JSPS Grants-in-Aid for Scientific Research (S)

Understanding Persistent Deflation in Japan

Working Paper Series

No. 042

June 2014

Beauty Contests and Fat Tails in Financial Markets

Makoto Nirei

Tsutomu Watanabe

UTokyo Price Project
702 Faculty of Economics, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
Tel: +81-3-5841-5595
E-mail: watlab@e.u-tokyo.ac.jp
http://www.price.e.u-tokyo.ac.jp/english/

Beauty Contests and Fat Tails in Financial Markets

Makoto Nirei* Tsutomu Watanabe

Hitotsubashi University University of Tokyo

June 25, 2014

Abstract

Using a simultaneous-move herding model of rational traders who infer other traders' private information on the value of an asset by observing their aggregate actions, this study seeks to explain the emergence of fat-tailed distributions of transaction volumes and asset returns in financial markets. Without making any parametric assumptions on private information, we analytically show that traders' aggregate actions follow a power law distribution. We also provide simulation results to show that our model successfully reproduces the empirical distributions of asset returns. We argue that our model is similar to Keynes's beauty contest in the sense that traders, who are assumed to be homogeneous, have an incentive to mimic the average trader, leading to a situation similar to the indeterminacy of equilibrium. In this situation, a trader's buying action causes a stochastic chain-reaction, resulting in power laws for financial fluctuations.

Keywords: Herd behavior, transaction volume, stock return, fat tail, power law

JEL classification code: G14

*Address: 2-1 Naka, Kunitachi, Tokyo 186-8603, Japan. Email: nirei@iir.hit-u.ac.jp.

1 Introduction

Since Mandelbrot [25] and Fama [13], it has been well established that stock returns exhibit fat-tailed and leptokurtic distributions. Jansen and de Vries [19], for example, have shown empirically that the power law exponent for stock returns is in the range of 3 to 5, which guarantees that the variance is finite but the distribution deviates substantially from the normal distribution in terms of the fourth moment. Such an anomaly in the tail shape, as well as kurtosis, has been regarded as one reason for the excess volatility of stock returns.

Efforts to explain this anomaly have been long ongoing. A traditional economic explanation for the excess volatility of transaction volumes and returns relies on rational herd behavior by traders. In a situation where traders' action space is coarser than their state space, traders' actions only partially reveals their private information on the value of an asset. This property makes it possible for a trader's action to cause an avalanche of similar actions by other traders. This idea of a chain reaction through the revelation of private information has been extensively studied in the literature on herd behavior, informational cascades, and information aggregation. However, there have been few attempts to explain fat tails in stock return distributions based on this idea. This paper is the first to develop an economic model in order to show that the chain reaction of information revelation generates fat tails in asset return distributions.

To this end, we consider a model consisting of informed and uninformed traders. There are a large number of informed traders who receive imperfect private signals on the true value of an asset. The informed traders simultaneously choose whether to buy one unit of the asset or not to buy at all. To simplify the model, and unlike Glosten and Milgrom [15] or Smith [31], we assume that informed traders cannot short-sell. We consider a rational expectations equilibrium in which each trader submits their demand schedule conditional on the price of the asset. The rational choice made by an informed trader is based on the private signal they receive as well as the information revealed by other traders' actions through the equilibrium price. The price is set by an auctioneer, who aggregates informed traders' demand and matches it with the supply schedule submitted by uninformed traders. We show that, in this setting, the more informed traders choose to buy the asset, the higher will be the price, which in turn is regarded as a signal of higher asset value. As a result, some of the traders who decided not to buy at the previous stage change their mind. In this way, traders' strategies exhibit complementarity, and their actions are positively correlated.

The main contribution of our study to the literature is that we characterize the probability distribution of the equilibrium number of buying traders and show analytically that it has a power law tail with an exponent of 0.5, which is defined for a cumulative distribution. The power law tail for the number of buying traders implies the presence of power law tails for the equilibrium transaction volume as well as for the equilibrium asset price.

In terms of the way it describes herding behavior, our model is similar to Keynes's beauty contest. Each trader recognizes that the other traders receive private signals that are as valuable as their own. Therefore, each trader seeks to mimic the average trader. However, when each trader tries to match the behavior of the average trader, the resulting equilibrium is fragile due to perfect strategic complementarity. In addition, since traders' actions are discrete in our setting, the equilibrium is locally unique, which allows us to quantitatively characterize fluctuations in transaction volumes and prices due to randomness in private signals. Our analysis therefore formalizes the idea of perfect strategic complementarity among traders with private signals and shows that a power law distribution of transaction volumes and prices emerges naturally in this setup.

Our study is related to the substantial theoretical and empirical literature on imitative behavior in financial markets. Important theoretical studies in this field include those by Scharfstein and Stein [30], Banerjee [5], and Bikhchandani, Hirshleifer, and Welch [6], who have developed models of herd behavior and informational cascades. These models have been employed in a number of studies to examine financial market crashes, including Caplin and Leahy [10], Lee [23] and Chari and Kehoe [11]. However, most of these studies feature all-or-nothing herding due to the type of information structure they assume, such as sequential trading. As a result, few studies in this literature address the issue of stochastic financial fluctuations. An important exception is the study by Gul and Lundholm [16], who demonstrated the emergence of stochastic herding by endogenizing traders' choice of waiting time. We follow this approach and focus on the stochastic aspect of financial fluctuations, but deviate from it by employing a model in which traders move simultaneously, and the equilibrium number of traders exhibits stochastic fluctuations and follows a power law distribution. Our model of stochastic herding contributes to the literature by showing that informational cascades can generate not only extremely large financial fluctuations in transaction volumes and prices but also an empirically relevant regularity regarding the frequency distribution of these fluctuations, which is summarized by power law distributions.

While there are many statistical models that can replicate the power law distribution of asset returns, few economic models have been developed that explain it. One exception is the model developed by Gabaix et al. [14]. They provide a model in which the power laws for asset returns and transaction volumes are accounted for by a power law in a different context, namely Zipf's law for firm sizes. Specifically, they argue that if traders' size follows a power law, transaction volume and price changes also follow a power law. Given the presence of extensive evidence on power laws in firm size distributions, the mechanism investigated by Gabaix et al. [14] at least partially accounts for power laws in financial fluctuations. However, our paper differs from theirs in that we do not rely on heterogeneity across traders in accounting for power laws in financial fluctuations; instead, we assume that traders are homogeneous in size and in other respects. We show that, even in this symmetric setting, the interaction of a large number of traders generates stochastic herding with different degrees (i.e., the number of traders who decide to buy the asset differs), thereby generating power laws in financial fluctuations. By showing this, we provide a new explanation for power laws in financial fluctuations which is complementary to the one advocated by Gabaix et al. [14].[1]

The remainder of the study is organized as follows. Section 2 presents our baseline model. Section 3 then analytically shows that a power law distribution emerge for transaction volumes and provides an intuition for the mechanism behind it. Section 4 presents numerical simulations to show that the equilibrium volumes follow a power

[1] Another area to which this study, especially the technical part, is related is the literature on critical phenomena in statistical physics. A number of statistical physicists have investigated the empirical fluctuations of financial markets (surveys of these studies can be found in Bouchaud and Potters [8] and Mantegna and Stanley [26]), and some studies in this literature reproduce the empirical power laws by applying a methodology often used for the analysis of critical phenomena to herd behavior models (Bak, Paczuski, and Shubik [3]; Cont and Bouchaud [12]; Stauffer and Sornette [34]). However, these studies do not model traders' purposeful behavior and rational learning, and therefore fail to link their analysis to the existing body of financial economics literature. More importantly, these studies do not address why market activities exhibit criticality. This issue is important because, according to these studies, power laws in financial fluctuations typically occur only when the parameter that governs the connectivity of the networked traders takes a critical value. These two issues will be addressed in this study.

law and that the equilibrium return distribution matches its empirical counterpart. Section 5 concludes.

2 Model

2.1 Model and equilibrium

We consider an economy with two possible states, which are denoted by H and L. There is a large number N of informed traders, each of which receives an imperfect and private signal x_i of the state. The signal is imperfect in the sense that x_i does not fully reveal the true state. Also, the signal is private in the sense that each trader does not observe other traders' signals.

Traders simultaneously choose whether to buy one unit of an asset or not to buy at all. The asset is worth 1 in state H and 0 in state L. Shares of the asset are issued and traded. Indivisibility is introduced in that the transaction unit is fixed. Specifically, the transaction unit is given by $1/N$. We consider a rational expectations equilibrium in which each trader submits their demand schedule conditional on the price of the asset, p. The demand function is denoted by $a_i = d(p, x_i)$ for trader i, where $a_i = 1$ indicates buying and $a_i = 0$ not-buying. Aggregate demand expressed in terms of the transaction unit is $D(p) = \sum_{i=1}^{N} d(p, x_i)/N$.

Uninformed traders decide on whether to supply the asset depending only on p.[2] The aggregate supply function, which is denoted by $S(p)$, is assumed to be continuous and upward sloping with a bounded domain, $0 \leq p < 1$, and satisfies $S(p_0) = 0$ for a

[2]The informational asymmetry between informed and uninformed traders in this model is similar to event uncertainty, which was introduced by Avery and Zemsky [2] as a condition for herding to occur in financial markets.

given initial price p_0. The equilibrium price is determined so that it clears the market, i.e., $D(p) = S(p)$. We define a sequence of exogenous price points p_k, $k = 1, 2, \ldots, N$ at which the demand from k informed traders is met by the supply from the uninformed traders, i.e., $S(p_k) = k/N$.

Transactions are implemented by an auctioneer, who receives the demand and supply schedules D and S from the informed and uninformed traders, and chooses price p_m such that $D(p_m) = S(p_m)$.[3] Note that m represents the equilibrium number of buying traders. The belief of informed trader i that state H occurs, after observing p_m, is denoted by r_i. Informed traders are assumed to be risk-neutral and maximize their subjective expected payoff. The expected payoff of a trader is 0 when $a_i = 0$ regardless of the belief, whereas it is $r_i - p$ when $a_i = 1$. Thus, trader i buys the asset if and only if $r_i \geq p$. The optimality condition for buying, $r_i \geq p$, can be rewritten as $\rho_i \leq (1-p)/p$ using the likelihood ratio $\rho_i \equiv (1 - r_i)/r_i$.

For each realization of a profile of private signals $(x_i)_{i=1}^N$, a rational expectations equilibrium consists of the number of buying informed traders m, price p_m, allocation $(a_i)_{i=1}^N$, demand schedules d, and the posterior likelihood ratios $(\rho_i)_{i=1}^N$, such that (i) for any p, $d(p, x_i)$ maximizes trader i's expected payoff evaluated at ρ_i for any i, (ii) ρ_i is consistent with the realized private signal x_i and p_m for any i, and (iii) the auctioneer clears the market, i.e., $S(p_m) = \sum_{i=1}^N a_i/N$, and delivers the orders $a_i = d(p_m, x_i)$ where $m = \sum_{i=1}^N a_i$.

[3]This mechanism of implementing a rational expectations equilibrium through the submission of demand schedules follows Bru and Vives [9]. Without information aggregation by the auctioneer, the model would become similar to that of Minehart and Scotchmer [27], who showed that traders cannot agree to disagree in a rational expectations equilibrium, i.e., an equilibrium may not exist, or if it exists, it is a herding equilibrium where all traders choose the same action.

2.2 Information structure and optimal strategy

We now specify the information structure. The private signal x_i is drawn independently across i from a known distribution F in state H and from distribution G in state L. We impose the standard assumption on private signals that they satisfy a property called the monotone likelihood ratio property (MLRP). We define an odds function $\ell(x_i) = g(x_i)/f(x_i)$, where f and g are derivatives of F and G, respectively. MLRP requires ℓ to be monotone. Without loss of generality, we assume that ℓ is strictly decreasing. This implies that a greater level of x_i is associated with a higher likelihood of state H. We also assume that the prior belief that H will occur is common across i with equal probabilities for H and for L. The likelihood ratio for the common prior belief is denoted by $\lambda = 1$. Furthermore, we specify that the initial price reflects the common prior belief as $p_0 = 0.5$. This assumption is imposed for the sake of simplicity and is relaxed in Appendix B, where the belief is allowed to be heterogeneous across traders in a dynamic setting.

We derive the optimal demand schedule of trader i as a threshold rule. The optimal threshold rule is given by

$$d(p_k, x_i) = \begin{cases} 1 & \text{if } x_i \geq \bar{x}(k), \\ 0 & \text{otherwise}, \end{cases} \tag{1}$$

for $k = 1, 2, \ldots, N$, where $\bar{x}(k)$ denotes the threshold level for the private signal at which a buying trader is indifferent between buying and not buying after observing p_k.

We solve for the optimal threshold $\bar{x}$ as follows. Under the threshold rule, the likelihood ratios revealed by inaction ($a_i = 0$) and by buying ($a_i = 1$) given p_k are respectively derived as follows:

$$A(\bar{x}) \equiv \frac{\Pr(x_i < \bar{x} \mid L)}{\Pr(x_i < \bar{x} \mid H)} = \frac{G(\bar{x})}{F(\bar{x})}, \tag{2}$$

$$B(\bar{x}) \equiv \frac{\Pr(x_i \geq \bar{x} \mid L)}{\Pr(x_i \geq \bar{x} \mid H)} = \frac{1 - G(\bar{x})}{1 - F(\bar{x})}, \tag{3}$$

where $\bar{x}$ is shorthand for $\bar{x}(k)$. As shown by Smith and Sørensen [32], MLRP implies that, for any value of x in the interior of the support of F and G,

$$A(x) > \ell(x) > B(x) > 0, \tag{4}$$

and that $A(x)$ and $B(x)$ are strictly decreasing in x:

$$\frac{dA(x)}{dx} = \frac{g(x)}{F(x)} - \frac{G(x)f(x)}{F(x)^2} = \frac{f(x)}{F(x)}\left(\ell(x) - A(x)\right) < 0, \tag{5}$$

$$\frac{dB(x)}{dx} = -\frac{g(x)}{1 - F(x)} + \frac{(1 - G(x))f(x)}{(1 - F(x))^2} = \frac{f(x)}{1 - F(x)}\left(B(x) - \ell(x)\right) < 0. \tag{6}$$

Consider a trader making a buying bid at price p_1. If this bid is struck by the auctioneer, this implies that the other $N-1$ informed traders do not bid at p_1. Thus, their inaction reveals the likelihood ratio $A(\bar{x}(1))^{N-1}$. The posterior likelihood ratio in this case is $\rho_i = A(\bar{x}(1))^{N-1}\ell(x_i)\lambda$. Thus, the threshold is determined by

$$1/p_1 - 1 = A(\bar{x}(1))^{N-1}\ell(\bar{x}(1))\lambda. \tag{7}$$

Similarly, each trader knows that, if the bid is executed at p_k, there are $k-1$ traders buying at p_k and $N-k$ traders not buying at p_k. Then, the threshold $\bar{x}(k)$ is obtained by solving

$$1/p_k - 1 = A(\bar{x}(k))^{N-k}B(\bar{x}(k))^{k-1}\ell(\bar{x}(k))\lambda. \tag{8}$$

Given the threshold behavior shown above, we obtain aggregate demand $D(p_k)$ by counting the number of informed traders with $x_i \geq \bar{x}(k)$ and dividing it by N. For the case of $k = 0$, we exogenously set $D(p_0) = D(p_1)$. If $D(p_1) = 0$ realizes as a result of realized private signals, then p_0 clears the market since $D(p_0) = D(p_1) = S(p_0) = 0$. If $D(p_1) > 0$, then p_0 cannot clear the market. With this setup, we obtain the following property:

Lemma 1 *There exists an $\bar{N}$ such that for any $N > \bar{N}$, the threshold level of signal $\bar{x}(k)$ is strictly decreasing in k and the aggregate demand function $D(p_k)$ is non-decreasing in k.*

Proof: By taking the log-difference of (8), we obtain

$$\log \frac{A(\bar{x}(k))}{B(\bar{x}(k))} + \log \frac{1/p_{k+1} - 1}{1/p_k - 1} = (N-k-1) \log \frac{A(\bar{x}(k+1))}{A(\bar{x}(k))} + k \log \frac{B(\bar{x}(k+1))}{B(\bar{x}(k))} + \log \frac{\ell(\bar{x}(k+1))}{\ell(\bar{x}(k))}. \tag{9}$$

Note that $\inf(\log A - \log B)$ is strictly positive and independent of N, while $\log(1/p_{k+1} - 1) - \log(1/p_k - 1)$ converges to 0 as $N \to \infty$, since $S(p_k) = k/N$. Hence, the left-hand side is strictly positive for a sufficiently large N. The right-hand side is strictly positive only if $\bar{x}(k+1) < \bar{x}(k)$, since $A' < 0$, $B' < 0$, and $\ell' < 0$. Thus, $\bar{x}(k)$ is strictly decreasing in k. Since $D(p_k)$ is the number of traders with $x_i \geq \bar{x}(k)$ for $k = 1, 2, \ldots, N$, divided by N, and since $D(p_0) = D(p_1)$, a decreasing $\bar{x}$ implies that $D(p_k)$ is non-decreasing in k for any realization of $(x_i)_{i=1}^N$. $\square$

According to Lemma 1, the more informed traders are buying, the more signals in favor of H are revealed, and hence, the more likely each informed trader is to buy.[4] This implies the presence of strategic complementarity in informed traders' buying decision.

[4] A similar result was presented in Nirei [29]. However, this study differs from Nirei [29] in that Nirei [29] used a Nash equilibrium, while the present study uses a rational expectations equilibrium. With the Nash formulation, Nirei [29] was not able to establish the existence of equilibrium with a finite number of traders, which is accomplished in this study, as shown in Proposition 1.

Lemma 1 also shows that as a higher price indicates that there are more informed traders who receive high signals, the aggregate demand curve is upward sloping when N is sufficiently large. On the other hand, when N is small, the increment in price p_{k+1}/p_k is substantial due to limited supply, thus leading to a higher purchasing cost and a downward sloping demand curve.

Finally, we show the existence of equilibrium. We define the aggregate reaction function as a mapping from the number of buying traders k inferred from p_k, to the number of buying traders determined by informed investors' choices after observing p_k in addition to their private signals. Specifically, the aggregate reaction function is given by $\Gamma : \mathcal{S} \mapsto \mathcal{S}$, where $\mathcal{S} = \{0, 1, 2, \ldots, N\}$, for each realization of $(x_i)_{i=1}^N$ such that $\Gamma(k) \equiv D(p_k)N$ for any $k \in \mathcal{S}$. With the upward sloping demand function, we obtain the existence of equilibrium in a finite economy as follows.

Proposition 1 *There exists an $\bar{N}$ such that for any $N > \bar{N}$ there exists an equilibrium outcome m for each realization of $(x_i)_{i=1}^N$.*

Proof: Since Γ is a non-decreasing mapping of a finite discrete set $\mathcal{S}$ onto itself, there exists a non-empty closed set of fixed points of Γ as implied by Tarski's fixed point theorem. Since $S(p_m) = m/N$, such fixed points m of Γ satisfy $D(p_m) = S(p_m)$. $\square$

In this economy, multiple equilibria may exist for each realization of $(x_i)_{i=1}^N$. We focus on the case where the auctioneer selects the minimum number of buying traders, m^*, among possible equilibria for each $(x_i)_{i=1}^N$. This assumption that the auctioneer selects the minimum number of buying traders means that we exclude fluctuations that arise purely from informational coordination such as in sunspot equilibria. Even with this assumption, we can show that the equilibrium price, $\log p_{m^*} - \log p_0$, shows large fluctuations. Note that this equilibrium selection uniquely maps each realization of $(x_i)_{i=1}^N$ to m^*. Thus, m^* is a random variable whose probability distribution is determined by the probability distribution of $(x_i)_{i=1}^N$ and the equilibrium selection mapping.

3 Analytical derivation of the power law

In this section, we characterize the minimum equilibrium aggregate action m^* and show that it follows a power law distribution. The power law distribution for m^* implies a fat tail and large volatility for the transaction volume. Since the asset price in this model is determined by the equilibrium condition $S(p_{m^*}) = m^*/N$, the power law for the transaction volume also implies a fat-tailed distribution of the equilibrium price p_{m^*}.

We start by showing that the equilibrium m^* is obtained making use of the best response dynamics adopted in Vives [36]. The best-response dynamics for the number of buying traders are defined by $m_u = \Gamma(m_{u-1})$ for $u = 1, 2, \ldots, T$, where m_u is the number of buying traders in step u, Γ is the aggregate reaction function, and T is stopping time, which is defined as the smallest step u such that $m_u - m_{u-1} = 0$. We set $m_0 = 0$. Throughout the study, we will assume that the true state of the economy is H unless stated otherwise. A similar analysis can be conducted for the case that the true state of the economy is L. Based on the above setting, we obtain the following result.

Lemma 2 *m_u converges to the minimum equilibrium m^* for each realization of $(x_i)_{i=1}^N$. Moreover, the threshold for buying, $\bar{x}$, decreases over u, i.e., $\bar{x}(m_{u+1}) < \bar{x}(m_u)$, for any realization of $(x_i)_{i=1}^N$.*

Proof: Applying Vives [35], it can be shown that m_u always reaches a fixed point m_T

with respect to Γ, since Γ is increasing, $\mathcal{S}$ is finite, and $m_0 = 0$ is the minimum in $\mathcal{S}$. Further, m_T must coincide with the minimum fixed point m^* for the following reason. Suppose that there exists another fixed point m that is strictly smaller than m_T. Then we can pick $u < T$ such that $m_u < m < m_{u+1}$. Applying the non-decreasing function Γ, we obtain $\Gamma(m_u) \leq \Gamma(m)$. Thus, $m_{u+1} \leq m$. This contradicts $m < m_{u+1}$. Hence, $m_T = m^*$. Finally, $\bar{x}(m_u)$ decreases with u, since, as shown in Lemma 1, $\bar{x}(k)$ is strictly decreasing in k and m_u is non-decreasing. □

Note that since the best-response dynamics start from $m_0 = 0$ and converge to m^*, we can express m^* as the cumulative sum of increments in m_u. Moreover, the best-response dynamics can be regarded as a stochastic process, because the probability of m_u conditional on $m_1, \ldots, m_{u-1}$ is determined by the joint distribution of the private signal profile $(x_i)_{i=1}^N$. Thus, m^* can be expressed as the sum of a stochastic difference process that converges to zero. This approach of characterizing the equilibrium outcome by a stochastic process is similar to the one adopted by Kirman [22].

An important implication of Lemma 2 is that there exists a non-trivial chance of a chain reaction during the process, since the threshold $\bar{x}$ decreases as u increases. Specifically, a trader who chooses to buy in step u will continue to choose buying in $u+1$, since the threshold is lower. On the other hand, a trader who chooses not to buy in u might switch to buying in $u+1$. The conditional probability of a non-buying trader switching to buying in response to $m_u - m_{u-1}$ is defined as follows:

$$q_u \equiv \int_{\bar{x}_u}^{\bar{x}_{u-1}} f(x)dx/F(\bar{x}_{u-1}), \quad u = 1, 2, \ldots, N, \tag{10}$$

where $\bar{x}_u$ is shorthand for $\bar{x}(m_u)$. Note that q_u is always non-negative because of the decreasing threshold. Thus, $m_{u+1} - m_u$, the number of traders who buy in $u+1$ for the first time, conditional on the history of m up to u, follows a binomial distribution with population parameter $N - m_u$ and probability parameter q_u. The distribution of

m_1 follows a binomial distribution with population N and probability $q_0 \equiv 1 - F(\bar{x}_0)$. This completes the definition of the stochastic process $m_u - m_{u-1}$, which is summarized in the following lemma.[5]

Lemma 3 *Consider a stochastic process $m_u - m_{u-1}$, $u = 1, 2, \ldots, T$, where $m_0 = 0$. Suppose that $m_{u+1} - m_u$ conditional on $m_u - m_{u-1}$ follows a binomial distribution with population $N - m_u$ and probability q_u. Further, suppose that m_1 follows a binomial distribution with population N and probability q_0. Then, the minimum equilibrium number of buying traders m^* follows the same distribution as m_T, the cumulative sum of the process.*

This lemma establishes that the minimum equilibrium m^* is equal to the sum of a binomial process. A binomial distribution can be approximated by a Poisson distribution when the population is "large" and the probability is "small." The approximation holds if q_u decreases as N^{-1}. The next lemma shows that this is indeed the case for $u > 1$.

Lemma 4 *As $N \to \infty$, the binomial process $m_{u+1} - m_u$ asymptotically follows a branching process with a state-dependent Poisson random variable with mean*

$$\phi_u = \frac{A(\bar{x}_{u-1})}{\ell(\bar{x}_{u-1})} \frac{\log A(\bar{x}_{u-1}) - \log B(\bar{x}_{u-1})}{A(\bar{x}_{u-1})/B(\bar{x}_{u-1}) - 1}, \quad u = 2, 3, \ldots, T. \tag{11}$$

Proof: See Appendix A.1.

Note that the Poisson mean ϕ_u can be taken arbitrarily close to 1 for any $\bar{x}_u$ by setting $G(x)$ close to $F(x)$. Specifically, we consider a series of economies in which the number of informed traders is $N = N_0, N_0 + 1, \ldots$ and the distribution function G_N evolves depending on N. This yields the following lemma.

[5]This is an application of the result obtained by Nirei [28] in a simple model of interactions.

Lemma 5 *Suppose that the distribution functions G_N converge to F as $\lim_{N\to\infty}\sup_x|G_N(x)-F(x)|=0$. Then, $\phi_u=1$ for $u=2,3,\ldots,T$.*

It is known that when the Poisson mean takes a constant value, which is denoted by ϕ, the sum of the branching process m_T conditional on m_1 follows a Borel-Tanner distribution (Kingman [21]). That is,

$$\begin{aligned}\Pr(m_T=m\mid m_1=l) &= (l/m)e^{-\phi m}(\phi m)^{m-l}/(m-l)!, \quad m=l,l+1,\ldots &(12)\\ &\propto e^{-(\phi-1-\log\phi)m}m^{-1.5}, &(13)\end{aligned}$$

where the second line holds asymptotically as $m\to\infty$. This indicates that the sum of the Poisson branching process, conditional on the initial number of buying traders m_1, follows a power law distribution with an exponent of 0.5 with exponential truncation,[6] where the truncation point is determined by ϕ. Using this result, we can show that the distribution of m^* conditional on m_1, which is the number of traders who buy in the first round of the best-response dynamics (i.e., $m_1=\Gamma(0)$), has a power law tail. This is summarized by the following proposition.

Proposition 2 *If $\lim_{N\to\infty}\sup_x|G_N(x)-F(x)|=0$, the asymptotic distribution function (12) holds with $\phi=1$ for m^* conditional on $m_1=l$.*

Furthermore, we can characterize the unconditional distribution of m^* in the following environment. We assume that F follows an exponential distributions with positive mean $1/\mu$ and G_N follows an exponential distribution with mean $1/\mu_N$, where $\mu_N=\mu+1/\log N$. In this case, we can show that m^* asymptotically follows a distribution with a power law tail, which is summarized in the following proposition.

[6]Note that the power law exponent is defined for a cumulative distribution.

Proposition 3 *Given the specification of F and G_N above, there exists a constant c such that the minimum equilibrium number of buying traders m^* asymptotically follows*

$$\Pr(m^* = m) = \frac{ce^{-(m+c)}}{m!}(m+c)^{m-1}, \quad m = 0, 1, \ldots \tag{14}$$

$$\propto m^{-1.5}, \tag{15}$$

where the second line holds asymptotically as $m \to \infty$.

Proof: See Appendix A.2.

Proposition 3 shows the presence of a power law tail for the equilibrium number of buying traders multiplied by $1/N$, i.e., m^*/N, implying that the variance of m^*/N is very large. A necessary condition for this to happen is that m^*/N can take a wide range of different values in equilibrium. In other words, we need a property similar to equilibrium indeterminacy. Indeterminacy in signal inference games is best captured by Keynes's beauty contest, in which voters care about who is picked by other voters rather than who is actually beautiful, so that any candidate can win. Our model has a similar property as that of Keynes's beauty contest, which can be seen from the optimal threshold condition (8). This condition reduces to the simple form $(1 - \alpha) \log A(\bar{x}) + \alpha \log B(\bar{x}) = 0$, where $\alpha \equiv m/N$, if we take the limit as N approaches infinity while keeping α unchanged. The condition indicates that the log of the geometric average of A and B evaluated at $\bar{x}$, which can be regarded as a summary statistic for information on the true state revealed by traders' actions, does not change even when α takes different values.

To explain why this happens, suppose that a trader switches from not-buying to buying. Since α increases, the average of A and B declines, so that the optimal threshold declines. However, this in turn increases the average likelihood ratio, since traders learn that some of them must have received very bad signals from the fact that they still choose not to buy even when the threshold is lower. As a result, the impact of a change in α on the average of A and B is exactly canceled out, so that any value of α is compatible with the optimal threshold condition. Note that this means that a higher value of α in an equilibrium cannot necessarily be regarded as an indication that a larger number of traders now believe that the true state is more likely to be H.

It is important to note that the above mechanism to generate indeterminacy depends on the information structure adopted. Specifically, if there were substantial heterogeneity in the information structure as to who observes whose actions, a trader who is observed by many traders would give a strong cue for herding. A useful example is Banerjee's sequential herding model, in which agents observe only the actions of those agents who move before them. In this information structure, it is possible that the first mover's action cascades to all agents, with private signals of most agents unrevealed. An important implication of this is that intermediate outcomes between "herding" and "no herding" do not occur. This is in sharp contrast with our model in which a symmetric information structure is assumed, and that assumption makes it possible to have any degree of herding in equilibrium.

Proposition 3 claims not only that various levels of aggregate outcome is possible, but also that the frequency of the occurrence of large synchronized actions m^* has a particular regularity signified by a power law. With exponent 0.5, the power law implies that the variance of m^*/N decreases as $N^{-0.5}$ when the number of traders N becomes large. This makes contrast with the case when the traders act independently. When traders' choices (a_i) are independent, the central limit theorem predicts that m^*/N asymptotically follows a normal distribution, whose tail is thin and variance declines as fast as N^{-1}. The variance of m^*/N differs by factor $\sqrt{N}$ between our model and the model with independent choices. This signifies the effect of stochastic herding that magnifies the small fluctuations in the average of signals x_i. While a magnification effect occurs whenever traders' actions are correlated, it requires a particular structure in correlation for the magnification effect to cause the variance declining more slowly than N^{-1}. The magnification effect in our model is analogous to a long memory process, in which a large deviation from the long-run mean is caused by long-range autocorrelation. In our static model, the long-range correlation of traders' actions is captured by a best-response dynamics that converges to an equilibrium. We showed that the best-response dynamics generates the power law under the condition $\phi = 1$, which means that the mean number of traders induced to buy by a buying trader is 1. The indeterminacy of the beauty contest underlies the condition $\phi = 1$, because it implies that a trader's average action responds one-to-one to average actions. The power law exponent 0.5 in our model is closely related to the same exponent in the Inverse Gaussian distribution that characterizes the first passage time of a martingale.[7]

[7] The mathematical properties used in the derivations in this section have been known for long in probability theory. We embed the best-response dynamics in a branching process. The branching process is a stochastic integer process of population in which each parent in a generation bears a random number of children in the next generation. Lemma 4 shows that m^* is characterized as the sum of a branching process with state-dependent mean. Lemma 5 establishes that m^* follows the sum of a Poisson branching process with $\phi = 1$. The exponential tail holds for a large finite m_T either for the subcritical case $\phi < 1$ or the supercritical case $\phi > 1$ (Harris [17]). The speed of exponential truncation is determined by $|\phi - 1 - \log \phi|$ as in (13). The speed of the exponential decay slows down as ϕ becomes close to 1, and disappears when $\phi = 1$. At this critical level $\phi = 1$, the branching process becomes a martingale, and the distribution (12) has a power law tail.

4 Numerical results on volumes and returns distributions

In this section, we conduct numerical simulations of the model with a finite number of informed traders N. By the simulations, we confirm that the probability distribution of the number of buying traders m^* follows a power law, which was only asymptotically shown in the previous section. Moreover, we simulate equilibrium asset returns $\log p_{m^*} - \log p_0$. We show that the simulated distribution of returns exhibits a fat tail that matches well with an empirical returns distribution.

An important facet of the model to be specified is the supply function $S(p)$, which determines how the fluctuation of volumes is translated to the fluctuation of returns. In our model where informed traders' demands are absorbed by uninformed traders' supply, the elasticity of supply function determines the impact of demand shifts on the returns. The relation between an exogenous shift in transaction volume and a resulting shift in asset price is often called a price impact function. In this paper, we adopt a square-root specification of the price impact function. Namely, we specify the inverse supply schedule of uninformed traders as $p_k = p_0 + p_0(k/N)^\gamma$, for $k = 1, 2, \ldots, N$, with $\gamma = 0.5$. A micro-foundation for the square-root specification is provided by Gabaix et al. [14] in a Barra model of uninformed traders who have a mean-variance preference and zero bargaining power against informed traders. The square-root specification is commonly used for the price impact (e.g., Hasbrouck and Seppi [18]), and its parameter specification, $\gamma = 0.5$, falls within the empirically identified range of the price impact by Lillo et al. [24].

Other parts of the model are specified as follows. The distributions of signal, F and G, are specified as normal distributions. This provides an alternative specification to Proposition 3, wherein F and G_N were specified as exponential distributions. The mean of F and G is normalized to 1 and 0, respectively. F and G are further specified as having a common standard deviation σ. We set σ at 25 or 50, which is large relative to the difference in mean, 1. The large standard deviation relative to the mean difference captures the situation where the informativeness of signal x_i is small. We set the number of informed traders N at a finite but large value between 500 and 4000. The numerical simulation is implemented as follows. First, the optimal threshold $\bar{x}$ is computed. Second, a profile of private signals $(x_i)_{i=1}^N$ is randomly drawn for 100,000 times, and m^* and p_{m^*} are computed for each draw.

Figure 1 plots the inverse cumulative distribution of m^*/N for various parameter values of N and σ. The inverse distribution $\text{Pr}_>(m^*)$ is cumulated from above, and is thus 0 at $m^* = N$ and 1 at $m^* = 0$. The distribution is plotted in log-log scale, and thus, a linear line indicates a power law $\text{Pr}_>(m^*) \propto m^{*-\xi}$, where the slope of the linear line ξ is called the exponent of the power law. The simulated distributions appear linear for smaller values of m^*, and decay fast when m^*/N is close to 1. This conforms to the model prediction that m^* asymptotically follows a power law distribution. The simulated distribution is exponentially truncated due to the finiteness of N.

The asymptotic result in Propositions 2 predicted the exponent of power law ξ to be 0.5. As shown in the left panel of Figure 1, we observe that the power law exponent of the simulated m^* is roughly equal to 0.5 when $N = 1000$ and $\sigma = 25$. Proposition 3 also predicted that the power law distribution with exponent 0.5 asymptotically holds when $N \to \infty$ and $\sup |F - G_N| \to 0$. To confirm this, we simulate the model with a larger number of traders and the smaller informativeness of the signal. In order to set the informativeness smaller, we lower the mean difference $\mu_F - \mu_G$ between F and G from 1 to 0.7 or 0.5 with σ fixed. In Figure 1, we observe that the similar slope $\xi = 0.5$

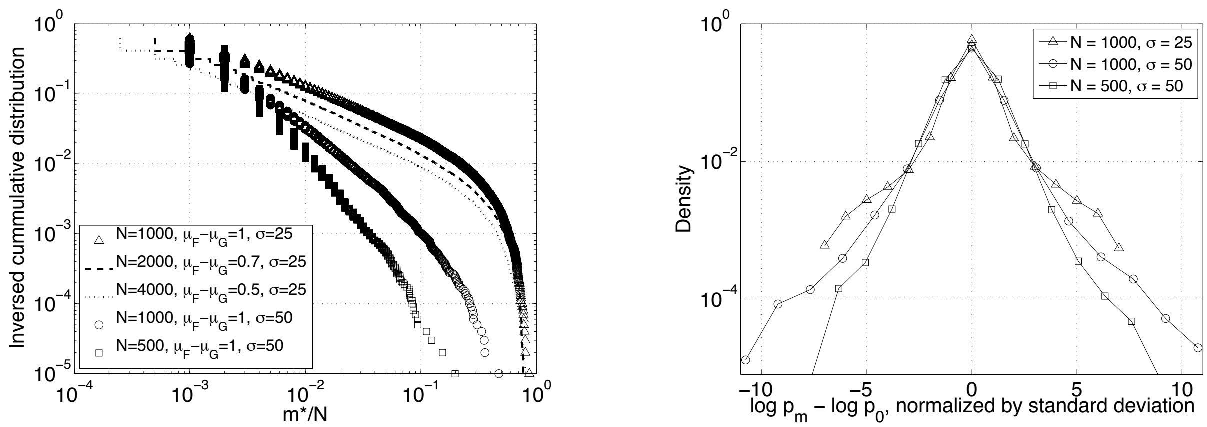

Figure 1: Left: Simulated inverse cumulative distributions of the minimum equilibrium number of buying traders m^*. N is the number of traders, σ is the standard deviation of the private information, and $\mu_F - \mu_G$ denotes the difference of the mean of the private information between F and G. Right: Simulated distributions of returns $\log p_m - \log p_0$.

for the power law holds for the cases with $N = 2000$ (dashed line) and 4000 (dotted line) when the informativeness of signal becomes smaller, confirming the prediction of Proposition 3. In the simulations under other parameter sets, however, we note that ξ can take larger values. This can be seen in the plot for a larger σ (circle-line) and a smaller N (square-line). This deviation in the exponent might result from the fact that the finite truncation occurs at a relatively small value of m^* in these cases. It is also possible that the state-dependence of ϕ_u is strong enough to cause a large deviation from the predicted exponent $\xi = 0.5$, since the power law exponent in (12) increases by 1 when the parameter ϕ fluctuates around the criticality value, 1, as shown by Sornette [33].

Our model also determines price p_{m^*} for each equilibrium number of buying traders m^*. We interpret the shifts in log price, $\log p_{m^*} - \log p_0$, as stock returns. We assume that the private signal is symmetric between two states, H and L. Thus, informed traders take demand side or supply side with probability 0.5. When the informed traders herd in the supply side, m^* is interpreted as the number of selling traders, and $\log p_m^* - \log p_0$ is interpreted as a negative of the associated stock return. We plot the distributions of the simulated returns in the right panel of Figure 1. The density is logarithmically scaled, and thus, a linear decline indicates an exponential distribution. Note that the returns are normalized by standard deviations. The normalized returns still span a wide range from -10 to 10. Thus, the plots well indicate that the simulated returns distributions exhibit the pattern of fat tails with exponential truncation.

The simulated distribution of returns is compared to the empirical distribution in Figure 2. The empirical distribution is generated using the daily returns data of TOPIX stock price index in the Tokyo Stock Exchange during 1998-2010. We define the daily return as the log difference from the opening price to the closing price. We use the opening-closing difference rather than the return in a business day in order to homogenize the time horizon of each observed return. The simulated distribution is generated under $N = 1000$. The standard deviation σ of the signal is set to 48.5, at which value the density estimate of simulated returns at 0 matches with that of the empirical distribution. The other parameters are set as before: $\gamma = 0.5$, $\mu_F = 1$, and $\mu_G = 0$.

In the left panel of Figure 2, the returns distributions are plotted in semi-log scale. The plot shows that the simulated distribution traces the empirical distribution rather well, especially in the left tail. In the same panel, we plot the standard normal density by a dotted line. Even though the simulated and empirical distributions are normalized by their standard deviations, the resulting distributions completely deviate from the normal distribution in the tail that is more than three standard deviations away from the mean. Note that we used σ as a free parameter in the simulation to match the empirical density at 0, but we did not use it to match the tail distribution. This indicates that our model is capable of generating the fat tail of empirical returns better than models that generate the normal distribution.

To further investigate the match between the simulated and empirical distributions, we show a Q-Q plot in the right panel of Figure 2. In the Q-Q plot, each quantile of the Topix returns data is plotted against the same quantile of the simulated returns data. Thus, the two distributions are identical if the Q-Q plot coincides with the 45 degree line, shown by a dashed line. Both quantiles are normalized by the standard deviations. In Figure 2, two quantiles resemble reasonably well in the overall support, although the simulated quantiles somewhat overshoot the empirical quantiles in the region greater than 2.

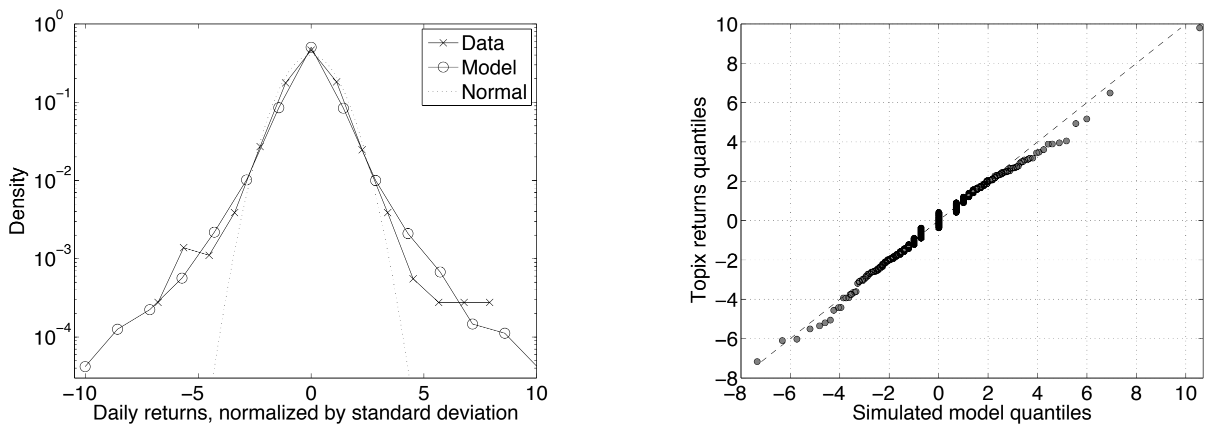

Figure 2: Distributions of TOPIX daily returns and simulated returns $\log p_m - \log p_0$. *Left:* Distributions plotted in semi-log scale, where returns are normalized by standard deviations. Empirical and model distributions are shown along with a standard normal distribution. *Right:* Quantile-quantile (Q-Q) plot. Each circle represents a pair of values, the simulated data value in the horizontal axis and the TOPIX data value in the vertical axis, under which the two distributions in comparison have the same fraction of the population.

5 Conclusion

This study analyzed aggregate fluctuations that arise from information inference behaviors among traders in financial markets. In a class of herd behavior models in which each trader infers other traders' private information only by observing their actions, we found that the number of traders who take the same action at equilibrium exhibits a large volatility with a regularity — a power law distribution. Furthermore, the model prediction was fitted to an empirical fat-tailed distribution of stock returns.

The power law distribution of aggregate actions emerges when the information structure of traders is symmetric. Every trader receives a private signal of the same magnitude of informativeness on the true value of an asset, and every trader observes the average action of all traders. Then, an action by a trader is as informative as an inaction by another. When some information is revealed by a trader's buying action, the inaction of other traders reveals their private information in favor of not buying. Thus, each trader's action is affected by the average action, resulting in a near-indeterminate equilibrium analogous to the Keynes's beauty contest. In this way, the information inference model provides an economic foundation for the criticality condition that generates a power law tailed fluctuations that were known in the models of critical phenomena.

This study provides a couple of directions for extension. A dynamically extended model is presented in Appendix B, which generates a time-series pattern similar to Lee [23] for sudden shifts in stock prices. As the present static model is shown to match with the quantitative properties of unconditional fluctuations, the natural next step is to bring the dynamic model to the time-series properties as pursued by, for example, Alfarano, et al. [1]. Another direction is to extend the model by incorporating more realistic market structures. Kamada and Miura [20] has taken a step toward this direction by extending this model to the case where both public and private signals exist and where informed traders can take both buying and selling sides.

Appendix

A Proofs

A.1 Lemma 4

Equation (9) implies that $\bar{x}(k) - \bar{x}(k+1)$ is of order $1/N$, by which we mean $O(1/N)$, namely that the term multiplied by N converges to a non-zero constant as $N \to \infty$. From Equation (10), we obtain that $q_u = (f(\bar{x}_{u-1})/F(\bar{x}_{u-1}))(\bar{x}(m_{u-1}) - \bar{x}(m_u)) + O(1/N^2)$. Thus, q_u is also of order $1/N$. The asymptotic mean of the binomial variable $m_{u+1} - m_u$ conditional on $m_u - m_{u-1} = 1$ is derived as follows:

$$\begin{aligned}
\phi_u &\equiv \operatorname*{plim}_{N\to\infty} q_u|_{m_u - m_{u-1}=1}(N - m_u) \\
&= \operatorname*{plim}_{N\to\infty} \frac{f(\bar{x}_{u-1})}{F(\bar{x}_{u-1})} \frac{\log A(\bar{x}_{u-1}) - \log B(\bar{x}_{u-1})}{\frac{N-m_{u-1}-1}{N}\frac{A'(\bar{x}_{u-1})}{A(\bar{x}_{u-1})} + \frac{m_{u-1}}{N}\frac{B'(\bar{x}_{u-1})}{B(\bar{x}_{u-1})}} \frac{-(N-m_u)}{N} &(16)\\
&= \operatorname*{plim}_{N\to\infty} \frac{\log A(\bar{x}_{u-1}) - \log B(\bar{x}_{u-1})}{\frac{N-m_{u-1}}{N}\left(1 - \frac{\ell(\bar{x}_{u-1})}{A(\bar{x}_{u-1})}\right) + \frac{m_{u-1}}{N}\frac{F(\bar{x}_{u-1})}{1-F(\bar{x}_{u-1})}\left(\frac{\ell(\bar{x}_{u-1})}{B(\bar{x}_{u-1})} - 1\right)} \frac{N-m_u}{N}, &(17)
\end{aligned}$$

where we used (9) and the fact that the difference of $\log p_k$ is of order $1/N$ for the second equation and (5) and (6) for the third equation. Note that m_u/N converges to $1 - F(\bar{x}_{u-1})$ with probability 1 for a fixed threshold $\bar{x}_{u-1}$ as $N \to \infty$ by the strong law of large numbers. Then, $(m_u/(N-m_u))(F(\bar{x}_{u-1})/(1-F(\bar{x}_{u-1})))$ converges to 1 with probability 1. Applying this to (17), we obtain the expression (11). Using that $\bar{x}(m_u) - \bar{x}(m_{u-1})$ is of order $1/N$, we obtain that $q_u(N - m_u) \to \phi_u(m_u - m_{u-1})$ for

$N \to \infty$. Hence, $m_{u+1} - m_u$ asymptotically follows a Poisson distribution with mean $\phi_u(m_u - m_{u-1})$, which is equivalently a $(m_u - m_{u-1})$-fold convolution of a Poisson distribution with mean ϕ_u. Thus, the binomial process asymptotically follows a branching process in which each parent bears a random number of children that follows a Poisson distribution with mean ϕ_u.

Define A_N, B_N, and ϕ_u^N similarly to A, B, and ϕ_u, respectively, by using G_N and F. Then, $A_N(\bar{x})/B_N(\bar{x}) \to 1$ holds since $A_N/B_N = (1/F - 1)/(1/G_N - 1)$. As the term A/B tends to 1, the first fraction in the right-hand side of (11) converges to 1 because of $A > \ell > B$, and the second fraction also converges to 1 by l'Hospital's rule. Then, $\phi_u^N \to 1$ obtains as $N \to \infty$.

A.2 Proposition 3

We show that m_1 asymptotically follows a Poisson distribution with some constant mean c as the number of traders N increases to infinity. Let $\bar{x}_N^1$ denote the threshold for traders to buy when there is no other trader buying in the economy with N traders. Using the exponential specification, the logarithm of Equation (7) that determines $\bar{x}_N^1$ is rewritten as:

$$\left(\frac{1/p_1 - 1}{\lambda}\right) / \left(1 + \frac{1}{\mu \log N}\right) = \left(1 + \frac{1 - e^{-\bar{x}_N^1/\log N}}{e^{\mu \bar{x}_N^1} - 1}\right)^{N-1} e^{-\bar{x}_N^1/\log N} \qquad (18)$$

Now define $R_N = \frac{1 - e^{-\bar{x}_N^1/\log N}}{e^{\mu \bar{x}_N^1} - 1}$. The first component of the right-hand side $(1 + R_N)^{N-1}$ converges to a constant if R_N declines as $1/N$. We make a guess that $\bar{x}_N^1/\log N$ converges to constant a. Suppose that $a < 1/\mu$ holds. Then, R_N declines slower than $1/N$, and thus, $(1 + R_N)^{N-1}$ diverges to infinity. Then, the right-hand side of (18) diverges, since its second component $e^{-\bar{x}_N^1/\log N}$ converges to a constant. Thus, this $\bar{x}_N^1$ cannot solve (18). Suppose that $a > 1/\mu$ holds. Then, R_N declines faster than $1/N$, and thus, $(1 + R_N)^{N-1}$ converges to 1. Noting that p_1 converges to p_0 as $N \to \infty$, (18) is solved at the limit of N if a satisfies $(1/p_0 - 1)/\lambda = e^{-a}$. However, this condition is not satisfied for $a > 1/\mu$, since $p_0 = 0.5$ and $\lambda = 1$. Finally, consider the case $a = 1/\mu$. We write $\bar{x}_N^1$ as $(\log N)/\mu + \epsilon_N$. Then, (18) at the limit becomes $((1/p_0 - 1)/\lambda)) \exp(1/\mu) = \exp((1 - e^{-1/\mu})/e^{\lim_{N\to\infty} \mu\epsilon_N})$. Thus, $\bar{x}_N^1 = (\log N)/\mu + \epsilon_N$ holds and ϵ_N converges to constant $(1/\mu)\log(\mu(1 - e^{-1/\mu}))$ as $N \to \infty$.

The asymptotic mean of m_1 is defined as $c \equiv \lim_{N\to\infty} N e^{-\mu\bar{x}_N^1}$. Substituting $\bar{x}_N^1 = (\log N)/\mu + \epsilon_N$, we obtain that $c = (\mu(1 - e^{-1/\mu}))^{-1} > 0$. Equation (14) is derived by mixing the Borel-Tanner distribution when $\phi = 1$ and the Poisson distribution with mean c. (For the explicit derivation of (14), see Nirei [28].) Finally, (15) is obtained by applying Stirling's formula.

B Dynamic extension of the model

The results in the basic model were derived under the assumption of homogeneous prior belief. The results hold even if the prior belief is heterogeneous. A particularly interesting case is when the belief evolves over time as private signals are drawn repeatedly. In this case, even though we maintain the assumption that the prior belief in the initial period is homogeneous, the belief in the subsequent periods will be heterogeneous due to the past private signals. In this sequence of static equilibria, we show that each static equilibrium in any period is characterized as in the basic static model.

The dynamic extension not only relaxes the assumption of common prior belief but also ensures that the propagation effect shown in the static model is triggered at some point of time. The limiting behavior of q_0, the mean number of traders who buy at p_0, when $N \to \infty$ was ambiguous in the static model except for the exponentially specified case of Proposition 3. This leaves a possibility that the chain reaction in the best-response dynamics is practically never triggered for a large N if $q_0 N \to 0$ holds. This is because the traders rarely react to the private signal when no other traders reveal their signals, if the prior belief is very low.

It turns out in the dynamic model that the traders eventually learn the true state as they accumulate private signals. This implies that, regardless of the level of the initial prior belief or N, the belief increases to the level at which traders start buying even though no other traders are buying. This triggers the chain reaction of buying. This dynamics is similar to the self-organized criticality advocated by Bak et al. [4] in the sense that traders' average belief converges to the state at which the size distribution of synchronized actions follows a power law.

B.1 Heterogeneous belief

We dynamically extend the basic model as follows. Each trader i draws private signal $x_{i,t}$ repeatedly over periods for $t = 1, 2, \ldots$. The private signal is identically and independently distributed across traders and periods. We consider the same asset as before that is worth 1 in H and 0 in L. Traders are given an opportunity to buy this asset in each period regardless of their past actions. Uninformed traders provide the supply function that has the same elasticity as in the static model and intercept $S(p_{t-1}) = 0$. The intercept p_{t-1} reflects the equilibrium price in the previous period, as it incorporates the information revealed to the public in that period. Informed traders submit their demand schedules to an auctioneer who clears the market each period as in the static model. To maximize the expected payoff of the transaction in t, trader i buys the asset in t if $\rho_{i,t} \geq (1-p_t)/p_t$ and does not buy otherwise. There is no dynamic aspect involved in the traders' decision other than updating the belief.

Let a_t and ρ_t denote the profiles of actions $(a_{i,t})_{i=1}^N$ and likelihood ratios $(\rho_{i,t})_{i=1}^N$, respectively. Informed traders observe their private signal history $x_i^t = (x_{i,1}, x_{i,2}, \ldots, x_{i,t})$. We study a sequence of static equilibria (p_t, a_t, ρ_t), $t = 1, 2, \ldots$, such that action $a_{i,t}$ maximizes trader i's expected period payoff under subjective belief $\rho_{i,t}$, which is consistent with the trader's observation.

The prior belief at the initial period is common and its likelihood ratio is denoted by λ. However, the belief is allowed to evolve stochastically as the traders draw signals repeatedly. Thus, the belief in each period $t > 1$ is heterogeneous across traders with a particular structure wherein the heterogeneity stems only from the distribution functions F and G that are ordered by MLRP.

Let a^t denote an action profile history $(a_1, a_2, \ldots, a_t)$. Let a_i^t denote $i's$ action history $(a_{i,1}, a_{i,2}, \ldots, a_{i,t})$. The set of possible a_i^t has 2^t elements, since $a_{i,\tau}$ is binary for any i and τ. Thus, for each a^{t-1}, all traders are divided into 2^{t-1} groups according to their action history a_i^{t-1}. Let $n_{k,t}$ denote the number of traders in the k-th group for $k = 1, 2, \ldots, 2^{t-1}$ (hence, $\sum_{k=1}^{2^{t-1}} n_{k,t} = N$), and $m_{k,t}$ denote the number of buying traders in the same group. Let $X_{k,s}$ for $s < t$ denote the domain of $x_{i,s}$ that is consistent with a_i^s under the threshold strategy defined below in (22) for trader i who belongs to group k. The likelihood ratios revealed by an action history of a non-buying trader i and a buying trader j in group k are written as follows:

$$A_{k,t} = \frac{\int_{X_{k,1}} \cdots \int_{X_{k,t-1}} G(\bar{x}_t(P_{k,t}, x_i^{t-1})) dG(x_{i,t-1}) \cdots dG(x_{i,1})}{\int_{X_{k,1}} \cdots \int_{X_{k,t-1}} F(\bar{x}_t(P_{k,t}, x_i^{t-1})) dF(x_{i,t-1}) \cdots dF(x_{i,1})}, \quad (19)$$

$$B_{k,t} = \frac{\int_{X_{k,1}} \cdots \int_{X_{k,t-1}} (1 - G(\bar{x}_t(P_{k,t}, x_j^{t-1}))) dG(x_{j,t-1}) \cdots dG(x_{j,1})}{\int_{X_{k,1}} \cdots \int_{X_{k,t-1}} (1 - F(\bar{x}_t(P_{k,t}, x_j^{t-1}))) dF(x_{j,t-1}) \cdots dF(x_{j,1})}, \quad (20)$$

where $P_{k,t}$ denotes the information inferred from observing $(m_{h,t})_h$ by a buying trader

i in group k,

$$P_{k,t} \equiv \left(\prod_h A_{h,t}^{n_{h,t}-m_{h,t}} B_{h,t}^{m_{h,t}} \right) / B_{k,t}. \tag{21}$$

Note that the right-hand side of (21) is divided by $B_{k,t}$ because $m_{k,t}$ includes the inferring trader i herself.

In the static model, informed traders submit the demand schedules conditional on p_m, where the conditioning on p_m is equivalent to conditioning on m. In the dynamically extended model, we assume that informed traders submit their demand schedules conditional on the vector of the number of buying traders in each group, $(m_{k,t})_k$.

We show that the equilibrium threshold strategy still exists in this setup.

Proposition 4 *For each realization of x^t, there exists an equilibrium outcome $(m_{k,t})$ and thresholds $\bar{x}_t$ such that the action profile a_t satisfies the optimal threshold rule:*

$$a_{i,t} = \begin{cases} 1 & \text{if } x_{i,t} \geq \bar{x}_t(P_{k,t}, x_i^{t-1}), \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

Proof: See Appendix B.3.

In the proof, we show that the threshold function is decreasing in the total number of buying traders in each period. Hence, as in the static model, we can define the best-response dynamics within each period, where a chain reaction of buying actions occurs. The best-response dynamics is characterised by a set of binomial distributions with population n_k and probability $q_{k,u}$ for each group k and step u in the best-response dynamics. When N is large, the sum of traders who buy in step u can be approximated by a Poisson with mean $\sum_k n_k q_{k,u}$. Thus, the best-response dynamics approximately follows a Poisson branching process as in the static model.

B.2 Self-organized criticality

In the dynamic extension, traders accumulate private signals that are independent across periods. Thus, through Bayesian learning by observing private signals and aggregate actions, traders eventually learn the true state almost surely.

Proposition 5 *The subjective belief $\rho_{i,t}$ converges to 0 as $t \to \infty$ almost surely.*

Proof: See Appendix B.4.

Proposition 5 means that the belief converges to the true state eventually. This is a natural consequence of the fact that traders have infinitely precise information in the long run as they accumulate their own private signals. The convergence of belief implies that there is no possibility for "wrong" herd behavior in the long run in the narrow sense that we have an infinite sequence of traders taking actions on the basis of a wrong belief or of traders completely neglecting their private information.

The convergence of belief to the true state H means that all traders will buy eventually. This implies that some traders start buying even without any other trader buying at some point of the process toward convergence. Such a buying action triggers the chain reaction of buying. Thus, the converging belief assures that the triggering actions eventually occur and almost surely cause the fat-tailed aggregate actions. The logic is analogous to the self-organized criticality proposed by Bak et al. [4]. In Bak's sand-pile model, the distribution of avalanche size depends on a slowly-varying variable (the slope of the sand pile), and the dynamics of the slope variable has a global sink exactly at the critical point at which the avalanche size exhibits a power law distribution. In our model, the average belief corresponds to the slope in the sand-pile model. The chain reaction is rarely triggered when the average belief is far below the threshold. As private information accumulates, the average belief increases toward the threshold.

This ensures that the triggering buying action will occur eventually.

B.3 Proof of Proposition 4

We define the threshold function $\bar{x}_t(P_{k,t}, x_i^{t-1})$ at which trader i is indifferent between buying and not buying. It is implicitly determined by

$$\frac{1}{p_t} - 1 = P_{k,t}\lambda\ell(\bar{x}_t)\prod_{\tau=1}^{t-1}\ell(x_{i,\tau}). \tag{23}$$

It follows that $\ell(\bar{x}_t(P_{k,t}, x_i^{t-1}))\prod_{\tau=1}^{t-1}\ell(x_{i,\tau})$ is equal to $(1/p_t - 1)/(\lambda P_{k,t})$, and thus, is constant across i in group k. Then, $A_{k,t} > (1/p_t - 1)/(\lambda P_{k,t}) > B_{k,t}$ can be shown as follows. The numerator of $A_{k,t}$ is expanded as

$$\int_{X_{k,1}} \cdots \int_{X_{k,t-1}} G(\bar{x}_t(P_{k,t}, x_i^{t-1}))\ell(x_{i,t-1})dF(x_{i,t-1})\cdots\ell(x_{i,1})dF(x_{i,1}) \tag{24}$$

$$> \int_{X_{k,1}} \cdots \int_{X_{k,t-1}} F(\bar{x}_t(P_{k,t}, x_i^{t-1}))\ell(\bar{x}_t(P_{k,t}, x_i^{t-1}))\Pi_{\tau=1}^{t-1}\left(\ell(x_{i,\tau})dF(x_{i,\tau})\right) \tag{25}$$

$$= \frac{1/p_t - 1}{\lambda P_{k,t}}\int_{X_{k,1}} \cdots \int_{X_{k,t-1}} F(\bar{x}_t(P_{k,t}, x_i^{t-1}))dF(x_{i,t-1})\cdots dF(x_{i,1}). \tag{26}$$

The integral in (26) is equal to the denominator of $A_{k,t}$, and thus, $A_{k,t} > (1/p_t - 1)/(\lambda P_{k,t})$ holds. Similarly, we obtain $B_{k,t} < (1/p_t - 1)/(\lambda P_{k,t})$.

Define $m_t \equiv \sum_h m_{h,t}$. Suppose that m_t increases due to an increase in $m_{k,t}$. $P_{k,t}$ decreases by $A_{k,t} > B_{k,t}$ and (21). Then, $\bar{x}_t$ needs to adjust in order to satisfy (23). By using $A_{k,t} > (1/p_t - 1)/(\lambda P_{k,t}) > B_{k,t}$, we obtain that the logarithms of $A_{k,t}$ and $B_{k,t}$ are decreasing in $\bar{x}_t$ as in the static model. Furthermore, MLRP ensures that $\ell(\bar{x}_t)$ is decreasing in $\bar{x}_t$. Thus, $\bar{x}_t$ in (23) decreases in response to the increase in m_t. The decreasing $\bar{x}_t$ entails a non-decreasing reaction function of m_{t+1} defined for each realization of x^t. Hence, the existence of an equilibrium is established by Tarski's fixed point theorem as in the static model. This completes the proof.

B.4 Proof of Proposition 5

A likelihood ratio for a private signal history, $\prod_{\tau=1}^{t} \ell(x_{i,\tau})$, converges to zero as in Billingsley [7]. The proof is outlined as follows. Likelihood ratio $\rho_{i,t}$ follows a martingale in the probability measure of the private signal under the true state: $E(\rho_{i,t} \mid \rho_{i,t-1}, H) = \rho_{i,t-1}$. Furthermore, the likelihood ratio is bounded from below at zero by construction. Then, the martingale convergence theorem asserts that the likelihood ratio converges in distribution to a random variable. Moreover, the probability measures for a sequence of private signal $(x_{i,1}, x_{i,2}, \ldots, x_{i,T})$ under H and L are mutually singular when $T \to \infty$, since $x_{i,t}$ is independent across t. Then, $\prod_{\tau=1}^{t} \ell_{i,\tau}$ converges to zero.

Hence, $\rho_{i,t}$ converges to zero if $P_{k,t}$ remains finite for $t \to \infty$. $P_{k,t}$ is finite for a finite $\bar{x}_t$ when N is finite. When $\bar{x}_t$ tends to a positive infinity, $P_{k,t}$ decreases to a finite value since $A_{k,t}$ and $B_{k,t}$ are decreasing in $\bar{x}_t$ and positive. When $\bar{x}_t$ tends to a negative infinity, all traders eventually choose to buy. Hence, $\prod_h A_{h,t}^{n_{h,t}-m_{h,t}}$ tends to one, and $P_{k,t}$ tends to $\prod_h B_{h,t}^{m_{h,t}}/B_{k,t}$. We showed that $B_{k,t} < (1/p_t - 1)/(\lambda P_{k,t})$ in the proof of Proposition 4. If $P_{k,t}$ tends to a positive infinity as $\bar{x}_t$ tends to a negative infinity, then, this inequality contradicts the fact that $P_{k,t}$ tends to $\Pi_h B_{h,t}^{m_{h,t}}/B_{k,t}$ for any finite N. Thus, $P_{k,t}$ is finite as $t \to \infty$. Hence, $\rho_{i,t}$ is dominated by private signal as $t \to \infty$ and converges to zero.

References

[1] Simone Alfarano, Thomas Lux, and Friedrich Wagner. Time variation of higher moments in a financial market with heterogeneous agents: An analytical approach. *Journal of Economic Dynamics & Control*, 32:101–136, 2008.

[2] Christopher Avery and Peter Zemsky. Multidimensional uncertainty and herd behavior in financial markets. *American Economic Review*, 88:724–748, 1998.

[3] P. Bak, M. Paczuski, and M. Shubik. Price variations in a stock market with many agents. *Physica A*, 246:430–453, 1997.

[4] Per Bak, Kan Chen, José Scheinkman, and Michael Woodford. Aggregate fluctuations from independent sectoral shocks: Self-organized criticality in a model of production and inventory dynamics. *Ricerche Economiche*, 47:3–30, 1993.

[5] Abhijit V. Banerjee. A simple model of herd behavior. *Quarterly Journal of Economics*, 107:797–817, 1992.

[6] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100:992–1026, 1992.

[7] Patrick Billingsley. *Probability and measure*. Wiley, NY, third edition, 1995.

[8] Jean-Philippe Bouchaud and Marc Potters. *Theory of Financial Risks*. Cambridge University Press, 2000.

[9] Lluís Bru and Xavier Vives. Informational externalities, herding, and incentives. *Journal of Institutional and Theoretical Economics*, 158:91–105, 2002.

[10] Andrew Caplin and John Leahy. Business as usual, market crashes, and wisdom after the fact. *American Economic Review*, 84:548–565, 1994.

[11] V.V. Chari and Patrick J. Kehoe. Financial crises as herds: overturning the critiques. *Journal of Economic Theory*, 119:128–150, 2004.

[12] Rama Cont and Jean-Philipe Bouchaud. Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic Dynamics*, 4:170–196, 2000.

[13] Eugene F. Fama. The behavior of stock-market prices. *Journal of Business*, 38:34–105, 1965.

[14] Xavier Gabaix, Parameswaran Gopikrishnan, Vasiliki Plerou, and H. Eugene Stanley. Institutional investors and stock market volatility. *Quarterly Journal of Economics*, 121:461–504, 2006.

[15] Lawrence R. Glosten and Paul R. Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14:71–100, 1985.

[16] Faruk Gul and Russell Lundholm. Endogenous timing and the clustering of agents' decisions. *Journal of Political Economy*, 103:1039–1066, 1995.

[17] Theodore E. Harris. *The Theory of Branching Processes*. Dover, NY, 1989.

[18] Joel Hasbrouck and Duane J. Seppi. Common factors in prices, order flows, and liquidity. *Journal of Financial Economics*, 59:383–411, 2001.

[19] Dennis W. Jansen and Casper G. de Vries. On the frequency of large stock returns: Putting booms and busts into perspective. *Review of Economics and Statistics*, 73:18–24, 1991.

[20] Koichiro Kamada and Ko Miura. Confidence erosion and herding behavior in bond markets: An essay on central bank communication strategy. *Bank of Japan Working Paper Series*, No.14-E6, 2014.

[21] J. F. C. Kingman. *Poisson Processes.* Oxford, NY, 1993.

[22] Alan Kirman. Ants, rationality, and recruitment. *Quarterly Journal of Economics*, 108:137–156, 1993.

[23] In Ho Lee. Market crashes and informational avalanches. *Review of Economic Studies*, 65:741–759, 1998.

[24] Fabrizio Lillo, J. Doyne Farmer, and Rosario N. Mantegna. Master curve for price-impact function. *Nature*, 421:129–130, 2003.

[25] Benoit Mandelbrot. The variation of certain speculative prices. *Journal of Business*, 36:394–419, 1963.

[26] Rosario N. Mantegna and H. Eugene Stanley. *An Introduction to Econophysics.* Cambridge University Press, 2000.

[27] Deborah Minehart and Suzanne Scotchmer. Ex post regret and the decentralized sharing of information. *Games and Economic Behavior*, 27:114–131, 1999.

[28] Makoto Nirei. Threshold behavior and aggregate fluctuation. *Journal of Economic Theory*, 127:309–322, 2006.

[29] Makoto Nirei. Self-organized criticality in a herd behavior model of financial markets. *Journal of Economic Interaction and Coordination*, 3:89–97, 2008.

[30] David S. Scharfstein and Jeremy C. Stein. Herd behavior and investment. *American Economic Review*, 80:465–479, 1990.

[31] Lones Smith. Do rational traders frenzy?, 1997. mimeo.

[32] Lones Smith and Peter Sørensen. Pathological outcomes of observational learning. *Econometrica*, 68:371–398, 2000.

[33] Didier Sornette. *Critical phenomena in natural sciences.* Springer, second edition, 2004.

[34] Dietrich Stauffer and Didier Sornette. Self-organized percolation model for stock market fluctuations. *Physica A*, 271:496–506, 1999.

[35] Xavier Vives. Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics*, 19:305–321, 1990.

[36] Xavier Vives. The speed of information revelation in a financial market mechanism. *Journal of Economic Theory*, 67:178–204, 1995.