.e., the market price of the underlying asset. We conjecture the price to take form of:

$$P_t \quad = \quad \rho P_{t-1} + P_V V_t + \frac{\rho}{\beta}\Theta_t + P_{\Delta^1}(\widehat{V}_{2,t}^1 - V_t^1) + P_{\Delta^2}(\widehat{V}_{3,t}^1 + \widehat{V}_{3,t}^2 - V_t^1 - V_t^2),$$

---

[12]See Jazwinski (1970) for a description.

where $\widehat{V}_{2,t}^1 = E[V_t^1|F_{2,t}]$, $\widehat{V}_{3,t}^1 = E[V_t^1|F_{3,t}]$, $\widehat{V}_{3,t}^1 = E[V_t^2|F_{3,t}]$.

Given Equation 42, the demands from all traders are:

$$X_t^1 = \omega_1 E[Q_{t+1}|F_{1,t}]$$

$$X_t^2 = \omega_2 E[Q_{t+1}|F_{2,t}]$$

$$X_t^3 = \omega_3 E[Q_{t+1}|F_{3,t}]$$

Using Equation 42, we have

$$
\begin{aligned}
Q_{t+1} &= \lambda V_{t+1} + (1-\lambda)P_{t+1} - (1+r)P_t \\
&= \lambda V_{t+1} + (1-\lambda)(\rho P_t + P_V V_{t+1} + P_\Theta \Theta_{t+1} + P_{\Delta^1}(\widehat{V}_{2,t+1}^1 - V_{t+1}^1) \\
&\quad + P_{\Delta^2}(\widehat{V}_{3,t+1}^1 + \widehat{V}_{3,t+1}^2 - V_{t+1}^1 - V_{t+1}^2)) - (1+r)P_t \\
&= ((1-\lambda)\rho - (1+r))P_t + (\lambda + (1-\lambda)P_V)(aV_t + b_V \epsilon_t) + (1-\lambda)P_\Theta \Theta_{t+1} \\
&\quad + (1-\lambda)(P_{\Delta^1}(\widehat{V}_{2,t+1}^1 - V_{t+1}^1) + P_{\Delta^2}(\widehat{V}_{3,t+1}^1 + \widehat{V}_{3,t+1}^2 - V_{t+1}^1 - V_{t+1}^2)
\end{aligned}
$$

(43)

Hence, we have:

$$
\begin{aligned}
E[Q_{t+1}|F_{3,t}] &= ((1-\lambda)\rho - (1+r))P_t + (\lambda + (1-\lambda)P_V)a(\widehat{V}_{3,t}^1 + \widehat{V}_{3,t}^2 + V_t^3) \\
&\quad + (1-\lambda)P_{\Delta^1}(E[\widehat{V}_{2,t+1}^1|F_{3,t}] - \widehat{V}_{3,t+1}^1) \\
&= ((1-\lambda)\rho - (1+r))P_t + (\lambda + (1-\lambda)P_V)a(\widehat{V}_{3,t}^1 + \widehat{V}_{3,t}^2 + V_t^3) \\
&= ((1-\lambda)\rho - (1+r))P_t + (\lambda + (1-\lambda)P_V)aV_t + \\
&\quad (\lambda + (1-\lambda)P_V)a(\widehat{V}_{3,t+1}^1 + \widehat{V}_{3,t+1}^2 - V_{t+1}^1 - V_{t+1}^2) \\
E[Q_{t+1}|F_{2,t}] &= ((1-\lambda)\rho - (1+r))P_t + (\lambda + (1-\lambda)P_V)a(\widehat{V}_t^1 + V_t^2 + V_t^3) \\
&\quad + (1-\lambda)P_{\Delta^2}E[(\widehat{V}_{3,t+1}^1 + \widehat{V}_{3,t+1}^2 - V_{t+1}^1 - V_{t+1}^2|F_{2,t})] \\
&= ((1-\lambda)\rho - (1+r))P_t + (\lambda + (1-\lambda)P_V)aV_t + (\lambda + (1-\lambda)P_V)a(\widehat{V}_{2,t}^1 - V_{2,t}^1)
\end{aligned}
$$

$$+(1-\lambda)P_{\Delta^2}E[(\widehat{V}^1_{3,t+1}+\widehat{V}^2_{3,t+1}-V^1_{t+1}-V^2_{t+1}|F_{2,t}]$$

$$-(1-\lambda)P_{\Delta^2}E[\widehat{V}_{2,t+1}-V^1_{t+1}|F_{2,t}]$$

$$E[Q_{t+1}|F_{1,t}] \;=\; ((1-\lambda)\rho-(1+r))P_t+(\lambda+(1-\lambda)P_V)aV_t+(1-\lambda)P_{\Delta^1}E[(\widehat{V}^1_{2,t+1}-V^1_{t+1})|F_{1,t}]$$

$$+(1-\lambda)P_{\Delta^2}E[(\widehat{V}^1_{3,t+1}+\widehat{V}^2_{3,t+1}-V^1_{t+1}-V^2_{t+1})|F_{1,t}] \tag{44}$$

Imposing market clearing condition and match the coefficients with Equation 42, we have:

$$\rho \;=\; \frac{\beta}{\Omega((1-\lambda)\rho-(1+r))}$$

$$P_V \;=\; -\frac{a\Omega(\lambda+(1-\lambda)P_V}{\Omega((1-\lambda)\rho-(1+r))}$$

$$P_\Theta \;=\; \frac{1}{\Omega((1-\lambda)\rho-(1+r))}$$

$$\tag{45}$$

Hence, we can solve all parameters except P_{Δ^1} and P_{Δ^2} and they are:

$$0 \;=\; (1-\lambda)\Omega\rho^2-(1+r)\Omega\rho-\beta$$

$$P_V \;=\; \frac{a\lambda}{-\frac{\beta}{\rho\Omega}-a(1-\lambda)}$$

$$P_\Theta \;=\; \frac{\rho}{\beta}$$

$$\tag{46}$$

Now, we need to calculate P_{Δ^1} and P_{Δ^2}. In order to do that, we need to model the filtering expectation problem for type 2 trader and type 3 trader explicitly. What type 2 trader can effectively observe is $(P_V-P_{\Delta^1}-P_{\Delta^2})V^1_t+P_\Theta\Theta_t$. And type 3 trader can effectively observe $(P_V-P_{\Delta^2})(V^1_t+V^2_t)+P_{\Delta^1}(\widehat{V}^1_{2,t})+P_\Theta\Theta_t$. Type 2 trader need to forecast V^1_t while type 3 trader need to forecast V^1_t, V^2_t and $\widehat{V}^1_{2,t}$. Notice that in hierarchical information structure, $F_{3,t}\subset F_{2,t}$. Therefore we have $E[E[V^1_t|F_{2,t}]|F_{3,t}]=E[V^1_t|F_{3,t}]=\widehat{V}^1_{3,t}$ which follows directly from law of iterated expectation.

Next, let us write the filter problem for type 2 trader first:

$$V_t^1 = aV_{t-1}^1 + b_V \epsilon_t^1$$

$$y_t = (P_V - P_{\Delta^1} - P_{\Delta^2})V_t^1 + P_\Theta \Theta_t$$

where ϵ_t^1 is iid normal with mean 0 and variance 1 and Θ_t is iid normal with mean 0 and variance σ_Θ^2. Applying Kalman-Bucy filter, we have:

$$\widehat{V}_t^1 = a(1 - k(P_V - P_\Delta))\widehat{V}_{t-1}^1 + k(P_V - P_\Delta)V_t^1 + kP_\Theta\Theta_t \tag{47}$$

where k solves

$$P_\Theta^2\sigma_\Theta^2 a^2(P_V - P_\Delta)k^2 + (P_\Theta^2\sigma_\Theta^2(1 - a^2) + b_V^2(P_V - P_\Delta)^2)k - b_V^2(P_V - P_\Delta) = 0$$

where $P_\Delta = P_{\Delta^1} + P_{\Delta^2}$.

Let $c = 1 - k(P_V - P_\Delta)$, we can rewrite Equation 47 into

$$\widehat{V}_t^1 - V_t^1 = ac(\widehat{V}_{t-1}^1 - V_{t-1}^1) - b_V c\epsilon_t^1 + kP_\Theta\Theta_t \tag{48}$$

We continue to investigate the filter problem faced by type 3 trader:

$$\begin{bmatrix} V_t^1 + V_t^2 \\ \widehat{V}_{2,t}^1 - V_t^1 \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & ac \end{bmatrix} \begin{bmatrix} V_{t-1}^1 + V_{t-1}^2 \\ \widehat{V}_{2,t-1}^1 - V_{t-1}^1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} b_V \epsilon_t \\ -b_V c\epsilon_t^1 + kP_\Theta\Theta_t \end{bmatrix} \tag{49}$$

The second line of Equation 49 means that type 3 trader do not make systematic expectation errors. It implies that in equilibrium, type 3 trader can forecast the right parameter values which determine the evolution process of the expectation errors for type 2 trader but not the errors themselves. If we apply Kalman-Bucy filter again, we have:

$$\begin{bmatrix} \widehat{V}_{3,t}^1 + \widehat{V}_{3,t}^2 \\ \widehat{V}_{3,t}^1 - \widehat{V}_{3,t}^1 \end{bmatrix} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} k_1(P_V - P_{\Delta^2}) & k_1P_{\Delta^1} \\ k_2(P_V - P_{\Delta^2}) & k_2P_{\Delta^1} \end{bmatrix} \right) \begin{bmatrix} a & 0 \\ 0 & ac \end{bmatrix} \begin{bmatrix} \widehat{V}_{3,t-1}^1 + \widehat{V}_{3,t-1}^2 \\ \widehat{V}_{3,t-1}^1 - \widehat{V}_{3,t-1}^1 \end{bmatrix}$$

46

$$+ \begin{bmatrix} k_1((P_V - P_{\Delta^2})(V_t^1 + V_t^2) + P_{\Delta^1}(\widehat{V}_{2,t}^1) + P_\Theta \Theta_t) \\ k_2((P_V - P_{\Delta^2})(V_t^1 + V_t^2) + P_{\Delta^1}(\widehat{V}_{2,t}^1) + P_\Theta \Theta_t) \end{bmatrix}, \tag{50}$$

where k_1 and k_2 are parameters describe the weights used in filtering. Notice that it is really complicated to solve the parameters matrix directly. But we do not need to solve for the parameters value for filtering. We only need to describe the evolution of expectation errors of type 3 trader. Notice that $\widehat{V}_{3,t}^1 - \widehat{V}_{3,t}^1 = 0$ by law of iterated expectation. Equation 50 can be reduced to:

$$\begin{bmatrix} \widehat{V}_{3,t}^1 + \widehat{V}_{3,t}^2 \\ 0 \end{bmatrix} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} k_1(P_V - P_{\Delta^2}) & k_1 P_{\Delta^1} \\ k_2(P_V - P_{\Delta^2}) & k_2 P_{\Delta^1} \end{bmatrix} \right) \begin{bmatrix} a & 0 \\ 0 & ac \end{bmatrix} \begin{bmatrix} \widehat{V}_{3,t-1}^1 + \widehat{V}_{3,t-1}^2 \\ 0 \end{bmatrix}$$

$$+ \begin{bmatrix} k_1((P_V - P_{\Delta^2})(V_t^1 + V_t^2) + P_{\Delta^1}(\widehat{V}_{2,t}^1) + P_\Theta \Theta_t) \\ k_2((P_V - P_{\Delta^2})(V_t^1 + V_t^2) + P_{\Delta^1}(\widehat{V}_{2,t}^1) + P_\Theta \Theta_t) \end{bmatrix}, \tag{51}$$

Explicit calculation shows:

$$\begin{aligned} \widehat{V}_{3,t}^1 + \widehat{V}_{3,t}^2 - V_t^1 - V_t^2 &= a(1 - k_1(P_V - P_{\Delta^2}))(\widehat{V}_{3,t-1}^1 + \widehat{V}_{3,t-1}^2 - V_{t-1}^1 - V_{t-1}^2) + k_1 P_\Theta \Theta_t \\ &\quad + k_1 P_{\Delta^1}(\widehat{V}_{2,t}^1 - V_t^1) + (1 - k_1(P_V - P_{\Delta^2}))b_V(\epsilon_t^1 + \epsilon_t^2) \\ P_{\Delta^1}(V_{2,t}^1 - V_t^1) &= a(P_V - P_{\Delta^2})(\widehat{V}_{3,t-1}^1 + \widehat{V}_{3,t-1}^2 - V_{t-1}^1 - V_{t-1}^2) \\ &\quad - b_V(P_V - P_{\Delta^2}(\epsilon_t^1 + \epsilon_t^2) - P_\Theta \Theta_t) \tag{52} \end{aligned}$$

Substitute the second line of Equation 52 into the first line of Equation 52, we have

$$(\widehat{V}_{3,t}^1 + \widehat{V}_{3,t}^2 - V_t^1 - V_t^2) = a(\widehat{V}_{3,t-1}^1 + \widehat{V}_{3,t-1}^2 - V_{t-1}^1 - V_{t-1}^2) - b_V(\epsilon_t^1 + \epsilon_t^2)$$

$$\tag{53}$$

Equation 48 and Equation 53 describe the evolution of expectation errors of type 2 trader and type 3 trader respectively. Using them with market clearing conditions, we have:

$$P_{\Delta^2} = \frac{a\omega_3(\lambda + (1 - \lambda)P_V)}{-\frac{\beta}{\rho} - a(\omega_1 + \omega_2)(1 - \lambda)}$$

$$P_{\Delta^1} \quad = \quad \frac{a\omega_2(\lambda + (1-\lambda)P_V)) - ac(1-\lambda)\omega_2 P_{\Delta^2}}{-\frac{\beta}{\rho} - ac(1-\lambda)\omega_1}$$

where c solves

$$\frac{\rho^2}{\beta^2}\sigma_\Theta^2(1-c)(1-a^2c) - cb_V^2\left(\frac{a\lambda}{-\frac{\beta}{\rho\Omega} - a(1+\lambda)} - P_{\Delta^1} - P_{\Delta^2}\right)^2 = 0$$

<div align="right">Q.E.D</div>

Proof to Proposition 8

From Proposition 2, $P_\Theta = \frac{\rho}{\beta}$. Hence, if $\beta < 0$ and $\rho > 0$, $P_\Theta < 0$. $P_\Theta < 0$ suggests that if there is a positive shock in noise traders' supply, price should decrease in response.

Then we only need to check the sign of P_Δ.

First, we prove that if $P_\Delta < 0$, then $P_V < 0$. Suppose not, $P_V > 0$.

From Proposition 2, $P_V = \frac{a\lambda}{-\frac{\beta}{\rho\Omega} - a(1-\lambda)}$, the sign of P_V is determined by the sign of $-\frac{\beta}{\rho\Omega} - a(1-\lambda)$. To see this, if $P_V > 0$, that is, $-\frac{\beta}{\rho\Omega} - a(1-\lambda) > 0$. We must have $\Omega(-\frac{\beta}{\rho\Omega} - a\frac{\omega_1}{\Omega}(1-\lambda)c) > 0$ and $a\omega_2(\lambda + (1-\lambda)P_V) > 0$. From Proposition 2, P_Δ is

$$\begin{aligned}
P_\Delta \quad &= \quad \frac{a\omega_2(\lambda + (1-\lambda)P_V)}{-\frac{\beta}{\rho} - a\omega_1(1-\lambda)c} \\
&= \quad \frac{a\omega_2(\lambda + (1-\lambda)P_V)}{\Omega[-\frac{\beta}{\rho\Omega} - a\frac{\omega_1}{\Omega}(1-\lambda)c]}
\end{aligned} \tag{54}$$

Recall $\Omega = \omega_1 + \omega_2$ and both ω_1 and ω_2 are positive. Hence, $\frac{\omega_1}{\Omega} < 1$. Therefore, both numerator and denominator are positive. Hence, $P_\Delta > 0$. That is, to have a $P_\Delta < 0$, P_V should be negative.

If P_V is negative, i.e., $-\frac{\beta}{\rho\Omega} - a(1-\lambda) < 0$.

There are two possible scenarios: first, the numerator in Equation 54 is negative, i.e., if $P_V < -\frac{\lambda}{1-\lambda}$. The denominator of Equation 54, $(-\frac{\beta}{\rho\Omega} - a\frac{\omega_1}{\Omega}(1-\lambda)c)$, should be positive in order for $P_\Delta < 0$. That is, $a\frac{\omega_1}{\Omega}(1-\lambda)c < -\frac{\beta}{\rho\Omega} < a(1-\lambda)$. Notice that this requires that $\frac{\omega_1}{\Omega}c < 1$. Because $\frac{\omega_1}{\Omega} < 1$, it is sufficient to have $c < 1$. From the proof to Proposition 2, we know that $0 < k < 1$ can guarantee $0 < c < 1$.

Second, the numerator in Equation 54 is positive, i.e., if $0 > P_V > -\frac{\lambda}{1-\lambda}$. However, it is not possible. To see that,

$$
\begin{aligned}
P_V + \frac{\lambda}{1-\lambda} &= \frac{a\lambda(1-\lambda) + \lambda[-\frac{\beta}{\rho\Omega} - a(1-\lambda)]}{[-\frac{\beta}{\rho\Omega} - a(1-\lambda)][1-\lambda]} \\
&= \frac{-\lambda\frac{\beta}{\rho\Omega}}{[-\frac{\beta}{\rho\Omega} - a(1-\lambda)][1-\lambda]}
\end{aligned}
\tag{55}
$$

We know that $\beta < 0$, $\rho > 0$, $\Omega > 0$ and $\lambda > 0$. Hence, the numerator of Equation 55 is positive. For the denominator, $[-\frac{\beta}{\rho\Omega} - a(1-\lambda)] < 0$ because of $P_V < 0$ and $[1-\lambda] > 0$. Thus, the denominator of Equation 55 is negative. That is, $P_V + \frac{\lambda}{1-\lambda} < 0$. In other word, $(\lambda + (1-\lambda)P_V) < 0$. That is, only the first scenario is possible.

Q.E.D

Appendix 2.B

Derivation of Equation 5

In order to get Equation 5, we need to rewrite Equation 4 into

$$
P_t - \rho P_{t-1} = -\frac{\Theta_t}{\omega(1+r)}\Delta + \frac{\Delta}{1+r}E[\lambda V_{t+1} + (1-\lambda)(P_{t+1} - \rho P_t)]
$$

where Δ is a constant. Hence,

$$
(1 + \frac{\Delta\rho(1-\lambda)}{1+r})P_t = \rho P_{t-1} - \frac{\Theta_t}{\omega(1+r)}\Delta + \frac{\Delta}{1+r}E[\lambda V_{t+1} + (1-\lambda)P_{t+1}]
$$

In order to match the coefficients of Equation 4, we need to have:

$$
\begin{aligned}
\frac{\rho}{\Delta} &= -\frac{\beta}{\omega(1+r)} \\
\Delta &= 1 + \frac{\Delta\rho(1-\lambda)}{1+r}
\end{aligned}
$$

The first line implies that $\Delta = -\frac{\beta}{\rho\omega(1+r)}$ and substitute it into second line. Then we have:

$$(1 - \lambda)\rho^2 + (1+r)\omega\rho + \beta = 0.$$

References

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

Grossman, S. and Stiglitz, J. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, **70**, 1393–408.

Harris, L. (2003). *Trading and exchange*. Oxford University Press.

Jazwinski, A. (1970). *Stochastic Processes and Filtering Theory*. Academic Press, New York.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, **48**, 65–91.

Kyle, A. S. (1985). Continous auctions and insider trading. *Econometrica*, **53**, 1315–1335.

Makarov, I. and Rytchkov, O. (2007). Forecasting the forecasts of others: implication for asset pricing. Working Paper.