Correlation Neglect and Financial Markets

Liying Wang[*]

Abstract

In this paper, I examine the impact of "correlation neglect" in a financial market, where naive traders neglect the correlation between signal errors. I develop a model including both naiveand rational traders. I find that the impact of naive traders on market quality, measured by liquidity and mispricing risk, depends on whether information is costly or not. If information acquisition is free of charge and correlation between signal errors is relatively low, mispricing risk decreases in the mass of naive traders; but when correlation is large enough, mispricing risk is U-shaped. Conversely, when information acquisition is costly, market liquidity deteriorates and mispricing risk increases in the mass of naive traders given that their mass is not too large to drive all informed rational traders out of the market; but market quality can improve afterwards after informed rational traders are entirely crowded out, depending on the correlation and the mass of naive traders.

1 Introduction

Information is valuable to market participants. But excess information availability can lead to availability bias. For example, "stereotyping can develop as a result of repeated news, resulting in representation bias, which encourages overconfidence or too little questioning or analysis of the situation." (Siegel and Yacht; 2009). In fact, many information structures in financial markets generate correlated rather than mutually independent signals. As proposed by Welch (2000), the correlation between signals can be caused by the common fundamental

[*]Liying Wang (liying.wang@bayes.city.ac.uk) is with Bayes Business School, City, University of London.

information, or it can result from“direct mutual imitation”. The latter suggests that there is a correlation between the biases of signals deviated from the common fundamental.

The motivating idea of this paper is to study how financial markets are influenced when the correlation between signal errors is neglected by some market participants. There is extensive literature documenting the existence of “correlation neglect”. For example, Enke and Zimmermann (2019) provide experimental evidence that people neglect the correlation in the updating process; Jiao, Veiga and Walther (2020) find evidence that there is a subset of “naive” traders that exists, who interpret the repeated signals of social networks as genuinely new information like the news media; Tetlock (2011) also find that stock market investors can not completely distinguish between new and old information about firms. In my paper, I model agents who cannot recognise the existence of a positive correlation between signal errors as naive traders.

In financial markets, the correlated signal errors can be captured by some stylized facts: one example is the repetition of media coverage. Some naiveinvestors interpret the repeated signals as genuinely new information (Jiao et al.; 2020). Another example is analyst herding, which is defined as the forecast errors with unusually high consensus in forecasts among analysts (e.g., De Bondt and Forbes (1999); Kim and Pantzalis (2003)). The difference between these two examples is whether information acquisition is costly. In the first example, the cost to learn from media, no matter whether news media or social media, is negligible, and even retail investors can freely acquire this kind of information. In the second case, the analyst service is costly, and this kind of information is always sold to financial institutions, who are relatively sophisticated.

I develop a theoretical model to conceptualise both the examples mentioned above and provide some insights on the economic impact. I build the model on the CARA-normal REE framework (e.g., Grossman and Stiglitz (1980); Admati and Pfleiderer (1986)), and I extend the classical models by introducing two types of traders, rational and naive, under the framework of costly and costless information acquisition. The correlation between signal errors measures the degree of information herding or repetition. Furthermore, another essential assumption emphasizes that, naive traders are unaware of the existence of rational traders due to “correlation neglect”, and rational traders are aware of the existence of their naivecounterparts. Some literature also assume the disagreement or uncertainty about the composition of market participants (e.g., Gao, Song and Wang (2013); Banerjee and Green (2015); Papadimitriou (2020)). Similarly, this paper assumes the market composition perceived by naiveand rational traders are different: the naive traders, who neglect the correlation between signal errors, unintentionally neglect the existence of rational traders, because they can not recognize there exists the other type of “smarter” traders who can better understand the information structure than them.

Based on the model, I find that the impact of “correlation neglect” on market quality depends on whether information is costly or not. If information is free of charge, mispricing risk decreases in the mass of naive traders when the correlation between signal errors is relatively low, but can be U-shaped when the correlation is relatively high. naive traders provide more liquidity than rational traders, but when there are too many naive traders in the market, their existence may amplify signal errors, and make the price too sensitive to public information, increasing the risk of mispricing and worsening market quality.

However, the story is different when information is costly. Under the consideration of information acquisition, the market quality, measured by liquidity and mispricing risk, can be worsened by naive traders even when their mass is small enough.

In this case, the information acquisition decision of naive traders is influenced by their behavioural bias, i.e., “correlation neglect”. There may be more or less of them willing to acquire information compared to the case if they were rational traders. However, no matter which case, the aggregate trading intensity of naive traders is always larger than that if they were rational traders, as the naive traders who choose to acquire costly information overestimate the precision of the information acquired and thus trade more aggressively than informed rational traders. In this way, informed naive traders, i.e. those who choose to acquire information, contribute to increasing the price informativeness about fundamentals. At the same time, as the mass of naive traders increases, the rational traders realize the price informativeness is improved and there is less profit margin for information acquisition, so less of them are willing to acquire costly information. It is highlighted that the existence of naive traders does not actually change the overall price informativeness, because rational traders always balance out the excess contribution of naive traders by reducing information acquisition.

The information acquisition model predicts that when there are rational traders who still choose to acquire information in the market, price informativeness is independent of the mass of naive traders. As the mass of naive traders increases, the total mass of informed traders decreases. This causes the aggregate response of agents to price declines, market liquidity deteriorates, and mispricing risk increases accordingly. The impact of naive traders on market quality with costly information acquisition is significantly different from the case without information cost. When the mass of naive traders is not sufficiently high, their existence tends to increase market liquidity and reduce mispricing risk if information is free of charge, but worsens market liquidity and increases mispricing risk if information is costly.

The model gives us implications regarding the empirical properties of financial markets. First, the findings in Jiao et al. (2020) suggest that the intense coverage on social media platforms such as Twitter led to "high volatility of returns and high trading volume" of stocks, because the contents of social media repeat that of news media. My results from the free information model can explain the potential mispricing risk triggered by the repetition of media coverage. Free information is available to retail investors, who are likely to lack the skills to interpret the structure of information and perceive the correlation between signal errors, so the fraction of naive traders among the retail traders may be relatively high, potentially bringing greater mispricing risk to the financial market.

Second, the past few years have seen the decline of active management: "The shift out of active and into passive has long been underway. Between 2014 and 2018, active funds had outflows of 738 billion, while passive funds saw inflows to the tune of 2.5 trillion." (CNBC, 10 Oct, 2019). This paper finds there is a crowding-out effect of "correlation neglect" on information acquisition of rational traders, which can be regarded as one potential explanation for the decline of active management. The excess information availability may amplify the cognitive limitation of some market participants, who trade much more aggressively on the costly information they acquire. Accordingly, financial markets become efficient enough to compress the profit margin of costly information acquisition under the rational perspective, making active management less attractive. Finally, this paper provides some insights into the long lasting debate on whether information efficiency is impaired by the decline of active management. As Qin and Singal (2015) state, "reduced incentives for information acquisition and arbitrage induced by indexing and passive trading" may lead to "degradation in price efficiency". As this paper predicts, the market liquidity worsens and mispricing risk increases with the decline of total informed mass when informed rational traders still exist, but price informativeness can be unacted even when the overall information acquisition declines.

The remainder of this paper is organised as follows. Section 2 reviews the relevant literature; Section 3 describes the basic framework of model with information cost; Section 4 characterises the equilibrium of the model with information acquisition when rational and naive traders coexist; Section 5 analyses the properties of market quality as well as the expected utility of traders; Section 6 studies the case without information acquisition cost; Section 7 concludes.

2 Literature Review

This paper is principally related to two strands of theoretical literature: information acquisition, and disagreement.

The first strand of literature, information acquisition, is mainly based on the rational expectation model in Grossman and Stiglitz (1980), where the fundamental issue of how costly information acquisition can be supported by the financial market is solved. The property of strategic substitution of information acquisition is robust in our model, which implies when more traders acquire information, information becomes less valuable (Verrecchia; 1982). In contrast, this paper emphasises an unilateral substitution effect of information acquisition between two types of traders, and only naive traders have a “crowding-out” effect on information acquisition of rational traders, but not vice versa. There are some papers, based on Grossman and Stiglitz (1980), investigating how information acquisition is impacted depending on the degree of “information linkage”, e.g., Goulding and Zhang (2018) find that the price is more informative when the degree of scattered information is low; Huang and Yueshen (2021) find that an improving information technology increases the mass of traders who trade faster and thus improves the efficiency of intermediate price. These papers, from different perspectives, demonstrate that more “information linkage” aggravates free riding, and may finally hurt price informativeness by discouraging information acquisition. Analogically, I find the excess contribution to price informativeness by naive traders reduces information acquisition of rational traders, but I find price informativeness can keep the constant under the interaction of the two groups of traders.

The second strand of related literature is about investor disagreement. This series of literature can explain various market anomalies that can beplausibly hard to explain by the rational model (REE), e.g., excess return volatility. Among the underlying mechanisms leading to the disagreement of investors (Hong and Stein; 2007), this paper has both the characteristics of “heterogenous beliefs” and “overconfidence”. On the one hand, “correlation neglect" leads the naive traders to overestimate the precision of aggregate signals, and overconfidence can also leadto the same outcome; on the other hand, "correlation neglect" causes the naive traders unintentionally ignore the role of rational traders in transmitting information to price, which coincides with the outcome due to "heterogenous beliefs".

In the literature of "heterogenous beliefs", the investors do not fully update their beliefs based on each other's trading decisions. For example, in Banerjee and Kremer (2010), investors disagree about the interpretations of public information and thus neglect others' interpretation; Eyster, Rabin and Vayanos (2019) considers the traders entirely or partially neglect the relationship between the price and other traders' information. Similarly, the naive traders in this paper do not update their belief and make the decision in response to the rational traders. But there are two main differences between this paper and the literature of "heterogenous beliefs": first, the naive traders make decisions without considering rational traders, not because they "agree to disagree" at equilibrium, but because the naive traders neglect the existence of rational traders; second, the naive traders still extract information and learn from price, but in a biased way due to cognitive limitation.

This paper is also strongly related to the literature on overconfidence, where "overconfidence" is modeled as a belief that the precision of signal perceived by a trader is higher than it actually is. In my model, the naive traders overestimate the precision of signals because of "correlation neglect" instead of psychological factors. Odean (1998) finds the influence of overconfidence on price quality depends on who is overconfident, e.g., price takers, the insider, or market maker. Conversely, in a perfectly competitive asset market of my model, how market quality is influenced by naive traders depends on the information cost and the mass of naive traders. Some literature that contains the coexistence of rational and overconfident participants, e.g., Benos (1998), and Kyle and Wang (1997), find that the market depth and price informativeness increase when there are overconfident informed investors participating in the market. My model with free information draws a similar conclusion: the existence of naive traders helps improves market depth when their mass is relatively low. But as the mass of naive traders becomes large enough, more mispricing risk may be introduced.

Some literature that investigates overconfidence considers information acquisition. Odean (1998) introduces a completely competitive model with information acquisition and they find the equilibrium obtained is not influenced by the level of overconfidence. I relax the assumption of Odean (1998) that all traders are overconfident by introducing rational traders, and find the information acquisition at the aggregate level is negatively influenced by the mass of naive traders when it is not large enough. García and Sangiorgi (2011) predicts the full participation of information acquisition by overconfident investors given the existence of informed rational traders, and concludes overconfidence has no effect on market quality. Their results rely on the assumption that overconfident traders agree to disagree with rational traders. My model, however, finds that the overestimation of signal precision does not always make all naive traders acquire information. Ko and Zhijian (2007) develop a variable cost model to study information acquisition of overconfident investors, and find that overconfidence generally improves market quality under some conditions. My model also predicts market quality can be improved by naive traders, not only depending on the overall degree of precision overestimation, but also depending on the cost of information.

3 Model with Information Cost

In this section, the model has three events. At time 1, agents decide whether to acquire information. At time 2, agents observe their signals if they pay and trade in a competitive asset market. At time 3, the assets pay off, and all agents consume.

There are two assets in the financial market: one risk-free asset and one risky asset. The payoff of the risky asset is v, which is a mean-zero normal random variable with precision τ.

Information market: At the beginning of period 1, the information market opens. The information seller provides n signals, denoted by s_i, $i \in \{1, ..., n\}$, to the traders who independently decide whether to pay for the information service at a fixed cost of c. Each signal has an error term from the fundamental: $s_i = v + \epsilon_i$, where $\epsilon_i \sim N(0, \frac{1}{\tau_\epsilon})$ and $\epsilon_i \perp v$. The error terms are multivariate normal distributed with correlation $\rho \in (0, 1)$: $\text{Corr}[\epsilon_i, \epsilon_j] = \rho$ for $i \neq j$, where ρ reflects the degree of information repetition.

Financial market: Once the traders have made their information acquisition decision, period 1 ends and the financial market opens in period 2. The participants in the financial market are:

- Risk-adverse traders: In the economy, there is a unit continuum of traders, each with constant absolute risk aversion (CARA) utility with risk-tolerance parameter $\gamma > 0$.

Suppose there is a fraction $\lambda \in [0,1]$ of the traders to acquire information, and the rest $(1-\lambda)$ of them keep uninformed.

- Noise traders: The demand of noise traders, x, is a mean-zero normal random variable with precision τ_x $(x \perp v, x \perp \epsilon)$.

In the market, there exists both naive traders and rational traders. The naive traders do not know the existence of rational traders, which implies that naive traders assume all traders in the market are homogenous and have the same prior belief as themselves.

Here are some main assumptions:

1. The mass of naive traders is β, and the rest $(1-\beta)$ of traders are rational.
2. The rational traders correctly anticipate the actual correlation of signal errors, ρ, as well as the mass of naive traders, β.
3. The naive traders neglect the existence of correlation (they take ρ for 0), as well as the existence of rational traders.
4. Both naiveand rational traders independently make their information acquisition decision at time 1.

3.1 All-rational Benchmark

In the benchmark, I derive the equilibrium of the trading game when there is no naive traders in the market, namely, $\beta = 0$. As in Grossman and Stiglitz (1980), I consider the rational expectation equilibrium (REE). Suppose that there is a linear price function of the form:

$$p = \eta \left(I \sum_{i=1}^{n} s_i + x \right) = nI\eta \left(v + \frac{I \sum_{i=1}^{n} \epsilon_i + x}{nI} \right) \tag{1}$$

where η and I are endogenous coefficients at equilibrium.

The demand of informed and uninformed investors are denoted by D_{inf} and D_{uninf} respectively. The CARA-normal setup implies the informed investor's demand function is

$$D_{inf}\left(\sum_{i=1}^{n} s_i, p \right) = \frac{\gamma\big(E\left[v \mid s_1, s_2, \ldots, s_n, p\right] - p\big)}{\mathrm{Var}[v \mid s_1, s_2, \ldots, s_n, p]} \tag{2}$$

Applying Bayes' rule, the conditional moments of the fundamental can be computed from the perspective of the informed trader as follows:

$$Var[v \mid s_1, s_2, \ldots, s_n, p] = \frac{1+(n-1)\rho}{\tau + (n-1)\rho\tau + n\tau_\epsilon} \tag{3}$$

$$E[v \mid s_1, s_2, \ldots, s_n, p] = \frac{\tau_\epsilon \sum_{i=1}^{n} s_i}{\tau + (n-1)\rho\tau + n\tau_\epsilon} \tag{4}$$

If there is no correlation between signal errors, $\rho = 0$, the posterior precision of the fundamental for informed traders, $\big(Var[v \mid s_1, s_2, \ldots, s_n]\big)^{-1}$, is $\tau + n\tau_\epsilon$. Otherwise, with positive correlation, the precision of fundamental increased by n signals is:

$$D = \Big(Var(v \mid \sum_{i=1}^{n} s_i)\Big)^{-1} - \tau = \frac{n\tau_\epsilon}{\Delta} \tag{5}$$

where $\Delta^{-1} = \frac{1}{1+(n-1)\rho}$, which can be regarded as the discount factor compared to that without correlation.

The actual precision of aggregate signals does not directly equal to the sum of precision of individual signal, $n\tau_\epsilon$, but equals to the discounted sum of individual precision. The larger ρ, the more precision of aggregate signals is discounted.

The uninformed trader only observes price p, and their demand function is

$$D_{uninf}(p) = \frac{\gamma\big(E[v \mid p] - p\big)}{\mathrm{Var}[v \mid p]} \tag{6}$$

Using Bayes' rule, we have

$$\mathrm{Var}[v \mid p] = \left(\tau + \frac{1}{\frac{1}{n^2 I^2 \tau_x} + \frac{1}{D}}\right)^{-1} \tag{7}$$

$$\mathrm{E}[v \mid p] = \left(\tau + \frac{1}{\frac{1}{n^2 I^2 \tau_x} + \frac{1}{D}}\right)^{-1} \frac{1}{\frac{1}{n^2 I^2 \tau_x} + \frac{1}{D}} \frac{1}{nI\eta} p \tag{8}$$

The market clearing condition is,

$$\lambda D_{inf}\left(\sum_{i=1}^{n} s_i, p\right) + (1-\lambda) D_{uninf}(p) + x = 0 \tag{9}$$

To derive the equilibrium price function, I insert the demand functions into the market clearing condition to solve the price in terms of $\sum_{i=1}^{n} s_i$ and x, and then compare with the conjectured price function in equation (1) to obtain a system defining the unknown coefficients of I and η.

Proposition 1 (*Financial market equilibrium*) *Given* $(\rho, \lambda, n, \gamma, \tau, \tau_\epsilon, \tau_x)$, *there exists a unique linear REE, in which*

$$p = \eta \left(I \sum_{i=1}^{n} s_i + x \right)$$

where η *and* I *are given in the appendix.* $\frac{1}{\eta}$ *measures the market depth, and* I *measures aggregate trading intensity.*

The aggregate trading intensity of the informed traders increases with the fraction of informed traders $\lambda \in [0, 1]$ uniquely in equilibrium according to the following closed form function

$$I = \lambda\gamma \frac{\tau_\epsilon}{1 + (n-1)\rho} = \frac{\lambda\gamma D}{n} \tag{10}$$

Corollary 1 *Given* $(\rho, n, \gamma, \tau, \tau_\epsilon, \tau_x)$, *price informativeness* $\big(Var(v \mid p)\big)^{-1}$ *increases with the fraction* λ *of informed traders.*

To achieve the overall equilibrium, I endogenise the information acquisition process, and consider the situation where traders can decide whether to subscribe to the information service by paying a fixed cost, c. I calculate the ex ante certainty equivalent of the expected utility of trading profit for the informed traders and uniformed traders, denoted by CE_{inf} and CE_{uninf}, respectively.

The difference between CE_{inf} and CE_{uninf} measures the benefit of being informed, which is given by

$$\begin{aligned} CE_{inf} - CE_{uninf} &= \frac{\gamma}{2} \log \frac{\mathrm{Var}[v \mid p]}{\mathrm{Var}[v \mid \sum_{i=1}^{n} s_i, p]} \\ &= \frac{\gamma}{2} \log \left(\frac{\tau + D}{\tau + \frac{1}{\frac{1}{n^2 I^2 \tau_x} + \frac{1}{D}}} \right) \end{aligned} \tag{11}$$

Lemma 1 *For given* $(\rho, n, \gamma, \tau, \tau_\epsilon, \tau_x)$, $CE_{inf} - CE_{uninf}$ *is a decreasing function of* λ.

If $CE_{inf} - CE_{uninf} > c$, traders decide to acquire information; otherwise, they do not. Thus, the equilibrium mass λ is determined by

$$CE_{inf} - CE_{uninf} = c \tag{12}$$

Proposition 2 *(Overall equilibrium)For given $(\rho, c, n, \gamma, \tau, \tau_\epsilon, \tau_x)$, there exists an unique information market equilibrium in which there is $\lambda \in [0, 1]$ fraction of traders acquiring information, where*

$$\lambda = \begin{cases} 1 & \textit{if } 0 < c \leqslant \underline{c} \quad \textit{(Corner equilibrium)} \\ \hat{\lambda} = \frac{n\hat{I}}{\gamma D} & \textit{if } \underline{c} < c < \bar{c} \quad \textit{(Interior equilibrium)} \\ 0 & \textit{if } c \geqslant \bar{c} \quad \textit{(Corner equilibrium)} \end{cases} \tag{13}$$

and the aggregate trading intensity I satisfies

$$nI = \begin{cases} \gamma D & \textit{if } 0 < c \leqslant \underline{c} \\ n\hat{I} = \sqrt{\frac{-A\tau D + D^2}{A\tau_x(\tau + D)}} & \textit{if } \underline{c} < c < \bar{c} \\ 0 & \textit{if } c \geqslant \bar{c} \end{cases} \tag{14}$$

where

$$A = e^{2c/\gamma} - 1 \tag{15}$$

$$\underline{A} = e^{2\underline{c}/\gamma} - 1 = \frac{1}{\gamma^2 \tau_x D + \tau/D + \tau_x \tau \gamma^2} \tag{16}$$

$$\bar{A} = e^{2\bar{c}/\gamma} - 1 = \frac{D}{\tau} \tag{17}$$

Proposition 3 *For given $(\rho, c, n, \gamma, \tau, \tau_\epsilon, \tau_x)$, the aggregate trading intensity I is monotonically increasing in the precision D of aggregate signals $\sum_{i=1}^{n} s_i$.*

Corollary 2 *For given $(\rho, c, n, \gamma, \tau, \tau_\epsilon, \tau_x)$, the aggregate trading intensity I is strictly decreasing in the signal correlation ρ for $c < \bar{c}$, and otherwise flat at level $I = 0$.*

From equation (7), the price informativeness, denoted by PI, is the inverse of $\mathrm{Var}[v \mid p]$, which is influenced by D through two channels:

$$
\begin{aligned}
\frac{dPI}{dD} &= \frac{d\left(\mathrm{Var}[v \mid p]\right)^{-1}}{dD} = \frac{d\left(\tau + \frac{1}{\frac{1}{n^2 I^2 \tau_x} + \frac{1}{D}}\right)}{dD} \\
&= \underbrace{\frac{\partial PI}{\partial D}}_{\text{direct effect } (+)} + \underbrace{\frac{\partial PI}{\partial I}\frac{\partial I}{\partial D}}_{\text{indirect effect}(+)}
\end{aligned}
\tag{18}
$$

The aggregate signals $\sum_{i=1}^{n} s_i$ not only include the fundamental information, but also bring noise $\sum_{i=1}^{n} \epsilon_i$ into price. The direct channel implies that as D increases, the informative content about fundamental v in $\sum_{i=1}^{n} s_i$ increases relative to the content of error $\sum_{i=1}^{n} \epsilon_i$.

The indirect effect of D on price informativeness is through the trading of informed traders: from Proposition 3, we know the informed traders trade more aggressively at the aggregate level as D increases. Thus the price reflects relatively more information about $\sum_{i=1}^{n} s_i$ comparing to the noise x of liquidity trading.

Proposition 4 *For given $(\rho, c, n, \gamma, \tau, \tau_\epsilon, \tau_x)$, price informativeness is monotonically increasing in D.*

From equation (13), we know that when all the traders are rational, their information acquisition decision is influenced by the precision of aggregate signals, regardless of what specifically τ_ϵ, n and ρ represent. It is intuitive because when all traders are rational, they can accurately perceive the value of aggregate signals taking them as a whole.

Proposition 5 *For given $(c, n, \gamma, \tau, \tau_\epsilon, \tau_x)$, when $D \in \left[A\tau, A\tau\left(1 + \sqrt{1 + \frac{1}{A}}\right)\right)$, λ is monotonically increasing in D; when $D \in \left[A\tau\left(1 + \sqrt{1 + \frac{1}{A}}\right), +\infty\right)$, λ is monotonically decreasing in D.*

4 Model Equilibrium

In the benchmark, I assume that there are only rational traders in the market, who can accurately perceive the correlation between signal errors, make an information acquisition decision and trade accordingly. In this section, I characterize the equilibrium with naive traders, who neglect the correlation between signal errors and are unaware of the existence of rational traders.

4.1 Equilibrium Concept

λ_1 is the informed fraction among the rational traders; and λ_2 is the fraction of informed traders in naive group. The total mass of informed traders in the market is

$$\lambda = (1-\beta)\lambda_1 + \beta\lambda_2 \tag{19}$$

Because the naive traders do not know of the existence of rational traders, they think λ_2 is the total mass of informed traders in the market, and they take the precision of aggregate signals for $n\tau_\epsilon$.

D denotes the precision of the aggregate signals perceived by rational traders, which is also the actual precision of aggregate signals: $D = \frac{n\tau_\epsilon}{1+(n-1)\rho}$. F denotes the precision of aggregate signals perceived by naive traders, where $F = n\tau_\epsilon$.

We have $F(n, \tau_\epsilon, \rho) > D(n, \tau_\epsilon, \rho)$, which implies the naive traders overestimate the precision of aggregate signal due to "correlation neglect".

4.2 Naive Traders' Perspective

From the assumptions, the naive traders not only overestimate the precision of the aggregate information, but also mistakenly perceive the mass of informed traders as λ_2 instead λ.

Let $\mathrm{Var}_i[\cdot]$ and $\mathrm{E}_i[\cdot]$ represent the variance and the expectation of variable from the perspective of naive traders. D_{inf1}, D_{uninf1}, D_{inf2}, D_{uninf2} denote the demand of informed rational, uninformed rational, informed naive, uninformed naive traders, respectively.

We conjecture a linear price function, and linear demand schedules of uninformed traders. The linear price function is given by

$$p = \eta I \sum_{i=1}^{n} s_i + \eta x \tag{20}$$

The demand function of uninformed rational traders is:

$$D_{uninf1}(p) = b_1 p \tag{21}$$

and the demand function of uninformed naive traders is:

$$D_{uninf2}(p) = b_2 p \tag{22}$$

The informed naive traders maximise their expected utility by trading:

$$\begin{aligned} D_{inf2}\left(\sum_{i=1}^{n} s_i, p\right) &= \frac{\gamma\big(E_i[v \mid s_1, s_2 \ldots s_n, p] - p\big)}{\mathrm{Var}_i[v \mid s_1, s_2 \ldots s_n, p]} \\ &= \gamma\tau_\epsilon \sum_{i=1}^{n} s_i - \gamma(\tau + n\tau_\epsilon)p \end{aligned} \tag{23}$$

where

$$\mathrm{Var}_i[v \mid s_1, s_2 \ldots s_n, p] = \frac{1}{\tau + n\tau_\epsilon} \tag{24}$$

and

$$\mathrm{E}_i[v \mid s_1, s_2 \ldots s_n, p] = \frac{\tau_\epsilon \sum_{i=1}^{n} s_i}{\tau + n\tau_\epsilon} \tag{25}$$

The naive traders incorrectly take "market clearing condition" as the following form:

$$\lambda_2 D_{inf2}\left(\sum_{i=1}^{n} s_i, p\right) + (1-\lambda_2) D_{uninf2}(p) + x = 0 \tag{26}$$

However, the actual market price does not satisfy equation (26), because the "market clearing condition" is conjectured by naive traders, who cannot accurately recognise the composition of market participants. In other words, the uniformed naive traders incorrectly interpret the market price and extract informative content. Let ω_2 denote the informative signal perceived by uninformed naive traders:

$$\omega_2 = (\lambda_2 \gamma \tau_\epsilon) \sum_{i=1}^{n} s_i + x \tag{27}$$

which is observationally equivalent to the noisy signal $\omega_2^{'}$:

$$\omega_2^{'} = v + \frac{\lambda_2 \gamma \tau_\epsilon \sum_{i=1}^{n} \epsilon_i + x}{n \lambda_2 \gamma \tau_\epsilon} \tag{28}$$

From the perspective of naive traders, they evaluate their ex-ante certainty equivalent, denoted by CE^*_{inf2} and CE^*_{uninf2} for informed and uninformed naive traders, respectively, and make the information acquisition decision. At equilibrium, λ_2 is the fraction of informed trader among them, which reflects their willingness to acquire information.

At equilibrium, the naive traders thinks it is equivalent to be informed or keep uninformed. The equilibrium is determined by

$$CE^*_{inf2} - CE^*_{uninf2} = c \tag{29}$$

where c is the cost of information, and

$$CE^*_{inf2} - CE^*_{uninf2} = \frac{\gamma}{2}\log\frac{\mathrm{Var}_i[v \mid p]}{\mathrm{Var}_i[v \mid \sum_{i=1}^n s_i, p]} \tag{30}$$

At equilibrium, the naive traders think acquiring information is equivalent to being uninformed.

Proposition 6 *For given $(\rho, c, n, \gamma, \tau, \tau_\epsilon, \tau_x)$, there exists a unique fraction $\lambda_2 \in [0, 1]$ of informed traders among naive traders:*

$$\lambda_2 = \begin{cases} 1 & \text{if } 0 < c \leqslant \underline{c}_2 \quad \text{(Corner equilibrium)} \\ \hat{\lambda}_2 = \sqrt{\frac{-A\tau + n\tau_\epsilon}{A\tau_x\gamma^2(\tau n\tau_\epsilon + n^2\tau_\epsilon^2)}} & \text{if } \underline{c}_2 < c < \bar{c}_2 \quad \text{(Interior equilibrium)} \\ 0 & \text{if } c \geqslant \bar{c}_2 \quad \text{(Corner equilibrium)} \end{cases} \tag{31}$$

Where $A = e^{2c/\gamma} - 1$.

$$\underline{A}_2 = e^{2\underline{c}_2/\gamma} - 1 = \frac{1}{\gamma^2\tau_x n\tau_\epsilon + \tau/(n\tau_\epsilon) + \tau_x\tau\gamma^2}$$
$$\bar{A}_2 = e^{2\bar{c}_2/\gamma} - 1 = \frac{n\tau_\epsilon}{\tau}$$

4.3 Financial Market Equilibrium

The existence of naive traders does not influence the posterior belief of informed rational traders. The demand of informed rational traders is:

$$\begin{aligned} D_{inf1}\left(\sum_{i=1}^n s_i, p\right) &= \frac{\gamma\big(\mathrm{E}[v \mid s_1, s_2 ... s_n, p] - p\big)}{\mathrm{Var}[v \mid s_1, s_2 ... s_n, p]} \\ &= \frac{\gamma D}{n}\sum_{i=1}^n s_i - \gamma(\tau + D)p \end{aligned} \tag{32}$$

Since the rational traders are aware of the existence of naive traders, they correctly conjecture market clearing conditions, which actually pins down the price:

$$\begin{aligned}(1-\beta)\lambda_1 D_{inf1}\left(\sum_{i=1}^{n} s_i, p\right) + (1-\beta)(1-\lambda_1)D_{uninf1}(p) \\ +\beta\lambda_2 D_{inf2}\left(\sum_{i=1}^{n} s_i, p\right) + \beta(1-\lambda_2)D_{uninf2}(p) + x = 0\end{aligned} \tag{33}$$

The uninformed rational traders correctly extract information from price:

$$\omega_1 = \left[(1-\beta)\lambda_1\gamma\frac{D}{n} + \beta\lambda_2\gamma\tau_\epsilon\right]\sum_{i=1}^{n} s_i + x \tag{34}$$

Let

$$\hat{I} = (1-\beta)\lambda_1\gamma\frac{D}{n} + \beta\lambda_2\gamma\tau_\epsilon, \tag{35}$$

which denotes the aggregate trading intensity on $\sum_{i=1}^{n} s_i$ and reflects how aggressively informed traders trade at the aggregate level.

ω_1 is observationally equivalent to the noisy signal ω_1':

$$\omega_1' = v + \frac{\sum_{i=1}^{n}\epsilon_i}{n} + \frac{x}{n\hat{I}} \tag{36}$$

The actual price informativeness, $\mathrm{PI} = \left(\mathrm{Var}(v \mid p)\right)^{-1}$, is interpreted as the posterior precision of the fundamental v given price at equilibrium, which can be correctly perceived by uninformed rational traders.

In the financial market with asymmetric information, informative trading contributes to price informativeness. The informative signal ω_1 extracted by rational traders can be divided into three components:

$$\omega_1 = \underbrace{(1-\beta)\lambda_1\gamma\frac{D}{n}\sum_{i=1}^{n} s_i}_{\text{by informed rational traders}} + \underbrace{\beta\lambda_2\gamma\tau_\epsilon\sum_{i=1}^{n} s_i}_{\text{by informed naive traders}} + \underbrace{x}_{\text{by noisy traders}}$$

The informative content in ω_1 is contributed by (i) the informed rational traders, and (ii) the informed naive traders. We write the price informativeness in the following form:

$$\begin{aligned} \text{PI} = & \big(\text{Var}(v \mid p)\big)^{-1} = \big(\text{Var}(v \mid \omega_1)\big)^{-1} \\ & = \tau + \frac{1}{\frac{1}{D} + \frac{1}{n^2 \hat{\text{I}}^2 \tau_x}} \end{aligned} \tag{37}$$

At the interior equilibrium, the difference of the ex-ante certainty equivalent between informed and uninformed rational traders equals to the cost of information:

$$\begin{aligned} CE_{inf1} - CE_{uninf1} & = \frac{\gamma}{2} \log \big(\text{Var}_1[v \mid p]\big) - \frac{\gamma}{2} \log \Big(\text{Var}_1[v \mid \sum_{i=1}^{n} s_i, p]\Big) \\ & = c \end{aligned} \tag{38}$$

Substituting equation (2) into equation (38) and rearrange, we get the expression of price informativeness:

$$\begin{aligned} PI(\tau, D, c, \gamma) = \big(\text{Var}_1(v \mid p, \beta, \rho)\big)^{-1} & = \exp\left\{ -\log \big(\text{Var}_1(v \mid p, \sum_{i=1}^{n} s_i)\big) - c \right\} \\ & = \frac{\tau + D}{A + 1} \end{aligned} \tag{39}$$

I interpret the information acquisition decision of rational traders by the following process that: more rational traders are willing to acquire information until price informativeness equals to $\frac{\tau+D}{A+1}$, which is independent of β. The constant can be regarded as the "ceiling of profit" for rational traders: if the actual price informativeness has not reached the "ceiling", there still exists profit margin to earn by information acquisition for rational traders, so λ_1 continuously increases until price informativeness equals to $\frac{\tau+D}{A+1}$.

Given (τ, D, c, γ), the "ceiling of profit" is the same with the price informativeness in benchmark, which implies that the price informativeness at equilibrium does not change in the mass of naive traders as long as there still exists informed rational traders. The general expression of price informativeness is given by:

$$\text{PI} = \begin{cases} \frac{\tau+D}{A+1} & \lambda_1 > 0 \\ \tau + \frac{1}{\frac{1}{D} + \frac{1}{n^2 \beta^2 \lambda_2^2 \gamma^2 \tau_\epsilon^2 \tau_x}} & \lambda_1 = 0 \end{cases} \tag{40}$$

Proposition 7 *There exists an equilibrium such that:*

1. *The fraction of rational traders who acquire information is λ_1; and the fraction of naive traders who acquire information is λ_2;*

2. *The coefficients (η, I, b_1, b_2) in price function and demand schedules are given in Appendix.*

Proposition 8 *Given $(n, \rho, c, \gamma, \tau, \tau_\epsilon, \tau_x)$, price informativeness keeps the constant as β increases until the informed rational traders are entirely crowded out of the market, and then price informativeness increases in the mass of naive traders β.*

The expression of λ_1 when $\lambda_1 \in (0,1)$ is given by:

$$\begin{aligned} \lambda_1 &= \frac{1}{1-\beta}\sqrt{\frac{D-\tau A}{A\tau_x\gamma^2(\tau D + D^2)}} - \frac{\beta\lambda_2 n\tau_\epsilon}{(1-\beta)D} \\ &= \lambda_0 - \frac{\beta}{1-\beta}\underbrace{\left(\lambda_2\Delta - \sqrt{\frac{D-\tau A}{A\tau_x\gamma^2(\tau D + D^2)}}\right)}_{\text{crowed-in (-) or crowed-out (+) effect}} \end{aligned} \tag{41}$$

where

$$\Delta = \frac{F}{D} = \frac{n\tau_\epsilon}{D} = 1 + (n-1)\rho \quad \Delta \in [1, +\infty) \tag{42}$$

λ_0 denotes the fraction of informed traders when $\beta = 0$ as in the benchmark, when $\lambda_0 \in (0,1)$:

$$\lambda_0 = \sqrt{\frac{D-\tau A}{A\tau_x\gamma^2(\tau D + D^2)}} \tag{43}$$

The first term of λ_1 in equaiton (41) equals to the fraction of informed traders in the absence of naive traders, and the second term is the effect of naive traders on information acquisition of rational traders.

Proposition 9 *Given $(n, \rho, c, \gamma, \tau, \tau_\epsilon, \tau_x)$, when $\rho \neq 0$ and $n \neq 1$,*

1. *If $\lambda_0 \in (0,1)$ when the cost of information is at a moderate level: $A \in \left(\frac{D}{\tau+\tau_x(\tau D+D^2)}, \frac{D}{\tau}\right)$:*

   (a) The naive traders always have "crowding-out" effect on information acquisition of rational traders, and the effect increases in ρ.

(b) *As β increasing, the "crowding-out" effect increases. There exists a β^*: all informed rational traders are crowded out of the market when $\beta \geqslant \beta^*$.*

2. *If $\lambda_0 = 0$ when information cost is sufficiently high: $A \in \left(\frac{D}{\tau}, +\infty\right)$, there are no informed rational traders regardless of whether naive traders exist or not.*

3. *If $\lambda_0 = 1$ when price is sufficiently low: $A \in (0, \frac{D}{\tau+\tau_x(\tau D+D^2)})$:*

   (a) *when $\beta < \frac{\sqrt{\frac{D-\tau A}{A\tau_x(\tau D+D^2)}}-1}{\lambda_2\Delta-1}$, all rational traders are informed;*

   (b) *when $\beta > \frac{\sqrt{\frac{D-\tau A}{A\tau_x(\tau D+D^2)}}-1}{\lambda_2\Delta-1}$, some rational traders choose not to acquire information.*

The trading intensity helps explain why the naive traders have a crowding-out effect on information acquisition of rational traders. I solve the aggregate trading intensity at equilibrium when $\lambda_1 > 0$:

$$\hat{I} = \sqrt{\frac{D^2 - \tau AD}{n^2\tau_x\gamma A(\tau + D)A}} \tag{44}$$

which is a constant unrelated to the mass of naive traders β. I decompose $\hat{I}$ as:

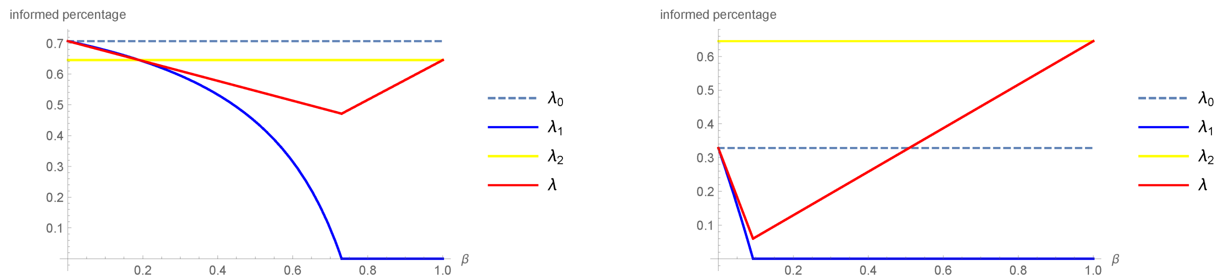
$$\hat{I} = \underbrace{(1-\beta)\lambda_1\gamma\frac{\tau_\epsilon}{\Delta}}_{\text{by informed rational traders}} + \underbrace{\beta\lambda_2\gamma\tau_\epsilon}_{\text{by informed naive traders}} \tag{45}$$

β influences the aggregate trading intensity in two ways:

$$\frac{dI(\beta,\lambda_1(\beta))}{d\beta} = \underbrace{\frac{\partial I(\beta,\lambda_1)}{\partial\beta}}_{\text{Direct effect (+)}} + \underbrace{\frac{\partial I(\beta,\lambda_1)}{\partial\lambda_1}\frac{\partial\lambda_1(\beta)}{\partial\beta}}_{\text{Indirect effect(-)}} \tag{46}$$

On the one hand, the increase of β has a positive direct effect on the aggregate intensity. As Proposition 3 shows, the larger the precision of signals perceived by traders, the larger the aggregate trading intensity they contribute. Because the naive traders overestimate the precision of aggregate signals, the aggregate trading intensity is larger than that if they were rational, leading a positive direct effect.

On the other hand, informed naive traders excessively contribute to the aggregate trading intensity and improve price effectiveness. When their contribution is out of proportion to their mass β, more informed rational traders are crowded out and contribute less to price informativeness, leading the negative indirect effect. At equilibrium, the direct effect and



(a) $\lambda_2 < \lambda_0$ ($A = 0.5, \gamma = 1, \tau_\epsilon = 0.5, \tau = 1, \tau_x = 1, n = 6, \rho = 0.1$)

(b) $\lambda_2 > \lambda_0$ ($A = 0.5, \gamma = 1, \tau_\epsilon = 0.5, \tau = 1, \tau_x = 1, n = 6, \rho = 0.9$)

Figure 1: Mass of informed traders

indirect effect cancel each other out, and the aggregate trading intensity as well as price informativeness does not change.

Corollary 3 *The aggregate trading intensity is maximised when $\beta = 1$; we have*

$$\frac{\lambda_0 \gamma D}{n} \leqslant \hat{I} \leqslant \lambda_2 \gamma \tau_\epsilon \tag{47}$$

The total mass of informed traders λ is given by:

$$\begin{aligned} \lambda &= (1-\beta)\lambda_1 + \beta\lambda_2 \\ &= \max\Big\{ \min\left\{\lambda_0 - \beta(n-1)\rho\lambda_2, (1-\beta) + \beta\lambda_2\right\}, \beta\lambda_2 \Big\} \end{aligned} \tag{48}$$

Proposition 10 *Given $(\rho, c, n, \gamma, \tau, \tau_\epsilon, \tau_x)$, when $n > 1$, and $\rho > 0$, the increase of naive traders monotonically reduces λ until informed rational traders are entirely crowded out of the market, and then λ rises gradually to λ_2 as β increases to 1.*

Comparing the total mass of informed traders λ with benchmark λ_0 when there are no naive traders in the market, I find that λ is smaller than λ_0 except that $\beta\lambda_2 > \lambda_0$.

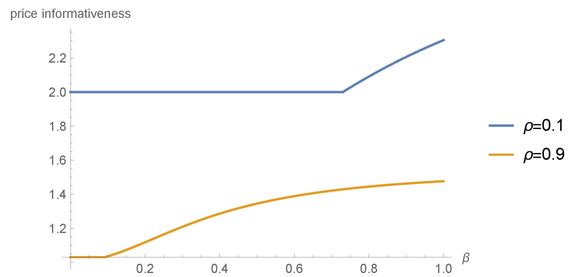


Figure 2: Price informativeness: ($A = 0.5, \gamma = 1, \tau_\epsilon = 0.5, \tau = 1, \tau_x = 1, n = 6$)

5 Model Implications

5.1 Market Depth

Formally, the measure of market liquidity is often referred to as Kyle's lambda (Kyle; 1985). The coefficient "Kyle lambda", which is η in equation (20), inversely measures market liquidity: a smaller η means that liquidity trading x has a smaller price impact, and thus the market is deeper and more liquid.

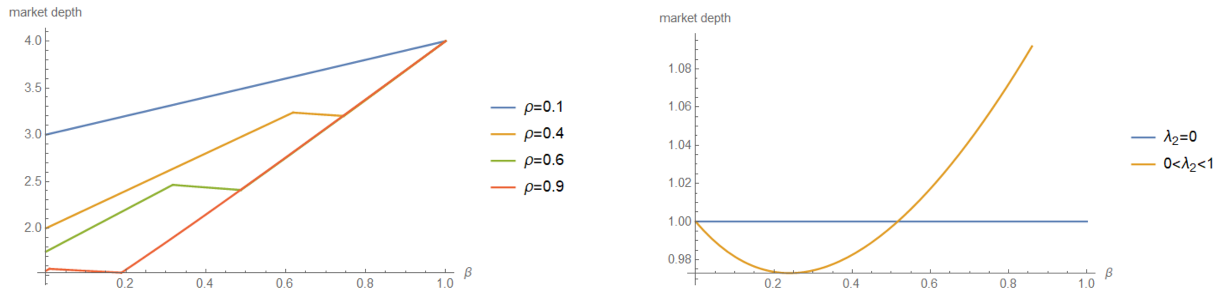
Market depth can be regared as the average responsiveness of market participants to price (Vives; 2008). Specially, in my framework, market depth equals to the average responsiveness of the following four types of traders: informed rational, uninformed rational, informed naive, uninformed naive traders:

$$\begin{aligned} \text{Market depth} &= \eta^{-1} \\ &= \lambda_2\beta\gamma(\tau + n\tau_\epsilon) + (1-\beta)\lambda_1\gamma(\tau + D) \\ &\quad - (1-\beta)(1-\lambda_1)b_1 - \beta(1-\lambda_2)b_2 \end{aligned} \tag{49}$$

To study how naive traders influences market depth, we discuss in the following three cases: (1) when $\lambda_0 \in (0,1)$; (2) when $\lambda_0 = 1$; (3) when $\lambda_0 = 0$. The first case is the most common when there are only rational traders in the market.

Proposition 11 *At equilibrium when* $\lambda_0 \in (0,1)$*,*

1. *Market depth* (η^{-1}) *decreases in the mass of naive traders until all informed rational traders are crowded out of the market.*



(a) $\lambda_0 = 1$ ($A = 0.5, \gamma = 1, \tau_\epsilon = 0.5, \tau = 1, \tau_x = 0.1, n = 6$)

(b) $\lambda_0 = 0$ ($\gamma = 1, \tau_\epsilon = 0.5, \tau = 1, \tau_x = 1, n = 6, \rho = 0.4, A \in \{2.5, 3\}$)

Figure 3: Market Depth in Corner Cases

2. *When all informed rational traders are driven out, if ρ is small enough, market depth always increases in β; when β is large enough, if λ_2 is sufficiently high, market depth increases in β; otherwise, there exists ρ^*, if $\rho > \rho^*$, market depth decreases in β.*

3. *Uninformed rational traders react more sensitively to price than uninformed naive traders, $|b_2| < |b_1|$.*

The effect of naive traders on market depth is ambiguous, depending on whether informed rational traders exist or not. When $\lambda_1 > 0$, the increase of β reduces λ_1, making more rational traders uninformed. Because the informed rational traders who are faced with less inventory risk, have stronger responsiveness to price than the uninformed traders, the decrease of λ_1 reduces the responsiveness of rational traders to price, leading the negative effect on market depth.

After informed rational traders are entirely crowded out of the market, the increase of β improves price informativeness, but aggravates adverse selection for rational traders. If the correlation between signal errors is small enough, the negative effect of adverse selection on market depth is dominated because the price system is efficient enough, so the market depth increases in β. However, if the informed fraction in naive traders is sufficiently small, and the correlation between signal errors is large enough, the negative effect induced by the increase of naive traders dominates when their mass is large enough, leading the decrease of market liquidity. My conclusions echo with Kyle (1985), where the market depth decreases in trading aggressiveness when it is relatively low, and increases when aggressiveness is sufficiently large.

I also study the corner cases when all $\lambda_0 = 1$, and $\lambda_0 = 0$, see Figure 3(a) and Figure 3(b), respectively.

Proposition 12 *At the corner equilibrium of benchmark:*

1. *When $\lambda_0 = 1$ and $\lambda_2 = 1$, market depth increases in β when $\lambda_1 = 1$; and decreases when λ_1 decreases until informed rational traders are entirely crowded out.*

2. *When $\lambda_0 = 0$, if $\lambda_2 = 0$,there is no informed traders in the market and market depth keeps the constant; if $\lambda_2 \neq 0$, market depth decreases in β when it is small, and increases in β when it is sufficiently large.*

5.2 Mispricing Risk

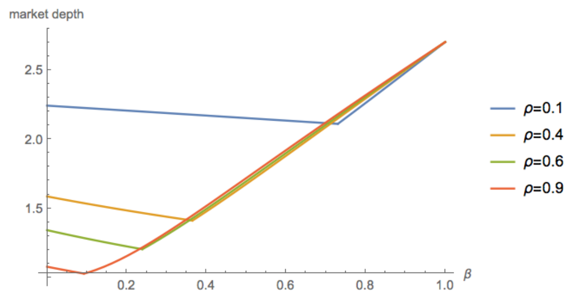
I use the mean-squared error between the asset's payoff and its price, $\mathrm{E}\big[(p-v)^2\big]$, to measure the mispricing risk that price is deviated from the fundamental (e.g., Odean (1998); Ko and Zhijian (2007); Goldstein and Yang (2017); Vives (2011)).

The expression of $\mathrm{E}\big[(p-v)^2\big]$ is given by:

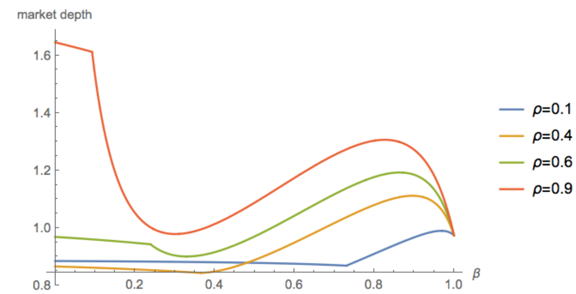
$$\begin{aligned}\mathrm{E}\big[(v-p)^2\big] &= \mathrm{Var}(v-p) \\ &= \frac{(1-\eta n I)^2}{\tau} + \frac{n^2 I^2 \eta^2}{D} + \frac{\eta^2}{\tau_x}\end{aligned} \tag{50}$$

When $\lambda_1 > 0$, as β increases, the total mass of uninformed traders increases, and the adverse selection is aggravated in the aggregate level due to the crowding out effect of naive traders. Thus the market liquidity decreases, and the mispricing risk increases in β. This finding is consistent with the behavioural models where more adverse selection increases mispricing (e.g., Daniel, Grinblatt, Titman and Wermers (1997); Hong and Stein (1999); Vives (2011)).

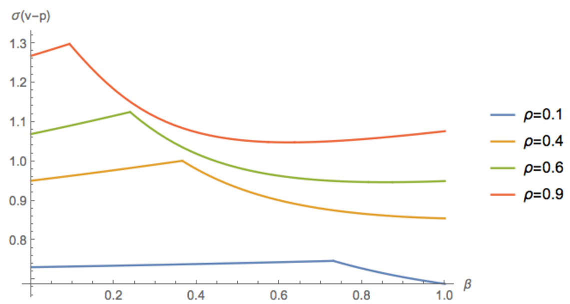
Proposition 13 *At equilibrium when $\lambda_0 \in (0,1)$, when there are informed rational traders, the mispricing risk increases in β.*



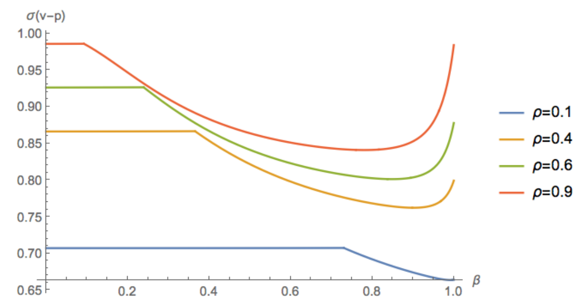
(a) *High* $\lambda_2 = 0.645$ ($\gamma = 1, \tau_\epsilon = 0.5, \tau = 1, \tau_x = 1, n = 6, A = 0.5$)



(b) *Low* $\lambda_2 = 0.02$ ($\gamma = 10, \tau_\epsilon = 0.5, \tau = 1, \tau_x = 10, n = 6, A = 0.5$)



(c) *High* $\lambda_2 = 0.645$ ($\gamma = 1, \tau_\epsilon = 0.5, \tau = 1, \tau_x = 1, n = 6, A = 0.5$)



(d) *Low* $\lambda_2 = 0.02$ ($\gamma = 10, \tau_\epsilon = 0.5, \tau = 1, \tau_x = 10, n = 6, A = 0.5$)

Figure 4: Market Depth and Mispricing Risk

5.3 Expected Utility

When $\lambda_2 \in (0, 1)$, the naive traders believe there is no difference to acquire information or not. However, the actual expected utility of informed naive traders is not the same with that of the uninformed, which is different from conventional literature about information acquisition (e.g., Grossman and Stiglitz (1980); Goldstein and Yang (2015)).

The expected certainty of uninformed naive traders is:

$$\begin{aligned} \mathrm{E}\left[-\exp\Big\{-\frac{1}{\gamma}(v-p)x\Big\}\right] &= \mathrm{E}\left[-\exp\Big\{-\frac{1}{\gamma}(v-p)D_{uninf2}(p)\Big\}\right] \\ &= -\mathrm{E}\left[\mathrm{E}\left[\exp\Big\{-\frac{1}{\gamma}(v-p)D_{uninf2}(p)\Big\}\Big| p\right]\right] \\ &\neq -E\left[\exp\left\{-\frac{(\mathrm{E}[v-p \mid p])^2}{2\mathrm{Var}[v \mid p]}\right\}\right] \end{aligned} \tag{51}$$

The inequality is induced by $D_{uninf2}(p) \neq \frac{\mathrm{E}[v-p|p]}{\mathrm{Var}[v|p]}$. We calculate the expected utility under the rational measure:

$$
\begin{aligned}
\mathrm{E}\left[-\exp\left\{-\frac{1}{\gamma}(v-p)x\right\}\right] &= \mathrm{E}\left[-\exp\left\{-\frac{1}{\gamma}(v-p)D_{uninf2}(p)\right\}\right] \\
&= \mathrm{E}\left[-\exp\left\{-\frac{b_2}{\gamma}(v-p)p\right\}\right]
\end{aligned}
\tag{52}
$$

The term of $(v-p)p$ in equation 51 is the product of two correlated normally distributed random variables. After standard normalization, we apply the function of MGF in Craig (1936) to calculate the expected utility of naive traders.

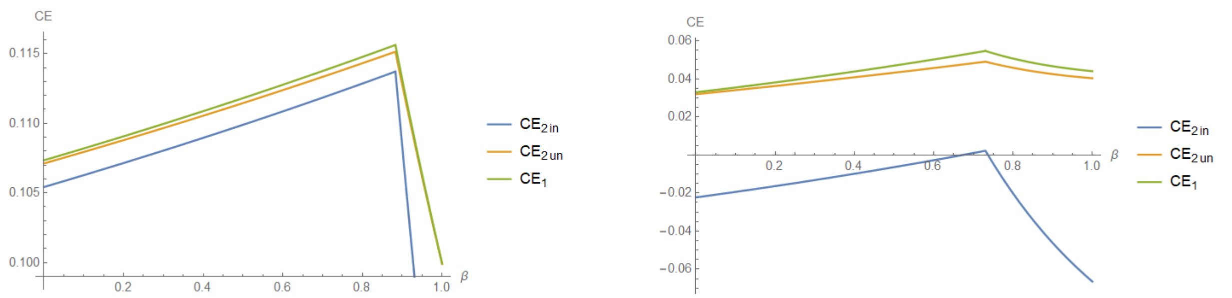
Proposition 14 *1. The expected utility of rational traders is always larger than that of the naive traders.*

2. The expected utility of rational traders can be improved by naive traders.

3. For many sets of the parameters specifying this economy, the expected utility of informed naive traders is lower than that of uninformed naive traders.

The actual expected utility of naive traders, namely, their welfare as defined in Goldstein and Yang (2015), is impaired by "correlation neglect". Moreover, although the naive traders believe acquiring information has no difference with being uninformed with respect to their expected utility, the welfare is actually weakened more if they choose to acquire information and trade more aggressively than keeping uninformed and trading less sensitively to price.

In Figure 5, the certainty equivalent of expected utility of rational traders are always higher than that of naive traders. In Figure 5(b), the certainty equivalent of informed naive traders can even be negative. This does not mean they are expected to lose money in the secondary market. In fact, the traders, regardless they are rational or naive, both exploit benefit from the noise traders. The negative certainty equivalent implies that the risk-adjusted wealth of informed naive traders is expected to be negative, but they may be expected to earn more than the rational traders (see Figure 5(b)). This can help explain why the naive traders can survive in financial markets. They can not recognize the fact that their actual utility is lower than others, but they may achieve higher expected return than their rational counterparts. The issue about the survival of naive traders has been analyzed in some literature, including Benos (1998); Kyle and Wang (1997); Hirshleifer and Luo (2001). My paper echoes with Hirshleifer and Hirshleifer and Luo (2001), which finds that the overconfident naive traders can better exploit noise traders and earn higher returns than their rational counterparts. My paper also further includes information acquisition, and finds that the survival of naive traders is still supported.



(a) $A = 0.5, \gamma = 1, \tau_\epsilon = 0.5, \tau = 1, \tau_x = 1, n = 2, \rho = 0.1$

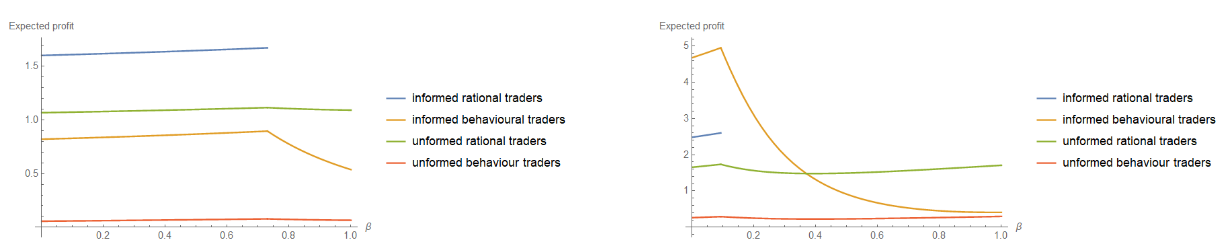
(b) $A = 0.5, \gamma = 1, \tau_\epsilon = 0.5, \tau = 1, \tau_x = 1, n = 6, \rho = 0.1$

Figure 5: Certainty equivalent

5.4 An Extension to Costly Information Acquisition Model

In the costly information acquisition model, it is assumed that the informed traders acquire the same bundle of information signals. In this subsection, I study the case that the informed traders acquire heterogenous information. All other model specifications remain the same as in Section 4. Under the alternative assumption, I assume that $s_{ji}(i \in (1, 2, ...n))$, which implies that the signal errors are entirely cancelled out when they are aggregated at equilibrium. This assumption can be rationalized by the different sources of information collected by informed traders. At equilibrium, the linear price function is given by

$$p = \eta n I v + \eta x \tag{53}$$



(a) $A = 0.5, \gamma = 1, \tau_\epsilon = 0.5, \tau = 1, \tau_x = 1, n = 6, \rho = 0.1$

(b) $A = 0.5, \gamma = 1, \tau_\epsilon = 0.5, \tau = 1, \tau_x = 1, n = 6, \rho = 0.9$

Figure 6: Expected profit

The conditional variance about fundamental for informed rational and naive traders are given by:

$$\mathrm{Var}_1\left[v\middle|\sum_{i=1}^{n}s_i, p\right] = \frac{1}{\tau + D + n^2 I^2 \tau_x} \tag{54}$$

$$\mathrm{Var}_2\left[v\middle|\sum_{i=1}^{n}s_i, p\right] = \frac{1}{\tau + F + n^2 I^2 \tau_x} \tag{55}$$

Based on the information acquisition decision of rational traders, the informed fraction of rational traders is given by:

$$\begin{aligned}\lambda_1 &= \frac{1}{1-\beta}\sqrt{\frac{D-\tau A}{A\tau_x\gamma^2 D^2}} - \frac{\beta\lambda_2 n\tau_\epsilon}{(1-\beta)D} \\ &= \lambda_0 - \frac{\beta}{1-\beta}\underbrace{\left(\lambda_2\Delta - \sqrt{\frac{D-\tau A}{A\tau_x\gamma^2 D^2}}\right)}_{\text{crowed-out }(+)\text{ effect}}\end{aligned} \tag{56}$$

And the price informativeness at equilibrium is given by D/A. The heterogeous private information of traders does not change the prior main conclusions: at equilibrium, price informativeness keeps the constant as β increases until the informed rational traders are entirely crowded out of the market, and then price informativeness increases in the mass of naive traders β. the existence of naive traders still has crowding-out effect on the information acquisition of informed rational traders.

6 Model without Information Cost

In this section, I turn to the case when the information signals are free of charge, and study how the market is influenced by "correlation neglect" of naive traders.

The information can be regarded as the public information, e.g., media, news, announcements. The model setting is the same as prior sections, except that each trader, either rational or naive, can observe n signals. Traders update their beliefs about the fundamental using the public information and trade accordingly.

The demand of each rational trader is given by:

$$D_1\left(\sum_{i=1}^{n} s_i, p\right) = \frac{\gamma D}{n}\sum_{i=1}^{n} s_i - \gamma(\tau + D)p \tag{57}$$

The demand of naivetrader is given by:

$$D_2\left(\sum_{i=1}^{n} s_i, p\right) = \gamma\tau_\epsilon\sum_{i=1}^{n} s_i - \gamma(\tau + n\tau_\epsilon)p \tag{58}$$

Based on the market clearing condition:

$$(1-\beta)D_{inf1}\left(\sum_{i=1}^{n} s_i, p\right) + \beta D_{inf2}\left(\sum_{i=1}^{n} s_i, p\right) + x = 0 \tag{59}$$

, price can be solved:

$$p = \eta\left(nIv + I\sum_{i=1}^{n}\epsilon_i + x\right) \tag{60}$$

where

$$I = (1-\beta)\gamma\frac{D}{n} + \beta\gamma\tau_\epsilon \tag{61}$$

and

$$\eta = \frac{1}{\gamma\tau + nI} \tag{62}$$

Because price does not reflect any private information in this case, it is meaningless to analyze price informativeness as before. We use the mean-squared error between the asset's payoff and its price, $\mathrm{E}\big[(p-v)^2\big]$, to measure the mispricing risk.

The expression of $\mathrm{E}\big[(p-v)^2\big]$ is given by:

$$\begin{aligned}\mathrm{E}\left[(p-v)^2\right] &= \mathrm{Var}(p-v)\\ &= \eta^2\left(\frac{\gamma^2}{\tau} + \frac{n^2I^2}{D} + \frac{1}{\tau_x}\right)\end{aligned} \tag{63}$$

As the mass of naive traders increases, on the one hand, because they underestimate the inventory risk and are willing to provide more liquidity than rational traders, the market depth increases, which reduces $\mathrm{Var}(p-v)$; on the other hand, errors in the public signals are reflected more in the price due to increasing aggressive trading at the aggregate level.

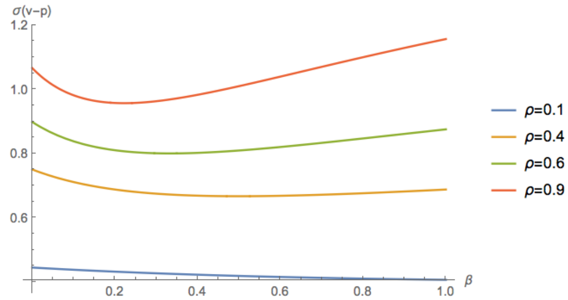


Figure 7: Mispricing risk without information acquisition cost: ($\gamma = 1, \tau_\epsilon = 0.5, \tau = 1, \tau_x = 1, n = 6$)

Proposition 15 *1. When $\rho < \frac{1}{\tau_x \gamma^2 \tau (n-1)}$, mispricing risk decreases in β.*

2. When $\rho > \frac{1}{\tau_x \gamma^2 \tau (n-1)}$, there exists a β^*, if $\beta \in (0, \beta^*)$, mispricing risk decreases in β; if $\beta > \beta^*$, mispricing risk increases in β.

The positive effect of naive traders on market quality dominates when the correlation between signal errors is sufficiently small, because the market benefits from the liquidity the naive traders provide. However, when the correlation between signals errors is high enough, and the mass of naive traders is sufficiently large, the negative effect on market quality dominates because the price is vulnerable to be deviated by the errors in public signals. The larger the correlation between errors, the more possibility that the market is exposed to mispricing risk.

7 Conclusion

In this paper, I develop a model where rational traders interact with naive traders who neglect the correlation between signal errors. I consider two alternative cases with and without information acquisition cost, and derive the implications about the market quality measured by price informativeness, market liquidity and mispricing risk.

First, I find that the impact of "correlation neglect" on financial markets depends on whether information is costly or not. If the information is free, mispricing risk can improve in the mass of naive traders, but may be impaired by them when their mass and the correlation are both high enough. This is because signal errors cause mispricing risk when the price is excessively sensitive to public signals. However, if information acquisition is costly, the existence of naive traders increases mispricing risk when their mass is not large enough to crowd out all informed rational traders out of the market.

Second, when information acquisition is costly, naive traders have a “crowding out” effect on the information acquisition of rational traders. The total mass of informed traders also declines in the mass of naive traders before informed rational traders are entirely crowded out, meanwhile, price informativeness keeps the same and market liquidity deteriorates. However, when informed rational traders do not exist, market quality, measured by liquidity and mispricing risk, can improve afterwards, depending on the correlation between signals and the mass of naive traders.

Finally, I am able to use the model to derive implications regarding the empirical properties of market quality. In particular, one distinct feature of my model is that the impact of “correlation neglect” on financial markets depends on information acquisition cost and the mass of naive traders. These implications can potentially serve as the explanation of the mispricing risk induced by the repetition of media. Moreover, the information acquisition model helps understand why active asset management has become less attractive in the past few years, and why market quality is potentially impaired.

A Appendix : Proofs

A.1 Proof of Proposition 1

Let $s = [s_1, ..., s_n]$ be the information the informed investor possesses at time 2. The price does not reflect more information about fundamental than s. The mean vector and variance-covariance matrix of the $n+1$ dimensional normal random variable $(v, s) \sim N(0, \Sigma)$, with the variance-covariance matrix $\Sigma \in \mathbb{R}^{(n+1)\times(n+1)}$. The mean vector and variance-covariance matrix can be partitioned as $\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, and

$$\Sigma = \begin{bmatrix} \Sigma_{v,v} & \Sigma_{v,s} \\ \Sigma_{s,v} & \Sigma_{s,s} \end{bmatrix} = \begin{bmatrix} 1/\tau & 1/\tau & ... & 1/\tau \\ 1/\tau & 1/\tau + 1/\tau_\epsilon & ... & 1/\tau + \rho/\tau_\epsilon \\ \vdots & \vdots & \ddots & \vdots \\ 1/\tau & 1/\tau + \rho/\tau_\epsilon & ... & 1/\tau + 1/\tau_\epsilon \end{bmatrix} \tag{A.1}$$

.

The conditional mean is

$$E[v \mid s_1, ..., s_n] = \Sigma_{v,s}\Sigma_{s,s}^{-1}s = \frac{\tau_\epsilon \sum_{i=1}^n s_i}{\tau + (n-1)\rho\tau + n\tau_\epsilon} \tag{A.2}$$

and the variance-covariance matrix

$$Var[v \mid s_1, ..., s_n] = \Sigma_{v,v} - \Sigma_{v,s}\Sigma_{s,s}^{-1}\Sigma_{s,v} = \frac{1 + (n-1)\rho}{\tau + (n-1)\rho\tau + n\tau_\epsilon} \tag{A.3}$$

Thus the demand of an informed trader is:

$$D_{inf}\left(\sum_{i=1}^n s_i, p\right) = \frac{\gamma\tau_\epsilon}{1 + (n-1)p}\sum_{i=1}^n s_i - \gamma\frac{\tau + (n-1)\rho\tau + n\tau_\epsilon}{1 + (n-1)p}p \tag{A.4}$$

The uniformed trader extract signal from price:

$$w = \lambda\frac{\gamma\tau_\epsilon}{1 + (n-1)p}\sum_{i=1}^n s_i + x$$

I define I as the aggregate trading intensity of informed traders, where

$$I = \lambda\gamma\frac{\tau_\epsilon}{1 + (n-1)\rho}$$

Thus we can write w as

$$w = In\left(v + \frac{I\sum_{i=1}^n \epsilon_i + x}{In}\right)$$

Because of the multivariate normal distribution of ϵ_i, I have

$$\text{Var}\left(\sum_{i=1}^n \epsilon_i\right) = \frac{n\big(1 + \rho(n-1)\big)}{\tau_\epsilon} = \frac{n^2}{D} \tag{A.5}$$

and the precision of public signal given the fundamental v is

$$\left(\text{Var}\left(\frac{I\sum_{i=1}^n \epsilon_i + x}{In}\right)\right)^{-1} = \frac{I^2n^2}{\frac{1}{\tau_x} + I^2\frac{n\big(1+\rho(n-1)\big)}{\tau_\epsilon}} = \frac{1}{\frac{1}{n^2I^2\tau_x} + \frac{1}{D}} \tag{A.6}$$

So the variance of fundamental v given the information set of the uninformed traders is:

$$\text{Var}[v \mid p] = \left(\tau + \frac{1}{\frac{1}{n^2 I^2 \tau_x} + \frac{1}{D}}\right)^{-1} \tag{A.7}$$

Under the conjecture that $p = nI\eta(v + \frac{I\sum_{i=1}^{n}\epsilon_i + x}{In})$, the expectation of fundamental v given the information set of the uninformed traders is:

$$\text{E}[v \mid p] = \left(\tau + \frac{I^2 n^2}{\frac{1}{\tau_x} + I^2 \frac{n\left(1+\rho(n-1)\right)}{\tau_\epsilon}}\right)^{-1} \frac{I^2 n^2}{\frac{1}{\tau_x} + I^2 \frac{n\left(1+\rho(n-1)\right)}{\tau_\epsilon}} \frac{\eta}{nI} p \tag{A.8}$$

Substitute the expressions above into $D_{uninf}(p) = \frac{\gamma(E[v|p]-p)}{\text{Var}[v|p]}$, and the market clearing condition. I get the coefficients:

$$\eta = \frac{\big(1 + (-1+n)\rho\big)\big(1 + (-1+n)\rho + n\gamma^2\lambda\tau_x\tau_\epsilon\big)}{\gamma\Big(\big(1 + (-1+n)\rho\big)^2\tau + n\lambda\big(1 + (-1+n)\rho\big)(1 + \gamma^2\lambda\tau\tau_x)\tau_\epsilon + n^2\gamma^2\lambda^2\tau_x\tau_\epsilon^2\Big)} \tag{A.9}$$

And thus I can get the demand of the uniformed trader as

$$D_{uninf}(p) = bp$$

where

$$b = -\frac{\gamma\tau}{1 + \gamma^2\lambda\tau_x D}$$

□

A.2 Proof of Proposition 5

According to Proposition 3, in the interior equilibrium,

$$\lambda = \sqrt{\frac{-A\tau + D}{A\tau_x\gamma^2(\tau D + D^2)}} \tag{A.10}$$

I take the derivative of λ with respect to D:

$$\frac{d\lambda}{dD} = \frac{1}{2A\tau_x\gamma^2}\left(\frac{-A\tau + D}{A\tau_x\gamma^2(\tau D + D^2)}\right)^{-1/2} \frac{-D^2 + 2A\tau D + A\tau^2}{(\tau D + D^2)^2} \tag{A.11}$$

The sign of $\frac{d\lambda}{dD}$ depends on the sign of $(-D^2 + 2A\tau D + A\tau^2)$. When $D \in \left[A\tau, A\tau\left(1 + \sqrt{1 + \frac{1}{A}}\right)\right)$, $\frac{d\lambda}{dD} > 0$; when $D \in \left[A\tau\left(1 + \sqrt{1 + \frac{1}{A}}\right), +\infty\right)$, $\frac{d\lambda}{dD} < 0$. □

A.3 Proof of Proposition 8

The demand of uninformed naive traders:

$$\begin{aligned} D_{uninf2}(p) &= \frac{\gamma\big(E_b[v \mid p] - p\big)}{\mathrm{Var}_b[v \mid p]} \\ &= b_2 p \end{aligned} \tag{A.12}$$

where

$$\mathrm{Var}_b[v \mid p] = \frac{1}{\tau + \frac{1}{\frac{1}{n\tau} + \frac{1}{n^2\lambda_2^2\gamma^2\tau_\epsilon^2\tau_x}}} \tag{A.13}$$

$$\mathrm{E}_b[v \mid p] = \frac{\frac{1}{\frac{1}{n\tau} + \frac{1}{n^2\lambda_2^2\gamma^2\tau_\epsilon^2\tau_x}}}{\tau + \frac{1}{\frac{1}{n\tau} + \frac{1}{n^2\lambda_2^2\gamma^2\tau_\epsilon^2\tau_x}}} \frac{\lambda_2\gamma(\tau + n\tau_\epsilon) - (1-\lambda_2)b_2}{n\lambda_2\gamma\tau_\epsilon} p \tag{A.14}$$

b_2 is solved from the following function:

$$\frac{b_2}{\gamma} = \frac{1}{\frac{1}{n\tau_\epsilon} + \frac{1}{n^2\lambda_2^2\gamma^2\tau_\epsilon^2\tau_x}} \frac{\lambda_2\gamma(\tau + n\tau_\epsilon) - (1-\lambda_2)b_2}{n\lambda_2\gamma\tau_\epsilon} - \tau - \frac{1}{\frac{1}{n\tau_\epsilon} + \frac{1}{n^2\lambda_2^2\gamma^2\tau_\epsilon^2\tau_x}} \tag{A.15}$$

I solve the equation above and get:

$$b_2 = -\frac{\gamma\tau}{1 + n\gamma^2\lambda_2\tau_x\tau_\epsilon}$$

The actual market clearing condition is:

$$\begin{aligned} \beta\lambda_2\gamma\tau_\epsilon \sum_{i=1}^n s_i - \lambda_2\beta\gamma(\tau + n\tau_\epsilon)p + (1-\beta)\lambda_1\gamma\frac{D}{n}\sum_{i=1}^n s_i - (1-\beta)\lambda_1\gamma(\tau + D)p \\ +(1-\beta)(1-\lambda_1)b_1 p + \beta(1-\lambda_2)b_2 p + x = 0 \end{aligned} \tag{A.16}$$

The price is pinned down by the market clearing condition, and the general form of price is:

$$\begin{aligned} p &= \eta I \sum_{i=1}^n s_i + \eta x \\ &= \underbrace{\eta I n v}_{\text{fundamental}} + \underbrace{\eta I \sum_{i=1}^n \epsilon_i}_{\text{Signal erros}} + \underbrace{\eta x}_{\text{Liquidity trading noise}} \\ &= W v + Y \sum_{i=1}^n \epsilon_i + \eta x \end{aligned} \tag{A.17}$$

where $W = \eta In$, $Y = \eta I$.

The demand function of uninformed rational traders extract noise signal:

$$\begin{aligned} D_{uninf1}(p) &= \frac{\gamma\big(E[v \mid p] - p\big)}{\mathrm{Var}[v \mid p]} \\ &= b_1 p \end{aligned} \tag{A.18}$$

Substitute the expression of $E[v \mid p]$ and $\mathrm{Var}[v \mid p]$ into the equation above, and solve b_1.

$$\mathrm{E}[v \mid p] = \frac{\frac{1}{\frac{1}{D}+\frac{1}{n^2I^2\tau_x}} \frac{\lambda_2\beta\gamma(\tau+n\tau_\epsilon)+(1-\beta)\lambda_1\gamma(\tau+D)-(1-\beta)(1-\lambda_1)b_1-\beta(1-\lambda_2)b_2}{nI}}{\tau + \frac{1}{\frac{1}{D}+\frac{1}{n^2I^2\tau_x}}} \tag{A.19}$$

$$\mathrm{Var}[v \mid p] = \frac{1}{\tau + \frac{1}{\frac{1}{D}+\frac{1}{n^2I^2\tau_x}}} \tag{A.20}$$

where

$$\mathrm{nI} = \begin{cases} \sqrt{\frac{-A\tau D + D^2}{A\tau_x(\tau+D)}} & \lambda_1 > 0 \\ \lambda_2\gamma n\tau_\epsilon & \lambda_1 = 0. \end{cases}$$

$$\begin{aligned} \frac{b_1}{\gamma} = &\frac{1}{\frac{1}{D}+\frac{1}{n^2I^2\tau_x}} \frac{\lambda_2\beta\gamma(\tau+n\tau_\epsilon)+(1-\beta)\lambda_1\gamma(\tau+D)-(1-\beta)(1-\lambda_1)b_1-\beta(1-\lambda_2)b_2}{nI} \\ &- \tau - \frac{1}{\frac{1}{D}+\frac{1}{n^2I^2\tau_x}} \end{aligned} \tag{A.21}$$

b_1 is solved from equation above. □

A.4 Proof of Proposition 8

I get the expression of price informativeness:

$$\mathrm{PI} = \begin{cases} \frac{\tau+D}{A+1} & \lambda_1 > 0 \\ \tau + \frac{1}{\frac{1}{D}+\frac{1}{n^2\beta^2\lambda_2^2\gamma^2\tau_\epsilon^2\tau_x}} & \lambda_1 = 0. \end{cases}$$

Only when $\tau + \frac{1}{\frac{1}{D}+\frac{1}{n^2\beta^2\lambda_2^2\gamma^2\tau_\epsilon^2\tau_x}} > \frac{\tau+D}{A+1}$, $\lambda_1 = 0$. And I have $\partial(\frac{\tau+D}{A+1})/\partial\beta = 0$, and

$$\partial\left(\tau + \frac{1}{\frac{1}{D}+\frac{1}{n^2\beta^2\lambda_2^2\gamma^2\tau_\epsilon^2\tau_x}}\right)/\partial\beta > 0$$

. □

A.5 Proof of Proposition 10

$$\lambda_2\Delta - \sqrt{\frac{D-\tau A}{A\tau_x\gamma^2(\tau D + D^2)}} = \frac{\Delta}{\sqrt{A\tau_x\gamma^2}}\left(\sqrt{\frac{n\tau_\epsilon - \tau A}{\tau n\tau_\epsilon + n^2\tau_\epsilon^2}} - \sqrt{\frac{D-\tau A}{\Delta\tau n\tau_\epsilon + n^2\tau_\epsilon^2}}\right) \tag{A.22}$$

Let

$$\Psi(\rho) = \sqrt{\frac{n\tau_\epsilon - \tau A}{\tau n\tau_\epsilon + n^2\tau_\epsilon^2}} - \sqrt{\frac{D-\tau A}{\Delta\tau n\tau_\epsilon + n^2\tau_\epsilon^2}}$$

Because $\frac{\partial\Psi(\rho)}{\partial\rho} > 0$, thus $\Psi(\rho) > \Psi(0) = 0$, and the second term is positive and increases in ρ. $\square$

A.6 Proof of Proposition 12

When $\lambda_1 > 0$, I have

$$\frac{b_1}{\gamma} = \frac{1}{\frac{1}{D} + \frac{1}{n^2I^2\tau_x}}\frac{\eta^{-1}}{nI} - \tau - \frac{1}{\frac{1}{D} + \frac{1}{n^2I^2\tau_x}} \tag{A.23}$$

The aggregate trading intensity I is unrelated to β, so $\frac{db_1}{d\beta}$ have the same sign with $\frac{d\eta^{-1}}{d\beta}$. To study how β influences market depth (η^{-1}), I can study how β influences b_1 instead.

When $\lambda_1 > 0$, I have $nI = \sqrt{\frac{-A\tau D + D^2}{A\tau_x(\tau + D)}}$.

According to the definition of I,

$$nI = \gamma\lambda_2\beta D\Delta + (1-\beta)\lambda_1\gamma D \tag{A.24}$$

, I rewrite

$$\lambda_2\beta\gamma(\tau + n\tau_\epsilon) + (1-\beta)\lambda_1\gamma(\tau + D) = nI + \frac{\tau nI}{D} - \gamma\tau\lambda_2\beta(n-1)\rho \tag{A.25}$$

Substitute equation (A.29) into equation (A.25), take the derivative of both sides of equation with respect to β, and rearrange it, I get

$$\left[\frac{(\frac{1}{D} + \frac{1}{n^2I^2\tau_x})nI}{\gamma} + (1-\beta)(1-\lambda_1)\right]\frac{\partial b_1}{\partial\beta} = (1-\lambda_1)b_1 - \gamma\tau\lambda_2(n-1)\rho - (1-\lambda_2)b_2 + (1-\beta)b_1\frac{\partial\lambda_1}{\partial\beta} \tag{A.26}$$

Take the derivative of both sides with respect to β, I solve

$$\frac{\partial \lambda_1}{\partial \beta} = -\frac{\lambda_2 \Delta - \lambda_1}{1-\beta} \tag{A.27}$$

Substitute equation (A.31) into equation (A.30), and I get

$$\begin{aligned}\left[\frac{(\frac{1}{D}+\frac{1}{n^2 I^2 \tau_x})nI}{\gamma} + (1-\beta)(1-\lambda_1)\right]\frac{\partial b_1}{\partial \beta} &= -\gamma\tau\lambda_2(n-1)\rho - \lambda_2 b_1(n-1)\rho \\ &\quad + (1-\lambda_2)b_1 - (1-\lambda_2)b_2\end{aligned} \tag{A.28}$$

I have $\gamma\tau > -b_1$, so the term $-\gamma\tau\lambda_2(n-1)\rho - \lambda_2 b_1(n-1)\rho < 0$. If I assume there exits a β satisfying $\frac{\partial b_1}{\partial \beta} > 0$, and I assume β^* is the minimum among all the value of β satisfying the condition. I can conclude $(1-\lambda_2)b_1 - (1-\lambda_2)b_2 > 0$. Because $\frac{\partial b_1}{\partial \beta} < 0$ when $\beta \in (0, \beta^*)$, thus $(1-\lambda_2)b_1 - (1-\lambda_2)b_2 < (1-\lambda_2)(-\frac{\gamma\tau}{1+\gamma^2\lambda_0\tau_x D} + \frac{\gamma\tau}{1+n\gamma^2\lambda_2\tau_x\tau_\epsilon}) < 0$, which contradicts to the original assumption that $\frac{\partial b_1}{\partial \beta} > 0$, I can prove $\frac{\partial \eta^{-1}}{\partial \beta} < 0$ when $\lambda_1 > 0$.

When $\lambda_1 = 0$, market depth is expressed by:

$$\eta^{-1} = \beta\lambda_2(\gamma\tau + \gamma n\tau_\epsilon) - \beta(1-\lambda_2)b_2 - (1-\beta)b_1 \tag{A.29}$$

Take the derivative with the respect to β:

$$\frac{\partial \eta^{-1}}{\partial \beta} = \lambda_2(\gamma\tau + \gamma n\tau_\epsilon) - (1-\lambda_2)b_2 + b_1 - (1-\beta)\frac{\partial b_1}{\partial \beta} \tag{A.30}$$

when $\rho \to 0$, because $\beta \geqslant \frac{\lambda_0}{\lambda_2\Delta}$, $\beta \to 1$, and $b_1 \to b_2$. $\frac{\partial b_1}{\partial \beta}$ have finite bounds. Thus $\frac{\partial \eta^{-1}}{\partial \beta} \to \lambda_2(\gamma\tau + \gamma n\tau_\epsilon + b_2)$, which is larger than zero.

Denote $q = nI = \beta\lambda_2$, I solve b_1 from equation (A.27),

$$b_1 = \frac{-\tau - \frac{1}{\frac{1}{D}+\frac{1}{q^2\tau_x}} + \frac{\beta\gamma\lambda_2(\tau+n\tau_\epsilon)}{q(\frac{1}{D}+\frac{1}{q^2\tau_x})} - \frac{\beta(1-\lambda_2)b_2}{q(\frac{1}{D}+\frac{1}{q^2\tau_x})}}{\frac{1}{\gamma} + \frac{1-\beta}{q(\frac{1}{D}+\frac{1}{q^2\tau_x})}} \tag{A.31}$$

When $\beta \to 1$, $b_1 \to -\gamma\tau + \gamma\frac{\gamma\lambda_2\tau-(1-\lambda_2)b_2}{q(\frac{1}{D}+\frac{1}{q^2\tau_x})}$, $\frac{\partial b_1}{\partial \beta}$ have finite bounds. Thus $\frac{\partial \eta^{-1}}{\partial \beta} \to \gamma\lambda_2\tau - (1-\lambda_2)b_2 - \gamma\tau + \gamma\frac{\gamma\lambda_2\tau-(1-\lambda_2)b_2}{q(\frac{1}{D}+\frac{1}{q^2\tau_x})}$.

$\gamma\lambda_2\tau - (1-\lambda_2)b_2 - \gamma\tau + \gamma\frac{\gamma\lambda_2\tau-(1-\lambda_2)b_2}{q(\frac{1}{D}+\frac{1}{q^2\tau_x})}$ is an increasing function about D, and D decreases in ρ. If $\gamma\lambda_2\tau - (1-\lambda_2)b_2 - \gamma\tau + \gamma\frac{\gamma\lambda_2\tau-(1-\lambda_2)b_2}{q(\frac{1}{\tau_\epsilon}+\frac{1}{q^2\tau_x})} < 0$, when $\rho > \rho^*$, where ρ^* is the solution of equation $\gamma\lambda_2\tau - (1-\lambda_2)b_2 - \gamma\tau + \gamma\frac{\gamma\lambda_2\tau-(1-\lambda_2)b_2}{q(\frac{1}{D}+\frac{1}{q^2\tau_x})} = 0$, and $q = \lambda_2\gamma n\tau_\epsilon$, market depth is decreasing in ρ as $\beta \to 1$.

A.7 Proof of Proposition 14

$$\begin{aligned}\mathrm{E}\left[(v-p)^2\right] &= \mathrm{Var}(v-p) \\ &= \frac{(1-\eta nI)^2}{\tau} + \frac{n^2I^2\eta^2}{D} + \frac{\eta^2}{\tau_x}\end{aligned}$$

(1) When $\lambda_1 > 0$, nI keeps the constant where $nI = \lambda_0\gamma D$. We have $\eta < \frac{1}{nI}$, and thus $1 > \eta nI$.

$$\begin{aligned}\frac{d\mathrm{Var}(v-p)}{d\eta} &= \frac{2(1-\eta nI)(-nI)}{\tau} + \frac{2n^2I^2\eta}{D} + \frac{2\eta}{\tau_x} \\ &= -\frac{2nI}{\tau} + \Big(\frac{2n^2I^2}{\tau} + \frac{2n^2I^2}{D} + \frac{2}{\tau_x}\Big)\eta\end{aligned} \tag{A.32}$$

When $\eta > \frac{1}{nI+\frac{nI\tau}{D}+\frac{\tau}{nI\tau_x}}$, $\frac{d\mathrm{Var}(v-p)}{d\eta} > 0$; otherwise, $\eta > \frac{1}{nI+\frac{nI\tau}{D}+\frac{\tau}{nI\tau_x}}$, $\frac{d\mathrm{Var}(v-p)}{d\eta} < 0$.

According to Proposition 5.2, η is increasing in β when $\lambda_1 > 0$. When $\beta = 0$, $\eta = \frac{1}{nI+\lambda_0\gamma\tau+\frac{(1-\lambda_0)\gamma\tau}{1+\gamma\tau_x nI}}$. Thus $\eta > \frac{1}{nI+\lambda_0\gamma\tau+\frac{(1-\lambda_0)\gamma\tau}{1+\gamma\tau_x nI}}$.

The threshold of η can be writen as $\frac{1}{nI+\frac{nI\tau}{D}+\frac{\tau}{nI\tau_x}} = \frac{1}{nI+\lambda_0\gamma\tau+\frac{\tau}{nI\tau_x}}$. Because $\frac{(1-\lambda_0)\gamma\tau}{1+\gamma\tau_x nI} < \frac{\tau}{nI\tau_x}$, thus $\eta > \frac{1}{nI+\frac{nI\tau}{D}+\frac{\tau}{nI\tau_x}}$. I can get $\frac{d\mathrm{Var}(v-p)}{d\eta} > 0$.

A.8 Proof of Proposition 15

The expected utility of uninformed naive traders is:

$$\mathrm{E}\left[-\exp\left\{-\frac{b_2\sqrt{\mathrm{Var}[v-p]\mathrm{Var}[p]}}{\gamma}\frac{(v-p)p}{\sqrt{\mathrm{Var}[v-p]\mathrm{Var}[p]}}\right\}\right] \tag{A.33}$$

Let ρ_{xy} be the correlation between the two standard normally distributed variables: $\frac{v-p}{\sqrt{\mathrm{Var}[v-p]}}$ and $\frac{p}{\sqrt{\mathrm{Var}[p]}}$.

$$M_{xy}(t) = \frac{1}{\sqrt{[1-(1+\rho_{xy})t][1+(1-\rho_{xy})t]}}$$

Let $t = -\frac{b_2\sqrt{\mathrm{Var}[v-p]\mathrm{Var}[p]}}{\gamma}$, $\rho_{xy} = \frac{\mathrm{Cov}(v-p,p)}{\sqrt{\mathrm{Var}[v-p]\mathrm{Var}[p]}}$. According to the linear expression of price function equation (A.21), $p = Wv + Y\sum_{i=1}^n \epsilon_i + \eta x$, we get $\mathrm{Var}[p] = \frac{W^2}{\tau}$, $\mathrm{Var}[v-p] = \frac{(1-W)^2}{\tau} + \frac{Y^2n^2}{D} + \frac{\eta^2}{\tau_x}$, and $\mathrm{Cov}(v-p,p) = \frac{(1-W)W}{\tau} + \frac{Y^2n^2}{D} + \frac{\eta^2}{\tau_x}$.

$$CE_{uninf2} = \frac{\gamma}{2}\log\Big(\big[1-(1+\rho_{xy})t\big]\big[1+(1-\rho_{xy})t\big]\Big) \tag{A.34}$$

Similarly, I calculate the expected certainty of informed naive traders under the rational measure:

$$\mathrm{E}\Big[-\exp\Big\{-\frac{1}{\gamma}\big[(v-p)x-c\big]\Big\}\Big]=\mathrm{E}\Big[-\exp\Big\{-\frac{1}{\gamma}\big[(v-p)D_{inf2}\Big(\sum_{i=1}^{n}s_i,p\Big)-c\big]\Big\}\Big] \tag{A.35}$$

I have

$$\begin{aligned}
D_{inf2}\left(\sum_{i=1}^{n}s_i,p\right)&=\gamma\tau_\epsilon\sum_{i=1}^{n}s_i-\gamma(\tau+n\tau_\epsilon)p\\
&=\gamma\tau_\epsilon\left(nv+\sum_{i=1}^{n}\epsilon_i\right)-\gamma(\tau+n\tau_\epsilon)\left(Wv+Y\sum_{i=1}^{n}\epsilon_i+\eta x\right)\\
&=\Big(\gamma\tau_\epsilon n-\gamma W(\tau+n\tau_\epsilon)\Big)v+\Big(\gamma\tau_\epsilon-Y\gamma(\tau+n\tau_\epsilon)\Big)\sum_{i=1}^{n}\epsilon_i-\gamma\eta(\tau+n\tau_\epsilon)x
\end{aligned} \tag{A.36}$$

,

which is a linear combination of v, $\sum_{i=1}^{n}\epsilon_i$ and p, thus $(v-p)D_{inf2}(\sum_{i=1}^{n}s_i,p)$ is the product of two correlated normally distributed variables.

$$CE_{inf2}=\frac{\gamma}{2}\log\Big(\big[1-(1+\rho_{xy})t\big]\big[1+(1-\rho_{xy})t\big]\Big) \tag{A.37}$$

where $t=-\frac{\sqrt{(v-p)D_{inf2}}}{\gamma}$, $\rho_{xy}=\frac{\mathrm{Cov}(v-p,D_{inf2})}{\sqrt{(v-p)D_{inf2}}}$, given that

$$\mathrm{Cov}(v-p,D_{inf2})=\frac{\big(\gamma\tau_\epsilon n-\gamma W(\tau+n\tau_\epsilon)\big)(1-W)}{\tau}-\frac{\big(\gamma\tau_\epsilon-Y\gamma(\tau+n\tau_\epsilon)\big)Yn^2}{D}+\frac{\gamma\eta^2(\tau+n\tau_\epsilon)}{\tau_x}$$

$$\mathrm{Var}(D_{inf2})=\frac{\big(\gamma\tau_\epsilon n-\gamma W(\tau+n\tau_\epsilon)\big)^2}{\tau}+\frac{\big(\gamma\tau_\epsilon-Y\gamma(\tau+n\tau_\epsilon)\big)n^2}{D}+\frac{\big(\gamma\eta(\tau+n\tau_\epsilon)\big)^2}{\tau_x}$$

□

A.9 Proof of Proposition 15

Let $q=nI\in[\gamma D,\gamma n\tau_\epsilon]$, q increases in β.

$$\frac{\partial\mathrm{Var}(p-v)}{\partial q}=\frac{\partial\frac{\gamma^2\tau+1/\tau_x+q^2/D}{(\gamma\tau+q)^2}}{\partial q}=\frac{q\gamma\tau/D-\gamma^2\tau-1/\tau_x}{(\gamma\tau+q)^3} \tag{A.38}$$

If $\rho < \frac{1}{\tau_x \gamma^2 \tau (n-1)}$, $\frac{\gamma^2\tau + 1/\tau_x + q^2/D}{(\gamma\tau+q)^2} \partial q < 0$ when $q \in [\gamma D, \gamma n \tau_\epsilon]$;

If $\rho > \frac{1}{\tau_x \gamma^2 \tau (n-1)}$, $\frac{\gamma^2\tau + 1/\tau_x + q^2/D}{(\gamma\tau+q)^2} \partial q < 0$ when $q \in [\gamma D, \gamma D + \frac{D}{\tau_x \gamma^2 \tau}]$; $\frac{\gamma^2\tau + 1/\tau_x + q^2/D}{(\gamma\tau+q)^2} \partial q > 0$ when $q \in (\gamma D + \frac{D}{\tau_x \gamma^2 \tau}, \gamma n \tau_\epsilon]$. □

References

Admati, A. R. and Pfleiderer, P. (1986). A monopolistic market for information, *Journal of Economic Theory* **39**(2): 400–438.

Banerjee, S. and Green, B. (2015). Signal or noise? Uncertainty and learning about whether other traders are informed, *Journal of Financial Economics* **117**(2): 398–423.

Banerjee, S. and Kremer, I. (2010). Disagreement and Learning: Dynamic Patterns of Trade, *The Journal of Finance* **65**(4): 1269–1302.

Benos, A. V. (1998). Aggressiveness and survival of overconfident traders, *Journal of Financial Markets* **1**(3-4): 353–383.

Craig, C. C. (1936). On the Frequency Function of xy, *The Annals of Mathematical Statistics* **7**(1): 1–15.

Daniel, K., Grinblatt, M., Titman, S. and Wermers, R. (1997). Measuring Mutual Fund Performance with Characteristic-Based Benchmarks, *The Journal of Finance* **52**(3): 1035–1058.

De Bondt, W. F. M. and Forbes, W. P. (1999). Herding in analyst earnings forecasts: evidence from the United Kingdom, *European Financial Management* **5**(2): 143–163.

Enke, B. and Zimmermann, F. (2019). Correlation Neglect in Belief Formation, *The Review of Economic Studies* **86**(1): 313–332.

Eyster, E., Rabin, M. and Vayanos, D. (2019). Financial Markets Where Traders Neglect the Informational Content of Prices, *The Journal of Finance* **74**(1): 371–399.

Gao, F., Song, F. and Wang, J. (2013). Rational expectations equilibrium with uncertain proportion of informed traders, *Journal of Financial Markets* **16**(3): 387–413.

García, D. and Sangiorgi, F. (2011). Information Sales and Strategic Trading, *Review of Financial Studies* **24**(9): 3069–3104.

Goldstein, I. and Yang, L. (2015). Information Diversity and Complementarities in Trading and Information Acquisition, *The Journal of Finance* **70**(4): 1723–1765.

Goldstein, I. and Yang, L. (2017). Information disclosure in financial markets, *Annual Review of Financial Economics* **9**: 101–125.

Goulding, C. L. and Zhang, X. (2018). The value of scattered information, *Available at SSRN 2846482* .

Grossman, S. J. and Stiglitz, J. E. (1980). On the Impossibility of Informationally Efficient Markets, *The American Economic Review* **70**(3): 393–408.

Hirshleifer, D. and Luo, G. Y. (2001). On the survival of overconfident traders in a competitive securities market, *Journal of Financial Markets* **4**(1): 73–84. Publisher: Elsevier.

Hong, H. and Stein, J. C. (1999). A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets, *The Journal of Finance* **54**(6): 2143–2184.

Hong, H. and Stein, J. C. (2007). Disagreement and the Stock Market, *Journal of Economic Perspectives* **21**(2): 109–128.

Huang, S. and Yueshen, B. Z. (2021). Speed acquisition, *Management Science* **67**(6): 3492–3518.

Jiao, P., Veiga, A. and Walther, A. (2020). Social media, news media and the stock market, *Journal of Economic Behavior & Organization* **176**: 63–90.

Kim, C. and Pantzalis, C. (2003). Global/Industrial Diversification and Analyst Herding, *Financial Analysts Journal* **59**(2): 69–79.

Ko, K. J. and Zhijian, J. H. (2007). Arrogance can be a virtue: Overconfidence, information acquisition, and market efficiency, *Journal of Financial Economics* **84**(2): 529–560.

Kyle, A. S. (1985). Continuous Auctions and Insider Trading, *Econometrica* **53**(6): 1315–1335.

Kyle, A. and Wang, F. A. (1997). Speculation Duopoly with Agreement to Disagree: Can Overconfidence Survive the Market Test?, *Journal of Finance* **52**(5): 2073–90.

Odean, T. (1998). Volume, Volatility, Price, and Profit When All Traders Are Above Average, *The Journal of Finance* **53**(6): 1887–1934.

Papadimitriou, D. (2020). Trading Under Uncertainty About Other Market Participants, *Available at SSRN 3497782* .

Qin, N. and Singal, V. (2015). Indexing and Stock Price Efficiency, *Financial Management* **44**(4): 875–904.

Siegel, R. and Yacht, C. (2009). *Personal Finance*, Saylor Foundation.

Tetlock, P. C. (2011). All the News That's Fit to Reprint: Do Investors React to Stale Information?, *The Review of Financial Studies* **24**(5): 1481–1512.

Verrecchia, R. E. (1982). Information Acquisition in a Noisy Rational Expectations Economy, *Econometrica* **50**(6): 1415–1430.

Vives, X. (2008). *Information and learning in markets: The impact of market microstructure*, Princeton University Press.

Vives, X. (2011). A large-market rational expectations equilibrium model, *CESifo Working Paper Series* .

Welch, I. (2000). Herding among security analysts, *Journal of Financial Economics* **58**(3): 369–396.