Information Aggregation and Fat Tails in Financial Markets

Makoto Nirei

Hitotsubashi University

November 30, 2010

Abstract

This paper demonstrates that fat-tailed distributions of trading volume and stock price emerge in a model of information aggregation. I consider a simultaneous-move model of traders who infer other traders' private information on the value of assets by observing aggregate actions. Without parametric assumptions on the private information, I show that the traders' aggregate actions follow a power-law distribution with exponential truncation. By combining the power law with the composition uncertainty that a market maker faces, the model is able to mimic the non-normal distribution of the stock returns.

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

JEL classification code: G14

1 Introduction

Since Mandelbrot [26] and Fama [14], it has been well established that the short-run stock returns exhibit a fat-tailed, leptokurtic distribution. Jansen and de Vries [20], for example, estimated the exponent of the power-law tail to be in the range of 3 to 5, which warrants a finite variance and yet deviates greatly from the normal distribution in the fourth moment. This anomaly in the tail and kurtosis has been considered as a reason for the excess volatility of stock returns.

There has been a long quest for the explanation for the anomaly. A traditional economic explanation for the excess volatility of the volumes and returns relies on traders' rational herd behavior. In a situation where a trader's private information on the asset value are partially revealed by her transaction, the trader's action can cause an avalanche of similar actions by the other traders. This idea of chain reaction through the revelation of private information has been extensively studied in the literature of herd behavior, informational cascade, and information aggregation. However, there have been few attempts to explain the fat tail in this framework. This paper shows that the chain reaction of information revelation leads to the fat-tail distributions of the traders' aggregate actions and asset returns.

I consider a model of a large number of informed traders who receive imperfect private information on the true value of an asset. The traders simultaneously choose whether to buy one unit of the asset or not to buy at all. I consider a rational expectations equilibrium in which each trader chooses an action based on her private information as well as the public information produced by a market maker. The market maker produces the public information by aggregating the traders' demand schedules submitted to the market maker. The equilibrium is a mapping from the space of private information of all traders to the aggregate actions. The larger the aggregate action becomes, the more the traders put subjective belief on the state of a high asset value, and the more likely the traders buy the asset. Hence, the traders' strategy exhibits complementarity, and their actions are positively correlated.

I derive the probability distribution of the equilibrium aggregate action and show that it decays as a power function with exponential truncation. The speed of the exponential truncation is determined by the degree of the strategic complementarity among traders. This analytical result is obtained by a new method that utilizes a fictitious stochastic tatonnement process to characterize the aggregate actions. The power-law distribution implies a large kurtosis. This illustrates that a significant magnitude of aggregate risk exists even when the uncertainty of the economy solely stems from the idiosyncratic private information drawn by a large number of traders.

I extend the model dynamically in which traders receive private information repeatedly. Suppose that the initial belief started far below the threshold belief. Then, traders buy only if they receive an extremely good news. As private information is accumulated over time, however, the average belief increases toward the threshold. At the threshold, an aggregate action follows a pure power-law distribution, and thus a herding occurs. This implies that a large amount of private information tends to be shared at once around the point of time when the average belief reaches the threshold. Thus, even though the subjective belief converges to the true value of asset in the long run, the price process toward the true value can deviate significantly from the smooth path which would occur if the private information is fully revealed each period. A sizable portion of the total price adjustment toward the true value is accounted for by rare events of synchronized actions of a large number of traders.

The model can account for the non-normal distribution of stock returns. I modify the basic model by incorporating a market maker who sets the price under the presence of the composition risk in the sense of Avery and Zemsky [1]. The market maker publishes the price which reflects all the information available to it, and absorbs any excess demand. The traders buy or sell the stock if their beliefs are above or below the price. I show that this model generates the distribution of stock returns which resembles the empirical distribution, although the current model is elementary as a trading model yet.

As a herding model, our model is analogous to the Keynes' beauty contest. Each trader recognizes that the other traders have the private information that is as valuable as her own. When each trader tries to match with the behavior of an average trader, the resulting equilibrium exhibits fragility due to the perfect strategic complementarity. This paper formalizes the idea of the perfect strategic complementarity among the traders with private information, and shows that a power law distribution of the aggregate actions emerges naturally in this setup.

An extensive array of literature addresses the issue of imitative behavior in financial markets. The models of herd behavior and informational cascade by Scharfstein and Stein [31], Banerjee [4], and Bikhchandani, Hirshleifer, and Welch [5] have been applied to financial market crashes by Lee [24] and Chari and Kehoe [10] among others. While the benchmark herd behavior model provides a robust intuition for the rational herding, it typically exhibits an all-or-nothing herding due to its particular information structure of a sequential trading. Some modification is in due in order to apply its intuition to stochastic fluctuations. Gul and Lundholm [17], for example, have demonstrated an emergence of stochastic herding by endogenizing traders' choice of waiting time. I extend this line of research by employing a simple simultaneous-move model of traders. Our approach is closely related to Caplin and Leahy [9] who argue that the aggregate revelation of dispersed information in the market tends to occur suddenly as the last straw that breaks the camel's back. Stretching their analogy, this paper claims that, when the camel's back breaks, the rupture size is distributed according to a power law. Another underlying theme of this paper is the aggregation of private information (Vives [37]) or idiosyncratic shocks (Jovanovic [21], Durlauf [13]). This paper shows that the aggregation of private information in the market leads to a non-trivial, structured fluctuation that is characterized by a power law.

The technical analysis I employ is linked to the field of critical phenomena in statistical physics. Recently, a number of statistical physicists investigated the empirical fluctuations of financial markets.[1]

[1] Survey of this attempt is provided by Bouchaud and Potters [7] and Mantegna and Stanley [27].

Some papers in this literature reproduce the empirical power-laws by introducing the methodology used for critical phenomena to the herd behavior models (Bak, Paczuski, and Shubik [2]; Cont and Bouchaud [11]; Stauffer and Sornette [35]). Two questions have been raised for these models of critical phenomena. One is that they lack the model of traders' purposeful behavior and rational learning, which hinders the integration of their methodology to the existing body of financial studies. The other is a fundamental question as to why at all the market has to exhibit criticality. The power-law fluctuation occurs typically only at the critical value of a parameter which governs the connectivity of the networked traders. Gabaix, Gopikrishnan, Plerou, and Stanley [15] address these questions by incorporating the trader's optimal behavior and by relating the power laws for the volumes and returns to Zipf's law for the size distribution of firms. This paper proposes an alternative by showing that the market necessarily converges to the critical point as a result of the purposeful behavior of individual traders who learn information from each other.

The remainder of the paper is organized as follows. In Section 2, a simple static model is presented. Section 3 analytically derives the power-law distribution and provides an intuition for the mechanism behind the fat tail. The model is also extended dynamically, and the power-law distribution is shown to occur at the state into which heterogeneous beliefs of traders evolve. Section 4 shows by numerical simulations that the equilibrium volumes follow a power law and the equilibrium returns distribution matches with the empirical counterpart. Section 5 discusses the role of discrete actions and symmetric information structure in the model. Section 6 concludes.

2 Model

2.1 Model and equilibrium

In this section, I consider the simplest case in which each trader receives private information just once. I will extend the analysis to the case of repeated information in the next section. I consider a financial market with a large number N of traders. Each trader receives private information on the state of the economy. The state takes one of two possible values: H and L. The value of an asset is high in state H and low in state L. Traders do not know which is the true state, but have a prior subjective probability $b_{i,0}$ for H to occur. Trader i receives private information x_i, which is a random variable identically and independently distributed across traders. Traders cannot observe the other traders' private information. The information x_i is drawn from a known distribution F in state H and from G in state L. I assume that the true state is H throughout this paper.

I consider an asset which is worth 1 in H and 0 in L. A share of the asset is indivisible. Traders choose either to buy it or not. The situation where traders choose to sell or not can be symmetrically analyzed. Trader i's buying action is denoted by $a_i = 1$ and the non-buying action is denoted by $a_i = 0$. Let m denote the number of the buying traders, $m = \sum_{i=1}^{N} a_i$, and let α denote the fraction of the buying traders, $\alpha = m/N$. The action profile and the information profile are denoted by $a = (a_1, a_2, \ldots, a_N)$ and $x = (x_1, x_2, \ldots, x_N)$, respectively.

The price of the asset $\bar{b}$ is set by a market maker. The market maker is committed to provide the asset at the price before it recognizes that the new private information is being drawn by the traders. Namely, there is an event uncertainty in the sense of Avery and Zemsky [1] at which the market maker is unaware of the news. I consider that the market maker sets the price at the expected value of the asset evaluated by the information available publicly. Thus, the market maker sets the price at the common prior belief, b_0, which is shared by the market maker and the traders.

Each trader i submits her demand schedule s_i to the market maker. The demand schedule specifies action a_i conditional on public information P. The public information P is the likelihood ratio for an action profile a to occur evaluated at state L divided by that evaluated at state H. P is produced by the market maker inferring the information revealed by the traders' actions at equilibrium. Thus, in this market, traders provide information to the market maker, and in return traders benefit by the fixed price committed by the market maker. I assume that the traders do not believe that their demand schedule affects P, since there are a large number of traders. Upon observing x_i and conditional on P, each trader rationally updates their belief according to Bayes' rule, and obtains the posterior belief $b_{i,1}$. The traders are risk-neutral and maximize their subjective expected payoff. The expected payoff of trader i is 0 when $a_i = 0$ regardless of the belief, whereas it is equal to $b_{i,1} - \bar{b}$ when $a_i = 1$.

I define a rational expectations equilibrium as a mapping from the space of private information profile x to the space of a pair of an action profile and public information (a, P) such that:

1. the action a_i maximizes i's expected payoff

2. the public information P is consistent with a

3. the belief $b_{i,1}$ is consistent with x_i and P according to Bayes' rule.

Trader i buys the asset if and only if $b_{i,1} \geq \bar{b}$. The rational expectations equilibrium is implemented by the market maker who chooses the equilibrium α at which the equilibrium outcome $\sum_{i=1}^{N} a_i(P)/N$ is consistent with the submitted demand schedule. If there are multiple equilibria, I assume that the market maker selects the smallest α among them.

Our equilibrium is implemented by the market maker who aggregates the demand schedule submitted by traders. This implementation is a version of a herding model with rational expectations flavor (Bru and Vives [8]). Without the market maker, our model is similar to Minehart and Scotchmer [29], who showed that the traders cannot agree to disagree in a rational expectations equilibrium, i.e., the equilibrium may not exist, or if it exists it is a herding equilibrium where all the traders choose the same action. A key difference from Minehart and Scotchmer [29] is that traders in our model do not believe their actions can affect the public information. This assumption is appropriate when there are a large number of traders in the market. With the market maker, our model captures auctioning situations where the market maker announces the public information, the traders respond by buying or not, the market maker announces an adjusted public information, and then the process is iterated until no traders respond by buying.

2.2 Information structure and optimal strategy

I impose two assumptions on the information structure. First, the prior belief is common across traders and the market maker: $b_{i,0} = b_0$. This assumption is made for the sake of simplicity, and it is relaxed later in Section 3.2 where the belief is allowed to evolve heterogeneously over periods. Secondly, we assume that the private information has the monotone likelihood ratio property (MLRP). Define an odds function $\delta(x_i) = g(x_i)/f(x_i)$, where f and g are density functions of F and G, respectively. MLRP requires δ to be monotone. Without loss of generality, we assume that δ is strictly decreasing. Namely, a larger x_i implies a larger likelihood of H. This assumption is also standard.

Bayes' rule is equivalently expressed in terms of a likelihood ratio $\theta_{i,t} \equiv (1 - b_{i,t})/b_{i,t}$. The traders update the likelihood ratio as:

$$\theta_{i,1} = P\delta(x_i)\theta_{i,0} \tag{1}$$

The optimality condition for a buying action $b_{i,1} \geq \bar{b}$ is equivalent to $\theta_{i,1} \leq 1/\bar{b} - 1$. Thus, trader j

optimally obeys a threshold rule:

$$a_j = \begin{cases} 1 & \text{if } x_j \geq \bar{x}(P) \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $\bar{x}(P)$ denotes the value of private information at which trader j is indifferent between buying and not buying conditional on P:

$$\bar{x} = \delta^{-1}\left((1/\bar{b} - 1)/(P\theta_{i,0})\right). \tag{3}$$

Under the threshold rule, the likelihood ratios revealed by an inaction ($a_j = 0$) and by a buying action ($a_j = 1$) are derived as follows, respectively,

$$A(\bar{x}) \equiv \frac{\Pr(x_i < \bar{x} \mid L)}{\Pr(x_i < \bar{x} \mid H)} = \frac{G(\bar{x})}{F(\bar{x})} \tag{4}$$

$$B(\bar{x}) \equiv \frac{\Pr(x_j \geq \bar{x} \mid L)}{\Pr(x_j \geq \bar{x} \mid H)} = \frac{1 - G(\bar{x})}{1 - F(\bar{x})}. \tag{5}$$

As in Smith and Sørensen [33], MLRP implies that, for any value of $\bar{x}$ in the interior of the support of F and G,

$$A(\bar{x}) > \delta(\bar{x}) > B(\bar{x}) > 0, \tag{6}$$

and that $A(\bar{x})$ and $B(\bar{x})$ are strictly decreasing in $\bar{x}$:

$$\frac{dA}{d\bar{x}} = \frac{g(\bar{x})}{F(\bar{x})} - \frac{G(\bar{x})f(\bar{x})}{F(\bar{x})^2} = \frac{f(\bar{x})}{F(\bar{x})}\left(\delta(\bar{x}) - A(\bar{x})\right) < 0 \tag{7}$$

$$\frac{dB}{d\bar{x}} = -\frac{g(\bar{x})}{1 - F(\bar{x})} + \frac{(1 - G(\bar{x}))f(\bar{x})}{(1 - F(\bar{x}))^2} = \frac{f(\bar{x})}{1 - F(\bar{x})}\left(B(\bar{x}) - \delta(\bar{x})\right) < 0. \tag{8}$$

Provided with the threshold rule, the public information revealed at equilibrium outcome α is:

$$P = A(\bar{x})^{N(1-\alpha)} B(\bar{x})^{N\alpha} \tag{9}$$

where $A^{N(1-\alpha)}$ represents the likelihood ratio revealed by $N(1-\alpha)$ not-buying traders and $B^{N\alpha}$ represents that by buying traders.

The optimal threshold $\bar{x}$ and public information P are simultaneously determined by (3,9) for each α. The threshold $\bar{x}$ is strictly increasing in P in (3), since δ is strictly decreasing. P in (9) is strictly decreasing in $\bar{x}$ due to the inequalities (7,8). Therefore, the threshold $\bar{x}$ exists uniquely for each α. Let $\bar{x}(\alpha)$ denote the optimal threshold for the level of public information produced when the fraction of buying traders is α.

The existence of an equilibrium is shown as follows. Define $\mathcal{S} = \{0, 1/N, 2/N, \ldots, 1\}$ as a set of possible equilibrium outcome α, and define a reaction function $\Gamma : \mathcal{S} \mapsto \mathcal{S}$ for each realization of x such that $\alpha' = \Gamma(\alpha)$ is the fraction of traders with $x_i \geq \bar{x}(\alpha)$.

Proposition 1 *For each realization of x, there exists an equilibrium outcome α in which the action profile satisfies the optimal threshold rule (2).*

Proof: Combining (3) and (9) yields:

$$A(\bar{x})^{N(1-\alpha)} B(\bar{x})^{N\alpha} \delta(\bar{x}) \theta_{i,0} = 1/\bar{b} - 1. \tag{10}$$

By taking logarithm of the both sides of this equation and then taking a total derivative, I obtain:

$$\frac{d\bar{x}}{d\alpha} = \frac{\log A(\bar{x}) - \log B(\bar{x})}{(1-\alpha)A'(\bar{x})/A(\bar{x}) + \alpha B'(\bar{x})/B(\bar{x}) + \delta'(\bar{x})/(\delta(\bar{x})N)}. \tag{11}$$

The right hand side is strictly negative, since $A > \delta > B > 0$, $A' < 0$, $B' < 0$, and $\delta' < 0$. Then $\bar{x}$ is decreasing in α. Since $\Gamma(\alpha)$ is the fraction of traders with $x_i \geq \bar{x}(\alpha)$, the decreasing $\bar{x}$ implies that Γ is non-decreasing in α for any realization of x. 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 by Tarski's fixed point theorem. □

The threshold strategy (2) allows multiple equilibria for each realization of x. Here I focus on the equilibrium selected by the market maker, which has a minimum fraction of buying traders, α^*, among possible equilibria for each x. By this equilibrium selection, I show that even the minimum shift in equilibrium exhibits large fluctuations. The equilibrium selection maps each realization of x to α^*. Thus α^* is a random variable if viewed unconditionally on x, and its probability distribution is determined by the probability measure of x and the equilibrium selection mapping. I further characterize the probability distribution of α^* in the next section.

3 Analytical results

3.1 Derivation of the power law

In this section, I analytically derive the power-law distribution of the minimum aggregate action α^* defined in the model. I propose a method to characterize the distribution of α^* by using a fictitious tatonnement process. In so doing, I clarify the condition for the power law of α^* and provide an economic interpretation for the mechanism that generates the power law.

The minimum equilibrium α^* is known to be reached by the best response dynamics of traders, which Vives [39] called an informational tatonnement. The best response dynamics requires the traders to know only the "aggregate" information, α in our case. Cooper [12] argued that the parsimonious informational requirement of the best response dynamics makes it a desirable equilibrium selection algorithm when there exist multiple equilibria and when it is difficult for traders to coordinate their actions.

I start by showing that the informational tatonnement converges to α^*. I continue to use the notation developed for Proposition 1 such as $\mathcal{S}$ for the set of α and Γ for the aggregate reaction function on α.

Proposition 2 *Consider an informational tatonnement process α^u, where $\alpha^0 = 0$ and $\alpha^u = \Gamma(\alpha^{u-1})$ for $u = 1, 2, \ldots, T$, where the stopping time T is the smallest u such that $m^u - m^{u-1} = 0$. Then, α^u converges to the minimum equilibrium α^* for each realization of x. Also, the threshold decreases over the informational tatonnement process: $\bar{x}(\alpha^{u+1}) < \bar{x}(\alpha^u)$ for any realization of x.*

Proof: I can apply Vives' result [38] to show that this tatonnement always reaches a fixed point α^T of Γ, since Γ is increasing, $\mathcal{S}$ is finite, and $\alpha^0 = 0$ is the minimum in $\mathcal{S}$. Also, α^T must coincide with the minimum fixed point α^* for the following reason. Suppose that there exists another fixed point α that is strictly smaller than α^T. Then I can pick $u < T$ such that $\alpha^u < \alpha < \alpha^{u+1}$. Applying the non-decreasing function Γ, we obtain $\Gamma(\alpha^u) \leq \Gamma(\alpha)$. Then $\alpha^{u+1} \leq \alpha$. This contradicts to $\alpha < \alpha^{u+1}$. □

Since the informational tatonnement starts from $\alpha^0 = 0$ and converges to α^*, I can express α^* as the sum of increments of the tatonnement process. Moreover, the informational tatonnement can be regarded as a stochastic process, once it is viewed unconditionally on the private information x. Thus, α^* can be expressed as the sum of a stochastic process that starts from and converges to zero. The idea of characterizing an equilibrium outcome by a stochastic process is similar in spirit to Kirman [23]. Note that the tatonnement does not impose any restriction on α^*, since I have uniquely defined it as the minimum equilibrium aggregate action. I simply use the tatonnement as an instrument to characterize α^*.

By showing that the threshold $\bar{x}$ decreases over the stochastic process, Proposition 2 establishes that there exists a non-trivial chance of chain reaction during the tatonnement. A trader who chooses to buy in u will continue to choose to buy in $u+1$, since the threshold is decreased. A trader who does not buy in u might choose to buy in $u+1$. The conditional probability of a non-buying trader switching to buying in response to $\alpha^u - \alpha^{u-1}$ is defined as follows:

$$p^u \equiv \int_{\bar{x}^u}^{\bar{x}^{u-1}} f(x)dx / F(\bar{x}^{u-1}) \tag{12}$$

where $\bar{x}^u$ is a shorthand for $\bar{x}(\alpha^u)$. The p^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 tatonnement history up to u, follows a binomial distribution with population parameter $N - m^u$ and probability parameter p^u. The distribution of m^1 follows a binomial distribution with population N and probability $p^0 \equiv 1 - F(\bar{x}^0)$. This completely defines the stochastic tatonnement process, as summarized in the following proposition.

Proposition 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 p^u which is determined by (12) and $\bar{x}^u = \bar{x}(m^u/N)$. Also suppose that m^1 follows a binomial distribution with population N and probability p^0. Then, the minimum equilibrium number of buying traders $m^* = \alpha^* N$ follows the same distribution as the cumulative sum of the process: m^T.*

Proposition 3 establishes that the minimum equilibrium $m^* = \alpha^* N$ is equal to the sum of a binomial process. A binomial distribution permits Poisson approximation when the population is "large" and the probability is "small". The approximation holds in our tatonnement if the probability p^u is of order $1/N$. I show this being the case.

Proposition 4 *The binomial process $m^{u+1} - m^u$ asymptotically follows a branching process with a state-dependent Poisson random variable with mean:*

$$\phi^u = \frac{A(\bar{x}^{u-1})}{\delta(\bar{x}^{u-1})} \frac{\log A(\bar{x}^{u-1}) - \log B(\bar{x}^{u-1})}{A(\bar{x})^{u-1}/B(\bar{x}^{u-1}) - 1}. \tag{13}$$

Moreover, $\phi^u \to 1$ when $\sup_{\bar{x}}(G(\bar{x}) - F(\bar{x})) \to 0$.

Proof: Equation (11) indicates that $d\bar{x}/dm$ is of order $1/N$. From Equation (12), I obtain that $p^u = (f(\bar{x}^{u-1})/F(\bar{x}^{u-1}))(d\bar{x}/dm)|_{m^{u-1}}(m^{u-1} - m^u) + O(1/N^2)$. Thus, p^u is also of order $1/N$ and its leading term is proportional to $m^u - m^{u-1}$. The asymptotic mean of the binomial variable $m^{u+1} - m^u$ conditional on $m^u - m^{u-1} = 1$ is derived as follows:

$$\begin{aligned}
\phi^u &\equiv \text{plim}_{N\to\infty} p^u|_{m^u - m^{u-1}=1}(N - m^u) \\
&= \text{plim}_{N\to\infty} \frac{f(\bar{x}^{u-1})}{F(\bar{x}^{u-1})} \frac{\log A(\bar{x}^{u-1}) - \log B(\bar{x}^{u-1})}{(1-\alpha^{u-1})A'(\bar{x}^{u-1})/A(\bar{x}^{u-1}) + \alpha^{u-1}B'(\bar{x}^{u-1})/B(\bar{x}^{u-1})} \frac{-(N - m^u)}{N} \quad (14) \\
&= \text{plim}_{N\to\infty} \frac{\log A(\bar{x}^{u-1}) - \log B(\bar{x}^{u-1})}{(1-\alpha^{u-1})\left(1 - \frac{\delta(\bar{x}^{u-1})}{A(\bar{x}^{u-1})}\right) + \alpha^{u-1}\frac{F(\bar{x}^{u-1})}{1-F(\bar{x}^{u-1})}\left(\frac{\delta(\bar{x}^{u-1})}{B(\bar{x}^{u-1})} - 1\right)}(1-\alpha^u) \quad (15)
\end{aligned}$$

where I used (11) for the second line and (7,8) for the third line. Note that m^{u+1}/N converges to $1 - F(\bar{x}^u)$ with probability 1 for a fixed threshold $\bar{x}^u$ as $N \to \infty$ by the strong law of large numbers. Then, $(\alpha^{u+1}/(1-\alpha^{u+1}))(F(\bar{x}^u)/(1-F(\bar{x}^u)))$ converges to 1 with probability 1. Also $\alpha^u - \alpha^{u-1} = 1/N$ when $m^u - m^{u-1} = 1$. Applying these to (15), I obtain (13). Hence, $m^{u+1} - m^u$ asymptotically follows a Poisson distribution with mean $\phi^u(m^u - m^{u-1})$, which is equivalently a $(m^u - m^{u-1})$-times 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. The asymptotic mean (13) is dependent on $\bar{x}^{u-1}$.

When the distribution G is taken closer to F, $A(\bar{x})/B(\bar{x}) \to 1$ holds since $A/B = (1/F-1)/(1/G-1)$. As I take $A(\bar{x})/B(\bar{x}) \to 1$, the first fraction in the right hand side of (13) converges to 1 because of $A > \delta > B$, and the second fraction also converges to 1 by L'Hospital's rule. Thus $\phi^u \to 1$. □

A branching process is a stochastic integer process of population in which each individual ("parent") in a generation bears a random number of "children" in the next generation. Proposition 4 shows that the number of newly buying traders in each step $u+1$ asymptotically follows a Poisson distribution with mean $\phi^u(m^u - m^{u-1})$. Since Poisson distribution is infinitely divisible, the tatonnement process asymptotically follows the branching process in which each newly buying trader in u induces a random number of the other traders to buy in $u+1$ according to the Poisson distribution with mean ϕ^u.

Proposition 4 shows that ϕ^u does not depend on N for a given level of threshold $\bar{x}$. This means that p^u, the probability of a non-buying trader switching to buying in u, is of order $1/N$. It also implies that the decrease in the threshold $\bar{x}$ in each step is of order $1/N$. This property is important for the tatonnement process to generate a non-degenerate distribution of the total number of buying traders m^*. If p^u is of order less than $1/N$, then m^* converges to zero as $N \to \infty$. If p^u is of order more than $1/N$, the process explodes to infinity with probability one as $N \to \infty$. Only when p^u is of

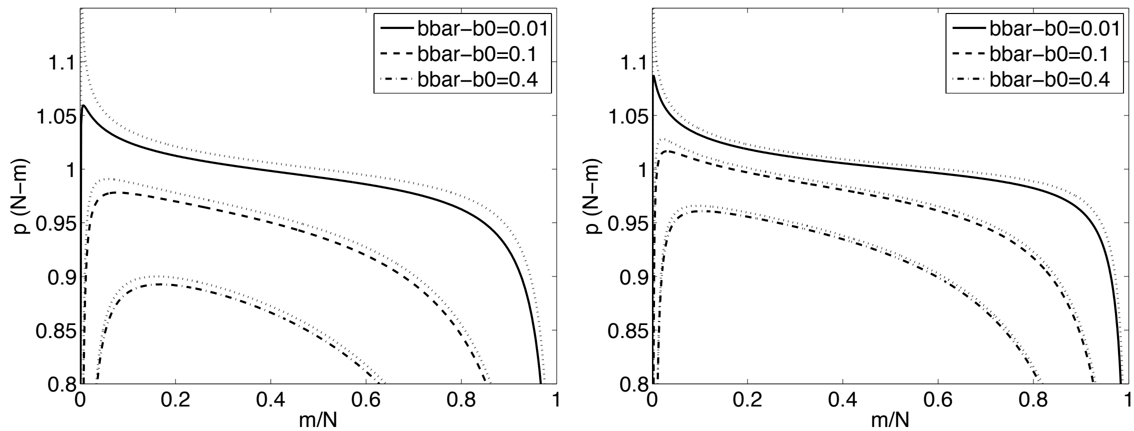

Figure 1: Mean number of newly buying traders in $u+1$ induced by a buying trader in u. The left panel is for $N = 500$ and the right for $N = 1000$. Dot lines show the approximations by ϕ^u for associated plots.

order $1/N$, m^* exhibits a well-defined fluctuation.

Figure 1 plots the exact mean $p^u(N - m^u)$ and the asymptotic mean ϕ^u of the number of traders induced to buy by a single buying trader. It is plotted as a function of m^u/N for various values of $\bar{b}-b_0$. The asymptotic mean is shown by dots. For the computation, F and G are set to an exponential distribution with mean 1.05 and 1, and $\bar{b} = 0.9$. The left panel shows the case of $N = 500$ and the right panel $N = 1000$. I observe that ϕ^u approximates the exact mean well and the approximation becomes better as N increases.

Now I derive the distribution of m^* as the cumulative sum of the informational tatonnement. Proposition 4 and Figure 1 show that the tatonnement can be approximated by a branching process with state-dependent Poisson mean, and Proposition 4 states that the Poisson mean is close to 1 when the informativeness of the private information is vanishingly small. When the Poisson mean is constant, I obtain the following distribution for the total size of the branching process.

Proposition 5 *Consider a branching process $m^u - m^{u-1}$ in which $m^0 = 0$, m^1 follows a Poisson distribution with mean μ, and $m^{u+1} - m^u$ conditional on $m^u - m^{u-1} = 1$ follows a Poisson distribution with mean ϕ. Then, the sum of the branching process m^T has the following distribution:*

$$\Pr(m^T = m) \quad = \quad \frac{\mu e^{-(\phi m+\mu)}}{m!}(\phi m + \mu)^{m-1} \tag{16}$$

$$\propto \quad (\phi e^{1-\phi})^m m^{-1.5} \tag{17}$$

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

Proof: The sum m^T conditional on $m^1 = 1$ follows a closed form distribution known as Borel-Tanner distribution in the queuing theory [22]. Equation (16) is derived by mixing the Borel-Tanner distribution and the Poisson distribution for m^1, and (17) is obtained by applying Stirling's formula [30]. □

Proposition 5 states that the sum of the Poisson branching process follows a power-law distribution with exponent 0.5 with exponential truncation (the exponent is defined for a cumulative distribution). In fact, such a tail distribution obtains not only for the Poisson branching process but also for any branching process.[2] The speed of exponential truncation is determined by $1-\phi$. The branching process is called critical, subcritical, or supercritical, when ϕ is above, equal to, or below 1, respectively. The distribution (16) becomes a pure power law at the criticality $\phi = 1$, at which the branching process becomes a martingale.

Since our informational tatonnement does not have a constant branching mean during the process for a finite N, Proposition 5 applies only asymptotically when $N \to \infty$ and $\sup(G-F) \to 0$. However, the distribution of simulated α^* in Figure 2 in the next section confirms that Proposition 5 provides an accurate prediction for the distribution. The exponent of the power law shown in Figure 2 is 0.5, which corresponds to the exponent derived analytically here.

What is the economic intuition for such a large fluctuation of m^*? The key to the fluctuation is that each trader responds to the *average* behavior of other traders. This point is seen by taking logarithm of the equilibrium condition (24):

$$(1-\alpha)\log A(\bar{x}) + \alpha \log B(\bar{x}) = 0 \tag{18}$$

which must hold both in the basic model and the model with the market maker at the limit of $N \to \infty$. When a trader buys, the other traders adjust their beliefs not only due to the observed action but also due to the observed *inactions* of the other traders. The buying action decreases the likelihood ratio of the other traders by a factor of B/A, and thus decreases their threshold $\bar{x}$. Inactions of those traders under the revised threshold reveal in turn their private information partially in favor of not buying. As a result of these two forces, the threshold is shifted so that the impact on the public information

[2] See Harris [18]. Also see Sornette [34] for robustness of this result for generalized branching processes.

caused by the triggering buying action is canceled out by the disbelief subsequently revealed by the inaction of bear traders. Since there are N traders in the market, the impact of one buying action on the threshold is of order $1/N$. Hence, the branching mean ϕ becomes of order 1. This is a natural order of magnitude for ϕ in the case of the symmetric information structure as I assume here, because the traders have no reason to imagine that the information revealed by an action should weigh more or less than that revealed by an inaction. Hence, it is a robust feature of herd behavior models that the chain-reaction mechanism operates at the criticality when the size of traders in terms of information weights is not greatly diversified.

The power law emerges at the criticality $\phi = 1$. Consider a deterministic process in which the population grows at the rate ϕ. The population of a generation decreases exponentially in a subcritical phase, $\phi < 1$, and increases exponentially in a supercritical phase $\phi > 1$, while it stays constant if $\phi = 1$ holds exactly. If $\phi \leq 1$, the branching process has a zero probability of continuing infinitely, whereas it has a positive probability of continuing infinitely if $\phi > 1$. Thus, $\phi = 1$ is the border case of the process that stops in a finite step.

As seen in Equation (17), the distribution of the total population has an exponential tail when the process has either $\phi < 1$ or $\phi > 1$. The speed of the exponential decay slows down as ϕ becomes close to 1, and disappears when $\phi = 1$. At this point, the slower power decay dominates the tail distribution. The power-law exponent 0.5 is closely related to the same exponent that appears for the distribution of the first return time of a random walk. I can see that the exponent has to be less than 1 as in the following argument. The branching process has a recursive characteristic, in which the total offsprings originated from a child has the same probability distribution as the total offsprings originated from the parent of the child. Let $H(s)$ denote the probability generating function of the total number of offsprings generated by one parent, and let $J(s)$ denote the probability generating function of the number of children each parent bears. Then the relation $H(s) = sJ(H(s))$ must hold, where $J(H(s))$ is the probability generating function for the offsprings originated by all children of the parent, and s is multiplied to $J(H(s))$ because $H(s)$ counts the parent itself in. By taking a derivative and evaluate it at $s = 1$, I obtain $H'(1) = 1 + J'(1)H'(1)$. $J'(1) = 1$ if the mean number of children per parent is one. Hence $H'(1)$ does not have a finite solution, implying that the total population does not have a finite mean. The diverging mean occurs when the power decay of the probability has an exponent less than 1.

3.2 Dynamic extension

The results in the static model were derived under the homogeneous prior belief. Our results hold even if the prior belief is heterogeneous. A particularly interesting case is when the belief evolves over time as the private information is repeatedly drawn. In this case, even though I 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 information. I show that our characterization of informational tatonnement continues to hold in the sequence of static equilibria.

The dynamic extension not only relaxes the assumption of common prior belief but also ties the loose end left in the static model. The limiting behavior of p^0 when $N \to \infty$ is ambiguous in the static model, leaving a possibility that the chain reaction is practically never triggered for a large N because of an extremely high threshold at the beginning of the tatonnement. In Figure 2, a large mass of probability is observed at $\alpha = 0$. This corresponds to the observation in Figure 1 that the branching mean decreases rapidly near $\alpha = 0$ when b_0 is lowered. This is because the traders rarely react to the private information in the situation when no other traders reveal their information and when the prior belief is small.

It turns out in the dynamic model that the belief naturally approaches to the threshold level, since the traders learn the true state eventually as they draw private information repeatedly. This implies that, regardless of the level of the initial prior belief or N, the belief eventually increases to the level at which traders start buying even though the other traders are not buying. This triggers the herd behavior. I may interpret this dynamics as a self-organized criticality following Bak et al. [3]: the distribution of traders' belief converges to the critical state at which the size distribution of herd behavior is characterized by a power law.

3.2.1 Heterogeneous belief

I dynamically extend the basic model with fixed price. Each trader i draws the private information $x_{i,t}$ repeatedly over periods for $t = 1, 2, \ldots$. The private information is identically and independently distributed across traders and periods. I consider the same asset as before which is worth 1 in H and 0 in L. Traders are given an opportunity to buy this asset for a fixed price $\bar{b}$ in each period, regardless of their past actions. Hence trader i buys the stock in period t if $b_{i,t} \geq \bar{b}$ and does not buy otherwise. There is no dynamic aspect involved in traders' decision other than updating the belief.

Traders observe their private information history $x_i^t = (x_{i,1}, x_{i,2}, \ldots, x_{i,t})$ as well as the aggregate information P_t which is the likelihood ratio revealed by the action profile history a^{t-1}. As in the static model, I assume that the traders believe that their actions cannot affect the aggregate information and thus they regard P_t as being independent of x_i^t.[3] I study a sequence of static equilibria (a_t, b_t), $t = 1, 2, \ldots$, such that the action $a_{i,t}$ maximizes trader i's expected period payoff under the subjective belief $b_{i,t}$ (defined for the state H as before) which is updated according to Bayes' rule upon i's observation of (x_i^t, P_t).

I continue to assume the common initial prior belief $b_{i,0} = b_0$ and MLRP. However, the belief is allowed to evolve stochastically as the traders draw information repeatedly. Thus, the belief in each period $t > 0$ is heterogeneous across traders with a particular structure that the heterogeneity stems only from the distribution functions F and G which are ordered by MLRP.

The aggregate information P_t is constructed as follows. Given an action profile history 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 $m_{k,t}$ the number of buying traders in the same group. Let $X_{k,s}$ for $s < t$ denote the domain of $x_{i,s}$ which is consistent with a_i^s under the threshold strategy supposed in Proposition 6 for trader i who belongs to the k-th group. The likelihood ratios revealed by an action history of a non-buying trader i and a buying trader j in the k-th group are written as follows.

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

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

I define the aggregate information as the likelihood ratio revealed by the action profile history:

$$P_t = \prod_k A_{k,t}^{n_{k,t} - m_{k,t}} B_{k,t}^{m_{k,t}} \tag{21}$$

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

Proposition 6 *For each realization of x^t, there exists an equilibrium outcome $(m_{k,t})$ and thresholds*

[3] As in the static model, it is possible to eliminate this assumption and construct a sequence of static Nash equilibria. It does not alter the characterization of the fluctuations at the limit of $N \to \infty$.

$\bar{x}_t$ *such that the action profile* a_t *satisfies the optimal threshold rule:*

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

Proof is deferred to Appendix A. The tatonnement process is characterized by a mixture of binomial distributions, which asymptotically follows a Poisson distribution. The threshold decreases over steps during the tatonnement within each node of the history, hence secures a well-defined tatonnement process with a chance of the chain reaction of buying actions.

3.2.2 Self-organized criticality

Traders in our model learn the true state by Bayesian learning by observing private information and aggregate actions. Since the private information is independent across periods, the Bayesian learning converges to the true state almost surely.

Proposition 7 *The subjective belief* $b_{i,t}$ *converges to 1 as* $t \to \infty$ *almost surely.*

Proof: The likelihood ratio for the private information history, $\prod_{\tau=1}^{t} \delta(x_{i,\tau})$, converges to zero as in Billingsley [6]. The proof is outlined as follows. The likelihood ratio $\theta_{i,t}$ follows a martingale in the probability measure of the private information under the true state: $E(\theta_{i,t} \mid \theta_{i,t-1}, H) = \theta_{i,t-1}$. Also, 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 represented by the distributions F^T and G^T for a sequence of private information $(x_{i,1}, x_{i,2}, \ldots, x_{i,T})$ are mutually singular when $T \to \infty$, since $x_{i,t}$ is independent across t. Then $\prod_{\tau=1}^{t} \delta_{i,\tau}$ converges to zero.

Hence, $\theta_{i,t}$ converges to zero if P_t remains finite for $t \to \infty$. P_t is finite for a finite $\bar{x}_t$ when N is finite. When $\bar{x}_t$ tends to a positive infinity, P_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, $\prod_k A_{k,t}^{n_{k,t}-m_{k,t}}$ tends to one, and P_t tends to $\prod_k B_{k,t}^{m_{k,t}}$. I proved that $B_{k,t} < (1/\bar{b} - 1)/(\theta_0 P_t)$ in the proof of Proposition 6. This inequality violates that P_t tends to $B_{k,t}^{m_{k,t}}$ if P_t tends to a positive infinity. Thus P_t is finite as $t \to \infty$. Hence $\theta_{i,t}$ is dominated by the private information as $t \to \infty$ and converges to zero, and $b_{i,t}$ converges to 1 almost surely. □

Proposition 7 means that the belief converges to the true state eventually. This is a natural consequence of that traders have infinitely precise information in the long run as they accumulate their own private information repeatedly. The convergence of belief implies that there is no possibility for the herd behavior in the long run in the narrow sense that the infinite sequence of traders take actions on the basis of wrong belief or that traders completely neglect their private information. The herd behavior can occur in a stage game in which the private information is not infinitely informative and thus can be hoarded.

The convergence of belief to the true state means that all the traders will buy eventually. This implies that some traders start buying even without any other trader buying during the process of the convergence. Such a buying action triggers the chain reaction of buying. The loose end in the previous section was that the triggering action may not occur with a meaningful probability when N is large. The converging belief assures that the triggering actions eventually occur and cause the fat-tail distributed aggregate actions almost surely.

The logic is analogous to the self-organized criticality proposed by Bak et al. [3]. 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 the private information accumulates, the average belief increases toward the threshold. It ensures that the triggering buying action occurs eventually.

4 Numerical results

4.1 Power law for aggregate actions

In this section, I show by numerical simulations that the probability distribution of the minimum equilibrium aggregate action α^* follows a power law when the distributions of the private information, F and G, are specified as exponential distributions. In Section 3.1, I derived the power-law distribution analytically without specifying F and G, but only asymptotically when N is taken to infinity and F is taken to G. In the simulation, the mean private information is set at 1.05 for H and 1 for L. Also I set $N = 1000$ and $\bar{b} = 0.9$. A profile of private information x is randomly drawn for 5000 times, and α^* is computed for each draw. Figure 2 plots the inverted cumulative distribution of α^* for various levels of the prior belief b_0. The inverted distribution $\Pr_{>}(\alpha^*)$ is cumulated from above, thus it takes

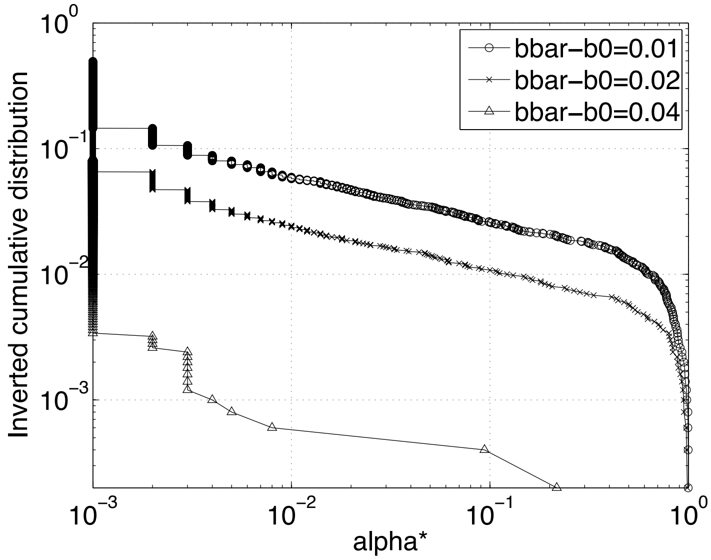

Figure 2: Inverted cumulative distribution of the minimum equilibrium number of buying traders α^*

0 at $\alpha^* = 1$ and 1 at $\alpha^* = 0$. The distribution is plotted in log-log scale, thus a linear line indicates a power law $\Pr_>(\alpha^*) \propto \alpha^{*-\xi}$, where the slope of the linear line ξ is called the exponent of the power law. I observe that the distribution shows a power law with exponent around 0.5 for a wide range of α^*.

The power-law distribution implies a fat tail, i.e., the concentration of a considerable probability in extreme events. In Figure 2 for the case $\bar{b} - b_0 = 0.01$, there is 6% probability for more than 1% of traders to buy the asset, 3% probability for more than 10% of traders, and 1% probability for more than 80% traders. The distribution exhibits the concentration of probability in the tail, considering that α^* in the plot has standard deviation 9.2% and mean 1.5%. This contrasts sharply to the case of the central limit theorem in which the average of the idiosyncratic shocks asymptotically follows a normal distribution with variance decreasing linearly in the number of the shocks. Suppose that the traders are not strategically complement and their actions solely depend on their own private information. The variance of a_i is less than 1. Then the aggregate action α^* converges to a normal distribution with variance less than $1/N$. With $N = 1000$, α^* would have less than 4% standard deviation, and hence more than 99% probability would be concentrated in the range of $\pm 12\%$ of the mean. Simulations also show that the variance of α^* does not decline as N increases. The simulated variances are 8%, 9%, and 11% for the cases of $N = 500$, 1000, and 2000, respectively. This contrasts again to the case of no strategic complementarity, where the variance should decline linearly in N according to the law of large numbers.

A power-law tail also implies that the distribution belongs to a domain of attraction of a stable law with the same exponent. This means that a sum of the random variables distributed according to a power law also follows the power law with the same exponent. Thus, if the herd size of traders follows a power law independently in each period, the cumulated size of the herds also follows a power law with the same exponent. If the power law is truncated exponentially, the central limit theorem takes effect in the accumulation and thus the cumulated herd size would converge to a normal distribution as the time horizon increases. Hence the distribution of the herd size would show a transition from a fat tail to a normal as I take a longer time horizon.

The power-law distribution of the aggregate action can be derived analytically. Before I move on to the analysis, let us see how the fat tail for the aggregate action can lead to the leptokurtosis of the stock returns.

4.2 Stock returns distribution

Fat-tailed distributions in financial markets are most prominent in the stock returns. In this section, I argue that our model can mimic the empirical stock returns distribution. To do so, I modify our basic model in order to relate the equilibrium aggregate actions to the stock returns.

The transformation of the aggregate actions to the stock returns raises two challenges to the standard herd behavior model. First, the benchmark herd behavior model assumes that the cost of acquiring an asset is constant, such as $\bar{b}$ in our basic model. In other words, the asset is infinitely supplied at the constant price. This is not a realistic assumption for stock markets where the price is constantly changing [36]. In this section, we incorporate a market maker who sets the price such that the expected returns of the market maker is zero conditional on the information available. Traders can now choose either "buy" or "sell". By abuse of notation, the action $a_i = 0$ now denotes "sell" instead of "inaction", while $a_i = 1$ remains to denote "buy". Traders buy if their subjective belief is above the price that the market maker posts, and sell otherwise. The market maker absorbs any excess demand from the traders. In the standard market microstructure models, the risk-taking behavior of the market maker is compensated by the revenue from the bid-ask spread that the market maker can set. Here I do not incorporate the spread in order to keep the model minimum.

The second difficulty is the no-trade theorem [28] which states that I may not observe any stable relation between the price and trading actions if the price reflects the publicly available information instantly. Smith [32] has shown that the trader's timing of trade will not be affected by public information either, since the price movement that reflects the public information will cancel out the effect of the public information on the trader's belief. This effect is seen in our model as follows. Suppose that the of price asset is determined by the market maker's belief, $1/(1+P)$, where P is the likelihood ratio revealed by traders' actions: $P = A^{N-m}B^m\theta_0$. The threshold likelihood $1/\bar{b} - 1$ is now P. Then, a change of action by trader i affects both the price and the subjective belief of trader j in the same manner, leaving j's incentive unaffected.

Considering this difficulty, Avery and Zemsky [1] proposed that another dimension of uncertainty is needed for the herd behavior models to deal with the stock price fluctuations. With the new dimension of uncertainty other than the asset value, there is a room for traders and the market maker to have different interpretation of observed actions. Avery and Zemsky exemplify two kinds of such uncertainty: event uncertainty and composition uncertainty. Here we consider a simple case of composition uncertainty in which the market maker cannot observe the actions of all the traders. This is the case when the market maker cannot fully identify the informed traders with private information from noise traders.

I suppose that the market maker observes the actions of randomly chosen traders of the number $\tilde{N} = qN$ for some constant $0 < q < 1$. Then, a trader's action is incorporated in the price only with probability q. I can interpret $1 - q$ as the noise level of the market. Let $\tilde{m}$ denote the number of buying traders among the observed traders. The total number of buying traders is denoted by m as before. I also assume that the market maker knows the threshold belief of the observed traders. This can be the case, for instance, if the market maker can observe the bidding price of some participants in the limit order book.

The price is set by the market maker at the expected value of the asset evaluated with the subjective belief of the market maker. Thus, the initial price is set as $b_0 = 1/(1+\theta_0)$ where θ_0 is the common prior likelihood ratio. The equilibrium price is set as $1/(1+P)$, where P is the likelihood ratio revealed to the market maker: $P = A(\bar{x})^{\tilde{N}-\tilde{m}}B(\bar{x})^{\tilde{m}}\theta_0$. Traders determine the optimal threshold given the

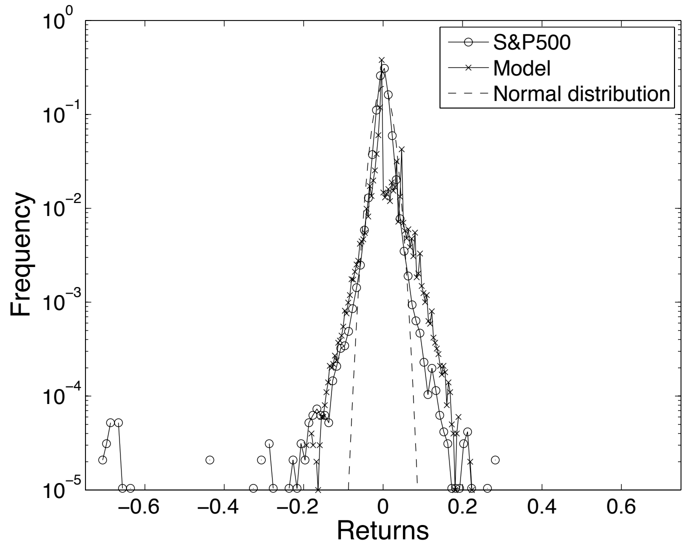

Figure 3: Returns distribution of S&P 500 stocks and simulated distribution of Q

price $1/(1+P)$ instead of $\bar{b}$. The threshold $\bar{x}$ thus satisfies:

$$(A(\bar{x})^{1-\alpha}B(\bar{x})^{\alpha})^{N}\delta(\bar{x})\theta_0 = P = (A(\bar{x})^{1-\tilde{\alpha}}B(\bar{x})^{\tilde{\alpha}})^{\tilde{N}}\theta_0 \tag{23}$$

The stock return Q is defined by the log difference between the equilibrium price $1/(1+P)$ and the initial price $1/(1+\theta_0)$.

Figure 3 shows that the distribution of the simulated returns Q can mimic the empirical distribution of stock returns. The empirical distribution is constructed from the daily returns of S&P500 stocks for a year starting from July 1, 2004. The dashed line shows the normal distribution with the mean and standard deviation computed from the S&P500 returns data. The horizontal axis is scaled linearly whereas the vertical axis is scaled in log. Thus, a linear tail of the distribution means an exponential decay and a parabola means a normal distribution. It is clear from the plot that the normal distribution fails to account for the fat tail and the leptokurtosis (i.e., high probabilities in the tail and at the center) of the empirical distribution. The simulated data are generated under the following specification. The noise level is set at $q = 0.1$ and the prior belief is set at $b_0 = 0.5$. The number of traders is $N = 1000$. The distributions of the private information, F and G, are exponential with mean $\lambda = 1.015$ and $\mu = 1$. Simulations show that Q has a larger standard deviation when λ is larger under fixed μ. I

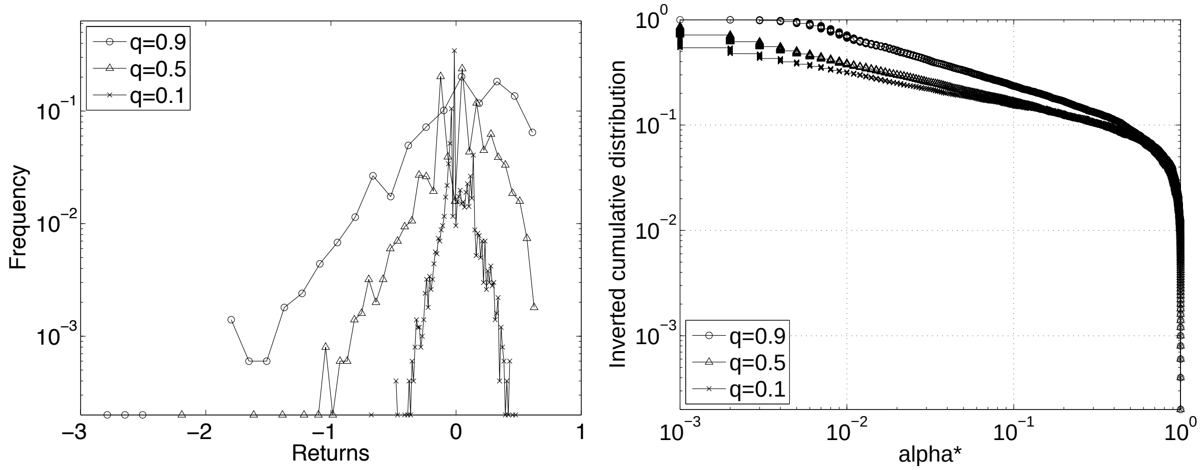

Figure 4: Simulated distributions of Q (left) and α^* (right) for various noise levels $1-q$

chose the value of λ so that the standard deviation of Q matches that of the empirical stock returns.

The left panel of Figure 4 plots the histogram of Q for various noise levels $1-q$. The parameters are the same as in Figure 3 except that λ is reset at 1.05 as in Figure 2. The distribution has a larger variance when q is larger. This is because more traders are observed by the market maker and thus the price reflects more information revealed. It is also skewed to the left when q is large, as I discuss later. I observe that the distributions are slightly biased to the right, since the true state of the economy is set to H and thus the private information brings a good news on average.

The right panel of Figure 4 plots the distribution of the fraction of buying traders, α^*, in the simulations that generate the left panel. I observe the same pattern of the power law with exponent around 0.5 as in Figure 2, although the exponent and the range of the power decay seem to slightly depend on q. The power law for various q can be understood as follows. Suppose that N approaches to infinity in Equation (23). The fraction of observed buying actions, $\tilde{\alpha} \equiv \tilde{m}/\tilde{N}$, converges to the fraction of the total buying actions, α. Then, for a fixed $0 < \alpha < 1$ at the limit of $N \to \infty$, $\bar{x}$ solves:

$$A(\bar{x})^{1-\alpha} B(\bar{x})^{\alpha} = \lim_{N\to\infty} \delta(\bar{x})^{-1/((1-q)N)} = 1 \tag{24}$$

This is equivalent to Equation (10) for the basic model in the limit of N. The asymptotic property of $\bar{x}$ inherits our basic model, and hence the distribution of α^* is characterized by a power-law distribution. The intuition is that the uncertainty regarding the observation of the market maker does not affect the

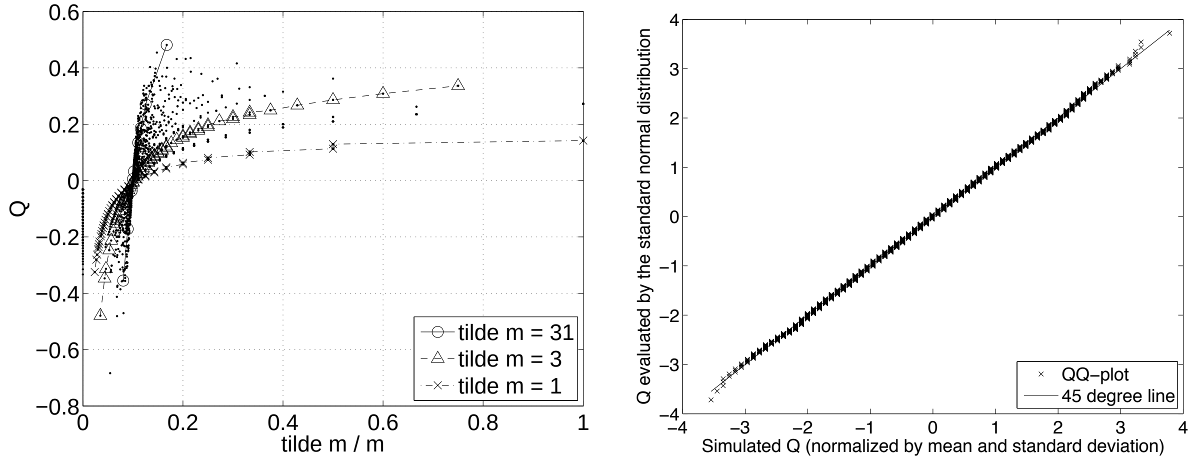

Figure 5: Left: Scatter plot of $\tilde{m}/m$ and Q. Right: QQ-plot of the log price against a normal distribution when $q = 1$.

mean number of traders that are affected by a trader's action, even though its variance is affected. It turns out in our analysis in Section 3.1 that the power-law relation only depends on the mean impact of a trader's action on the others'.

The left panel of Figure 5 shows the scatter plots of Q against $\tilde{m}/m$ in the simulated data for $q = 0.1$. Note that Q is close to 0 around $\tilde{m}/m = q$. That is, the price shows little movement if there is no composition uncertainty. The effect of composition uncertainty is shown by the large deviations of Q from 0 when $\tilde{m}/m$ is away from q. The stock price shows a large deviation from the informationally efficient price when there is a large gap in observations between the market maker and traders. The stock price also shows a large deviation when the number of buying traders is large. This is shown in the figure by the lines connecting the sample points that have the same number of observed buying traders $\tilde{m}$. For any fixed level of $\tilde{m}/m$, a larger $\tilde{m}$ (and thus a proportionally larger m) has a larger impact on Q. This reflects the herding effect shown by the basic model. The traders' actions tend to synchronize, and the synchronization of the information revelation causes a large change in the price. Our plot shows that the herd behavior magnifies the effect of composition uncertainty.

The model with a market maker nests the basic model as a special case $q = 0$, in which the market maker observes no traders and the price does not move at all. The other polar case is $q = 1$, when the price reflects all the revealed information and thus a trader's action has no spill-over effect on the other traders. When $q = 1$, the equilibrium threshold satisfies $\delta(\bar{x}) = 1$ regardless of m. The equilibrium outcome m^* follows a binomial distribution with population N and probability $1 - F(\bar{x})$. Then the log-likelihood ratio $\log P = (N - m^*) \log A(\bar{x}) + m^* \log B(\bar{x})$ asymptotically follows a normal distribution. If the support of P is very large, the log price $(-\log(1 + P))$ approximately follows a normal distribution. If the support of P is near zero, the log price is approximately $-P$, where P follows a lognormal distribution. Simulations confirm this point. The right panel of Figure 5 shows a QQ-plot of simulated series of the log-price Q. Each Q is computed for 50000 sets of randomly drawn private information x, when $\lambda = 1.01$, $\mu = 1$, and $N = 200$. The fit by the normal distribution for Q is perfect if the QQ-plot coincides with the 45 degree line. I see a good fit in the range of three standard deviations away from the mean. The left panel of Figure 4, on the other hand, shows the case of a high q ($q = 0.9$) when the resulting Q is relatively centered around 0. The distribution is skewed to the left, which is the shape of a negative lognormal distribution. Hence, our model explains the empirical distribution of stock returns only when the market observes maker a small fraction of the informed traders.

5 Discussions

5.1 Discrete actions

The discreteness of the action space plays an important role in our results of herd behavior. In our model, traders can choose either to buy or not (or sell in the model with a market maker). Suppose, instead, that each trader chooses an action from a continuous action space and that the action corresponds to the private information one-to-one. The private information is drawn from an exponential distribution and the state space is the possible mean of the distribution, $\{\lambda, \mu\}$ where $\lambda > \mu$, and the true state is λ. Then, the likelihood ratio of observing an average information profile $\langle x \rangle$ is $(\lambda/\mu)^N e^{(1/\lambda - 1/\mu)N\langle x \rangle}$. Since $\langle x \rangle \to \lambda$ and $(\lambda/\mu)e^{(1-\lambda/\mu)} < 1$, the likelihood ratio converges to 0 as N increases. Thus, for a large N, one round of information draw is sufficient for the market to learn the true state.

Our result on the aggregate effect of the binary choice has an implication on the effectiveness of the Tobin tax scheme. The tax on short-term asset transactions raises the transaction costs and thus can suppress speculative trades. A byproduct of the increased transaction costs and the decreased trades is the inhibition of the revelation of private information. Our model suggests that the inhibition can lead to a larger aggregate fluctuation. An increase in transaction costs will increase the amount of assets that a trader buys in one transaction. Since the transaction volume is the amount of assets bought in one transaction times the number of buying traders, the variance of the transaction volume will increase, and hence the variance of the asset price will increase. This corresponds to the empirical finding by Hau [19] that the volatility of stock prices is increased by an increase in transaction costs.

The effect of the transaction costs on the aggregate fluctuations can be put alternatively. Suppose that traders need to trade a certain amount of assets in each year. Suppose that traders can divide the total amount into monthly transactions if the transaction cost is low, whereas they can afford only one transaction a year if the transaction cost is high. If traders tend to herd in transactions due to the revelation of private information associated to the transactions, then I observe a small monthly herding when the transaction cost is low, whereas I observe a large yearly herding when the transaction cost is high. If the herd involves all the traders at once, the second moment of daily volume is $12(X/12)^2/365$ in the case of monthly herdings while $X^2/365$ in yearly herdings, where X is the total volume traded per year. Thus, the Tobin tax shifts the frequency domain of the herding from high to low, but it can increase the volatility of the trades. The aggregate fluctuation is suppressed by inducing the traders to trade frequently when the fluctuation is driven by information inferences among traders. In the limiting case when the traders trade continuously, the stochastic fluctuations of the herd size ceases to take effect as I discussed with the case of the continuous action space.

This paper mainly aims at explaining the fat-tailed fluctuations observed in the high frequency domain for which the empirical evidence is most supportive. However, some empirical studies such as Jansen and de Vries [20] and Longin [25] suggest that the largest crashes and booms in the history can be understood as an extreme event within the same power-law tail, instead of being outliers. Our model is consistent with their view that the market crashes are caused by the same mechanism that causes the price fluctuations in normal times. The natural time horizon for the uncertainty of the value of an asset to resolve just depends on the nature of the uncertainty. I can imagine a kind of uncertainty that can affect traders' long-run portfolio choice significantly. Such an uncertainty can linger for an extended time until it finally resolves by a collective conviction among traders and exerts a big impact on the price.

5.2 Information structure

The model can be generalized to allow asymmetric impacts of an action on the other traders' beliefs. I can define an information weight that each trader puts on the revealed likelihood ratio of the other traders. It is thus an N-by-N matrix $[w_{i,j}]$ for a decentralized market or an $N+1$-by-N matrix for a market with a market maker. The geometric sum of the traders' revealed likelihood in the right hand side of (9) can be weighted by this matrix. The model I have considered is the case of a symmetric information weight in which $w_{i,j} = 1$ for all i, j.

The noisy observation of a market maker provides a simple example of the information weight. For the market maker, $w^m_{N+1,i}$ takes either 1 or 0 randomly with probability q for 1. In a similar fashion, I can make the traders' reference groups heterogeneous or localized. An important example is the standard herd behavior model with sequential trading. In this case, a trader can only observe the actions of traders who have taken actions earlier. Thus the weight matrix is $w_{i,j} = 1$ for $j \leq i$ and $w_{i,j} = 0$ for $j > i$, when the trader i is sorted in the order of the trade timing.

Traders' network can be incorporated in the model by w. This alters the nature of the branching process, and interestingly, it can affect the exponent of the power law. For example, Bak et al. [3] showed that the critical branching process in a two-dimensional directed lattice results in avalanches characterized by a power-law distribution with exponent 1/3. The exponent is known to increase up to 1/2 as the dimension increases [16]. The exponent in our model achieves the upper bound, since our network is dimensionless and can be treated as a case of infinite dimensions. The model can incorporate the models of critical phenomena in the market with networked traders such as Cont and Bouchaud [11] and Stauffer and Sornette [35], since it allows to incorporate an exogenous random process which determines w.

The model suggests that the information structure is a crucial factor that determines the nature of the aggregate fluctuations. In a sequential trading, the first trader exerts overwhelming influence on the subsequent traders, whereas the power law distribution emerges in the symmetric information structure. In the model, the aggregate fluctuations can be suppressed by any means in which traders can share their private information, such as the survey of traders' sentiments or the public provision of market analysis.

6 Conclusion

This paper analyzed aggregate fluctuations that arise from the information inference among traders in financial markets. In a class of herd behavior models in which each trader infers other traders' private information only through observing their actions, I found that the number of traders who take the same action at equilibrium could exhibit large variation. The distribution of the size of the synchronized actions follows a power-law distribution with exponential truncation. The model prediction is fitted to the empirical fat-tail distribution of the stock returns. The parameters that determine the power-law distribution and its exponential truncation is identified by a new analytical method that utilizes a fictitious tatonnement process. I also showed that such chain reactions are eventually triggered almost surely in the situation where the private information is drawn repeatedly over time. This implies that the model features a self-organized criticality: traders' belief converges to the point at which the fluctuations of the aggregate actions follow a power law.

The power-law distribution of aggregate actions emerges when the information structure of traders is symmetric. Every trader receives a private information of the same degree of informativeness on the true value of asset. Thus an action by a trader is as informative as an inaction by another. When an information is revealed by a trader's buying action, the inaction of the other traders reveal their private information in favor of not buying. This counter-revelation is facilitated by a change in the threshold private information of order $1/N$, which leads to the criticality condition for the aggregate fluctuations. This model can nest the classic herd behavior model in sequential trading and the recent models of critical phenomena in the market with networked traders in which the network topology determines the exponent of the power-law fluctuations of the traders' herd size. The model implies that an increase in transaction costs raises the lumpiness of discrete actions and thus increases aggregate volatility.

Appendix

A Proof of Proposition 6

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

$$1/\bar{b} - 1 = P_t \theta_0 \delta(\bar{x}_t) \prod_{\tau=1}^{t-1} \delta(x_{i,\tau}) \tag{25}$$

I observe that $\delta(\bar{x}_t(P_t; x_i^{t-1})) \prod_{\tau=1}^{t-1} \delta(x_{i,\tau})$ is equal to $(1/\bar{b} - 1)/(\theta_0 P_t)$ and thus constant across i. Then $A_{k,t} > (1/\bar{b} - 1)/(\theta_0 P_t) > B_{k,t}$ can be shown as follows. The numerator of $A_{k,t}$ is expanded as:

$$\begin{aligned}
&\int_{X_{k,1}} \cdots \int_{X_{k,t-1}} G(\bar{x}_t(P_t; x_i^{t-1}))\delta(x_{i,t-1}) dF(x_{i,t-1}) \cdots \delta(x_{i,1}) dF(x_{i,1}) && (26)\\
> &\int_{X_{k,1}} \cdots \int_{X_{k,t-1}} F(\bar{x}_t(P_t; x_i^{t-1}))\delta(\bar{x}_t(P_t; x_i^{t-1}))\delta(x_{i,t-1}) \cdots \delta(x_{i,1}) dF(x_{i,t-1}) \cdots dF(x_{i,1}) && (27)\\
= &((1/\bar{b} - 1)/(\theta_0 P_t)) \int_{X_{k,1}} \cdots \int_{X_{k,t-1}} F(\bar{x}_t(P_t; x_i^{t-1})) dF(x_{i,t-1}) \cdots dF(x_{i,1}) && (28)
\end{aligned}$$

The integral in (28) is equal to the denominator of $A_{k,t}$, and thus $A_{k,t} > (1/\bar{b} - 1)/(\theta_0 P_t)$ holds. Similarly I obtain that $B_{k,t} < (1/\bar{b} - 1)/(\theta_0 P_t)$.

Suppose that m_t increases. P_t decreases by (21) and $A_{k,t} > B_{k,t}$. Then $\bar{x}_t$ needs to adjust in order to satisfy (25). By using $A_{k,t} > (1/\bar{b} - 1)/(\theta_0 P_t) > B_{k,t}$, I can obtain that the logarithm of $A_{k,t}$ and $B_{k,t}$ are decreasing in $\bar{x}_t$ as in the static model. Also $\delta(\bar{x}_t)$ is decreasing in $\bar{x}_t$. Thus, $\bar{x}_t$ in (25) 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 equilibrium is established by Tarski's fixed point theorem. This completes the proof.

The informational tatonnement process is characterized by a set of binomial distributions with probability p_k^u and population n_k. When m_k and $\bar{x}^u - \bar{x}^{u+1}$ are small and n_k are large, the binomial allows a Poisson approximation. Thus the sum of traders who buy in step u is approximated by a Poisson with mean $\sum_k n_k p_k^u$. Hence, the tatonnement approximately follows a Poisson branching process.

References

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

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

[3] 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.

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

[5] 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.

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

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

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

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

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

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

[12] Russell Cooper. Equilibrium selection in imperfectly competitive economies with multiple equilibria. *Economic Journal*, 104:1106–1122, 1994.

[13] Steven N. Durlauf. Path dependence in economics: The invisible hand in the grip of the past. *American Economic Association, Papers and Proceedings*, 81:70–74, 1991.

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

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

[16] Geoffrey Grimmett. *Percolation*. Springer, NY, second edition, 1999.

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

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

[19] Harald Hau. The role of transaction costs for financial volatility: Evidence from the Paris Bourse. *Journal of the European Economic Association*, 4:862–890, 2006.

[20] 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.

[21] Boyan Jovanovic. Micro shocks and aggregate risk. *Quarterly Journal of Economics*, 102:395–410, 1987.

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

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

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

[25] François M. Longin. The asymptotic distribution of extreme stock market returns. *Journal of Business*, 69:383–408, 1996.

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

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

[28] Paul Milgrom and Nancy Stokey. Information, trade, and common knowledge. *Journal of Economic Theory*, 26:17–27, 1982.

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

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

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

[32] Lones Smith. Private information and trade timing. *American Economic Review*, 90:1012–1018, 2000.

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

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

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

[36] Elias M. Stein and Jeremy C. Stein. Stock price distributions with stochastic volatility: An analytic approach. *Review of Financial Studies*, 4:727–752, 1991.

[37] Xavier Vives. Aggregation of information in large Cournot markets. *Econometrica*, 56:851–876, 1988.

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

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