CARF Working Paper

CARF-F-434

Herding and Power Laws in Financial Markets

Makoto Nirei
The University of Tokyo

John Stachurski
Australian National University

Koichiro Takaoka
Hitotsubashi University

Tsutomu Watanabe
The University of Tokyo

May 4, 2018

CARF is presently supported by The Dai-ichi Life Insurance Company, Limited, Nomura Holdings, Inc., Sumitomo Mitsui Banking Corporation, MUFG Bank, The Norinchukin Bank, The University of Tokyo Edge Capital Co., Ltd., Finatext Ltd., Sompo Holdings, Inc., Tokio Marine & Nichido Fire Insurance Co., Ltd. and Nippon Life Insurance Company. This financial support enables us to issue CARF Working Papers.

Herding and Power Laws in Financial Markets

Makoto Nirei*
University of Tokyo

John Stachurski
Australian National University

Koichiro Takaoka
Hitotsubashi University

Tsutomu Watanabe
University of Tokyo

May 4, 2018

Abstract

This study provides an explanation of the emergence of power laws in trading volume and asset returns. In the model, traders infer other traders' private signals regarding the value of an asset from their actions and adjust their own behavior accordingly. When the number of traders is large and the signals for asset value are noisy, this leads to power laws for equilibrium volume and returns. We also provide numerical results showing that the model reproduces observed distributions of daily stock volume and returns.

Keywords: Herd behavior; trading volume; stock returns; fat tail; power law

JEL classification code: G14

*Email: nirei@e.u-tokyo.ac.jp. Working paper versions of this study were circulated under titles "Herd behavior and fat tails in financial markets," "Information aggregation and fat tails in financial markets," and "Beauty contests and fat tails in financial markets" [45].

1 Introduction

Recently, the literature on empirical finance has converged on a broad consensus: Daily returns in equities, foreign exchange and commodities obey a power law. This striking property of high frequency returns has been found across both space and time through a variety of statistical procedures, from conditional likelihood methods and nonparametric tail decay estimation to straightforward log-log regression.[1] A power law has also been obtained for trading volume by Gopikrishnan et al. [23] and Plerou et al. [49].

While the heavy-tailed nature of returns is well understood (see, e.g., Fama [19], Mandelbrot [40] and many other authors[2]), these new power law findings are highly significant, mainly because extreme outcomes are by definition rare, so attempts to estimate any prices or quantities with tail risk sensitivity through nonparametric methods are deeply problematic (Salhi et al. [50]). Thus, information on the specific functional form of the tails of these distributions has great value for theorists, econometricians and practitioners. In addition, even elementary concepts from financial and economic theory—such as the benefits of diversification in the presence of risk—are sensitive to the precise nature of the tail properties of returns (see, e.g., Ibragimov [28]).

In this paper we respond to the developing empirical consensus by building a simultaneous-move herding model of asset markets that generates a power law in both

[1] For examples of recent empirical studies see Jansen and de Vries [29], Lux [37], Cont et al. [16], Gopikrishnan et al. [22], Ibragimov et al. [27] and Ankudinov et al. [3]. For overviews of the literature see Lux and Alfarano [38] or Gabaix [20].

[2] In the financial econometrics literature, for example, de Haan et al. [18] and Stein and Stein [54] incorporated high kurtosis under GARCH processes, Salhi et al. [50] proposed a regime switching model, and Cont and Tankov [17] and Kyprianou et al. [34] adopted jump-diffusion processes and Lévy processes to asset pricing.

volume and price. The underlying driver of this power law is asymmetric information on the value of assets. In particular, private information on the value of an asset is dispersed among many traders. The action of buying suggests a positive private signal. However, when traders' action space is coarser than their private signal, their actions only partially reveal private information. As a result of discreteness of the action space, a single trader's action can cause an avalanche of similar actions by other traders. This avalanche leads to power laws in volume and returns.

The mathematical framework behind our power law is as follows: Consider the first passage time of a stochastic process to zero, where zero represents market equilibrium and the first passage time represents the resulting number of buyers. It is known that, while a supermartingale dies down exponentially and a submartingale diverges with a positive probability, a martingale exhibits a criticality: its first passage time follows a power law distribution. We show that, in the model developed below, herding does indeed obey a martingale, due to the fact that one buying trader induces on average one other buying trader.

The intuition for this last fact is similar to Keynes' beauty contest: When an investor has an incentive to imitate the average behavior of n traders, the act of buying by one trader has $1/n$ impact on the other traders' behavior. While this would lead to a continuum of equilibria if the trader's action were continuous, discreteness of the action space implies that the equilibrium is locally unique, with the property that buying by one trader raises the other traders' likelihood of buying by $1/n$.

Our model builds on the herd behavior literature, which connects asymmetric information to excess volatility and kurtosis in asset pricing. The models of herding and information cascades proposed by Banerjee [7] and Bikhchandani, Hirshleifer and

Welch [8] have been employed to examine financial market fluctuations.[3] In particular, Gul and Lundholm [25] demonstrated the emergence of stochastic clustering by endogenizing traders' choice of waiting time. Moreover, signal properties leading to herding behavior in sequential trading were identified by Smith and Sørensen [51] and Park and Sabourian [48]. We inherit the spirit of these models in which the asymmetric information among traders results in clustering behavior and market volatility. However, none of these models generate a power law of financial fluctuations specifically.

Other market microstructure models have served as inspiration for components of our research. For example, our model draws on Minehart and Scotchmer [41], where a large number of informed traders simultaneously choose between buying one unit of an asset or not buying the asset at all. Informed traders submit demand schedules conditional on all possible prices, rather than choosing an action unconditionally on prices. This type of market competition was formulated as Nash equilibria in supply functions by Grossman [24] and Klemperer and Meyer [32], which have been introduced to the analysis of asset markets with private signals by Kyle [33], Vives [56], and Attar, Mariotti and Salanié [4]. However, none of these models lead to a power law. We employ the supply function equilibrium with private signals in asset markets, and extend it to the case where the action is discrete—buy or not buy. This restriction on action space leads to power laws in equilibrium.

There are other models that generate a power law of returns. For example, models of critical phenomena in statistical physics have been applied to herding behavior in financial markets,[4] in which a power law emerges if traders' connectivity parameter falls

[3]See, e.g., Caplin and Leahy [11], Lee [35], Chari and Kehoe [13], and Cipriani and Guarino [14]. For extensive surveys, see Brunnermeier [9], Chamley [12], and Vives [55].

[4]Studies in this literature include Bak, Paczuski, and Shubik [6]; Cont and Bouchaud [15]; Stauffer and Sornette [53].

at criticality. These studies do not, however, address why trader connectivity should exhibit criticality. In contrast, herding follows a martingale in our model as a result of an equilibrium process. Nirei [44] sketched out the basic idea that herd behavior can generate power-law size of cascades in the environment similar to Orléan [47], but fell short of substantiating the claim with rigorous analysis.

In another strand of the literature, Lux and Sornette [39] show that a stochastic rational bubble can generate a power law. Gabaix et al. [21] also show the power laws of trading volume and price changes if the amount of funds managed by traders follows a power law. In contrast to these explanations, we focus on the role of asymmetric information that results in herding behavior of investors. Many studies have associated the asymmetric information with financial phenomena such as crises, cascades and herding, from a historical account of crises by Mishkin [42] to the estimation of information content of trading volume on prices by Hasbrouck [26]. The latter noted, "Central to the analysis of market microstructure is the notion that in a market with asymmetrically informed agents, trades convey information and therefore cause a persistent impact on the security price." The present study seeks to link the investor behavior under asymmetric information to the ubiquitously observed power-law fluctuations.

The remainder of the study is organized as follows. Section 2 presents the model. Section 3.1 analytically shows that a power-law distribution emerges for trading volume when the number of traders tends to infinity, and provides an intuition for the mechanism behind it. Section 3.2 shows that a power law is obtained for returns. Section 3.3 numerically confirms that the equilibrium volumes follow a power law under a finite number of traders, and that the equilibrium return distribution matches its empirical counterpart. Section 4 concludes.

2 Model

In this section we describe the basic features of the model, including the nature of the asset market and the definition of equilibrium.

2.1 Market

The asset market consists of n informed traders, a continuum of uninformed traders and an auctioneer. The uninformed traders supply a single asset and informed traders demand it.[5] The asset has common intrinsic value 1 in state H and 0 in state L. The true state is not known to any market participant. All hold a common prior belief $\Pr(H) = \Pr(L) = 0.5$.

Let $S(p)$ denote aggregate supply of the uninformed traders at price p. We assume that S is continuously differentiable and strictly increasing with $S(0.5) = 0$, so that aggregate supply is zero at the price level that reflects common prior belief. We also assume that $\bar{p} := S^{-1}(1) < 1$, implying an upper bound on equilibrium price below the maximum value of the asset.

Each informed trader chooses whether or not to buy a single trading unit, set to $1/n$ so as to normalize maximum total demand to unity. Hence aggregate demand takes values from discrete set $\{0, 1/n, \ldots, 1\}$. The equilibrium price P^* takes values in $\mathcal{P} := \{p_0, p_1, \ldots, p_n\}$, where p_m for $m \in \{0, 1, \ldots, n\}$ is determined by the market-clearing condition $S(p_m) = m/n$.

Informed traders receive private signals that are correlated with the state. Signal

[5]We later discuss the case when both uninformed and informed traders can buy and sell. The informational asymmetry between informed and uninformed traders in this model is similar to event uncertainty, as introduced by Avery and Zemsky [5] as a condition for herding to occur in financial markets.

X_i is IID across i with conditional cumulative distribution function F^s for $s = H, L$. The two distributions have common bounded support $\mathcal{X} := [\underline{x}, \bar{x}] \subset \mathbb{R}$.

Each informed trader $i = 1, 2, \ldots, n$ submits his demand function to the auctioneer. The demand function describes his action for each realization of price, given his private signal. The permissible set of demand functions is $\mathcal{D} = \{d_i : \mathcal{X} \mapsto \{0, 1\}^{\mathcal{P}_1}\}$, where $\mathcal{P}_1 := \mathcal{P} \backslash \{p_0\} = \{p_1, p_2, \ldots, p_n\}$. Note that p_0 is excluded from $\mathcal{P}_1$ because p_0 cannot be realized in equilibrium if any trader, including i, chooses buying at p_0. Thus, traders are allowed to demand assets only at prices strictly greater than p_0. The demand function indicates buying at p when $d_i(p \mid x_i) = 1$ and not-buying when $d_i(p \mid x_i) = 0$. Let $\boldsymbol{x} = (x_i)_{i=1}^n$ denote a profile of private signals. Aggregate demand expressed in terms of the number of buying traders is $D(p \mid \boldsymbol{x}) := \sum_{i=1}^n d_i(p \mid x_i)$, which maps $\mathcal{X}^n$ to $\{0, 1, \ldots, n\}^{\mathcal{P}_1}$.

Decisions take place under the following timing. First, nature sets the state $s \in \{H, L\}$. Second, a signal profile $\boldsymbol{x}$ is drawn from conditional distribution $(F^s)^n$. Third, informed traders submit demand function $d_i(p \mid x_i)$ to the auctioneer. Fourth, the auctioneer determines equilibrium price p^* through the following protocol: If $D(p_1 \mid \boldsymbol{x}) = 0$, then the auctioneer sets $p^* = p_0$, since no informed trader is willing to buy given that all other traders do not buy. If $D(p_1 \mid \boldsymbol{x}) > 0$, then the auctioneer determines $p^* > p_0$ such that $S(p^*) = D(p^*, \boldsymbol{x})/n$. Finally, transactions take place. A unit of asset is delivered to informed trader i with $d_i(p^* \mid x_i) = 1$. The equilibrium number of buying traders is determined as $m^* := D(p^* \mid \boldsymbol{x})$.[6]

[6]The model may be extended to the case where informed traders can sell as well as buy or not buy the asset. In the extended model, there are uninformed traders on both supply and demand sides. An informed trader submits a demand function d that can take values 1, 0, or -1. The auctioneer stipulates that no transactions take place if there exist non-zero traders buying at p_1 and selling at p_{-1}. In this way, informed traders always transact against uninformed traders as in the original model.

Since the asset has common value 1 in H and 0 in L while its purchase cost is p, a buying trader ($d_i(p \mid x_i) = 1$) obtains payoff $1 - p$ in state H and $-p$ in L, whereas a not-buying trader ($d_i(p \mid x_i) = 0$) obtains 0 in either state. We assume that informed traders are risk-neutral and maximize their expected payoff. The expected payoff for the choice $d_i(p \mid x_i) = 0$ is 0 regardless of i's belief. The expected payoff for the choice $d_i(p \mid x_i) = 1$ is $r_i(p, x_i) - p$, where $r_i(p, x_i) := \Pr(H \mid p, x_i, d_i(p \mid x_i) = 1)$ denotes the probability of $s = H$ conditional on that trader i receives signal x_i and buys at p and that p is an equilibrium price. Given x_i and $d_i(p_m \mid x_i) = 1$, p_m is an equilibrium price if and only if there are $m-1$ other traders buying at p_m, i.e., $\sum_{j \neq i} d_j(p_m \mid x_j) = m-1$. Let $\Omega_{m,i}$ denote such an event: $\Omega_{m,i} := \{(x_j)_{j \neq i} : \sum_{j \neq i} d_j(p_m \mid x_j) = m-1\}$. Moreover, x_i is independent of other traders' decisions d_j conditional on m. Therefore, we can write the conditional probability as follows for any $m \in \{1, 2, \dots, n\}$.

$$r_i(p_m, x_i) = \frac{\Pr(\Omega_{m,i}, x_i, H)}{\Pr(\Omega_{m,i}, x_i)} = \frac{\Pr(\Omega_{m,i} \mid H)}{\Pr(\Omega_{m,i}, x_i)} \Pr(x_i \mid H) \Pr(H) \tag{1}$$

A Bayesian Nash equilibrium consists of a profile of informed traders' demand functions $d_i : \mathcal{X} \mapsto \{0,1\}^{\mathcal{P}_1}$, a profile of conditional probabilities $r_i(p, x_i)$, and equilibrium price correspondence p^* for $(\mathcal{X}, \mathcal{P})$ such that (i) for any $i = 1, 2, \dots, n$, and given d_j for $j \neq i$, there is no other function $d_i' \in \mathcal{D}$ that achieves an expected payoff greater than that achieved by d_i, (ii) for any $i = 1, 2, \dots, n$, $r_i(p, x_i)$ is consistent with demand functions $\{d_j\}_j$ and equilibrium price correspondence p^*, and (iii) p^* clears the market: $S(p^*) = \sum_{i=1}^{n} d_i(p^*, x_i)/n$, $\forall \boldsymbol{x} \in \mathcal{X}^n$.

2.2 Signal

We are concerned with the case where there are many informed traders who receive private signals of the state, and where the informativeness of the signal is small. Thus, we consider a series of markets indexed by the number of informed traders, n. The density function of F_n^s, the distribution of the private signal, is denoted by f_n^s, for $s = H, L$. We order the signal based on the monotone likelihood ratio property (MLRP) such that the likelihood ratio $\ell_n := f_n^H / f_n^L$ is strictly increasing. This property holds for a signal with two states without loss of generality. We further assume that f_n^s is continuous, bounded, strictly positive over $\mathcal{X}$, and has a bounded derivative. The likelihood ratio $\ell_n(x)$ represents how likely the state is H when the signal is x.

We also define likelihood ratios $\Lambda_n(x) := (1 - F_n^H(x))/(1 - F_n^L(x))$ and $\lambda_n(x) := F_n^H(x)/F_n^L(x)$. $\Lambda_n(x)$ expresses the likelihood when the signal is greater than x. Thus, a trader's bidding action reveals the information $\Lambda_n(x)$ to observers of the action under a decision rule that a trader buys only if the signal is greater than x. Similarly, $\lambda_n(x)$ is the likelihood when the signal is smaller than x, and it is the information revealed by *inaction* of the trader.

The likelihood ratios satisfy $\Lambda_n(\underline{x}) = \lambda_n(\bar{x}) = 1$, $\lim_{x \to \underline{x}} \lambda_n(x) = \ell_n(\underline{x})$, and $\lim_{x \to \bar{x}} \Lambda_n(x) = \ell_n(\bar{x})$ (obtained using L'Hopital's rule). Also, MLRP implies $0 < \lambda_n(x) < \ell_n(x) < \Lambda_n(x)$ for any x interior of $\mathcal{X}$ as in Smith and Sørensen [51], and strictly increasing likelihood ratios: $\Lambda_n'(x) > 0$ and $\lambda_n'(x) > 0$ (see Appendix for proof). Figure 1 depicts these properties of the likelihood ratios.

2.3 Traders' optimal strategy

As we saw previously, choice $d_i(p, x_i) = 1$ yields expected payoff $r_i(p, x_i) - p$, whereas choice $d_i(p, x_i) = 0$ results in expected payoff 0 regardless of r_i. Therefore, trader i chooses $d_i(p, x_i) = 1$ if and only if $r_i(p, x_i) \geq p$. This condition is equivalent to $\rho_i(p, x_i) \geq p/(1-p)$, where $\rho_i(p, x_i) := r_i(p, x_i)/(1 - r_i(p, x_i))$ is a conditional likelihood ratio for i with private signal x_i and decision $d_i(p, x_i) = 1$.

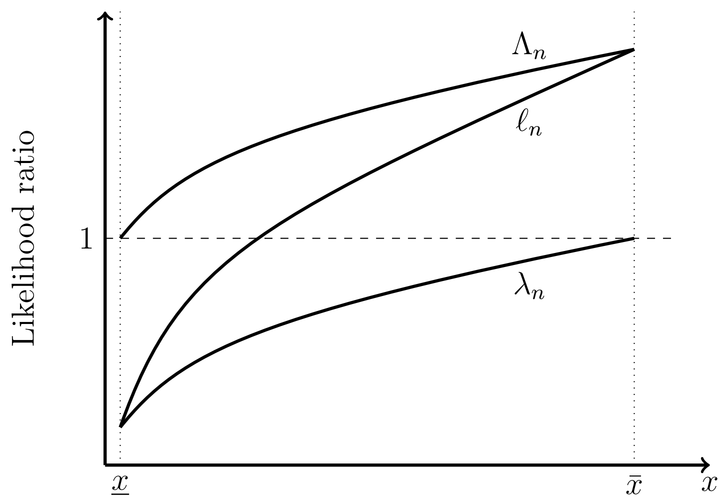

Figure 1: Likelihood ratios $\ell_n = f_n^H/f_n^L$, $\lambda_n = F_n^H/F_n^L$, and $\Lambda_n = (1 - F_n^H)/(1 - F_n^L)$

The decision rule is characterized by a threshold under our two-state environment. Using (1) and $\Pr(H) = \Pr(L) = 0.5$, we obtain

$$\rho_i(p, x_i) = \frac{\Pr(\Omega_{m,i} \mid H)}{\Pr(\Omega_{m,i} \mid L)} \ell_n(x_i). \tag{2}$$

Since $\ell_n(x_i)$ is continuous and strictly increasing, $\rho_i(p_m, x_i)$ is continuous and strictly increasing in x_i for any p_m. Therefore, for each $p_m \in \mathcal{P}_1$, there exists threshold $\sigma \in \mathcal{X}$ such that it is optimal for trader i to buy if and only if $x_i \geq \sigma$. The threshold σ indicates either an indifference level of signal $\rho_i(p_m, \sigma) = p_m/(1 - p_m)$ or a corner solution $\sigma = \underline{x}, \bar{x}$ for each m. We denote a threshold function as $\sigma : \{1, 2, \ldots, n\} \mapsto \mathcal{X}$. Trader i's optimal demand function follows a threshold rule

$$d_i(p_m, x_i) = \begin{cases} 1 & \text{if } x_i \geq \sigma(m), \\ 0 & \text{otherwise,} \end{cases}$$

for each $p_m \in \mathcal{P}_1$.

A buying trader at price p_m can infer that there are $m - 1$ other buying traders at p_m if the price is realized under the stipulated rule for the auctioneer. Moreover, the threshold function $\sigma(m)$ is common for all informed traders. Thus, a buying trader can infer that, for p_m to occur, there must be $m-1$ other traders who receive signals greater than $\sigma(m)$ and $n-m$ traders who receive signals smaller than $\sigma(m)$. Such an event occurs with probability

$$\Pr(\Omega_{m,i} \mid s) = \binom{n-1}{m-1}(1 - F_n^s(\sigma(m)))^{m-1} F_n^s(\sigma(m))^{n-m}$$

for each state $s \in \{H, L\}$, since the probability of receiving signal smaller than σ is $F_n^s(\sigma)$. Therefore, the likelihood ratio for p_m to occur is

$$\begin{aligned}\frac{\Pr(\Omega_{m,i} \mid H)}{\Pr(\Omega_{m,i} \mid L)} &= \left(\frac{1 - F_n^H(\sigma(m))}{1 - F_n^L(\sigma(m))}\right)^{m-1} \left(\frac{F_n^H(\sigma(m))}{F_n^L(\sigma(m))}\right)^{n-m} \\ &= \Lambda_n(\sigma(m))^{m-1}\lambda_n(\sigma(m))^{n-m}.\end{aligned}$$

Substituting into (2), we obtain $\rho_i(p_m, x_i) = \Lambda_n(\sigma(m))^{m-1}\lambda_n(\sigma(m))^{n-m}\ell_n(x_i)$. Thus, the threshold $\sigma(m)$ for $m \in \{1, 2, \ldots, n\}$ is implicitly determined by

$$\frac{p_m}{1 - p_m} = \lambda_n(\sigma)^{n-m}\Lambda_n(\sigma)^{m-1}\ell_n(\sigma) \tag{3}$$

if an interior solution σ exists.

Equation (3) is the key to the subsequent analysis. The right-hand side shows the likelihood ratio of the posterior belief of a trader who receives signal $x_i = \sigma(m)$ and buys at p_m. This equation determines the threshold level of signal $\sigma(m)$ at which a trader is indifferent between buying and not-buying given p_m. Due to the threshold behavior shown above, we obtain aggregate demand $D(p_m, \boldsymbol{x})$ by counting the number of informed traders with $x_i \geq \sigma(m)$.

With this setup, the more informed traders buy, the more signals in favor of H are revealed. This further encourages informed traders to buy. The resulting aggregate demand curve is upward sloping if the signal revealed by larger demand has a greater effect on expected payoff than an increase in purchasing costs caused by the demand does.

We formalize this environment as follows. Let δ_n denote the minimum distance between Λ_n and λ_n in logarithm: $\delta_n := \min_{x\in\mathcal{X}}(\log\Lambda_n(x) - \log\lambda_n(x))$. We assume the following property on the series of likelihood ratio functions.

Assumption 1. *There exists a triplet (n_o, ξ, δ) such that $n_o < \infty$, $\xi \in (0,1)$, $\delta \in (0,\infty)$, and $\delta_n > \delta/n^{\xi}$ for any $n > n_o$.*

Assumption 1 sets the lower bound on the informativeness of signal. It allows the signal to deteriorate to pure noise as $n \to \infty$, but the speed of convergence to the pure noise is strictly slower than $1/n$.

With this setup, we obtain the following proposition stating that the aggregate demand curve is upward sloping when n is sufficiently large.

Proposition 1. *Under Assumption 1, there exists an integer n_o such that for any $n > n_o$, the threshold level of signal $\sigma(m)$ is decreasing in m and the aggregate demand $D(p_m, \boldsymbol{x})$ is increasing in m.*

Proofs are deferred to Appendix unless otherwise stated.

Figure 2 depicts an aggregate demand curve $D(p, \boldsymbol{x})$. The upward-sloping aggregate demand indicates the presence of strategic complementarity in informed traders' buying decisions through the information revealed by price: a higher price indicates that there are more informed traders who receive high signals.[7]

The number of informed traders n needs to be large in order to obtain the upward sloping demand curve in our model. The increment in price p_{m+1}/p_m caused by an

[7] The mechanism in which demand feeds on itself is reminiscent of Bulow and Klemperer [10]'s "rational frenzies."

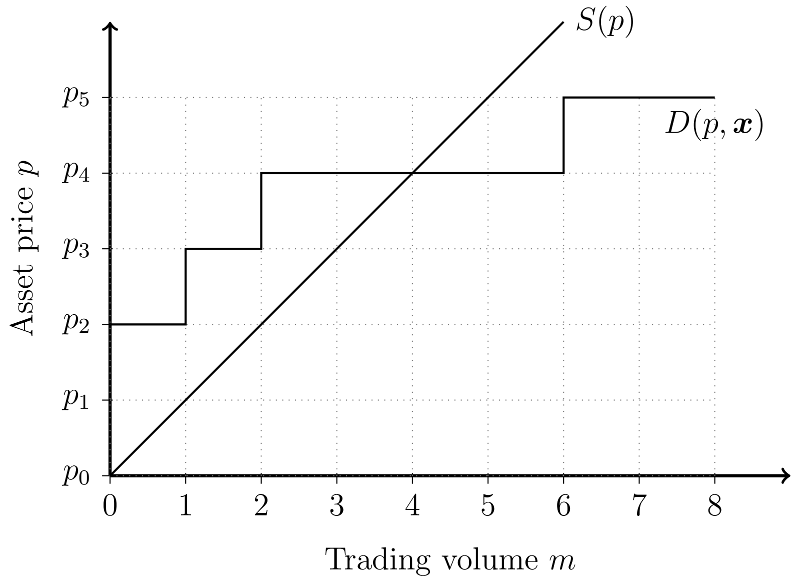

Figure 2: An equilibrium outcome for a realization of signal profile $\boldsymbol{x}$

increase in demand m is of order $1/n$. Thus, when n is small, purchasing costs can increase quickly due to limited supply and overwhelm the effect of signal revealed by the increase in demand, leading to a downward sloping demand curve. Given that the signal does not deteriorate as fast as $1/n$, the aggregate demand curve is upward sloping for sufficiently large n. With the upward sloping demand function, we establish the existence of equilibrium in a finite economy as follows.

Proposition 2. *Under Assumption 1, for any $n > n_o$, there exists an equilibrium outcome (p^*, m^*) for each realization of $\boldsymbol{x}$.*

While multiple equilibria may exist for each realization of $\boldsymbol{x}$, we focus on the case where the auctioneer selects the minimum number of buying traders among possible equilibria, $m^\dagger$, for each $\boldsymbol{x}$.[8] Note that this equilibrium selection uniquely maps each

[8]By assuming that the auctioneer selects the minimum number of buying traders, we exclude

realization of $\boldsymbol{x}$ to $m^{\dagger}$. Thus, $M_n^{\dagger}$ is a random variable whose probability distribution is determined by the probability distribution of $(X_{n,i})_{i=1}^{n}$ and the equilibrium selection mapping.

3 Results

3.1 Derivation of the power law

Next we characterize the minimum equilibrium aggregate trading volume $M_n^{\dagger}$ and show that it follows a power law distribution asymptotically in n. Since the asset price in this model is determined by the equilibrium condition $S(p^*) = m^{\dagger}/n$, the power law for the trading volume also implies a fat-tailed distribution of the equilibrium price P^*.

In order to characterize $M_n^{\dagger}$, we introduce a stochastic process that counts the number of traders who receive signal greater than x. Such a process is expressed as $\sum_{i=1}^{n} I_{X_{n,i} \geq x}$, where I is an indicator function: $I = 1$ if $X_{n,i} \geq x$ and $I = 0$ otherwise. As x travels from maximum $\bar{x}$ to minimum $\underline{x}$, this process generates an increasing number of buying traders. Now, we replace x with the threshold level of signal, $\sigma(m)$. Then, $\sum_{i=1}^{n} I_{X_{n,i} \geq \sigma(m)}$ indicates the number of traders with private information greater than threshold $\sigma(m)$. For each realization of $\boldsymbol{x}$, $\sum_{i=1}^{n} I_{x_i \geq \sigma(m)}$ is increasing in m because $\sigma_n(m)$ is decreasing in m by Proposition 1. Equilibrium $m^{\dagger}$ is determined as the point where this counting process achieves $m^{\dagger}$ for the level of signal $\sigma(m^{\dagger})$ for the first time. Namely, by appropriately defining the counting process, $M_n^{\dagger}$ can be formulated as a first passage time for the process to cut through the diagonal where time and counts coincide. We construct such a counting process below. Throughout this analysis, we

fluctuations that arise purely from informational coordination such as in sunspot equilibria. Even with this rule, we show that the equilibrium price exhibits large fluctuations.

specify that the underlying random variable $X_{n,i}$ follows F_n^H, assuming that the true state is H.

Equation (3) implicitly determines threshold σ continuously when m is a real variable. By using the continuous threshold function, we define a change of variable as $t = \sigma^{-1}(x)$. Note that $t = m$ for $m \in \{1, 2, \ldots, n\}$. Using $t = \sigma^{-1}(x)$ and $f_n^H(x)$, the probability density function defined over t is obtained as $f_n^H(\sigma(t))|\sigma'(t)|$ for sufficiently large $n > n_o$, because $\sigma(t)$ is monotone in t for such n. Then, we construct a counting process $\Gamma(t) := \sum_{i=1}^{n} I_{\sigma^{-1}(X_{n,i}) \geq t}$.[9]

When t increases from t to $t + dt$, the threshold $\sigma(t)$ decreases. Thus, a trader who chooses to buy before t continues to buy at $t + dt$, whereas a trader who chooses not to buy before t might switch to buying at $t + dt$. The conditional probability of a non-buying trader switching to buying between t and $t + dt$ for a small dt is equal to $\pi_n(t)dt := f_n^H(\sigma(t))|\sigma'(t)|dt/F_n^H(\sigma(t))$. Thus, the number of traders who buy between t and $t + dt$ for the first time, conditional on $\Gamma(t)$, follows a binomial distribution with population parameter $n - \Gamma(t)$ and probability parameter $\pi_n(t)dt$. $\Gamma(1)$ indicates the number of traders with $x_i \geq \sigma(1)$. Thus, the distribution of $\Gamma(1)$ follows a binomial distribution with population n and probability $\pi_n^o := 1 - F_n^H(\sigma(1))$. This completes the definition of the stochastic process $\Gamma(t)$ for $t \in [1, n]$.

Let $\phi_n(t)dt$ denote the mean of $\Gamma(t + dt) - \Gamma(t)$ for a small dt. Thus, $\phi_n(t) := \pi_n(t)(n - \Gamma(t))$. For a finite $\Gamma(t)$, the binomial distribution of $\Gamma(t + dt) - \Gamma(t)$ converges to a Poisson distribution with mean $\phi_n(t)dt$ as $n \to \infty$. Hence, for sufficiently large n, $\Gamma(t)$ asymptotically follows a Poisson process with time-dependent intensity $\phi_n(t)$.

[9] $\Gamma(t)$ differs from $\Gamma_{\boldsymbol{x}}(m)$ defined in Appendix (Proof of Proposition 2) in two regards. First, $\Gamma(t)$ is not conditional on $\boldsymbol{x}$. Thus, $\Gamma(t)$ is a random variable. Second, $\Gamma(t)$ is defined over a transformed variable of signal, $t = \sigma^{-1}(x)$. Despite these differences, both $\Gamma(t)$ and $\Gamma_{\boldsymbol{x}}(m)$ share the property that they count the number of traders with private signal greater than some threshold.

In line with our concern with high-frequency fluctuations in volume and price, we formalize the idea that the signal tends to pure noise in this series of markets as n increases as follows:

Assumption 2. *As $n \to \infty$, $\ell_n(x)$ converges to 1 uniformly in $\mathcal{X}$.*

Assumption 2 holds in short time intervals when the signal received by traders tends to be noisy. Along with Assumption 1, we consider an asymptotic case where the signal contains vanishingly small information on the fundamental value of an asset, and yet the informativeness is larger than the impact of increasing purchasing costs.

Under this environment, it turns out that the intensity function ϕ_n converges to 1 as $n \to \infty$, as stated in the following lemma.

Lemma 1. *Under Assumptions 1 and 2, $\Gamma(t)$ asymptotically follows a Poisson process with intensity 1 as $n \to \infty$.*

The unitary intensity $\phi_n = 1$ implies that the mean number of informed traders who switch to buying from non-buying after observing an informed trader buying is equal to 1.

Since $\Gamma(1) = 0$ indicates that no trader receives private signal greater than $\sigma(1)$, the equilibrium volume in this case is $m^\dagger = 0$. When $\Gamma(1) = 1$, one trader is willing to buy at p_1. Thus, the equilibrium volume is $m^\dagger = 1$. When $\Gamma(1) > 1$, the minimum equilibrium volume $m^\dagger$ is the minimum integer that satisfies $\Gamma(m^\dagger) = m^\dagger$. Thus, when $\Gamma(1) > 1$, $m^\dagger$ can be interpreted as the first passage time t at which $\Gamma(t)$ achieves the level t.

We focus on the first passage time conditional on $\Gamma(1) > 1$. It is convenient to shift the time variable so that it starts from 0. We define $G(t) := \Gamma(t+1)$ and $\varphi_n(t) := \phi_n(t+1)$ for $t \in [0, n-1]$. Note that, when $\Gamma(m^\dagger) = m^\dagger$ is achieved,

$m^\dagger - \Gamma(1) = \Gamma(m^\dagger) - \Gamma(1) = G(m^\dagger - 1) - G(0)$ holds. Thus, $m^\dagger - 1$ corresponds to the first passage time of $G(t)$ reaching t with initial condition $G(0) = \Gamma(1) - 1 > 0$. Let a positive integer $c_o > 0$ denote the initial value $G(0)$.

$G(t)$ asymptotically follows a Poisson process with intensity $\varphi_n(t)$ and $G(0) = c_o$, as n becomes large. Let $\tau_{\varphi_n(\cdot)}$ denote the first passage time of $G(t)$ reaching t. Then, $\tau_{\varphi_n(\cdot)}$ is also the first passage time of $G(t) - G(0)$ reaching $t - c_o$. Let us define $N(t)$ as the Poisson process with constant intensity 1 and $N(0) = 0$. Then, τ_1 denotes the first passage time of $N(t)$ reaching $t - c_o$. An inhomogeneous Poisson process with intensity $\varphi_n(t)$ for $t \geq 0$ can be transformed by a change of time to a homogeneous Poisson process as $N(\int_0^t \varphi_n(u)du)$. Thus, the first passage time we consider is

$$\tau_{\varphi_n(\cdot)} := \inf \left\{ t \geq 0 \mid N\left(\int_0^t \varphi_n(u)du \right) \leq t - c_o \right\},$$

where $\inf \emptyset := \infty$ by convention.

We consider a case where ℓ_n uniformly converges to 1 as $n \to \infty$ (Assumption 2), which implies that signal $X_{n,i}$ is close to a pure noise. This case occurs, for example, in short-term tradings in which the information content traders obtain from signals during a trading period is quite small. With this setup, the following lemma establishes that the first passage time of the inhomogeneous Poisson process $G(t)$ converges in distribution to the first passage time of the standard Poisson process $N(t)$.

Lemma 2. *Under Assumptions 1 and 2, $\tau_{\varphi_n(\cdot)}$ converges in distribution to τ_1 as $n \to \infty$.*

We have shown that the minimum equilibrium number of buying traders $M_n^\dagger$, conditional on $M_n^\dagger > 1$, has the same distribution as the stopping time: $\inf\{t > 1 \mid \Gamma(t) = t\}$. This $M_n^\dagger$ corresponds to $\tau_{\varphi_n} + 1$, since G is shifted from Γ in time by 1. Lemma 2 then shows that τ_{φ_n} converges in distribution to τ_1 for large n. Hence, we have shown that the distribution of $\tau_1 + 1$ characterizes the distribution of $M_n^\dagger$ conditional on $M_n^\dagger > 1$ asymptotically for large n.

We can further derive the distribution function of τ_1 explicitly, using the fact that τ_1 has the same distribution function as the sum of a branching process. The stopping time τ_1 follows the same distribution as $M_n^\dagger - 1$ conditional on that there are $\Gamma(1) = c_o + 1 > 1$ traders who receive private signal $x_i \geq \sigma(1)$. Hence, we obtain the conditional distribution of the equilibrium number of buying traders, $M_n^\dagger \mid \Gamma(1)$, for sufficiently large n.

Proposition 3. *Under Assumptions 1 and 2, $M_n^\dagger$ conditional on $\Gamma(1) = c > 1$ follows asymptotically as $n \to \infty$,*

$$\Pr\left(M_n^\dagger = m \mid \Gamma(1) = c\right) = \frac{(c-1)(m-1)^{m-c-1}e^{-m+1}}{(m-c)!}, \tag{4}$$

for $m = c, c+1, \ldots$. Moreover, the tail of the asymptotic distribution follows a power law with exponent 0.5, i.e., $\Pr(M_n^\dagger > m) \propto m^{-0.5}$ for sufficiently large values of m.

Proposition 3 shows that the distribution of $M_n^\dagger$ conditional on $\Gamma(1)$ has a power-law tail. This implies that, given there are $\Gamma(1)$ traders who receive favorable private signals $x_i \geq \sigma(1)$, their buying actions may trigger a stochastic herd, and the size of the herd follows a power-law distribution.

We can pin down the distribution of $\Gamma(1)$ under a certain condition, in which case we can explicitly derive an unconditional asymptotic distribution of $M_n^\dagger$. For finite n, $\Gamma(1)$ follows a binomial distribution with population n and probability $\pi_n^o := 1 - F_n^H(\sigma(1))$. The behavior of the asymptotic mean $\phi_o := \lim_{n\to\infty} n\pi_n^o$ depends on the specification of signal $X_{n,i}$. If ϕ_o is finite, $\Gamma(1)$ follows a Poisson distribution with mean ϕ_o asymptotically as $n \to \infty$. In this case, we obtain the unconditional distribution of $M^\dagger$ explicitly as follows.

Proposition 4. *Under Assumptions 1 and 2, and if $\phi_o < \infty$, the asymptotic distribution of $M_n^\dagger$ as $n \to \infty$ follows, for $m = 2, 3, \ldots$,*

$$\Pr(M_n^\dagger = m) = \frac{e^{-\phi_o - m + 1}(m-1)^{m-1}}{m!} \left[1 + (\phi_o - 1) \left(1 + \frac{\phi_o}{m-1} \right)^{m-1} \right],$$

and $\Pr(M_n^\dagger = m) = \phi_o^m e^{-\phi_o}$ for $m = 0, 1$. Moreover, $M_n^\dagger$ has a power-law tail distribution with exponent 0.5.

Finally, we provide an example where the bounded ϕ_o exists under a particular class of signals.

Assumption 3. *The distributions of signal $X_{n,i}$ satisfy $F_n^s(x) = e^{a_n^s(x-\bar{x})}$ for $s = H, L$ in $(\bar{x} - \epsilon, \bar{x}]$ for some $\epsilon > 0$, where $(a_n^H, a_n^L)_{n=n_o}^\infty$ is a pair of bounded series of real numbers which converges to (a, a) for some $a > 0$ and satisfies $a_n^H > a_n^L > 0$ and $\log(a_n^H / a_n^L) > \delta / n^\xi$ for any $n > n_o$.*

In Appendix, we show that the distributions specified by this assumption are well behaved and generate $\phi_o = 1$. Thus, under this specific class of signals, we obtain a simple distribution for the volume fluctuation as follows.

Proposition 5. *Under Assumptions 1, 2 and 3, the asymptotic distribution of $M_n^\dagger$ as $n \to \infty$ follows, for $m = 1, 2, \ldots$,*

$$\Pr(M_n^\dagger = m) = \frac{e^{-m}(m-1)^{m-1}}{m!},$$

and $\Pr(M_n^\dagger = 0) = e^{-1}$. $M_n^\dagger$ has a power-law tail distribution with exponent 0.5.

Propositions 3, 4 and 5 indicate the emergence of a power-law tail for the equilibrium number of buying traders. Proposition 5 shows that the asymptotic distribution of $M_n^\dagger$ does not depend on any model parameter. This implies that our result of power law is quite robust to the details of model specification.

In general, a power law with exponent α implies that any k-th moment for $k \geq \alpha$ is infinite. Thus, with exponent 0.5, $M_n^\dagger$ does not have a finite asymptotic variance or mean as $n \to \infty$. This implies that the variance of the fraction of buying traders, M_n^*/n, can be quite large even when n is large. By integrating $(M_n^\dagger/n)^2$ up to $M_n^\dagger = n$ with a power-law tail exponent 0.5, we find that the variance of $M^\dagger/n$ decreases as $n^{-0.5}$ when n becomes large. This contrasts with the case when the traders act independently. If traders' choices $(d_{n,i})_{i=1}^n$ were independent, the central limit theorem predicts that $M_n^\dagger/n$ would asymptotically follow a normal distribution, whose tail is thin and variance declines as fast as n^{-1}. Thus, the variance of $M_n^\dagger/n$ differs by factor $n^{0.5}$ between our model and the model with independent choices. This signifies the effect of stochastic herding that amplifies the small fluctuations in the received signals $X_{n,i}$.[10]

Even though such amplification effects can occur whenever traders' actions are correlated, it requires a particular structure in the correlation for the amplification effect to cause the variance to decline more slowly than n^{-1}, i.e., the speed the central limit theorem predicts. Mathematically, the amplification 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 the asymptotic martingale process $\Gamma(t)$. In fact, the power law exponent 0.5 obtained in our model is closely related to the same exponent in the Inverse Gaussian distribution that characterizes the first passage time of the Wiener process.

The economic meaning of $\Gamma(t)$ being a martingale in our model is that the mean number of traders induced to buy by a buying trader is 1. This property can be

[10] On the implications of a tail distribution on aggregate fluctuations, see, for example, Nirei [46] and Acemoglu, Ozdaglar and Tahbaz-Salehi [1].

seen from optimal threshold condition (3). This condition reduces to a simple form $(1-\mu)\log\lambda_n(\sigma)+\mu\log\Lambda_n(\sigma)=0$, where $\mu := m/n$, if we take the limit $n\to\infty$ while fixing μ. The condition indicates that the geometric average of λ and Λ evaluated at σ, which can be regarded as revealed likelihood on the true state revealed by traders' actions, does not change even when μ takes different values. Suppose that a trader switches from not-buying to buying. This increases μ, which leads to an increase in the revealed likelihood that traders observe, and lowers the optimal threshold. This in turn decreases the revealed likelihood, because traders learn that the signals received by non-buying traders must have been below the decreased level of threshold. As a result, the impact of an increase in μ on the geometric average of λ and Λ is counteracted by a decrease in σ. These effects turn out to cancel out with each other when the signal is vanishingly small (i.e., $\log\Lambda_n - \log\lambda_n \approx 0$) and m is finite ($\mu \approx 0$). Hence, under the noisy signal, any finite m satisfies the equilibrium condition above.

This environment is analogous to the fable of Keynes's beauty contest, in which the average action of a single trader responds one-to-one to the average actions of traders. The beauty contest would lead to indeterminate equilibria, if there are a continuum of traders or if the traders' actions are continuous. This type of local indeterminacy is avoided in our model with a finite number of traders and discrete actions. However, the indeterminacy described above provides an intuition why our model can generate equilibrium trading volumes at any order of magnitude as demonstrated by the power law.

Our model depicts the situation where a large number of informed traders try to glean information from other traders' actions under noisy signal. A power law of herd size emerges asymptotically when traders' interaction leads to indeterminacy. However, this does not necessarily imply that traders fail to learn. We can extend our model dynamically where traders draw private signal repeatedly and eventually learn the true state.[11] A power law in this case implies that the collective learning does not occur smoothly over time. The noisy signal generates little transaction and is hoarded privately for most of the times, but once in a while, a large herd occurs and the accumulated private information is revealed. Thus, a power law of herd size implies that the revelation happens at once in the collective learning of traders.

Propositions 3, 4 and 5 claim not only that various levels of aggregate trading volumes $M_n^\dagger$ are possible, but also that the distribution of $M_n^\dagger$ has a particular regularity signified by a power law. The power law for $M_n^\dagger$ implies a particular fat-tailed distribution of the equilibrium price P_n^*. In the following sections, we show a case where the return distribution also follows a power law in our model.

3.2 Return distributions

Having established the power law for volume, we now turn to the power law for returns. An important facet of the model to be specified is the supply function $S(p)$, which determines how the fluctuation of volume is translated to the fluctuation of returns. In our model, informed traders' demands are absorbed by uninformed traders' supply. Thus, the supply function of uninformed traders $m^* = S(p^*)$ determines the impact of volume m^* on the return $q := \log p^* - \log p_0$. The relation between an exogenous shift in trading volume and a resulting shift in asset price, i.e. S^{-1}, is called a price impact function. Following the literature (see, e.g., Hasbrouck [26] and Lillo et al. [36]), we specify the price impact function as a concave power function $q = \beta(m/n)^\gamma$ with $\beta > 0$ and $0 < \gamma < 1$ for $m = 1, 2, \ldots, n$.

The following proposition establishes that our model generates a power law for the

[11] See the working paper version [45] for the extension.

returns distribution when the price impact is specified as the power function.

Proposition 6. *Suppose that the volume m/n follows a power law* $\Pr(m/n) \propto (m/n)^{-\alpha-1}$ *and the supply function S satisfies*

$$\log(S^{-1}(m)) - \log(S^{-1}(0)) = \beta(m/n)^{\gamma} \tag{5}$$

for $\beta, \gamma > 0$. Then, returns q follow a power law,

$$\Pr(q) \propto q^{-\alpha/\gamma-1}.$$

Proof. By applying the change of variable for m/n using the supply function, we obtain the power law for q as: $\Pr(q) \propto \Pr(m/n)|d(m/n)/dq| \propto ((q/\beta)^{1/\gamma})^{-\alpha-1}(1/\gamma)q^{1/\gamma-1} \propto q^{-\alpha/\gamma-1}$. □

As Gabaix et al. [21] and Lux and Alfarano [38] demonstrate, there is a growing consensus among empiricists of financial data that stock returns generally obey the "cubic law," in which the return distribution follows a power law with exponent 3. The cubic law corresponds to $\alpha/\gamma = 3$ in the above equation. The analysis in the previous section established that $\alpha = 0.5$ holds asymptotically. Hence, the cubic law holds in our asymptotic case if $\gamma = 1/6$. This value is consistent with the empirically estimated range of $\gamma \in [0.1, 0.5]$ (Lillo et al. [36]).

Gabaix et al. [21] provides a micro-foundation for a square-root specification of price impact function, $\gamma = 0.5$, and combines it with the "half-cubic law" of trading volume found by Plerou et al. [49], $\alpha = 1.5$, to obtain the cubic law $\alpha/\gamma = 3$. However, Lux and Alfarano [38] note that it is not clear whether the half-cubic law of volume is "of a similarly universal nature" as the cubic law of returns. While our model predicts $\alpha = 0.5$ in an asymptotic case with a diverging number of informed traders n and a vanishingly small signal, we will show that the model is also capable of generating a different exponent when n is finite. In the next section, we explore through numerical analyses the model's prediction on the power law for returns and volume under finite n and an empirically plausible price-impact function.

3.3 Quantitative analysis with finite n

In this section, we conduct numerical analysis of the model with a finite number of informed traders n. The purpose of this exercise is to confirm that, even with finite n, the number of buying traders $M_n^\dagger$ follows a power law, which was shown as an asymptotic property when n tends to infinity in the previous sections. Moreover, we show that the fluctuation of equilibrium asset returns $q = \log P_n^* - \log p(0)$ exhibits a power law that matches with the returns distribution observed in empirical data.

The model is specified as follows. The signal distribution F^s for $s \in \{H, L\}$ is normal with common standard deviation σ and different mean $\mu_H = 1$ and $\mu_L = 0$. We set σ at between 30 and 50. This large standard deviation relative to the difference in mean 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 2000. The supply function of uninformed traders is specified as in (5), and its parameters are set at our estimates $\gamma = 0.4642$ and $\beta = 0.768$ as explained later. With these parameter values, the optimal threshold function $\sigma(\cdot)$ is computed. Using the threshold function, we conduct Monte Carlo simulations by randomly drawing a profile of private signals $(x_i)_{i=1}^n$ for 10 million times and computing $m_n^\dagger$ and p_n^* for each draw.

The top panel of Figure 3 plots the histograms of $M_n^\dagger$ for various parameter values of n and σ. Since the histogram is plotted in log-log scale, a linear line indicates a power law $\Pr(M_n^\dagger = m) \propto m^{-\alpha-1}$, where the slope of the linear line reflects $-\alpha - 1$. As can be seen, the simulated log-log histograms appear linear for a wide range of $M_n^\dagger$.

This conforms to the model prediction that $M_n^\dagger$ follows a power law distribution. Note that the simulated histogram decays exponentially when $M_n^\dagger/n$ is close to 1, due to the finiteness of n.

The asymptotic results in Propositions 3, 4 and 5 predicted the exponent of power law α to be 0.5. The top panel of Figure 3 confirms this pattern for finite n, when $(n, \sigma) = (800, 30)$ or $(2000, 50)$. We also observe that the power law exponent can take larger values when the parameter alignment differs, as observed in the case when n is decreased (the circle-line compared to the cross-line) or when σ is increased (the cross-line compared to the triangular-line). This deviation of exponent from the asymptotic case $\alpha = 0.5$ can result from the state-dependence of the intensity φ_n in our model when n is finite (see, e.g., Sornette [52]). This property renders flexibility to our model in fitting to various exponents for trading volume.

Given this flexibility, we bring our model to fit the empirical distributions of daily volume and returns. We intend this exercise to be a proof of concept for the capacity of our model as an explanation of the observed power laws. Our direct target for comparison is the time-series fluctuations of a single stock volume and returns in daily frequency. We use Nikkei Financial Quest dataset that includes daily volume and prices for the firms listed in the first section of Tokyo Stock Exchange (TSE) from March 1988 to March 2018.

The bottom panel of Figure 3 shows histograms of daily trading volumes for four single stocks. The volume is divided by the time-series average volume for each stock. The four firms are selected at the quintiles of market capitalization size among all manufacturing firms listed in the first section of TSE. The plotted histogram exhibits a fat tail for each stock. However, the number of observations (7401) for each stock is not sufficiently large to investigate the tail in detail, and the sample period (30 years)

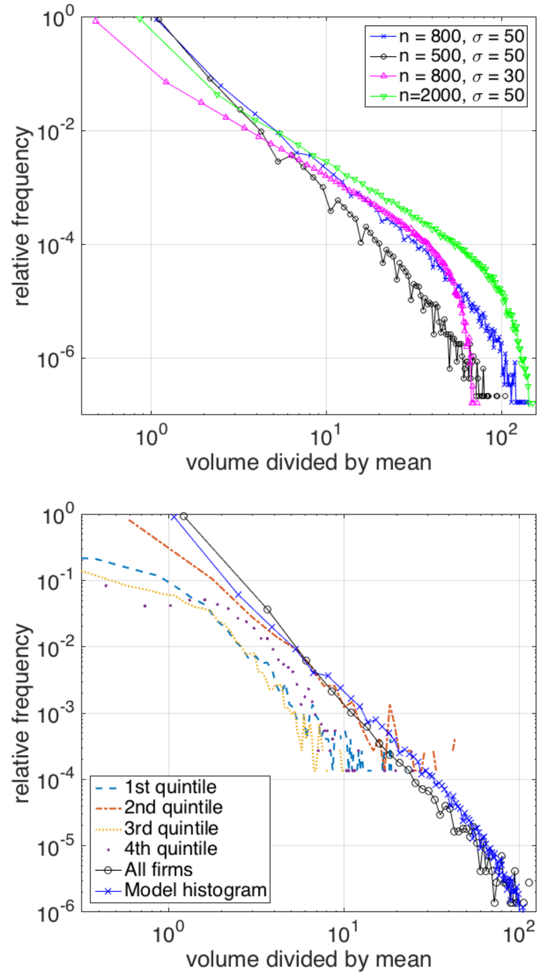

Figure 3: Histograms of volume. The volume is normalized by its time-series average. *Top*: Histograms of equilibrium volume $M_n^\dagger$ for various parameter values, where n is the number of traders and σ is the standard deviation of the private information. *Bottom*: Daily volume histograms for individual stocks and a pooled sample, along with the model histogram. Data for individual firms cover 30 years (from 1988 to 2018), while data for the pooled sample cover one year (2016). The individual firms are selected at the quintiles of market capitalization size in TSE. The circle-line shows the histogram of pooled data for all listed firms. The cross-line shows a simulated histogram generated by our model with $n = 800$ and $\sigma = 50$.

is too long to assure the invariance of the daily volume distribution. To deal with this data limit, we prepare a pooled dataset of a large number of stocks for a shorter sample period. We collect daily trading volumes for all (2250) listed firms for year 2016, and divide the volume by the average volume for each firm during the year. The circle-line in the plot shows the histogram of the normalized volume for the pooled sample. We now observe a longer tail, whose exponent is similar to the tails for individual stocks. The pooled data shows that the exponent for volume is about 2 (the slope of the histogram in log-log scale is 3). We then superimpose the volume histogram generated by our model for the case $(n, \sigma) = (800, 50)$, shown as the cross-line. As can be seen, the simulated histogram reasonably agrees with the empirical histogram.

The top panel of Figure 4 shows the histograms of daily returns for the same samples. We define the daily return as a logarithmic difference from the opening to closing price. The open-close difference is used rather than a business day return so that the time horizon of each observed return is homogenized. We subtract time-series average returns and divide by standard error of the returns for each stock, and take an absolute value for returns, pooling both positive and negative returns across stocks. The empirical histograms for the individual stocks and the pooled sample show a power law with exponent about 3, which is consistent with the literature (Lux and Alfarano [38]). The model-generated histogram also shows a fat tail, which is slightly thinner than data but clearly exhibits a power law.

The bottom panel of Figure 4 shows a scatter plot of daily volume and absolute returns for all listed firms, along with the price-impact function specified in the simulated model (5). The parameter values ($\gamma = 0.4642$, $\beta = 0.768$) are estimated by fitting (5) to the pooled sample by the non-linear least squares.[12] In sum, Figures 3 and 4 in-

[12]The plotted sample is truncated at the volume divided by mean being 50, in order to enhance

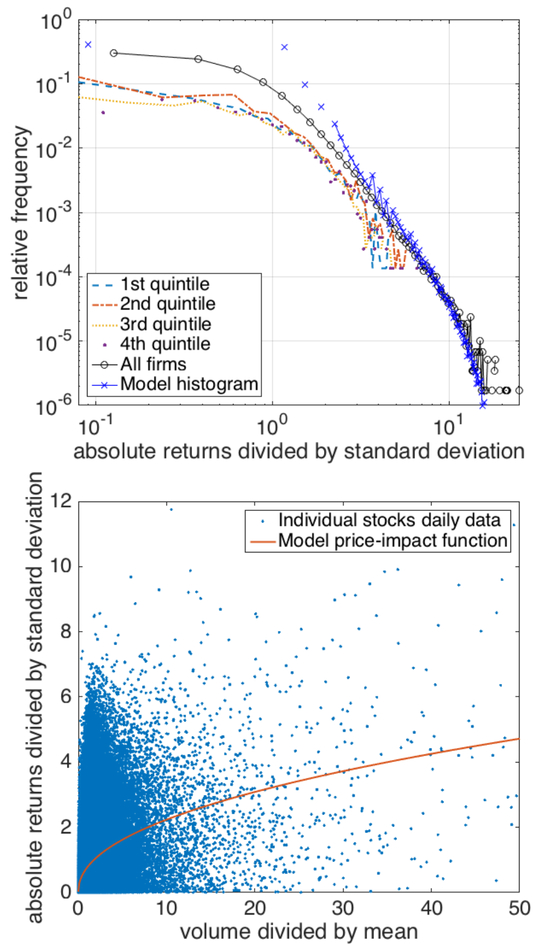

Figure 4: *Top*: Histograms of the absolute values of daily returns. The horizontal axis shows the daily return, which is the difference in logarithm of close and open prices for each business day. The samples are the same as in Figure 3. *Bottom*: A scatter plot of daily volume and absolute returns for all listed firms in 2016. The red line shows the price-impact function fitted by the non-linear least squares.

dicate that our model is capable of generating the power laws of volume and absolute returns with a price-impact function consistently estimated with the high-frequency sample observed in TSE.

4 Conclusion

This study analyzed aggregate fluctuations of trading volume and prices that arise from asymmetric information among traders in financial markets. In a herding model in which each trader infers the private information of other traders only by observing their actions, we found that the number of traders taking the same action at equilibrium exhibits large volatility with a statistical regularity—a power-law distribution. Furthermore, we showed that the model is capable of generating a power-law distribution of asset returns. The simulated distributions of equilibrium returns and volume were demonstrated to match the distributions of observed stock returns and volume.

The power law distribution of herding size emerges when the number of traders is large and the signal is noisy, consistent with the empirical observation that the power-law fluctuation of returns manifests in high frequency data. In our model, an action by one trader is as informative as inaction by another. When some information is revealed by a trader's buying action, the inaction of other traders despite of their observation of the initial buying action, reveals their private information in favor of not buying. Thus, each trader's action is influenced by the average action, resulting in a near-indeterminate equilibrium, analogous to Keynes' beauty contest. In this way, our model of asymmetric information provides an economic reasoning for the criticality condition that generates power law fluctuations.

visibility, while all the data are used for the parameter estimation.

This study suggests several directions for extension. One would be to develop a dynamic model that accounts for the time-series properties as pursued by, for example, Alfarano et al. [2]. Another direction would be to extend the model by incorporating more realistic market structure. Kamada and Miura [30] have taken a step in 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: Proofs

Properties of λ_n and Λ_n

Taking derivatives of λ_n and Λ_n, we have

$$\frac{d\lambda_n(x)}{dx} = \frac{f_n^H(x)}{F_n^L(x)} - \frac{F_n^H(x) f_n^L(x)}{(F_n^L(x))^2} = \frac{f_n^L(x)}{F_n^L(x)} \left(\ell_n(x) - \lambda_n(x)\right), \tag{6}$$

$$\begin{aligned} \frac{d\Lambda_n(x)}{dx} &= -\frac{f_n^H(x)}{1 - F_n^L(x)} + \frac{(1 - F_n^H(x)) f_n^L(x)}{(1 - F_n^L(x))^2} \\ &= \frac{f_n^L(x)}{1 - F_n^L(x)} \left(\Lambda_n(x) - \ell_n(x)\right). \end{aligned} \tag{7}$$

Thus, inequality $\lambda_n(x) < \ell_n(x) < \Lambda_n(x)$ implies that $\lambda_n'(x) > 0$ for $x \in (\underline{x}, \bar{x}]$ and $\Lambda_n'(x) > 0$ for $x \in [\underline{x}, \bar{x})$. At $x = \underline{x}$, we obtain $\lambda_n'(\underline{x}) = \ell_n'(\underline{x})/2 > 0$ by applying L'Hopital's rule for (6) and rearranging terms. Similarly, we obtain $\Lambda_n'(\bar{x}) = \ell_n'(\bar{x})/2 > 0$ by evaluating (7) at $x = \bar{x}$. Hence, we obtain $\lambda_n'(x) > 0$ and $\Lambda_n'(x) > 0$ for any $x \in \mathcal{X}$.

Proof of Proposition 1

The market-clearing condition, $S(p_m) = m/n$, implicitly determines p_m not only for integers but also for any real number m. Thus, Equation (3) implicitly determines

$\sigma(m)$ for real numbers of m. In this proof, we extend p_m and $\sigma(m)$ to real numbers. To be precise, we define real variables $t \in [1, n]$, p_t and $\sigma(t)$, such that p_t is determined by the market-clearing condition $S(p_t) = t/n$ and $\sigma(t)$ is implicitly determined by

$$0 = \Phi(\sigma, t) := (n-t) \log \lambda_n(\sigma) + (t-1) \log \Lambda_n(\sigma) + \log \ell_n(\sigma) - \log \frac{p_t}{1-p_t}, \qquad (8)$$

which is a logarithmic transformation of (3) with m being replaced by t.

We first show that an interior solution σ of $\Phi(\sigma, t) = 0$ exists at the boundaries $t = 1$ and $t = n$. $\Phi(\sigma, t)$ is increasing in σ, since λ_n, Λ_n, and ℓ_n are increasing functions. It achieves minimum at $\sigma = \underline{x}$, and the minimum value is $\Phi(\underline{x}, t) = (n-t+1) \log \lambda_n(\underline{x}) - \log(p_t/(1-p_t))$, where we used $\Lambda_n(\underline{x}) = 1$ and $\lambda_n(\underline{x}) = \ell_n(\underline{x})$. Noting that $\lambda_n(\underline{x}) < 1$ and $\log(p_t/(1-p_t)) > 0$, we obtain $\Phi(\underline{x}, t) < 0$ for any $t \in [1, n]$.

$\Phi(\sigma, t)$ achieves maximum at $\bar{x}$, and the maximum value is obtained as $\Phi(\bar{x}, t) = t \log \Lambda_n(\bar{x}) - \log(p_t/(1-p_t))$, using $\lambda_n(\bar{x}) = 1$ and $\Lambda_n(\bar{x}) = \ell_n(\bar{x})$. When $t = 1$, the maximum is $\Phi(\bar{x}, 1) = \log \Lambda_n(\bar{x}) - \log(p_1/(1-p_1))$. $\log \Lambda_n(\bar{x}) > \delta/n^\xi$ holds under Assumption 1, since $\log \Lambda_n(x) - \log \lambda_n(x) \geq \delta_n$ and $\lambda_n(\bar{x}) = 1$. In contrast, $\log(p_1/(1-p_1))$ declines to 0 as fast as $1/n$, as shown below. The market-clearing condition implies that $S'(p_t)dp_t = dt/n$. Using this, we obtain

$$\frac{d \log(p_t/(1-p_t))}{dp_t} \frac{dp_t}{dt} = \frac{1}{p_t(1-p_t)} \frac{1}{nS'(p_t)}.$$

Then, there exists some $c_1 > 0$ such that $\log(p_1/(1-p_1)) < c_1/n$, because

$$\log \frac{p_1}{1-p_1} = \log \frac{p_0}{1-p_0} + \int_0^1 \frac{1}{p_t(1-p_t)} \frac{1}{nS'(p_t)} dt,$$

where $1/S'$ is bounded since $S(\cdot)$ is strictly increasing. Thus, we obtain $\Phi(\bar{x}, 1) > \delta/n^\xi - c_1/n$, which is strictly positive for sufficiently large n since $\xi < 1$.

When $t = n$, the maximum of $\Phi(\sigma, n)$ is $n \log \Lambda_n(\bar{x}) - \log(p_n/(1-p_n))$. The second term is bounded, because $p_n/(1-p_n) < \bar{p}/(1-\bar{p})$. The first term tends to positive infinity as $n \to \infty$, since $n \log \Lambda_n(\bar{x}) > \delta n^{1-\xi}$. Thus, $\Phi(\bar{x}, n) > 0$ for sufficiently large n. Since $\Phi(\underline{x}, t) < 0$ and $\Phi(\bar{x}, t) > 0$ for $t = 1$ and $t = n$ and since Φ is continuous in σ, an interior solution σ exists for both $t = 1, n$ when n is sufficiently large.

Next, we show that the interior solution σ is decreasing in t. The total derivative of $\Phi(\sigma, t) = 0$ is

$$\begin{aligned}
&\frac{1}{p_t(1-p_t)} \frac{1}{nS'(p_t)} dt \\
&= \log \frac{\Lambda_n(\sigma)}{\lambda_n(\sigma)} dt + \left((n-t) \frac{\lambda_n'(\sigma)}{\lambda_n(\sigma)} + (t-1) \frac{\Lambda_n'(\sigma)}{\Lambda_n(\sigma)} + \frac{\ell_n'(\sigma)}{\ell_n(\sigma)} \right) d\sigma.
\end{aligned}$$

This determines the derivative of σ with respect to t as

$$\frac{d\sigma}{dt} = \left. \frac{-\log\left(\Lambda_n(x)/\lambda_n(x)\right) + \{p_t(1-p_t)S'(p_t)n\}^{-1}}{(n-t)\lambda_n'(x)/\lambda_n(x) + (t-1)\Lambda_n'(x)/\Lambda_n(x) + \ell_n'(x)/\ell_n(x)} \right|_{x=\sigma(t)}. \tag{9}$$

The denominator is strictly positive, since λ_n, Λ_n, and ℓ_n are strictly positive and strictly increasing. In the numerator, the first term is strictly negative, and $-\log(\Lambda_n(x)/\lambda_n(x)) < -\delta/n^{\xi}$ by Assumption 1. The second term in the numerator is positive and of order $1/n$, as shown above. Thus, the numerator is negative for large n. Hence, for sufficiently large values of n, we obtain that $d\sigma/dt \leq 0$.

Since an interior solution σ for $\Phi(\sigma, t)$ exists for $t = 1$ and $t = n$ and since an interior solution σ is decreasing in t, an interior solution of (3) exists for any $m \in \{1, 2, \ldots, n\}$.

Finally, since $D(p_m, \boldsymbol{x})$ is the number of traders with $x_i \geq \sigma(m)$ for $m = 1, 2, \ldots, n$, the decreasing function $\sigma(m)$ implies that $D(p_m, \boldsymbol{x})$ is increasing in m for any realization of $\boldsymbol{x}$.

Proof of Proposition 2

We define an aggregate reaction function as a mapping from the number of buying traders m to the number of buying traders determined by traders' choices given p_m

and their private signals. Specifically, the aggregate reaction function is given by $\Gamma_{\boldsymbol{x}} : \{0, 1, \ldots, n\} \mapsto \{0, 1, \ldots, n\}$ for each realization of $\boldsymbol{x}$. It coincides with D for $m > 0$, i.e., $\Gamma_{\boldsymbol{x}}(m) := D(p_m, \boldsymbol{x})$ for $m \in \{1, 2, \ldots, n\}$. For $m = 0$, we let $\Gamma_{\boldsymbol{x}}(0) = D(p_1, \boldsymbol{x})$. Then, $\Gamma_{\boldsymbol{x}}$ is an increasing mapping of $\{0, 1, \ldots, n\}$ onto itself. Moreover, $\{0, 1, \ldots, n\}$ is a finite totally ordered set, and thus it is a complete lattice. Therefore, by Tarski's fixed point theorem, there exists a non-empty closed set of fixed points of $\Gamma_{\boldsymbol{x}}$.

The auctioneer chooses $m^* = 0$ if $D(p_1, \boldsymbol{x}) = 0$, and chooses $m^* > 0$ such that $D(p_{m^*}, \boldsymbol{x})/n = S(p_{m^*}) = m^*/n$ if $D(p_1, \boldsymbol{x}) > 0$. Hence, the fixed points of $\Gamma_{\boldsymbol{x}}$ coincide with a set of equilibrium ouctome m^*. This establishes the existence of m^* and equilibrium price $p^* = p_{m^*}$.

Proof of Lemma 1

We transform $\phi_\infty := \operatorname{plim}_{n\to\infty} \phi_n$ using change of variable for density of $t = \sigma^{-1}(x)$:

$$\begin{aligned}\phi_\infty(t) &= \operatorname*{plim}_{n\to\infty} \pi_n(t)(n - \Gamma(t)) \\ &= \operatorname*{plim}_{n\to\infty} \left(1 - \frac{\Gamma(t)}{n}\right) n|\sigma'(t)| \left.\frac{f_n^H(x)}{F_n^H(x)}\right|_{x=\sigma(t)}.\end{aligned} \tag{10}$$

Using Equations (6) and (9) for $\sigma'(t)$, we obtain

$$n|\sigma'(t)|\frac{f_n^H(x)}{F_n^H(x)} = \left|\frac{\log(\Lambda_n(x)/\lambda_n(x)) - \{p_t(1-p_t)S'(p_t)n\}^{-1}}{\left(1 - \frac{t}{n}\right)\left(1 - \frac{\lambda_n(x)}{\ell_n(x)}\right) + \frac{1}{n}\frac{F_n^H(x)}{f_n^H(x)}\left(\frac{(t-1)\Lambda_n'(x)}{\Lambda_n(x)} + \frac{\ell_n'(x)}{\ell_n(x)}\right)}\right|. \tag{11}$$

We examine the right-hand side of (11) evaluated at $x = \sigma(t)$ as $n \to \infty$. Since $\{p_t(1-p_t)S'(p_t)\}^{-1}$ is bounded, the second term in the numerator is of order $1/n$. The second term in the denominator is also of order $1/n$ as can be shown below. First, f_n^H, Λ_n, and ℓ_n are strictly positive. Second, $F_n^H \leq 1$, and ℓ_n' is bounded, because f_n^s is assumed to have a bounded derivative. Finally, $\Lambda_n'(x)$ is bounded for $x \in \mathcal{X}$, as shown in (7).

We next examine $\Lambda_n(x)/\lambda_n(x)$ and $\lambda_n(x)/\ell_n(x)$ in the right-hand side of (11). To do so, we show that $\sigma(t) \to \bar{x}$ as $n \to \infty$ for finite t. We note that

$$\begin{aligned}\log \Lambda_n(\sigma) &= \log \frac{1 - F_n^H(\sigma)}{1 - F_n^L(\sigma)} = \log \frac{1/F_n^L(\sigma) - \lambda_n(\sigma)}{1/F_n^L(\sigma) - 1} \\ &= \log\left(1 + \frac{1 - \lambda_n(\sigma)}{1/F_n^L(\sigma) - 1}\right).\end{aligned}$$

Since $\log(1+y) \leq y$ and $1 + \log y \leq y$ for any $y \geq 0$, we have, for $\sigma < \bar{x}$,

$$\log \Lambda_n(\sigma) \leq \frac{1 - \lambda_n(\sigma)}{1/F_n^L(\sigma) - 1} \leq \frac{-\log \lambda_n(\sigma)}{1/F_n^L(\sigma) - 1}.$$

Hence, we obtain

$$\log \Lambda_n(\sigma) - \log \lambda_n(\sigma) \leq \frac{-\log \lambda_n(\sigma)}{1 - F_n^L(\sigma)}.$$

Assumption 1 implies $\log \Lambda_n - \log \lambda_n > \delta n^{-\xi}$. Thus, for sufficiently large n,

$$-\log \lambda_n(\sigma) \geq (1 - F_n^L(\sigma))\delta n^{-\xi}. \tag{12}$$

Now, Equation (8) can be modified as:

$$n \log \lambda_n(\sigma) = \log \frac{p_t}{1 - p_t} + t \log \frac{\lambda_n(\sigma)}{\Lambda_n(\sigma)} + \log \frac{\Lambda_n(\sigma)}{\ell_n(\sigma)}. \tag{13}$$

The right-hand side of (13) is finite for any finite t. The left-hand side of (13) would diverge toward negative infinity as $n \to \infty$ if $F_n^L(\sigma)$ were bounded by a value strictly below 1, as implied by inequality (12) and $\xi < 1$. Hence, (13) holds only if $F_n^L(\sigma)$ tends to 1, which is equivalent to that $\sigma(t) \to \bar{x}$ as $n \to \infty$ for any finite t. This implies that $\Lambda_n(\sigma(t))/\lambda_n(\sigma(t))$ tends to $\ell_n(\sigma(t))/\lambda_n(\sigma(t))$ as $n \to \infty$, since $\Lambda(\bar{x}) = \ell(\bar{x})$.

Thus, using $z_n := \log(\Lambda_n(\sigma(t))/\lambda_n(\sigma(t)))$, the right-hand side of Equation (11) is expressed as

$$\lim_{n\to\infty} \frac{z_n - O(1/n)}{(1 - t/n)\,(1 - e^{-z_n}) + O(1/n)} = \lim_{n\to\infty} \frac{z_n - O(1/n)}{(1 - t/n)\,(z_n + O(z_n^2)) + O(1/n)},$$

where we used $\lim_{n\to\infty} z_n = 0$ and Taylor expansion of $e^{z_n} - 1$ around $z_n = 0$, and we used notation $y_n = O(x_n)$ if there exist c_o and n_o such that $|y_n| \le c_o x_n$ for any $n \ge n_o$. Dividing both the denominator and numerator by z_n and applying $nz_n > \delta n^{1-\xi}$ with $\xi < 1$ (Assumption 1), we obtain

$$\lim_{n\to\infty} \frac{1 - O(1/(nz_n))}{(1 - t/n)\,(1 + O(z_n)) + O(1/(nz_n))} = 1.$$

Substituting the above result in (10), we obtain $\phi_\infty(t) = \operatorname{plim}_{n\to\infty} 1 - \Gamma(t)/n$. This implies that $\phi_\infty(t)$ is bounded, and hence, the asymptotic variance of $\Gamma(t+dt) - \Gamma(t)$ is also bounded. Hence as $n \to \infty$, $\Gamma(t)/n$ converges in the L^2-norm, and thus in probability, to 0. Hence, we obtain that $\phi_\infty(t) = 1$ for finite t.

Proof of Lemma 2

We show that the random variable $\tau_{\varphi_n(\cdot)}$ defined over $[0, \infty]$ converges in distribution to τ_1 as n tends to ∞. We prove this by showing that the Laplace transform of $\tau_{\varphi_n(\cdot)}$ converges to that of τ_1 as $n \to \infty$. Namely, we show that, for any $\eta > 0$,

$$\lim_{n\to\infty} \mathbb{E}\left[\exp(-\eta\tau_{\varphi_n(\cdot)})\right] = \mathbb{E}\left[\exp(-\eta\tau_1)\right]. \tag{14}$$

Note that $e^{-\eta\tau}$ is set at 0 for the events where $\tau = \infty$ by convention.

In (10), we observe that $\phi_n(t) = \varphi_n(t-1)$ is a product of (11) and a stochastic term $1 - \Gamma(t)/n$. The former term converges to 1 uniformly over any finite interval $[0, T]$, and the latter term converges in probability to 1 as $n \to \infty$. Thus, the probability of events in which $\Gamma(t)/n$ exceeds $n^{-\nu_0}$ for some $t \in [0, T]$ for a fixed $\nu_0 \in (0, 1)$ declines to 0 as $n \to \infty$.[13] Since $e^{-\eta\tau}$ is bounded, such events have vanishingly small contribution to the expectation in the left-hand side of (14). Combining with the fact that (11) is

[13] See Technical Appendix for the construction of ν_0.

uniformly convergent to 1, there exists a sequence ϵ_n such that $1-\epsilon_n < \varphi_n(t) < 1+\epsilon_n$ for finite t excluding those events where $\Gamma(t)/n$ exceeds $n^{-\nu_0}$.

Since an inhomogeneous Poisson process can be transformed to a homogeneous Poisson process with a change of time, inequalities $\tau_{1-\epsilon_n} \le \tau_{\varphi_n(\cdot)} \le \tau_{1+\epsilon_n}$ hold for each realization of $\boldsymbol{x}$. Thus, in order to establish (14), it is sufficient to show that $\mathbb{E}[\exp(-\eta\tau_\psi)]$ is continuous with respect to $\psi > 0$. We also note that

$$\begin{aligned}\tau_\psi &= \inf\{t \ge 0 \mid N(\psi t) \le t - c_o\} \\ &= \inf\{t \ge 0 \mid t - N(\psi t) \ge c_o\} \\ &= \frac{1}{\psi} \inf\left\{t \ge 0 \mid \frac{t}{\psi} - N(t) \ge c_o\right\} \\ &= \frac{1}{\psi}\tilde{\tau}_\psi\end{aligned}$$

where $\tilde{\tau}_\psi := \inf\{t \ge 0 \mid N(t) \le t/\psi - c_o\}$.

Let ζ be a constant in $(0,1)$. Consider a stochastic differential equation:

$$dZ(t) = -\zeta Z(t\text{-})\{dN(t) - dt\}, \quad Z(0) = 1,$$

where $Z(t\text{-})$ denotes the value of $Z(t)$ before a jump occurs at t if any. The solution of the stochastic differential equation is a martingale and satisfies

$$Z(t) = e^{\zeta t}(1-\zeta)^{N(t)} = \left(\frac{1}{1-\zeta}\right)^{\frac{t}{\psi}-N(t)} \exp\left\{\left(\zeta + \frac{\log(1-\zeta)}{\psi}\right)t\right\},$$

where the second equation is obtained by multiplying and dividing by $(1-\zeta)^{t/\psi}$.

Now, for fixed η and ψ, there exists a unique ζ that satisfies an equation

$$\zeta\psi + \log(1-\zeta) = -\eta.$$

Let $\zeta(\eta,\psi)$ denote the unique solution. Note that $\zeta(\eta,\psi)$ is continuous and monotonically increasing with respect to both η and ψ. Then, Z is written as

$$Z(t) = \left(\frac{1}{1-\zeta(\eta,\psi)}\right)^{\frac{t}{\psi}-N(t)} \exp\left(-\frac{\eta}{\psi}t\right).$$

Note that $t/\psi - N(t) = c_o$ at the stopping time $t = \tilde{\tau}_\psi$. Thus, $Z(t)$ is positive and takes a value less than or equal to $\{1 - \zeta(\eta, \psi)\}^{-c_o}$ at and before the stopping time $\tilde{\tau}_\psi$. Namely, $Z(t)$ is bounded. Therefore, $\mathbb{E}[Z(\tilde{\tau}_\psi)] = 1$ holds by the optional sampling theorem. (Note that $Z = 0$ for the events where $\tilde{\tau}_\psi = \infty$.) Moreover, noting that $N(t)$ does not jump at the point of time $\tilde{\tau}_\psi$, we obtain that

$$Z(\tilde{\tau}_\psi) = \left(\frac{1}{1 - \zeta(\eta, \psi)}\right)^{c_o} \exp\left(-\frac{\eta}{\psi}\tilde{\tau}_\psi\right),$$

for both cases of $\tilde{\tau}_\psi < \infty$ and $\tilde{\tau}_\psi = \infty$. Thus,

$$\mathbb{E}[\exp(-\eta\tau_\psi)] = \mathbb{E}\left[\exp\left(-\frac{\eta}{\psi}\tilde{\tau}_\psi\right)\right] = \{1 - \zeta(\eta, \psi)\}^{c_o}.$$

Since $\zeta(\eta, \psi)$ is continuous with respect to ψ, this completes the proof.

Proof of Proposition 3

Consider the Poisson process $N(t)$ with intensity 1 and $N(0) = c_o > 0$. The first passage time τ_1 of $N(t)$ reaching $t - c_o$ must be greater than or equal to c_o. Now we introduce a process b with $b(0) = c_o$. During the time interval c_o, the increment $N(c_o) - N(0)$, denoted as $b(1)$, follows a Poisson distribution with mean c_o. Since a Poisson random variable is infinitely divisible, a Poisson random variable with mean c_o is equivalent to c_o-fold convolution of the Poisson with mean 1. Thus, we can regard $b(1)$ as the sum of "children" borne by $c_o = b(0)$ "parents," where each parent bears a number of children following the Poisson with mean 1. If $b(1) = 0$, the process b stops, and the first passage time is $b(0) = c_o$. If $b(1) > 0$, the first passage time is greater than or equal to $b(0) + b(1)$. During the time interval $(b(0), b(0) + b(1)]$, new increment $b(2) := N(b(0) + b(1)) - N(b(0))$ follows the Poisson distribution with mean $b(1)$, which is equivalent to $b(1)$-fold convolution of the Poisson with mean 1 and regarded as the number of children borne by $b(1)$ parents (note that the increment $b(1)$ of a Poisson process is always an integer). This process $b(u)$ continues for $u = 1, 2, \ldots, U$, where U denotes the stopping time at which $b(U)$ is equal to 0 for the first time. Thus, the first passage time τ_1 is equal to $\sum_{u=0}^{U} b(u)$, the total number of population generated in the so-called Poisson branching process $b(u)$ in which each parent bears a number of children according to the Poisson distribution with mean 1.

It is known that the sum of the Poisson branching process, cumulated over time until the process stops, follows a Borel-Tanner distribution (Kingman [31]; see also Nirei [43]). When the Poisson mean of the branching process $b(u)$ is $\phi > 0$ generally, the Borel-Tanner distribution is written as

$$\Pr\left(\sum_{u=0}^{U} b(u) = m \mid b(0) = c_o\right) = \frac{c_o}{m}\frac{e^{-\phi m}(\phi m)^{m-c_o}}{(m-c_o)!}, \tag{15}$$

for $m = c_o, c_o + 1, \ldots$. By applying Stirling's formula, we obtain the tail characterization:

$$\Pr\left(\sum_{u=0}^{U} b(u) = m \mid b(0) = c_o\right) \propto e^{-(\phi-1-\log\phi)m} m^{-1.5} \tag{16}$$

as $m \to \infty$. Since τ_1 follows the same distribution as the sum of the Poisson branching process with mean 1, it follows (15) and (16) with $\phi = 1$.

Finally, we change variables using $M_n^\dagger = \tau_1 + 1$ and $c = \Gamma(1) = G(0) + 1 = c_o + 1$. With $m' := m + 1$, (15) is rewritten as

$$\Pr(M_n^\dagger = m' \mid \Gamma(1) = c) = \frac{c-1}{m'-1}\frac{e^{-\phi(m'-1)}(\phi(m'-1))^{m'-c}}{(m'-c)!}.$$

Using $\phi = 1$, we obtain the desired result.

Proof of Proposition 4

Under finite ϕ_o, $\Gamma(1)$ asymptotically follows a Poisson distribution with mean ϕ_o. Hence, for $m = \{0, 1\}$, $\Pr(M_n^\dagger = m)$ asymptotically follows $\Pr(\Gamma(1) = m) = \phi_o^m e^{-\phi_o}/m!$.

For $m > 1$, the unconditional distribution of $M_n^\dagger$ is derived by combining the distribution (4) and the Poisson distribution with mean ϕ_o for $\Gamma(1)$ as follows.

$$\begin{aligned}
&\sum_{c=2}^{m} \Pr\left(M^\dagger = m \mid \Gamma(1) = c\right) \Pr(\Gamma(1) = c) \qquad (17)\\
&= \sum_{c=2}^{m} \frac{(c-1)(m-1)^{m-c-1}e^{-m+1}}{(m-c)!} \frac{\phi_o^c e^{-\phi_o}}{c!}\\
&= \sum_{c=1}^{m} \frac{(c-1)(m-1)^{m-c-1}e^{-m+1}}{(m-c)!} \frac{\phi_o^c e^{-\phi_o}}{c!}\\
&= \frac{e^{-\phi_o-m+1}(m-1)^{m-1}}{m!} \left[\sum_{c=1}^{m} \frac{(\phi_o/(m-1))^c m!}{(m-c)!(c-1)!} - \sum_{c=1}^{m} \frac{(\phi_o/(m-1))^c m!}{(m-c)!c!}\right].
\end{aligned}$$

Using the binomial theorem, we obtain

$$\begin{aligned}
\sum_{c=1}^{m} \frac{(\phi_o/(m-1))^c m!}{(m-c)!(c-1)!} &= \frac{\phi_o m}{m-1} \sum_{c=1}^{m} \frac{(\phi_o/(m-1))^{c-1}(m-1)!}{(m-c)!(c-1)!}\\
&= \frac{\phi_o m}{m-1} \sum_{c'=0}^{m-1} \frac{(\phi_o/(m-1))^{c'}(m-1)!}{(m-1-c')!c'!}\\
&= \frac{\phi_o m}{m-1} \left(1 + \frac{\phi_o}{m-1}\right)^{m-1}
\end{aligned}$$

and

$$\begin{aligned}
\sum_{c=1}^{m} \frac{(\phi_o/(m-1))^c m!}{(m-c)!c!} &= \sum_{c=0}^{m} \frac{(\phi_o/(m-1))^c m!}{(m-c)!c!} - 1\\
&= \left(1 + \frac{\phi_o}{m-1}\right)^{m} - 1.
\end{aligned}$$

Substituting back to (17), we obtain

$$\begin{aligned}
&\frac{e^{-\phi_o-m+1}(m-1)^{m-1}}{m!} \left[\frac{\phi_o m}{m-1}\left(1 + \frac{\phi_o}{m-1}\right)^{m-1} - \left(1 + \frac{\phi_o}{m-1}\right)^{m} + 1\right]\\
&= \frac{e^{-\phi_o-m+1}(m-1)^{m-1}}{m!} \left[(\phi_o - 1)\left(1 + \frac{\phi_o}{m-1}\right)^{m-1} + 1\right].
\end{aligned}$$

Thus, we obtain the desired result. Applying Stirling's formula, we obtain that the tail follows a power law with exponent 0.5.

Proof of Proposition 5

The proof proceeds in two steps. First, we show that the distributions under Assumption 3 are compatible with Assumption 1. Second, we show that $\phi_o = 1$. Then, Proposition 5 is immediately obtained by substituting $\phi_o = 1$ into the result in Proposition 4.

The distributions in the class specified by Assumption 3 have positive density $f_n^s(x) = a_n^s e^{a_n^s(x-\bar{x})} > 0$ for $s = H, L$ and increasing likelihood $\ell_n(x) = (a_n^H/a_n^L)e^{(a_n^H - a_n^L)(x-\bar{x})}$ for $x \in (\bar{x}-\epsilon, \bar{x}]$. Moreover, the distributions satisfying Assumption 3 also satisfy Assumption 1 in $(\bar{x}-\epsilon, \bar{x}]$, as shown below. We have

$$\log \frac{\Lambda_n(x)}{\lambda_n(x)} = \log \frac{1 - e^{a_n^H(x-\bar{x})}}{1 - e^{a_n^L(x-\bar{x})}} - (a_n^H - a_n^L)(x - \bar{x}).$$

L'Hopital's rule yields:

$$\log \frac{\Lambda_n(\bar{x})}{\lambda_n(\bar{x})} = \log \left. \frac{1 - e^{a_n^H(x-\bar{x})}}{1 - e^{a_n^L(x-\bar{x})}} \right|_{x=\bar{x}} = \log \left. \frac{a_n^H e^{a_n^H(x-\bar{x})}}{a_n^L e^{a_n^L(x-\bar{x})}} \right|_{x=\bar{x}} = \log \frac{a_n^H}{a_n^L}.$$

It can be also shown that $\log(\Lambda_n(x)/\lambda_n(x))$ is decreasing in x as follows:

$$\begin{aligned}
&\frac{d}{dx} \log \frac{\Lambda_n(x)}{\lambda_n(x)} = \frac{\Lambda_n'(x)}{\Lambda_n(x)} - \frac{\lambda_n'(x)}{\lambda_n(x)} \\
&= \frac{f_n^L(x)}{1 - F_n^L(x)} \left(1 - \frac{\ell_n(x)}{\Lambda_n(x)}\right) + \frac{f_n^L(x)}{F_n^L(x)} \left(1 - \frac{\ell_n(x)}{\lambda_n(x)}\right) \\
&= \frac{a_n^L e^{a_n^L(x-\bar{x})}}{1 - e^{a_n^H(x-\bar{x})}} \left(\frac{1 - e^{a_n^H(x-\bar{x})}}{1 - e^{a_n^L(x-\bar{x})}} - \frac{a_n^H}{a_n^L} e^{(a_n^H - a_n^L)(x-\bar{x})} \right) + a_n^L \left(1 - \frac{a_n^H}{a_n^L}\right) \\
&= \frac{a_n^L}{1 - e^{a_n^L(x-\bar{x})}} - \frac{a_n^H}{1 - e^{a_n^H(x-\bar{x})}} < 0,
\end{aligned}$$

where the last inequality uses the fact,

$$\frac{\partial}{\partial a}\frac{a}{1-e^{a(x-\bar{x})}} = \frac{1-(1-a(x-\bar{x}))e^{a(x-\bar{x})}}{(1-e^{a(x-\bar{x})})^2} > 0,$$

noting $\log(1+y) - y < 0$ for any $y > 0$. Hence, we have $\log(\Lambda_n(x)/\lambda_n(x)) \geq \log(\Lambda_n(\bar{x})/\lambda_n(\bar{x})) = \log a_n^H/a_n^L > \delta/n^{\xi}$ for any $x \in (\bar{x}-\epsilon, \bar{x}]$.

In the next step, we show that $\phi_o = 1$ holds under Assumption 3. Equation (3) implies that $\sigma(1)$ must satisfy $p_1/(1-p_1) = \lambda_n(\sigma(1))^{n-1}\ell_n(\sigma(1))$. This equation is solved under Assumption 3 as

$$a_n^H(\sigma(1)-\bar{x}) = \frac{a_n^H}{n(a_n^H - a_n^L)}\log\left(\frac{p_1}{1-p_1}\frac{a_n^L}{a_n^H}\right).$$

Let z_n denote the right-hand side of the above equation. Then, we obtain

$$1 - F_n^H(\sigma(1)) = 1 - e^{a_n^H(\sigma(1)-\bar{x})} = 1 - e^{z_n}.$$

Taylor expansion around $z_n = 0$ generates

$$\begin{aligned} 1 - F_n^H(\sigma(1)) &= -\frac{a_n^H}{n(a_n^H - a_n^L)}\log\left(\frac{p_1}{1-p_1}\frac{a_n^L}{a_n^H}\right) + O(z_n^2) \\ &= \frac{a_n^H}{a_n^L}\frac{\log(a_n^H/a_n^L)}{n(a_n^H/a_n^L - 1)} - \frac{a_n^H}{a_n^L}\frac{\log(p_1/(1-p_1))}{n(a_n^H/a_n^L - 1)} + O(z_n^2). \end{aligned} \tag{18}$$

Since $S(p_1) = 1/n$, $S(p_0) = 0$ and $p_0 = 0.5$, we have $p_1 = 0.5 + (nS'(p_0))^{-1} + O(n^{-2})$. This implies that $\log(p_1/(1-p_1)) = O(n^{-1})$. Also, $a_n^H/a_n^L - 1 \geq \log(a_n^H/a_n^L) > \delta/n^{\xi}$. Hence, the absolute value of the second term in (18) is bounded by $O(n^{\xi-2})$. Also by L'Hopital's rule we have

$$\lim_{(a_n^H/a_n^L)\to 1}\frac{\log(a_n^H/a_n^L)}{a_n^H/a_n^L - 1} = 1.$$

Applying these results to (18), we obtain $\phi_o = \lim_{n\to\infty} n(1 - F_n^H(\sigma(1))) = 1$.

References

[1] Daron Acemoglu, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. Microeconomic origins of macroeconomic tail risks. *American Economic Review*, 107:54–108, 2017.

[2] 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.

[3] Andrei Ankudinov, Rustam Ibragimov, and Oleg Lebedev. Heavy tails and asymmetry of returns in the russian stock market. *Emerging Markets Review*, 32:200–219, 2017.

[4] Andrea Attar, Thomas Mariotti, and François Salanié. Nonexclusive competition under adverse selection. *Theoretical Economics*, 9:1–40, 2014.

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

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

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

[8] 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.

[9] Markus K. Brunnermeier. *Asset Pricing under Asymmetric Information—Bubbles, Crashes, Technical Analysis, and Herding*. Oxford University Press, 2001.

[10] Jeremy Bulow and Paul Klemperer. Rational frenzies and crashes. *Journal of Political Economy*, 102:1–23, 1994.

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

[12] Christophe P. Chamley. *Rational Herds—Economic Models of Social Learning.* Cambridge University Press, 2004.

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

[14] Marco Cipriani and Antonio Guarino. Herd behavior in a laboratory financial market. *American Economic Review*, 95:1427–1443, 2005.

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

[16] Rama Cont, Marc Potters, and Jean-Philippe Bouchaud. Scaling in stock market data: Stable laws and beyond. In *Scale Invariance and Beyond*, pages 75–85. Springer, 1997.

[17] Rama Cont and Peter Tankov. *Financial Modelling with Jump Processes*. Chapman and Hall, 2004.

[18] Laurens de Haan, Sidney I. Resnick, Holger Rootzén, and Casper G. de Vries. Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Processes and their Applications*, 32:213–224, 1989.

[19] Eugene F Fama. Mandelbrot and the stable paretian hypothesis. *Journal of Business*, 36(4):420–429, 1963.

[20] Xavier Gabaix. Power laws in economics and finance. *Annual Review of Economics*, 1(1):255–294, 2009.

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

[22] Parameswaran Gopikrishnan, Martin Meyer, Luís A. Nunes Amaral, and H. Eugene Stanley. Inverse cubic law for the distribution of stock price variations. *European Physical Journal B*, 3(2):139–140, 1998.

[23] Parameswaran Gopikrishnan, Vasiliki Plerou, Xavier Gabaix, and H. Eugene Stanley. Statistical properties of share volume traded in financial markets. *Physical Review E*, 62:4493–4496, 2000.

[24] Sanford J. Grossman. Nash equilibrium and the industrial organization of markets with large fixed costs. *Econometrica*, 49:1149–1172, 1981.

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

[26] Joel Hasbrouck. Measuring the information content of stock trades. *Journal of Finance*, 46:179–207, 1991.

[27] Marat Ibragimov, Rustam Ibragimov, and Paul Kattuman. Emerging markets and heavy tails. *Journal of Banking & Finance*, 37(7):2546–2559, 2013.

[28] Rustam Ibragimov. Heavy-tailed densities. *The New Palgrave Dictionary of Economics Online*, 2009.

[29] 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.

[30] 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.

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

[32] Paul D. Klemperer and Margaret A. Meyer. Supply function equilibria in oligopoly under uncertainty. *Econometrica*, 57:1243–1277, 1989.

[33] Albert S. Kyle. Informed speculation with imperfect competition. *Review of Economic Studies*, 56:317–356, 1989.

[34] Andreas Kyprianou, Wim Schoutens, and Paul Wilmott. *Exotic Option Pricing and Advanced Lévy Models.* Wiley, 2005.

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

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

[37] Thomas Lux. The stable paretian hypothesis and the frequency of large returns: an examination of major german stocks. *Applied Financial Economics*, 6(6):463–475, 1996.

[38] Thomas Lux and Simone Alfarano. Financial power laws: Empirical evidence, models, and mechanisms. *Chaos, Solitons & Fractals*, 88:3–18, 2016.

[39] Thomas Lux and Didier Sornette. On rational bubbles and fat tails. *Journal of Money, Credit and Banking*, 34:589–610, 2002.

[40] Benoit Mandelbrot. The variation of some other speculative prices. *Journal of Business*, 40(4):393–413, 1967.

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

[42] Frederic S. Mishkin. Asymmetric information and financial crises: A historical perspective. In R. Glenn Hubbard, editor, *Financial Markets and Financial Crises*, chapter 3, pages 69–108. The University of Chicago Press, 1991.

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

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

[45] Makoto Nirei. Beauty contests and fat tails in financial markets. *Research Center for Price Dynamics Working Paper*, 76, 2011.

[46] Makoto Nirei. An interaction-based foundation of aggregate investment fluctuations. *Theoretical Economics*, 10:953–985, 2015.

[47] André Orléan. Bayesian interactions and collective dynamics of opinion: Herd behavior and mimetic contagion. *Journal of Economic Behavior and Organization*, 28:257–274, 1995.

[48] Andreas Park and Hamid Sabourian. Herding and contrarian behavior in financial markets. *Econometrica*, 79:973–1026, 2011.

[49] Vasiliki Plerou, Parameswaran Gopikrishnan, Xavier Gabaix, Luís A. Nunes Amaral, and H. Eugene Stanley. Price fluctuations, market activity, and trading volume. *Quantitative Finance*, 1(2):262–269, 2001.

[50] Khaled Salhi, Madalina Deaconu, Antoine Lejay, Nicolas Champagnat, and Nicolas Navet. Regime switching model for financial data: Empirical risk analysis. *Physica A*, 461:148–157, 2016.

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

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

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

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

[55] Xavier Vives. *Information and Learning in Markets.* Princeton, 2008.

[56] Xavier Vives. Strategic supply function competition with private information. *Econometrica*, 79:1919–1966, 2011.