Theoretical Economics 15 (2020), 1365–1398 1555-7561/20201365

Trade clustering and power laws in financial markets

Makoto Nirei
Graduate School of Economics, University of Tokyo

John Stachurski
Research School of Economics, Australian National University

Tsutomu Watanabe
Graduate School of Economics, University of Tokyo

This study provides an explanation for the emergence of power laws in asset trading volume and returns. We consider a two-state model with binary actions, where traders infer other traders' private signals regarding the value of an asset from their actions and adjust their own behavior accordingly. We prove that this leads to power laws for equilibrium volume and returns whenever the number of traders is large and the signals for asset value are sufficiently noisy. 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. G14.

1. Introduction

Recently, the literature on empirical finance has converged on a broad consensus: Daily returns on 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 found for trading volume by Gopikrishnan et al. (2000) and Plerou et al. (2001).

Makoto Nirei: nirei@e.u-tokyo.ac.jp
John Stachurski: john.stachurski@anu.edu.jp
Tsutomu Watanabe: watanabe@e.u-tokyo.ac.jp
We have benefited from comments by the anonymous referees, Daisuke Oyama, and especially Koichiro Takaoka. Nirei acknowledges financial support from Japanese Society for the Promotiion of Science (JSPS) Kakenhi Grant 15K03350. Stachurski acknowledges financial support from ARC Grant FT160100423. Watanabe acknowledges financial support from JSPS Kakenhi Grant 18H05217. Working paper versions of this study were circulated under titles "Herding and power laws in financial markets," "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" (Nirei 2011).

[1]See, for example, Jansen and de Vries (1991), Lux (1996), Cont et al. (1997), Gopikrishnan et al. (1998), Ibragimov et al. (2013), and Ankudinov et al. (2017). Earlier work on heavy-tailed returns can be found in Fama (1963) and Mandelbrot (1967). For overviews of the literature, see Lux and Alfarano (2016) or Gabaix (2009).

 https://doi.org/10.3982/TE3523

These power-law findings are highly consequential, mainly because extreme outcomes are by definition rare, so attempts to estimate prices or quantities with tail risk sensitivity through nonparametric methods are deeply problematic (Salhi et al. 2016). Thus, information on the specific functional form of the tails of these distributions has great value for 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 2009).

In this paper, we respond to the developing empirical consensus by building a model of asset markets that generates a power law in both volume and price. The underlying driver of this power law is private asymmetric information on the value of assets, dispersed among many traders. Suppose that traders choose between buying and not buying. The action of buying suggests a positive private signal. As a result, a single trader's action can cause clustering of similar actions by other traders. This trade clustering leads to power laws in volume and returns. In particular, we consider a series of markets in which the number of traders increases and the informativeness of signals diminishes, and we show that the equilibrium size of trade clustering asymptotically exhibits power-law fluctuations.

To further understand our power-law result, suppose that, for each realization of private signals, informed traders are sorted in descending order according to their signals and then classified as follows: The first group of traders buys regardless of the actions of other traders. The second group buys if there is at least one trader buying. The third group buys if there are at least two other traders buying, and so forth. Now consider a fictitious best response dynamic where traders choose whether to buy after viewing the decisions of previous traders. We show that under reasonable assumptions on the informativeness of the private signal, the decision to buy on the part of one trader induces on average one new trader to buy. An analogy can be made with Keynes' beauty contest, where a voter's decision is affected by the average actions of n other voters. As a consequence, one vote has an impact of size $1/n$ on the decisions of others. In our model, when an investor has an incentive to imitate the average behavior of n traders, the act of buying by one trader has an impact of size $1/n$ on the other traders' behavior.

To understand the implications of this property, we view excess demand as a stochastic process, indexed by the number of buyers (rather than time) and generated by the fictitious best response dynamic discussed above. The first passage to zero for this process produces an equilibrium number of buying traders. Because the decision to buy by one trader induces on average one new trader to buy, this excess demand process is a martingale. As is well known, the first passage time to zero for a martingale follows a power-law distribution.[2] In this way, we derive a power-law distribution for the number of buying traders, which translates to the equilibrium trading volume.

The market environment of our model draws on Minehart and Scotchmer (1999), where a large number of informed traders receive private signals on a binary state of the

[2] For example, the first passage time of a Brownian motion with no drift follows a particular inverse Gaussian distribution, which has an asymptotic power-law tail with exponent 0.5. Further examples can be found in Redner (2001).

world, and simultaneously choose between buying one unit of an asset or not buying at all. Informed traders submit demand schedules conditional on all possible prices, rather than choosing an action unconditionally. This type of market competition was formulated as Nash equilibria in supply functions by Grossman (1981) and Klemperer and Meyer (1989), and has been introduced to the analysis of asset markets with private signals by Kyle (1989), Vives (2011), and Attar et al. (2014). However, none of these models leads to a power law.

Herd behavior models, which connect asymmetric information to excess fluctuations in asset pricing, have also served as inspiration for our research. The models of herding and information cascades proposed by Banerjee (1992) and Bikhchandani et al. (1992) have been employed to examine financial market fluctuations.[3] Gul and Lundholm (1995) demonstrated the emergence of stochastic clustering by endogenizing traders' choice of waiting time. Signal properties leading to herding behavior in sequential trading were identified by Smith and Sørensen (2000) and Park and Sabourian (2011). While none of these models generates a power law of financial fluctuations specifically, we inherit the spirit of these models, in which asymmetric information among traders results in trade clustering.

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 at criticality. Unlike the present study, these papers do not address why trader connectivity should exhibit criticality.[5]

In another strand of the literature, Lux and Sornette (2002) show that a stochastic rational bubble can produce a power law. Gabaix et al. (2006) generate power laws for trading volume and price changes when 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 clustering behavior by investors. This is in line with many previous studies that have linked asymmetric information in financial markets to phenomena such as crises, cascades, and herding. These studies range from a historical account of crises by Mishkin (1991) to the estimation of information content of trading volume on prices by Hasbrouck (1991). 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 this impact to the ubiquitously observed power-law fluctuations.

[3]See also Caplin and Leahy (1994), Lee (1998), Chari and Kehoe (2004), and Cipriani and Guarino (2005). For extensive surveys, see Brunnermeier (2001), Chamley (2004), and Vives (2008).

[4]Studies in this literature include Bak et al. (1997), Cont and Bouchaud (2000), Stauffer and Sornette (1999).

[5]In a similar vein, Nirei (2008) sketched out the basic idea that herd behavior can generate power–law-sized cascades in an environment similar to Orléan (1995), but fell short of substantiating his claim with rigorous analysis. This paper generates a power law in a standard market microstructure model, which allows us to relate the conditions necessary for generating power laws to a broad range of studies in financial economics.

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 intuition for the mechanism behind it. Section 3.2 elaborates on the power-law exponent for volume. Section 3.3 shows that a power law holds for returns. Section 3.4 numerically confirms that the equilibrium volumes follow a power law with a finite number of traders and that the equilibrium return distribution matches its empirical counterpart. Section 3.5 discusses some extensions of the model and Section 4 concludes. Long proofs are deferred to the Appendix.

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. Uninformed traders supply a single asset and informed traders demand it.[6] Informed traders are risk neutral and indexed by $i \in \{1, \dots, n\}$. There is an underlying state s that affects the value of the asset and takes values in $\{H, L\}$. We assume in particular that the asset has common intrinsic value 1 in state H and 0 in state L. While the true state is not known to any market participant, all agents hold a common prior for s given by $\Pr(H) = \Pr(L) = 1/2$. Moreover, informed traders receive a private signal X_i drawn independently from a common distribution F^s with finite support $[x_a, x_b]$. This signal is used to make inferences about the value of s, as described below.

Let $S(p)$ denote aggregate supply by uninformed traders at price p. We assume that S is continuously differentiable and strictly increasing with $S(1/2) = 0$, so that aggregate supply is zero at the price level that reflects the common prior. 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 to buy a single trading unit, set to $1/n$ so as to normalize maximum total demand to unity. Hence aggregate demand takes values in the discrete set $\{0, 1/n, \dots, 1\}$. The equilibrium price p^* takes values in $\{p_0, p_1, \dots, p_n\}$, where each p_m is determined by the market-clearing condition $S(p_m) = m/n$. Since $S(1/2) = 0$, we have $p_0 = 1/2$.

The demand function d_i of an informed trader describes his action for each realization of price, given his private signal. In particular, trader i buys at p when $d_i(p \mid x_i) = 1$ and refrains when $d_i(p \mid x_i) = 0$. Let $\mathcal{D}$ be the set of all such (binary) functions on

[6] We later discuss the case where 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 (1998) as a condition for herding to occur in financial markets.

$[x_a, x_b] \times \{p_1, p_2, \dots, p_n\}$.[7] Aggregate demand expressed in terms of trade volume is

$$D\big(p \mid \mathbf{X}\big) := \sum_{i=1}^{n} d_i(p \mid X_i),$$

where $\mathbf{X} = (X_i)_{i=1}^n$ denotes a profile of private signals.

Outcomes evolve as follows.

Step 1. Nature selects the state $s \in \{H, L\}$.

Step 2. A signal profile $\mathbf{X}$ is drawn from the joint distribution $\prod_{i=1}^n F^s$.

Step 3. Informed traders submit their demand functions to the auctioneer.

Step 4. The auctioneer determines the equilibrium price p^*.

Step 5. Transactions take place, with a unit of the asset delivered to each trader i with $d_i(p^* \mid X_i) = 1$.

Step 6. Equilibrium trading volume is realized as $m^* := D\big(p^* \mid \mathbf{X}\big)$.

In Step 4, the auctioneer obeys the following protocol: If $D(p_1 \mid \mathbf{X}) = 0$, then the auctioneer sets $p^* = p_0$, since no informed trader is willing to buy given that all other traders do not buy. If $D\big(p_1 \mid \mathbf{X}\big) > 0$, then the auctioneer determines $p^* > p_0$ such that

$$S\big(p^*\big) = \frac{D\big(p^* \mid \mathbf{X}\big)}{n}.$$

Since the asset has common value $\mathbb{1}\{s = H\}$ and its purchase cost is p, a trader who buys obtains payoff $\mathbb{1}\{s = H\} - p$. Therefore, the expected payoff of a trader who buys at signal X_i is $r_i(p, X_i) - p$, where $r_i(p, X_i)$ denotes the probability of $s = H$, conditional on signal X_i and equilibrium price p. A trader who refrains from buying obtains zero in either state.

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. Let $\Omega_{m,i}$ denote such an event. Since X_i is independent of other traders' decisions d_j given m, we have

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

Our equilibrium concept is defined as follows.

Definition 1 (Equilibrium). A Bayesian Nash equilibrium consists of a profile of informed traders' demand functions $d_i \in \mathcal{D}$, a profile of conditional probabilities r_i obeying (1), and an equilibrium price correspondence p^* such that (i) for any $i = 1, 2, \dots, n$ and at each information set (p, x_i), d_i maximizes expected payoff given d_j for $j \neq i$, (ii)

[7] Here p_0 is excluded because p_0 cannot be realized in equilibrium if any trader, including i, chooses buying at p_0.

for any $i = 1, 2, \ldots, n$, r_i is consistent with demand functions $\{d_j\}$ and equilibrium price correspondence p^*, and (iii) p^* clears the market. That is, $nS(p^*) = D(p^* \mid \mathbf{x})$ for all $\mathbf{x} \in [x_a, x_b]^n$.

2.2 *The signal*

To consider outcomes when the number of traders becomes large, we consider a sequence of markets indexed by the number of informed traders n. The supply function S is held constant as n changes, but the distribution of the private signal varies. At fixed n and state s, the private signal distribution is denoted by F_n^s, with density function f_n^s. Each f_n^s is continuously differentiable and strictly positive on $[x_a, x_b]$. We also define the functions

$$\ell_n := \frac{f_n^H}{f_n^L}, \qquad \Lambda_n := \frac{1 - F_n^H}{1 - F_n^L} \quad \text{and} \quad \lambda_n := \frac{F_n^H}{F_n^L}. \tag{2}$$

The likelihood ratio ℓ_n is taken to be strictly increasing on $[x_a, x_b]$ for each n. This monotone likelihood ratio property (MLRP) means that larger x_i is evidence in favor of $s = H$. The value $\Lambda_n(x)$ 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.

Purchases by informed traders reveal signals in favor of H, further encouraging informed traders to buy. The resulting aggregate demand curve is upward sloping if the signal effect dominates the scarcity effect of price. To implement this scenario, we assume the following property on the sequence of likelihood ratio functions, which guarantees that *aggregate* information for the informed traders increases without bound as $n \to \infty$.

Assumption 1. *There is an* $n_1 \in \mathbb{N}$, *a* $\xi \in (0, 1)$, *and a* $\delta > 0$ *such that*

$$n^\xi \log\left(\frac{\Lambda_n(x)}{\lambda_n(x)}\right) > \delta$$

for all $x \in [x_a, x_b]$ *whenever* $n > n_1$.

As we are concerned with high-frequency fluctuations in volume and price, we work in an environment where the informativeness of the signal is vanishingly small. We formalize this idea by requiring that the signal tends to pure noise.

Assumption 2. *The likelihood ratio* ℓ_n *converges to* 1 *uniformly on* $[x_a, x_b]$ *as* $n \to \infty$.

Assumption 2 holds in short time intervals when the signal received by traders tends to be noisy. Along with Assumption 1, this produces an asymptotic setting 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.

Assumption 3. *There is an* $x_c < x_b$ *such that, for each* $n \in \mathbb{N}$, *the signal satisfies* $\lambda_n''(x)\lambda_n(x) \leq \lambda_n'(x)^2$ *whenever* $x \in [x_c, x_b]$.

Assumption 3 is a regularity condition on behavior of the signal around the boundary of its domain. It is also possible to obtain heavy-tailed outcomes that replicate power laws in finite samples without this assumption, as discussed in the Appendix.[8]

Assumptions 1–3 are satisfied by a variety of signals. Examples include the linear distribution pair

$$f_n^H(x) = \frac{1}{2} + \epsilon_n x \quad \text{and} \quad f^L(x) = \frac{1}{2}, \quad -1 \leq x \leq 1,$$

where $\epsilon_n = n^{-\xi}/3$ and $0 < \xi < 1$, as well as the exponential distribution pair

$$f^H(x) = \frac{\mu e^{-\mu x}}{1 - e^{-\mu}} \quad \text{and} \quad f_n^L(x) = \frac{(\mu + \epsilon_n)e^{-(\mu+\epsilon_n)x}}{1 - e^{-(\mu+\epsilon_n)}}, \quad 0 \leq x \leq 1,$$

where $\epsilon_n = \delta_\epsilon n^{-\xi}$, $\delta_\epsilon > 0$, $\mu > 2$, and $0 < \xi < 1$. The Technical Appendix in the Supplemental Material (available in a supplementary file on the journal website, http://econtheory.org/supp/3523/supplement.pdf) verifies these claims.

2.3 *Strategies and equilibria*

We saw in Section 2.1 that 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

$$\rho_i(p, X_i) \geq \frac{p}{1-p}, \tag{3}$$

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 $d_i(p, X_i) = 1$. Using (1) and $\Pr(H) = \Pr(L) = 1/2$, we obtain

$$\rho_i(p, X_i) = \frac{\Pr(\Omega_{m,i} \mid H)}{\Pr(\Omega_{m,i} \mid L)} \ell_n(X_i). \tag{4}$$

Since $\ell_n(x)$ is continuous and strictly increasing, $\rho_i(p_m, x)$ is continuous and strictly increasing in x for any p_m. Therefore, for each $p_m \in \{p_1, p_2, \ldots, p_n\}$, there exists threshold $\sigma \in [x_a, x_b]$ such that it is optimal for trader i to buy if and only if $X_i \geq \sigma$. The threshold $\sigma = \sigma(m)$ indicates either an indifference level of signal $\rho_i(p_m, \sigma) = p_m/(1 - p_m)$ or a corner solution. With this notation, trader i's demand function follows the rule

$$d_i(p_m, x_i) = \mathbb{1}\{X_i \geq \sigma(m)\}.$$

A trader who buys at price p_m can infer that there are $m - 1$ other buying traders at p_m under the stipulated rule for the auctioneer. Moreover, the threshold function $\sigma(m)$ is

[8] See, in particular, the discussion after the proof of Proposition 3.

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 less than $\sigma(m)$. Such an event occurs with probability

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

Combining this expression with the definitions in (2), the likelihood ratio for p_m to occur can be expressed as

$$\frac{\Pr(\Omega_{m,i} \mid H)}{\Pr(\Omega_{m,i} \mid L)} = \Lambda_n\big(\sigma(m)\big)^{m-1} \lambda_n\big(\sigma(m)\big)^{n-m}.$$

Substituting into (4) and using equality in the decision rule (3), we find that the threshold $\sigma(m)$ at which a trader is indifferent between buying and not buying given p_m is implicitly determined by

$$\frac{p_m}{1-p_m} = \lambda_n(\sigma)^{n-m} \Lambda_n(\sigma)^{m-1} \ell_n(\sigma) \tag{5}$$

if an interior solution σ exists.

Equation (5) 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. Aggregate demand $D(p_m \mid \mathbf{x})$ can be obtained by counting the number of informed traders with $x_i \geq \sigma(m)$.

Proposition 1 (Properties of demand). *Under Assumption 1, there exists an $n_o \in \mathbb{N}$ such that, for any $n \geq n_o$ and any $\mathbf{x} \in [x_a, x_b]^n$, the threshold signal level $\sigma(m)$ is decreasing in m and aggregate demand $D(p_m \mid \mathbf{x})$ is increasing in m.*

Figure 1 depicts aggregate demand $D(p \mid \mathbf{x})$ as a function of m. 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.[9] The increment in price p_{m+1}/p_m along the supply curve is of order $1/n$ because each informed trader demands quantity $1/n$ of the asset.

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

The proof of Proposition 2 involves a straightforward application of Tarski's fixed point theorem. While multiple equilibria may exist for each realization of $\mathbf{x}$, we focus on the case where the auctioneer selects the minimum number of buying traders among

[9] The mechanism in which demand feeds on itself is reminiscent of Bulow and Klemperer's (1994) "rational frenzies."

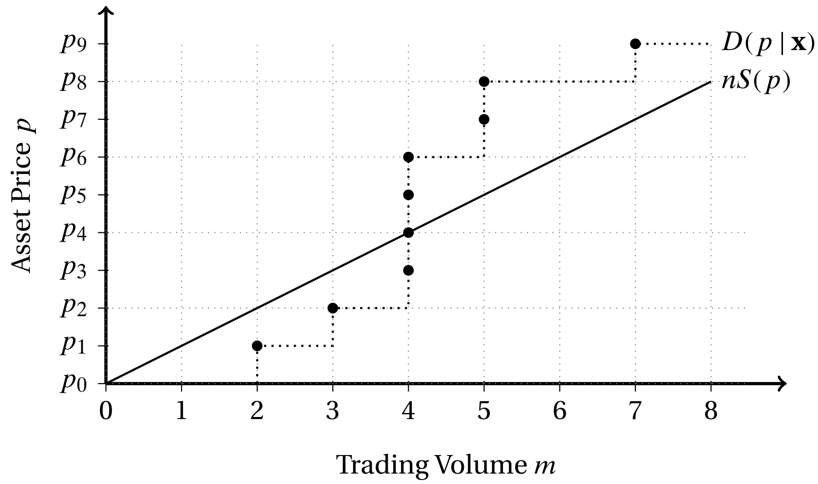

FIGURE 1. An equilibrium outcome for signal profile realization $\mathbf{x}$.

possible equilibria, $m^\dagger$, for each $\mathbf{x}$.[10] This equilibrium selection uniquely maps each realization of $\mathbf{x}$ to $m^\dagger$, rendering $m^\dagger$ a realization of a well defined random variable. This random variable, denoted henceforth by $M_n^\dagger$, represents equilibrium volume, the probability distribution of which is determined by the distribution of $\mathbf{X}$ and the equilibrium selection mapping.

3. Power law results

Next we turn to our main analytical results, including a power law for volume and returns. In addition to these results, which are asymptotic, we provide quantitative analysis investigating the case of finite n.

3.1 *Power law for volume*

The right tail of a random variable Y is said to obey a *power law with exponent* α if $\Pr(Y \geq y) \propto y^{-\alpha}$ for sufficiently large y. Our first step is to show that equilibrium aggregate trading volume $M_n^\dagger$ follows a power law asymptotically in n.

PROPOSITION 3 (Power law for volume). *If Assumptions 1–3 hold, then $M_n^\dagger$ converges in distribution as $n \to \infty$, with*

$$\lim_{n\to\infty} \Pr\bigl(M_n^\dagger = m\bigr) = \frac{e^{-m}(m-1)^{m-1}}{m!}$$

[10]An interpretation of the selection rule is that the auctioneer is mandated by the exchange to minimize the impact of transaction on prices (Hasbrouck 1991). By assuming that the auctioneer selects the minimum number of buying traders, we exclude fluctuations that arise purely from informational coordination such as in sunspot equilibria. Even with this rule of selecting minimum volume, we show that the equilibrium volume and price in the model exhibit large fluctuations.

for each integer $m \geq 1$ *and* $\lim_{n\to\infty} \Pr(M_n^\dagger = 0) = e^{-1}$. *In particular, the right tail of the asymptotic distribution obeys a power law with exponent* $1/2$.

That the second claim in Proposition 3 follows from the first can be shown via Stirling's formula $m! \sim (2\pi m)^{1/2}(m/e)^m$, which yields

$$\Pr\bigl(M_n^\dagger \geq m\bigr) \propto m^{-1/2} \quad \text{for large } m.$$

Note that, under the stated assumptions, the power-law exponent does not depend on the parametric specifications of signals.

The key to the proof of Proposition 3 is that $E_m := D\bigl(p_m, \mathbf{X}\bigr) - nS(p_m)$ is a martingale when considered as a stochastic process indexed by m. The reason this matters is twofold. First, recall that equilibrium trade volume $M_n^\dagger$ is, by definition, the smallest m such that $D\bigl(p_m, \mathbf{X}\bigr) = nS(p_m)$. In other words, $M_n^\dagger$ is the first passage time to zero for the excess demand process $\{E_m\}$. Second, it is well known that, for at least some kinds of martingales, the first passage time to zero follows a power law. We show that this result extends to the martingale $\{E_m\}$ under the stated assumptions.[11]

But why is $\{E_m\}$ a martingale in our model? The underlying reason is that the mean number of traders induced to buy by a trader who buys is 1. This one-for-one response is analogous to actions in Keynes' beauty contest, in which the average action of a single trader responds one-to-one to the average actions of traders. The beauty contest leads to indeterminate equilibria if there is a continuum of traders or if the traders' actions are continuous. This type of local indeterminacy is avoided in our model with finitely many traders and binary actions. However, the indeterminacy described above provides intuition as to why our model can generate equilibrium trading volumes at any order of magnitude, as demonstrated by the power law.[12]

The one-for-one response causes excess demand to obey the martingale property. To see this, suppose that $E_m > 0$. The auctioneer then bids up by 1 and thereby finds one more supplier. By observing this bidding up process, traders act so that aggregate

[11] Feller (1966) treats the cases of Brownian motion and random walks. Our proof extends this power-law finding to a class of Poisson processes. To give some intuition as to when first passage times follow power laws, suppose that $\{Y_m\}$ is a stochastic process indexed by m and starting at $Y_0 = 1$, say. If $\{Y_m\}$ drifts down, then first passage times to zero will typically be small, with large values having very low probability. If $\{Y_m\}$ drifts up, then the first passage times to zero are typically be infinite. If $\{Y_m\}$ is a martingale, however, we observe both small first passage times, which occur by chance, and also very long first passage times, as paths that initially deviated upward eventually return. This is the source of the heavy right tail.

[12] A deeper understanding of the one-for-one response by traders can be gained from examining the optimal threshold condition (5), which reduces to the simple form $(1-\mu)\log \lambda_n(\sigma) + \mu \log \Lambda_n(\sigma) = 0$ when $\mu := m/n$, 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 with μ. If a trader switches 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 nonbuying 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 cancel each other out when the signal is small (i.e., $\log \Lambda_n \approx \log \lambda_n$) and m is finite ($\mu \approx 0$). An increase in m by 1 lowers σ so that D is increased by 1.

demand increases. Since the mean increase is 1, the increased supply is equal to the mean increased demand. Hence $\{E_m\}$ is a martingale.

A power law implies nontrivial aggregate fluctuations even for large n. In general, a power law with exponent α implies that any kth moment for $k \geq \alpha$ is infinite. Thus, with exponent 0.5, $M_n^\dagger$ does not have a finite asymptotic mean or variance as $n \to \infty$. This implies that the variance of the fraction of buying traders, $M_n^\dagger/n$, 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 with probability $\delta n^{-\xi}$ of $d_{n,i} = 1$, the central limit theorem predicts that $M_n^\dagger/n$ would asymptotically follow a normal distribution, where the tail is thin and variance declines as fast as $n^{-1-\xi}$. Thus, the variance of $M_n^\dagger/n$ differs by factor $n^{0.5+\xi}$ between our model and the model with independent choices. This signifies the effect of stochastic clustering that amplifies the small fluctuations in the received signals X_i.[13]

Even if traders' actions are correlated, it requires a particular structure in this correlation for the amplification effect to cause the variance to decline more slowly than n^{-1}, i.e., the speed that 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 martingale property of excess demand $D(p_m, \mathbf{X}) - nS(p_m)$.

Proposition 3 has an implication relevant to the information aggregation literature (see, e.g., Vives 2008). Our model depicts the situation where a large number of informed traders try to learn the true state of the world by gleaning information from other traders' actions under noisy signals. Traders as a group have likelihood $\prod_{i=1}^n \ell_n(x_i)$. Since the values of $|\ell_n(x_i) - 1|$ near bounds $\{x_a, x_b\}$ are bounded from below by $\delta n^{-\xi}$, the collective likelihood diverges at the bounds. Hence, if all traders reveal their private signals, they can learn the true state asymptotically. In our model, traders learn the state only partially due to information asymmetry. Moreover, the extent of partial learning is determined by the number of buying traders, which follows a power law. To see the implication of the power law, we can extend our model to a dynamic setting where traders draw private signals repeatedly and eventually learn the true state.[14] A power law in this setup implies that collective learning does not occur smoothly over time. The noisy signal generates few transactions and is hoarded privately most of the time. However, once in a while, a large cluster of trades occurs and accumulated private information is revealed. Thus, the power law for volume implies that the revelation happens at once in the collective learning of traders in our setup.

3.2 *Comments on the power-law exponent*

At 0.5, the power-law exponent obtained in Proposition 3 is smaller than most empirical estimates for volume, which are summarized by Gabaix et al. (2006) as the half-cubic

[13] On the implications of a tail distribution on aggregate fluctuations, see, for example, Acemoglu et al. (2017) and Nirei (2015).

[14] See the working paper version (Nirei 2011) for the extension.

law (i.e., a power-law exponent of 1.5). However, our model can be modified to generate the half-cubic law, as we now describe.

The exponent 0.5 from Proposition 3 is obtained asymptotically when the number of traders n tends to infinity and the informativeness of the signal vanishes. In an economy with finite n, however, E_m may deviate from a martingale. Let ϕ denote the mean number of traders induced to buy by a trader who buys. The analysis in the Appendix shows that $\Pr(M^\dagger = m \mid D(p_1, \mathbf{X}))$ for $\phi < 1$ is proportional to $e^{-(\phi-1-\log\phi)m} m^{-1.5}$ for large m. With $\omega := \phi - 1 - \log\phi$, we observe that $\Pr(M^\dagger = m \mid D(p_1, \mathbf{X}))$ is approximately a power law with exponent 0.5 for $m < 1/\omega$ and exponentially truncated for $m > 1/\omega$. This distribution includes the pure power-law result as the limiting case $\omega = 0$, which corresponds to $\phi = 1$.

Now consider an extended model in which ω exhibits uniform variation within $(0, \epsilon)$ for some $\epsilon > 0$. Then the asymptotic tail distribution is an integral of the above probability distributions across ω,

$$\int_0^\epsilon \frac{\Pr(M^\dagger = m \mid D(p_1, \mathbf{X}), \omega)}{\epsilon} \, d\omega \propto \int_0^\epsilon \frac{e^{-\omega m} m^{-1.5}}{\epsilon} \, d\omega = \frac{1 - e^{-\epsilon m}}{\epsilon} m^{-2.5} \tag{6}$$

for large m. Note that the power-law exponent is now 1.5 (in the cumulative distribution), exactly matching the half-cubic law.

The property that the power-law exponent increases when the underlying parameter fluctuates around the critical value ($\phi = 1$ in our case) is known as *sweeping of a control parameter toward an instability* (see, e.g., Sornette 2004). Although the uniform variation argument used in (6) to obtain the half-cubic law was ad hoc, we return to this idea in Section 3.4, where we use Monte Carlo methods to show that the same property can explain the empirical power-law exponent for volume when n takes a finite value.

3.3 *Power law for returns*

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 fluctuation of volume is translated into fluctuation of returns. In our model, informed traders' demands are absorbed by uninformed traders' supply. Thus, the supply function of uninformed traders $m^*/n = S(p^*)$ determines the impact of volume m^* on the return $q := \log p^* - \log p_0$.[15]

The relation between an exogenous shift in trading volume and the resulting shift in asset price, i.e., S^{-1}, is called the *price impact function*. Empirical studies suggest that the price impact function is concave (see, e.g., Hasbrouck 1991 and Lillo et al. 2003).

[15]We define q as the logarithmic return of the asset under the interpretation that p_0 is the price that prevailed in the previous period -1. In this setup, p_0 reflects the prior belief $\Pr(H)$ under an extended model where informed traders can both buy and sell. 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 the informed traders are matched with uninformed traders only when either of the informed traders buying at p_1 or selling at p_{-1} is zero and the other is strictly positive. In this way, informed traders always transact with uninformed traders, as in the original model.

Keim and Madhavan (1996) proposed a model for concave price impact functions in which market makers incur search costs to find counterparty traders. In the model, an increase in volume does not lead to a linear increase in price impact, because the market maker reaches out to more traders who absorb the demand. In line with this research, we specify the price impact function as $q = \beta(m/n)^\gamma$ for $m = 1, 2, \dots, n$, where β is a positive constant and $0 < \gamma < 1$.

The following proposition establishes that our model generates a power law for the returns distribution when the price impact is specified as above.

Proposition 4 (Power law for returns). *If volume m/n follows a power law of the form $\Pr(m/n) \propto (m/n)^{-\alpha-1}$ and the supply function S satisfies*

$$\log\bigl(S^{-1}(m/n)\bigr) - \log\bigl(S^{-1}(0)\bigr) = \beta(m/n)^\gamma \tag{7}$$

for $\beta, \gamma > 0$, then returns q follow a power law with exponent α/γ.

Proof. By applying the change of variable for m/n and using the specified monotone supply function, we obtain

$$\Pr(q) = \Pr(m/n)\bigl|d(m/n)/dq\bigr| \propto \bigl((q/\beta)^{1/\gamma}\bigr)^{-\alpha-1}(1/\gamma)q^{1/\gamma-1} \propto q^{-\alpha/\gamma-1}.$$

Hence, q follows a power law with exponent α/γ. □

There is a growing consensus among empiricists that stock returns generally obey a cubic law, in which the return distribution follows a power law with exponent 3 (see, e.g., Gabaix et al. 2006, Lux and Alfarano 2016, and Gu and Ibragimov 2018). The cubic law corresponds to $\alpha/\gamma = 3$ in the above equation. The analysis in the previous section established that $\alpha = 0.5$ asymptotically. Hence, the cubic law holds in our asymptotic case if $\gamma = 1/6$. This value is consistent with empirical estimates for γ, which range between 0.1 and 0.5 (see, e.g., Lillo et al. 2003).

3.4 *Quantitative analysis with finite agents*

In this section, we conduct numerical analysis of the model with a finite number of informed traders n. One aim of this exercise is to confirm that, even with finite n, the number of buying traders $M_n^\dagger$ exhibits a power law, complementing the asymptotic result from Proposition 3. A second aim is to show that the model is in fact capable of generating a more empirically relevant power law exponent when n is taken to be finite. Finally, we show that the fluctuation of equilibrium asset returns $q = \log p_n^* - \log p_0$ exhibits a power law that matches the returns distribution observed in the 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 means $\mu_H = 1$ 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

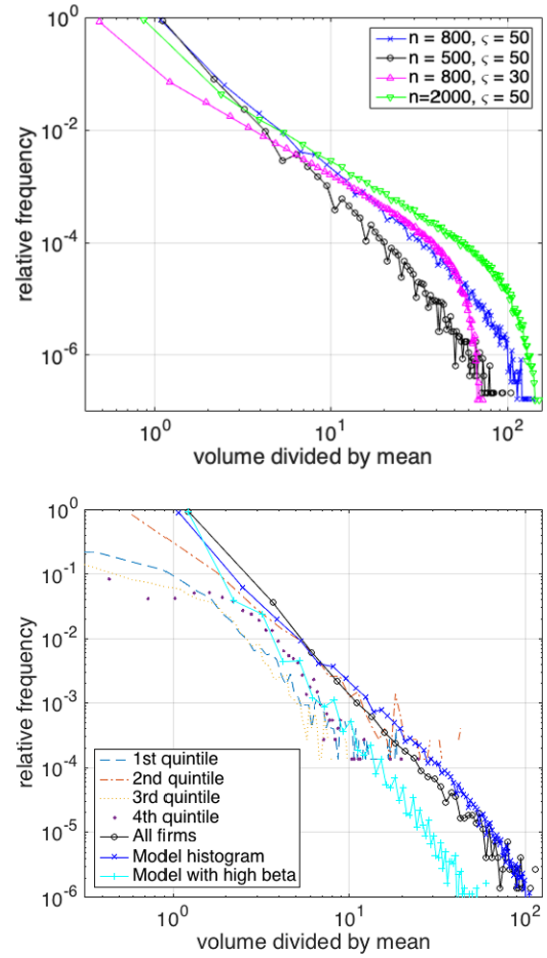

FIGURE 2. Histograms of volume normalized by its time-series average. Top: Histograms of equilibrium volume $M_n^\dagger$ for various parameter values, where n is the number of traders and ς is the standard deviation of private information. Bottom: Daily volume histograms for individual stocks and a pooled sample, plus simulated histograms. Data for individual firms cover 1988–2018. Data for the pooled sample are for 2016. Individual firms are selected at the quintiles of market capitalization size in the TSE. The circle-line shows the histogram of pooled data for all listed firms. The cross-line shows a simulated histogram with $n = 800$ and $\varsigma = 50$. The plus-line shows the case with a higher β.

of uninformed traders is specified as in (7), and its parameters are set at our estimates $\gamma = 0.4642$ and $\beta = 0.768$, as explained below. 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 $(x_i)_{i=1}^n$ 10^6 times and computing $m_n^\dagger$ and p_n^* for each draw.

The top panel of Figure 2 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 Proposition 3 predicted the exponent of the power law α to be 0.5. The top panel of Figure 2 confirms this pattern for finite n, when $(n, \varsigma) = (800, 30)$ 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 triangle-line). This deviation of the exponent from the asymptotic case $\alpha = 0.5$ can result from finite n, as we discussed previously. In this way, our model has flexibility in fitting various empirically observed exponents for trading volume.

Using this flexibility, we fit our model to 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 the Nikkei Financial Quest data set, which includes daily volume and prices for the firms listed in the first section of the Tokyo Stock Exchange (TSE) from March 1988 to March 2018.

The bottom panel of Figure 2 shows histograms of daily trading volumes for four single stocks. The volume is divided by the daily 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 the 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) is too long to assure invariance of the daily volume distribution. To deal with these limitations, we prepare a pooled data set of a large number of stocks for a shorter sample period. We collect daily trading volume for all (2250) listed firms for the year 2016, and divide volume by average volume for each firm during that 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 show 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 $(n, \varsigma) = (800, 50)$, shown as the cross-line. As can be seen, the simulated histogram effectively matches the empirical histogram.

The power-law exponent can be estimated using the Hill estimator. Since the power law applies only to the tail distribution, we augment the Hill estimator with an estimated lower threshold for the tail region, following the methodology proposed by Clauset et al. (2009). The estimated power-law exponents for volume for the first to fourth quintile stocks and for the pooled data are 1.73 (0.04), 1.31 (0.04), 2.50 (0.08), 3.68 (0.25), and 1.89 (0.02), respectively (standard errors in parentheses). The Hill estimate of the power-law exponent for simulated volume is within the above range at 2.03.

The top panel of Figure 3 shows the histograms of daily returns for the same samples. We define the daily return as the logarithmic difference from the opening to closing price. The open–close difference is used rather than a business day return so that the

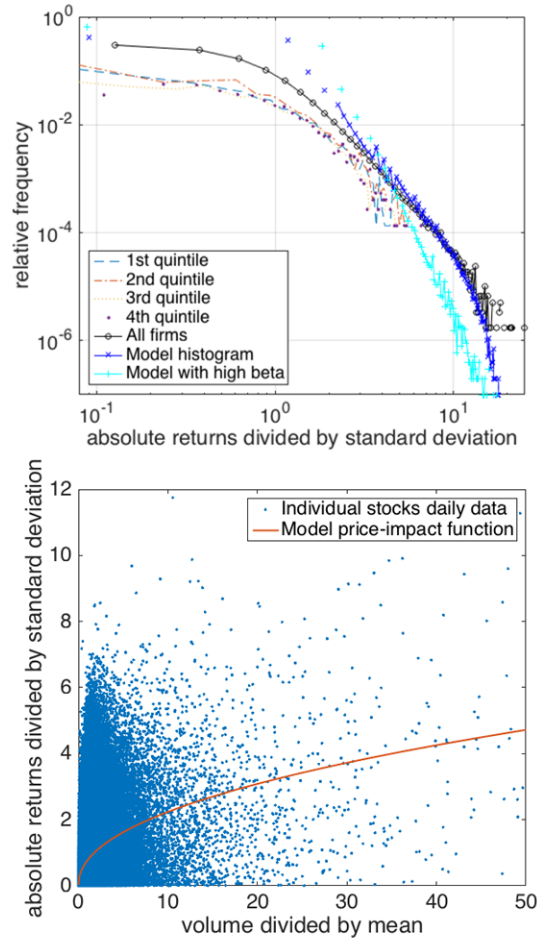

FIGURE 3. Top: Histograms of the absolute values of daily returns. The horizontal axis shows the daily return, which is the difference in logarithm of opening and closing prices for each business day. The samples are the same as in Figure 2. Bottom: A scatter plot of daily volume and absolute returns for all listed firms in 2016. The line shows the price-impact function fitted by nonlinear least squares.

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 2016). The model-generated histogram also shows a fat tail, which is slightly thinner than the data, but clearly exhibits a power law. The Hill estimates of the power-law exponent for returns for the quintile stocks and the pooled data are 4.52 (0.72), 2.49 (0.11), 3.26 (0.16), 2.83 (0.13), and 3.62 (0.03), whereas that for the simulated returns is 4.38.

The bottom panel of Figure 3 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. The parameter values ($\gamma = 0.4642$, $\beta = 0.768$) are estimated by fitting (7) to the pooled sample by nonlinear least squares.[16] In sum, Figures 2 and 3 indicate that our model is capable of generating power laws for volume and absolute returns while using a price-impact function consistently estimated via the high-frequency sample observed in the TSE.

3.5 *An extension*

In this section, we consider two extensions to the baseline model.

3.5.1 *Learning by uninformed traders* Our benchmark model assumes uninformed traders are unaware of the fact that price movements are driven by informed traders. This assumption can be relaxed. Suppose, for instance, that the uninformed traders know that some price movements contain information on the value of assets but cannot distinguish such movements from purely random ones. Thus, uninformed traders perceive that the price movements reflect revealed information of informed traders with some probability π. Let p^l_m denote the price the uninformed traders accept for supplying m in this environment, while they are willing to accept the price p_m (defined in our benchmark model) so as to fill their liquidity need for supplying m even if there is no information contained in the transaction. The price then satisfies $p^l_m - p_0 = (p^l_m - p_0 + p_m - p_0)\pi + (p_m - p_0)(1 - \pi)$, implying $(p^l_m - p_0) = (p_m - p_0)/(1 - \pi)$.

This result implies that the supply function $S(p)$ becomes steeper by $1/(1 - \pi)$ when uninformed traders can learn with probability π. Since supply elasticity does not affect our asymptotic results, the power laws obtained from Section 3.1 continue to hold in this environment.

Numerical results under a finite number of traders may be affected by the possibility of learning. In fact, Figures 2 and 3 show this to be the case. The steeper supply function corresponds to the higher β in the price-impact function (7). We specify that $\pi = 1/3$, which means that β is increased by 50%. The power laws for volume and returns under the high β are shown in the bottom panel of Figure 2 and the top panel of Figure 3. We observe that exponents for both cases become greater than the benchmark case.

3.5.2 *Variable transaction size* In actual markets, trading is not a binary choice. It is possible to extend our model to more general settings for trading size. For example, the fixed trading size can be heterogeneous across informed traders. Suppose that the distribution of heterogeneous trading size has a tail thinner than the power law with exponent 0.5. Also suppose that informed traders can observe the number of buying informed traders. In this case, the information inferred by informed traders through price will be the same as in the benchmark model. Since the size heterogeneity has a thinner tail than the number of buying traders, the same power law of aggregate volume as in the benchmark model continues to hold.[17]

[16] The plotted sample is truncated at the volume divided by mean being 50 to enhance visibility, while all the data are used for the parameter estimation.

[17] For a mathematical reference, see Jessen and Mikosch (2006).

We can also consider the case where traders can choose trading size depending on the signal they receive. Suppose that informed traders are risk averse. Then the trader has an incentive to buy a large amount when receiving a favorable signal. Thus, the optimal trading strategy is to buy when the belief surpasses a threshold and the purchasing amount is increasing in the belief. If the purchasing amount is chosen from a continuous set, the exact private signal $\ell_n(x)$ is revealed by the amount if a trader buys at all, in contrast to the benchmark model where only $\Lambda_n(x)/\lambda_n(x)$ is revealed by a buying action. Nonetheless, these two revealed bits of informations are asymptotically the same, and our power-law result still holds. However, we need to note that the revealed information is shared by all traders in a rational expectations equilibrium. This implies that the beliefs of all the buying traders will be equalized, while the beliefs of nonbuying traders remain heterogeneous. Thus, equilibrium trading size is constant across buying traders. To extend the rational expectations model to heterogeneous trading size correlated to signals, one would need to incorporate some noise, which prevents the signal from being exactly revealed. This would be the case if, say, the choice set is discrete with more than binary choices. While analytical characterization in this setup is complex, numerical investigation should be feasible.

4. Conclusion

This study analyzes aggregate fluctuations of trading volume and prices that arise from asymmetric information among traders in financial markets. In an asset market model in which each trader infers the private information of other traders only by observing their actions, we find that the number of traders taking the same action in equilibrium exhibits large volatility with a particular statistical regularity: a power-law distribution. We also show that the model is capable of generating a power-law distribution of asset returns. The simulated distributions of equilibrium returns and volume are demonstrated to match the distributions of observed stock returns and volume. In this way, we explicitly link the large and growing literature on asymmetric information and clustering with a well documented statistical regularity for volume and return distributions.

This study suggests several directions for future work. One would be to develop a dynamic model that accounts for time-series properties as pursued by, for example, Alfarano et al. (2008). Another direction would be to extend the model by incorporating more realistic market structure. Some extensions, such as learning by uninformed traders or variable transaction size, would seem to be easily incorporated. Other extensions, such as the case where the signal and trading size are correlated, where both public and private signals exist, and where informed traders can take both buying and selling sides, appear to be more involved. While some steps have been taken in these directions (see, e.g., Kamada and Miura 2014), we leave such explorations for future work.

Appendix

Properties of λ_n and Λ_n

We note for future reference that the likelihood ratios satisfy $\Lambda_n(x_a) = \lambda_n(x_b) = 1$, $\lim_{x\to x_a} \lambda_n(x) = \ell_n(x_a)$, and $\lim_{x\to x_b} \Lambda_n(x) = \ell_n(x_b)$ (obtained using l'Hôpital's rule).

Also, the MLRP implies that $0 < \lambda_n(x) < \ell_n(x) < \Lambda_n(x)$ for any x interior to $[x_a, x_b]$, as in Smith and Sørensen (2000), as well as strictly increasing likelihood ratios as shown below.

Taking derivatives of λ_n and Λ_n, we have

$$\frac{d\lambda_n(x)}{dx} = \frac{f_n^H(x)}{F_n^L(x)} - \frac{F_n^H(x) f_n^L(x)}{\big(F_n^L(x)\big)^2} = \frac{f_n^L(x)}{F_n^L(x)}\big(\ell_n(x) - \lambda_n(x)\big) \tag{8}$$

$$\frac{d\Lambda_n(x)}{dx} = -\frac{f_n^H(x)}{1 - F_n^L(x)} + \frac{\big(1 - F_n^H(x)\big) f_n^L(x)}{\big(1 - F_n^L(x)\big)^2} = \frac{f_n^L(x)}{1 - F_n^L(x)}\big(\Lambda_n(x) - \ell_n(x)\big). \tag{9}$$

Thus, inequality $\lambda_n(x) < \ell_n(x) < \Lambda_n(x)$ implies that $\lambda_n'(x) > 0$ for $x \in (x_a, x_b]$ and $\Lambda_n'(x) > 0$ for $x \in [x_a, x_b)$. At $x = x_a$, we obtain $\lambda_n'(x_a) = \ell_n'(x_a)/2 > 0$ by applying l'Hôpital's rule for (8) and rearranging terms. Similarly, we obtain $\Lambda_n'(x_b) = \ell_n'(x_b)/2 > 0$ by evaluating (9) at $x = x_b$. Hence, we obtain $\lambda_n'(x) > 0$ and $\Lambda_n'(x) > 0$ for any $x \in [x_a, x_b]$.

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, (5) implicitly determines $\sigma(m)$ for real 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

$$0 = \Phi(\sigma, t) := (n - t)\log\lambda_n(\sigma) + (t - 1)\log\Lambda_n(\sigma) + \log\ell_n(\sigma) - \log\frac{p_t}{1 - p_t}, \tag{10}$$

which is a logarithmic transformation of (5) 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 $t = n$. The function $\Phi(\sigma, t)$ is increasing in σ, since λ_n, Λ_n, and ℓ_n are increasing functions. It achieves minimum at $\sigma = x_a$, and the minimum value is $\Phi(x_a, t) = (n - t + 1)\log\lambda_n(x_a) - \log(p_t/(1 - p_t))$, where we used $\Lambda_n(x_a) = 1$ and $\lambda_n(x_a) = \ell_n(x_a)$. Noting that $\lambda_n(x_a) < 1$ and $\log(p_t/(1 - p_t)) > 0$, we obtain $\Phi(x_a, t) < 0$ for any $t \in [1, n]$.

The function $\Phi(\sigma, t)$ achieves maximum at x_b, and the maximum value is obtained as $\Phi(x_b, t) = t\log\Lambda_n(x_b) - \log(p_t/(1 - p_t))$, using $\lambda_n(x_b) = 1$ and $\Lambda_n(x_b) = \ell_n(x_b)$. When $t = 1$, the maximum is $\Phi(x_b, 1) = \log\Lambda_n(x_b) - \log(p_1/(1 - p_1))$. Since $\lambda_n(x_b) = 1$, Assumption 1 implies $\log\Lambda_n(x_b) > \delta/n^\xi$ for sufficiently large n. 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

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

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 $\Phi(x_b, 1) > \delta/n^{\xi} - c_1/n$, which is strictly positive for sufficiently large n since $\xi < 1$.

When $t = n$, the maximum of $\Phi(\sigma, n)$ is $n \log \Lambda_n(x_b) - \log(p_n/(1-p_n))$. 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 $n \log \Lambda_n(x_b) > \delta n^{1-\xi}$. Thus, $\Phi(x_b, n) > 0$ for sufficiently large n. Since $\Phi(x_a, t) < 0$ and $\Phi(x_b, t) > 0$ for $t = 1$ and $t = n$, and since Φ is continuous in σ, an interior solution σ exists for both $t \in \{1, n\}$ 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

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

This determines the derivative of σ with respect to t as

$$\frac{d\sigma}{dt} = \frac{-\log\bigl(\Lambda_n(x)/\lambda_n(x)\bigr) + \bigl\{p_t(1-p_t)S'(p_t)n\bigr\}^{-1}}{(n-t)\lambda_n'(x)/\lambda_n(x) + (t-1)\Lambda_n'(x)/\Lambda_n(x) + \ell_n'(x)/\ell_n(x)} \bigg|_{x=\sigma(t)}. \tag{11}$$

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, there exists some n_o such that for any $n > n_o$, inequality $d\sigma/dt \le 0$ holds.

Since an interior solution σ for $\Phi(\sigma, t)$ exists for both $t \in \{1, n\}$ and since an interior solution σ is decreasing in t, an interior solution of (5) exists for any $m \in \{1, 2, \ldots, n\}$.

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

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 $\Gamma_{\mathbf{x}} : \{0, 1, \ldots, n\} \mapsto \{0, 1, \ldots, n\}$ for each realization of $\mathbf{x}$. It coincides with D for $m > 0$, i.e., $\Gamma_{\mathbf{x}}(m) := D(p_m, \mathbf{x})$ for $m \in \{1, 2, \ldots, n\}$. For $m = 0$, we let $\Gamma_{\mathbf{x}}(0) = D(p_1, \mathbf{x})$. Then $\Gamma_{\mathbf{x}}$ is an increasing mapping of $\{0, 1, \ldots, n\}$ onto itself for $n > n_o$. Moreover, $\{0, 1, \ldots, n\}$ is a finite totally ordered set and, hence, a complete lattice. Therefore, by Tarski's fixed point theorem, there exists a nonempty closed set of fixed points of $\Gamma_{\mathbf{x}}$.

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

Preparation for the Proof of Proposition 3

So as 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 $\sum_{i=1}^n \mathbb{1}\{X_i \geq x\}$. As x travels from maximum x_b to minimum x_a, this process generates an increasing number of buying traders. Now we replace x with the threshold level of signal, $\sigma(m)$. Then $\sum_{i=1}^n \mathbb{1}\{X_i \geq \sigma(m)\}$ indicates the number of traders with private information greater than threshold $\sigma(m)$. For each realization of $\mathbf{x}$, $\sum_{i=1}^n \mathbb{1}\{x_i \geq \sigma(m)\}$ is increasing in m because $\sigma(m)$ 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. Equation (5) 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 $m \in \{1, 2, \dots, n\}$. Using $t = \sigma^{-1}(x)$ and $f_n^\theta(x)$, where θ denotes the true state, the probability density function defined over t is obtained as $f_n^\theta(\sigma(t))|\sigma'(t)|$ for sufficiently large $n > n_o$, because $\sigma(t)$ is monotone in t for such n. Then we construct a counting process $\Gamma(t) := \sum_{i=1}^n \mathbb{1}\{\sigma^{-1}(X_i) \geq t\}$.[18] Since our model is static, the "time" t is fictitiously introduced here so as to define a stochastic process $\Gamma(t)$. The fictitious notion of time turns out to be useful, as we employ analysis of first passage times below.

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 nonbuying trader switching to buying between t and $t + dt$ for small dt is equal to $\pi_n(t)\,dt := f_n^\theta(\sigma(t))|\sigma'(t)|\,dt/F_n^\theta(\sigma(t))$. 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 $1 - F_n^\theta(\sigma(1))$. 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 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 $\phi_n(t)$.

Under Assumption 2, it turns out that the intensity function ϕ_n converges to 1 as $n \to \infty$:

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

[18] The term $\Gamma(t)$ differs from $\Gamma_{\mathbf{x}}(m)$ defined in the previous section (Proof of Proposition 2) in two regards. First, $\Gamma(t)$ is not conditional on $\mathbf{x}$. 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 $\Gamma_{\mathbf{x}}(m)$ share the property that they count the number of traders with private signal greater than some threshold.

The proof is shown in the next section. The intensity $\phi_n = 1$ implies that the mean number of informed traders who switch to buying from nonbuying 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 $G(0)$.

The process $G(t)$ 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

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

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

We consider the case where ℓ_n converges uniformly to 1 as $n \to \infty$ (Assumption 2). 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. *If Assumptions 1 and 2 hold, $\tau_{\varphi_n(\cdot)}$ converges in distribution to τ_1 as $n \to \infty$.*

The proof is shown in the subsequent section. We have shown that $M_n^\dagger$ conditional on $M_n^\dagger > 1$ has the same distribution as the first passage time: $\inf\{t > 1 \mid \Gamma(t) = t\}$. The variable $M_n^\dagger$ corresponds to $\tau_{\varphi_n} + 1$, reflecting that 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 minimum equilibrium number of buying traders, $M_n^\dagger$, conditional on $M_n^\dagger > 1$ has the same asymptotic distribution as $\tau_1 + 1$.

We show that τ_1 follows the same distribution as the sum of a branching process $\sum_{u=0}^{U} b_u$, where the initial value for the branching process is $G(0) = \Gamma(1) - 1$. To do so, we consider a general Poisson process $N(t)$ with intensity parameter $\phi > 0$, where $N(0) = b_0$ is a positive integer. The first passage time of $N(t)$ reaching $t - b_0$ must be greater than or equal to b_0. Now we introduce a process b_u for $u = 0, 1, \dots$. During the time interval b_0, the increment $N(b_0) - N(0)$, denoted as b_1, follows a Poisson distribution with mean

ϕb_0. Since a Poisson random variable is infinitely divisible, a Poisson random variable with mean ϕb_0 is equivalent to b_0-fold convolution of the Poisson with mean ϕ. Thus, we can regard b_1 as the number of traders induced to buy by b_0 traders, where each trader in b_0 brings about a number of induced traders following the Poisson with mean ϕ. If $b_1 = 0$, the process b_u stops, and the first passage time is b_0. 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 ϕb_1, which is equivalent to b_1-fold convolution of the Poisson with mean ϕ and is regarded as the number of traders induced by b_1 traders (note that the increment b_1 of a Poisson process is always an integer). This process b_u continues for $u = 1, 2, \dots, U$, where U denotes the stopping time at which b_U is equal to 0 for the first time. Thus, the first passage time is equal to $\sum_{u=0}^{U} b_u$, i.e., the total number of population generated in the so-called Poisson branching process b_u in which each trader bears a number of induced traders according to the Poisson distribution with mean ϕ.

It is known that the sum of the Poisson branching process, cumulated over time until the process stops, follows a Borel–Tanner distribution (Kingman 1993; see also Nirei 2006). When the Poisson mean of the branching process b_u is $\phi > 0$, the Borel–Tanner distribution is written as

$$\Pr\left(\sum_{u=0}^{U} b_u = m \mid b_0\right) = \frac{b_0}{m} \frac{e^{-\phi m}(\phi m)^{m-b_0}}{(m-b_0)!} \tag{12}$$

for $m = b_0, b_0+1, \dots$. Applying Stirling's formula to the factorial term, we obtain the tail characterization

$$\Pr\left(\sum_{u=0}^{U} b_u = m \mid b_0\right) \propto e^{-(\phi-1-\log\phi)m} m^{-1.5} \quad \text{for sufficiently large } m. \tag{13}$$

Using $\phi = 1$ in our asymptotic characterization of $M_n^\dagger$, we obtain the distribution of $M_n^\dagger$ conditional on $\Gamma(1) = c > 1$ for sufficiently large n as follows.

Lemma 3. *If Assumptions 1 and 2 hold, then, as $n \to \infty$,*

$$\Pr\bigl(M_n^\dagger = m \mid D\bigl(p_1, \mathbf{X}\bigr) = c\bigr) \to \frac{(c-1)(m-1)^{m-c-1}e^{-m+1}}{(m-c)!} \tag{14}$$

for $m = c, c+1, \dots$. In particular, the right tail of the asymptotic distribution obeys a power law with exponent $1/2$:

$$\frac{(c-1)(m-1)^{m-c-1}e^{-m+1}}{(m-c)!} \sim \frac{c-1}{\sqrt{2\pi}} m^{-1.5} \quad \textit{for large } m.$$

Proof. As shown above, τ_1 follows (12) with $\phi = 1$. We change variables in (12) using $\tau_1 = M_n^\dagger - 1$ and $b_0 = G(0) = \Gamma(1) - 1$. With $m' := m+1$, (12) is rewritten as

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

Using $\phi = 1$, we obtain (14). Applying Stirling's formula to the factorial term, we obtain,

$$\begin{aligned}\frac{(c-1)(m-1)^{m-c-1}e^{-m+1}}{\sqrt{2\pi(m-c)}\big((m-c)/e\big)^{m-c}} &= \frac{c-1}{\sqrt{2\pi}}e^{1-c}\left(\frac{m-1}{m}\right)^{m-c-1}\left(\frac{m}{m-c}\right)^{m-c+0.5} m^{-1.5} \\ &\sim \frac{c-1}{\sqrt{2\pi}}m^{-1.5}\end{aligned}$$

for (14) for large m. □

Proof of Lemma 1

We transform ϕ_n using a change of variable for the density of $t = \sigma^{-1}(x)$:

$$\phi_n = \pi_n(t)\big(n - \Gamma(t)\big) = \left(1 - \frac{\Gamma(t)}{n}\right)n\big|\sigma'(t)\big|\frac{f_n^\theta(x)}{F_n^\theta(x)}\bigg|_{x=\sigma(t)}. \tag{15}$$

Using (8) and (11) for $\sigma'(t)$, we obtain

$$n\big|\sigma'(t)\big|\frac{f_n^H(x)}{F_n^H(x)} = \left|\frac{\log\big(\Lambda_n(x)/\lambda_n(x)\big) - \big\{p_t(1-p_t)S'(p_t)n\big\}^{-1}}{\left(1 - \frac{t}{n}\right)\left(1 - \frac{\lambda_n(x)}{\ell_n(x)}\right) + \frac{1}{n}\frac{F_n^H(x)}{f_n^H(x)}\left(\frac{(t-1)\Lambda_n'(x)}{\Lambda_n(x)} + \frac{\ell_n'(x)}{\ell_n(x)}\right)}\right| \tag{16}$$

$$n\big|\sigma'(t)\big|\frac{f_n^L(x)}{F_n^L(x)} = \left|\frac{\log\big(\Lambda_n(x)/\lambda_n(x)\big) - \big\{p_t(1-p_t)S'(p_t)n\big\}^{-1}}{\left(1 - \frac{t}{n}\right)\left(\frac{\ell_n(x)}{\lambda_n(x)} - 1\right) + \frac{1}{n}\frac{F_n^L(x)}{f_n^L(x)}\left(\frac{(t-1)\Lambda_n'(x)}{\Lambda_n(x)} + \frac{\ell_n'(x)}{\ell_n(x)}\right)}\right|. \tag{17}$$

We examine the right-hand side of (16) and (17) 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 seen below. First, Λ_n, ℓ_n, and f_n^θ for $\theta \in \{H, L\}$ are strictly positive. Second, $F_n^\theta \le 1$, and ℓ_n' is bounded because f_n^s is assumed to have a bounded derivative. Finally, $\Lambda_n'(x)$ is bounded for $x \in [x_a, x_b]$, as shown in (9).

We next examine $\Lambda_n(x)/\lambda_n(x)$ and $\lambda_n(x)/\ell_n(x)$ in the right-hand side of (16). To do so, we show that $\sigma(t) \to x_b$ as $n \to \infty$ for finite t. We note that

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

Since $\log(1+y) \le y$ and $1 + \log y \le y$ for any $y \ge 0$, we have, for $\sigma < x_b$,

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

Hence, we obtain

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

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

$$-\log \lambda_n(\sigma) \geq \big(1 - F_n^L(\sigma)\big)\delta n^{-\xi}. \tag{18}$$

Now (10) can be modified to

$$n \log \lambda_n(\sigma) = \log \frac{p_t}{1-p_t} + t \log \frac{\lambda_n(\sigma)}{\Lambda_n(\sigma)} + \log \frac{\Lambda_n(\sigma)}{\ell_n(\sigma)}. \tag{19}$$

The right-hand side of (19) is finite for any finite t. The left-hand side of (19) 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 (18) and $\xi < 1$. Hence, (19) holds only if $F_n^L(\sigma)$ tends to 1, which is equivalent to $\sigma(t) \to x_b$ 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 $\Lambda(x_b) = \ell(x_b)$.

Thus, using $z_n := \log(\Lambda_n(\sigma(t))/\lambda_n(\sigma(t)))$, the limit of the right-hand side of (16) as $n \to \infty$ is expressed as

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

where we used $\lim_{n\to\infty} z_n = 0$ and a Taylor expansion of $e^{z_n} - 1$ around $z_n = 0$, as well as notation $y_n = O(x_n)$ if there exist c_2 and n_2 such that $|y_n| \leq c_2 x_n$ for any $n \geq n_2$. Dividing both the denominator and the numerator by z_n, and applying $nz_n > \delta n^{1-\xi}$ with $\xi < 1$ (Assumption 1), we obtain

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

Similarly, the limit of the right-hand side of (17) as $n \to \infty$ is

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

Substituting this into (15), we obtain $\operatorname{plim}_{n\to\infty} \phi_n(t) = \operatorname{plim}_{n\to\infty} 1 - \Gamma(t)/n$. This implies that $\operatorname{plim}_{n\to\infty} \phi_n(t)$ is bounded. Hence the asymptotic variance of $\Gamma(t+dt) - \Gamma(t)$ is also bounded. Thus, as $n \to \infty$, $\Gamma(t)/n$ converges to zero in L^2-norm and, therefore, in probability. In this way, we obtain $\operatorname{plim}_{n\to\infty} \phi_n(t) = 1$ 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 $n \to \infty$. In other words, we show that, for any $\eta > 0$,

$$\lim_{n\to\infty} \mathbb{E}\big[\exp(-\eta \tau_{\varphi_n(\cdot)})\big] = \mathbb{E}\big[\exp(-\eta \tau_1)\big]. \tag{20}$$

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

In (15), we observe that $\phi_n(t) = \varphi_n(t-1)$ is a product of (16) or (17) 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 $n \to \infty$.[19] Since $e^{-\eta\tau}$ is bounded, such events have vanishingly small contribution to the expectation in the left-hand side of (20). Combining this with the fact that (16) and (17) are 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 $\mathbf{x}$. Thus, to establish (20), it is sufficient to show that $\mathbb{E}[\exp(-\eta\tau_\chi)]$ is continuous with respect to $\chi > 0$. We also note that $\tau_\chi = \inf\{t \ge 0 \mid N(\chi t) \le t - c_o\}$ is equal to

$$\inf\{t \ge 0 \mid t - N(\chi t) \ge c_o\} = \frac{1}{\chi} \inf\left\{t \ge 0 \,\middle|\, \frac{t}{\chi} - N(t) \ge c_o\right\}.$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Proof of Proposition 3

Let $\theta \in \{H, L\}$ denote the true state and let $\psi_\theta := \lim_{n\to\infty} n(1-F_n^\theta(\sigma(1)))$ denotes the asymptotic mean number of traders with $x_i > \sigma(1)$. Under finite ψ_θ, $\Gamma(1)$ asymptotically follows a Poisson distribution with mean ψ_θ. Hence, for $m = \{0, 1\}$, $\Pr(M_n^\dagger = m)$ asymptotically follows $\Pr(\Gamma(1) = m) = \psi_\theta^m e^{-\psi_\theta}/m!$.

For $m > 1$, the unconditional distribution of $M_n^\dagger$ is derived by combining the distribution (14) and the Poisson distribution with mean ψ_θ for $\Gamma(1)$ as

$$\begin{aligned}
&\sum_{c=2}^{m} \Pr\big(M^\dagger = m \mid \Gamma(1) = c\big)\Pr\big(\Gamma(1) = c\big) \\
&\quad= \sum_{c=2}^{m} \frac{(c-1)(m-1)^{m-c-1}e^{-m+1}}{(m-c)!}\frac{\psi_\theta^c e^{-\psi_\theta}}{c!} \\
&\quad= \sum_{c=1}^{m} \frac{(c-1)(m-1)^{m-c-1}e^{-m+1}}{(m-c)!}\frac{\psi_\theta^c e^{-\psi_\theta}}{c!} \\
&\quad= \frac{e^{-\psi_\theta-m+1}(m-1)^{m-1}}{m!}\left[\sum_{c=1}^{m}\frac{\big(\psi_\theta/(m-1)\big)^c m!}{(m-c)!(c-1)!} - \sum_{c=1}^{m}\frac{\big(\psi_\theta/(m-1)\big)^c m!}{(m-c)!c!}\right].
\end{aligned} \tag{21}$$

Using the binomial theorem, we obtain

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

and

$$\sum_{c=1}^{m}\frac{\big(\psi_\theta/(m-1)\big)^c m!}{(m-c)!c!} = \sum_{c=0}^{m}\frac{\big(\psi_\theta/(m-1)\big)^c m!}{(m-c)!c!} - 1 = \left(1+\frac{\psi_\theta}{m-1}\right)^m - 1.$$

Substituting back to (21) yields

$$\begin{aligned}&\frac{e^{-\psi_\theta-m+1}(m-1)^{m-1}}{m!}\left[\frac{\psi_\theta m}{m-1}\left(1+\frac{\psi_\theta}{m-1}\right)^{m-1}-\left(1+\frac{\psi_\theta}{m-1}\right)^{m}+1\right]\\ &\quad=\frac{e^{-\psi_\theta-m+1}(m-1)^{m-1}}{m!}\left[(\psi_\theta-1)\left(1+\frac{\psi_\theta}{m-1}\right)^{m-1}+1\right].\end{aligned}\tag{22}$$

Applying Stirling's formula for $m!$ and using $(1+\psi_\theta/(m-1))^{m-1}\to e^{\psi_\theta}$ as $m\to\infty$, we obtain the power-law result in the cumulative distribution:

$$\Pr(M_n^\dagger\geq m)\approx\sqrt{\frac{2}{\pi}}(e^{-\psi_\theta}+\psi_\theta-1)m^{-1/2}\quad\text{for large } m.\tag{23}$$

Finally, we show that $\psi_\theta=1$ under Assumption 3 for any θ. By substituting $\psi_\theta=1$ into (22) and (23), and substituting $\psi_\theta^m e^{-\psi_\theta}/m!$ for $m=\{0,1\}$, we obtain Proposition 3. Recall $\psi_\theta=\lim_{n\to\infty}n(1-F_n^\theta(\sigma_1))$, where $\sigma_1:=\sigma(1)$ is determined by (5) as

$$\frac{1}{1/p_1-1}=\lambda_n^{n-1}(\sigma_1)\ell_n(\sigma_1)\tag{24}$$

and p_1 is determined by $S(p_1)=1/n$. Since S is continuously differentiable, strictly increasing, and satisfies $S(0.5)=0$, we have $p_1=0.5+O(1/n)$. Therefore, $\log(1/p_1-1)$ is a negative term of order $1/n$. Moreover, we have shown that $\sigma_1\to x_b$ as $n\to\infty$ in the proof of Lemma 1. Defining $\epsilon_n:=x_b-\sigma_1$, we have $\epsilon_n\searrow 0$ as $n\to\infty$. By using Taylor's theorem for $\log\ell_n(\sigma_1)$ and $\log\lambda_n(\sigma_1)$ around $\sigma_1=x_b$, as well as using $\lambda_n(x_b)=1$, we have $a_{1n},a_{2n}\in[\sigma_1,x_b]$ such that

$$\begin{aligned}\log\ell_n(\sigma_1)&=\log\ell_n(x_b)-\frac{\ell_n'(a_{1n})}{\ell_n(a_{1n})}\epsilon_n\\ \log\lambda_n(\sigma_1)&=-\lambda_n'(x_b)\epsilon_n+\eta_n(a_{2n})\epsilon_n^2,\end{aligned}$$

where

$$\eta_n(x):=\frac{\lambda_n''(x)\lambda_n(x)-\lambda_n'(x)^2}{2\lambda_n(x)^2}.$$

Applying these results to (24), we have

$$(n-1)\bigl(-\lambda_n'(x_b)\epsilon_n+\eta_n(a_{2n})\epsilon_n^2\bigr)+\log\ell_n(x_b)-\frac{\ell_n'(a_{1n})}{\ell_n(a_{1n})}\epsilon_n=O(1/n).$$

Rearranging terms, we obtain

$$\left(n-1+\frac{\ell_n'(a_{1n})}{\lambda_n'(x_b)\ell_n(a_{1n})}\right)\epsilon_n=\frac{\log\ell_n(x_b)}{\lambda_n'(x_b)}+\frac{(n-1)\eta_n(a_{2n})\epsilon_n^2-O(1/n)}{\lambda_n'(x_b)}.\tag{25}$$

By using $\ell_n=f_n^H/f_n^L$, we have

$$\frac{\log\ell_n(x_b)}{\lambda_n'(x_b)}=\frac{\log\bigl(f_n^H(x_b)/f_n^L(x_b)\bigr)}{f_n^H(x_b)-f_n^L(x_b)},$$

where $\lambda_n'(x_b) = f_n^H(x_b) - f_n^L(x_b)$ obtains from $\lambda_n' = (f_n^H F_n^L - F_n^H f_n^L)/(F_n^L)^2$ and $F_n^H(x_b) = F_n^L(x_b) = 1$. By Assumption 2, $\lim_{n\to\infty} f_n^H(x_b)/f_n^L(x_b) = \lim_{n\to\infty} \ell_n(x_b) = 1$. Thus, l'Hôpital's rule implies that

$$\lim_{n\to\infty} \frac{\log \ell_n(x_b)}{\lambda_n'(x_b)} = \lim_{n\to\infty} \frac{\log\bigl(f_n^H(x_b)/f_n^L(x_b)\bigr)}{f_n^H(x_b)/f_n^L(x_b) - 1} \frac{1}{f_n^L(x_b)} = \lim_{n\to\infty} \frac{1}{f_n^L(x_b)},$$

which is a finite positive constant.

Note that $\Lambda_n(x_b) = \ell_n(x_b)$ and $\lambda_n(x_b) = 1$. Hence, Assumption 1 implies that

$$\frac{f_n^H(x_b)}{f_n^L(x_b)} = \ell_n(x_b) = \frac{\Lambda_n(x_b)}{\lambda_n(x_b)} > e^{\delta n^{-\xi}}.$$

Therefore,

$$\frac{1}{\lambda_n'(x_b)} = \frac{1/f_n^L(x_b)}{f_n^H(x_b)/f_n^L(x_b) - 1} < \frac{1/f_n^L(x_b)}{e^{\delta n^{-\xi}} - 1} = O\bigl(n^{\xi}\bigr).$$

We apply this result to terms in (25). First, $O(1/n)/\lambda_n'(x_b) < O(n^{\xi-1})$. Thus, this term is dominated by $\log(\ell_n(x_b)/\lambda_n'(x_b))$, which is an $O(1)$ term. Second, since $|\ell_n'/\ell_n| < \infty$, we have

$$\frac{\ell_n'(a_{1n})}{\lambda_n'(x_b)\ell_n(a_{1n})} < O\bigl(n^{\xi}\bigr).$$

Since $\xi < 1$, this term is dominated by $n-1$ for large n. Third, $\eta_n(a_{2n}) > -\infty$, since f_n^s is continuously differentiable. Also, $a_{2n} \to x_b$ as $n \to \infty$, since $a_{2n} \in [\sigma_1, x_b]$ and $\sigma_1 \to x_b$ as $n \to \infty$. Hence, $\eta_n(a_{2n}) \le 0$ for sufficiently large n by Assumption 3.

Collecting these results, (25) implies an asymptotic relation

$$O(n)\epsilon_n = \frac{\log \ell_n(x_b)}{\lambda_n'(x_b)} + \frac{O(n)\eta_n(a_{2n})\epsilon_n^2}{\lambda_n'(x_b)}. \tag{26}$$

On the one hand, if ϵ_n is dominated by $O(n^{-1})$, then $-O(n)\eta_n(a_{2n})\epsilon_n^2/\lambda_n'(x_b)$ is dominated by $O(n^{\xi-1})$. Hence, both $O(n)\epsilon_n$ and $-O(n)\eta_n(a_{2n})\epsilon_n^2/\lambda_n'(x_b)$ converge to 0 as $n \to \infty$, which contradicts that $(\log \ell_n(x_b))/\lambda_n'(x_b)$ converges to a positive constant in (26). On the other hand, if ϵ_n dominates $O(n^{-1})$, then $O(n)\epsilon_n - O(1)$ becomes positive for sufficiently large n. This contradicts $\eta_n(a_{2n}) \le 0$ in (26). Hence, $\epsilon_n = O(n^{-1})$. Substituting into (25), we obtain

$$\lim_{n\to\infty} (n-1)(x_b - \sigma_1) = \lim_{n\to\infty} \frac{1}{f_n^L(x_b)}.$$

Applying this to

$$n\bigl(1 - F_n^{\theta}(\sigma_1)\bigr) = n\bigl(f_n^{\theta}(x_b)(x_b - \sigma_1) - O(x_b - \sigma_1)^2\bigr),$$

we obtain, for any $\theta \in \{H, L\}$,

$$\lim_{n\to\infty} n\bigl(1 - F_n^{\theta}(\sigma_1)\bigr) = \lim_{n\to\infty} \frac{f_n^{\theta}(x_b)}{f_n^L(x_b)} = 1.$$

This completes the proof.

Finally, we note that Assumption 3 is not essential for heavy-tailed outcomes that replicate power laws in finite samples. Lemma 3 established a power-law tail for volume conditional on initial buying traders $D(p_1, \mathbf{X})$. The proof of Lemma 3 showed that $D(p_1, \mathbf{X})$ follows the binomial distribution with probability $\pi_n^\theta := 1 - F_n^\theta(\sigma_1)$ and population n. The analysis in the previous paragraph implies that if Assumption 3 fails, the mean of the binomial, $n\pi_n^\theta$, may diverge as $n \to \infty$. However, for finite n, the binomial distribution for $D(p_1, \mathbf{X})$ is well defined. Combining it with (14), we obtain the unconditional probability of $M_n^\dagger = m$ for $m = 2, 3, \ldots$ as

$$\sum_{c=2}^{\infty} \binom{n}{c} (\pi_n^\theta)^c (1 - \pi_n^\theta)^{n-c} \frac{(c-1)(m-1)^{m-c-1} e^{-m+1}}{(m-c)!}.$$

Note that π_n^θ is determined independently of m. Hence, applying Stirling's formula for $(m-c)!$ as in Lemma 3, we obtain an approximate power law with exponent $1/2$ for $M_n^\dagger$.

References

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi (2017), "Microeconomic origins of macroeconomic tail risks." *American Economic Review*, 107, 54–108. [1375]

Alfarano, Simone, Thomas Lux, and Friedrich Wagner (2008), "Time variation of higher moments in a financial market with heterogeneous agents: An analytical approach." *Journal of Economic Dynamics and Control*, 32, 101–136. [1382]

Ankudinov, Andrei, Rustam Ibragimov, and Oleg Lebedev (2017), "Heavy tails and asymmetry of returns in the Russian stock market." *Emerging Markets Review*, 32, 200–219. [1365]

Attar, Andrea, Thomas Mariotti, and François Salanié (2014), "Nonexclusive competition under adverse selection." *Theoretical Economics*, 9, 1–40. [1367]

Avery, Christopher and Peter Zemsky (1998), "Multidimensional uncertainty and herd behavior in financial markets." *American Economic Review*, 88, 724–748. [1368]

Bak, P., M. Paczuski, and M. Shubik (1997), "Price variations in a stock market with many agents." *Physica A: Statistical Mechanics and its Applications*, 246, 430–453. [1367]

Banerjee, Abhijit V. (1992), "A simple model of herd behavior." *Quarterly Journal of Economics*, 107, 797–817. [1367]

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1992), "A theory of fads, fashion, custom, and cultural change as informational cascades." *Journal of Political Economy*, 100, 992–1026. [1367]

Brunnermeier, Markus K. (2001), *Asset Pricing Under Asymmetric Information—Bubbles, Crashes, Technical Analysis, and Herding*. Oxford University Press, London. [1367]

Bulow, Jeremy and Paul Klemperer (1994), "Rational frenzies and crashes." *Journal of Political Economy*, 102, 1–23. [1372]

Caplin, Andrew and John Leahy (1994), "Business as usual, market crashes, and wisdom after the fact." *American Economic Review*, 84, 548–565. [1367]

Chamley, Christophe P. (2004), *Rational Herds—Economic Models of Social Learning*. Cambridge University Press, Cambridge. [1367]

Chari, V. V. and Patrick J. Kehoe (2004), "Financial crises as herds: Overturning the critiques." *Journal of Economic Theory*, 119, 128–150. [1367]

Cipriani, Marco and Antonio Guarino (2005), "Herd behavior in a laboratory financial market." *American Economic Review*, 95, 1427–1443. [1367]

Clauset, Aaron, Cosma Rohilla Shalizi, and M. E. J. Newman (2009), "Power-law distributions in empirical data." *SIAM Review*, 51, 661–703. [1379]

Cont, Rama and Jean-Philippe Bouchaud (2000), "Herd behavior and aggregate fluctuations in financial markets." *Macroeconomic Dynamics*, 4, 170–196. [1367]

Cont, Rama, Marc Potters, and Jean-Philippe Bouchaud (1997), "Scaling in stock market data: Stable laws and beyond." In *Scale Invariance and Beyond*, Vol. 7 (B. Dubrulle, F. Graner, and D. Sornette, eds.), 75–85, Springer, Berlin, Heidelberg. [1365]

Fama, Eugene F. (1963), "Mandelbrot and the stable Paretian hypothesis." *Journal of Business*, 36, 420–429. [1365]

Feller, William (1966), *An Introduction to Probability Theory and Its Applications*, second edition, Vol. II. Wiley, New York, New York. [1374]

Gabaix, Xavier (2009), "Power laws in economics and finance." *Annual Review of Economics*, 1, 255–294. [1365]

Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and H. Eugene Stanley (2006), "Institutional investors and stock market volatility." *Quarterly Journal of Economics*, 121, 461–504. [1367, 1375, 1377]

Gopikrishnan, Parameswaran, Martin Meyer, Luís A. Nunes Amaral, and H. Eugene Stanley (1998), "Inverse cubic law for the distribution of stock price variations." *European Physical Journal B*, 3, 139–140. [1365]

Gopikrishnan, Parameswaran, Vasiliki Plerou, Xavier Gabaix, and H. Eugene Stanley (2000), "Statistical properties of share volume traded in financial markets." *Physical Review E*, 62, 4493–4496. [1365]

Grossman, Sanford J. (1981), "Nash equilibrium and the industrial organization of markets with large fixed costs." *Econometrica*, 49, 1149–1172. [1367]

Gu, Zhiye and Rustam Ibragimov (2018), "The 'Cubic Law of the Stock Returns' in emerging markets." *Journal of Empirical Finance*, 46, 182–190. [1377]

Gul, Faruk and Russell Lundholm (1995), "Endogenous timing and the clustering of agents' decisions." *Journal of Political Economy*, 103, 1039–1066. [1367]

Hasbrouck, Joel (1991), “Measuring the information content of stock trades.” *Journal of Finance*, 46, 179–207. [1367, 1373, 1376]

Ibragimov, Marat, Rustam Ibragimov, and Paul Kattuman (2013), “Emerging markets and heavy tails.” *Journal of Banking & Finance*, 37, 2546–2559. [1365]

Ibragimov, Rustam (2009), *Heavy-Tailed Densities*. Palgrave Macmillan, London. [1366]

Jansen, Dennis W. and Casper G. de Vries (1991), “On the frequency of large stock returns: Putting booms and busts into perspective.” *Review of Economics and Statistics*, 73, 18–24. [1365]

Jessen, Anders Hedegaard and Thomas Mikosch (2006), “Regularly varying functions.” *Publications de L'Institut Mathématique*, 80, 171–192. [1381]

Kamada, Koichiro and Ko Miura (2014), “Confidence erosion and herding behavior in bond markets: An essay on central bank communication strategy.” Working paper, Bank of Japan Working Paper Series No. 14-E6. [1382]

Keim, Donald B. and Ananth Madhavan (1996), “The upstairs market for large-block transactions: Analysis and measurement of price effects.” *Review of Financial Studies*, 9, 1–36. [1377]

Kingman, John F. C. (1993), *Poisson Processes*. Oxford University Press, New York. [1387]

Klemperer, Paul D. and Margaret A. Meyer (1989), “Supply function equilibria in oligopoly under uncertainty.” *Econometrica*, 57, 1243–1277. [1367]

Kyle, Albert S. (1989), “Informed speculation with imperfect competition.” *Review of Economic Studies*, 56, 317–355. [1367]

Lee, In Ho (1998), “Market crashes and informational avalanches.” *Review of Economic Studies*, 65, 741–759. [1367]

Lillo, Fabrizio, J. Doyne Farmer, and Rosario N. Mantegna (2003), “Master curve for price-impact function.” *Nature*, 421, 129–130. [1376, 1377]

Lux, Thomas (1996), “The stable Paretian hypothesis and the frequency of large returns: An examination of major German stocks.” *Applied Financial Economics*, 6, 463–475. [1365]

Lux, Thomas and Simone Alfarano (2016), “Financial power laws: Empirical evidence, models, and mechanisms.” *Chaos, Solitons & Fractals*, 88, 3–18. [1365, 1377, 1380]

Lux, Thomas and Didier Sornette (2002), “On rational bubbles and fat tails.” *Journal of Money, Credit and Banking*, 34, 589–610. [1367]

Mandelbrot, Benoit (1967), “The variation of some other speculative prices.” *Journal of Business*, 40, 393–413. [1365]

Minehart, Deborah and Suzanne Scotchmer (1999), “Ex post regret and the decentralized sharing of information.” *Games and Economic Behavior*, 27, 114–131. [1366]

Mishkin, Frederic S. (1991), "Asymmetric information and financial crises: A historical perspective." In *Financial Markets and Financial Crises* (R. Glenn Hubbard, ed.), 69–108, University of Chicago Press. Chicago. [1367]

Nirei, Makoto (2006), "Threshold behavior and aggregate fluctuation." *Journal of Economic Theory*, 127, 309–322. [1387]

Nirei, Makoto (2008), "Self-organized criticality in a herd behavior model of financial markets." *Journal of Economic Interaction and Coordination*, 3, 89–97. [1367]

Nirei, Makoto (2011), "Beauty contests and fat tails in financial markets." Working paper, Research Center for Price Dynamics, Institute of Economic Research, Hitotsubashi University, No. 76. [1365, 1375]

Nirei, Makoto (2015), "An interaction-based foundation of aggregate investment fluctuations." *Theoretical Economics*, 10, 953–985. [1375]

Orléan, André (1995), "Bayesian interactions and collective dynamics of opinion: Herd behavior and mimetic contagion." *Journal of Economic Behavior & Organization*, 28, 257–274. [1367]

Park, Andreas and Hamid Sabourian (2011), "Herding and contrarian behavior in financial markets." *Econometrica*, 79, 973–1026. [1367]

Plerou, Vasiliki, Parameswaran Gopikrishnan, Xavier Gabaix, Luís A. Nunes Amaral, and H. Eugene Stanley (2001), "Price fluctuations, market activity, and trading volume." *Quantitative Finance*, 1, 262–269. [1365]

Redner, Sidney (2001), *A Guide to First-Passage Processes*. Cambridge University Press, New York, New York. [1366]

Salhi, Khaled, Madalina Deaconu, Antoine Lejay, Nicolas Champagnat, and Nicolas Navet (2016), "Regime switching model for financial data: Empirical risk analysis." *Physica A: Statistical Mechanics and its Applications*, 461, 148–157. [1366]

Smith, Lones and Peter Sørensen (2000), "Pathological outcomes of observational learning." *Econometrica*, 68, 371–398. [1367, 1383]

Sornette, Didier (2004), *Critical Phenomena in Natural Sciences*, second edition. Springer-Verlag, Berlin, Heidelberg. [1376]

Stauffer, Dietrich and Didier Sornette (1999), "Self-organized percolation model for stock market fluctuations." *Physica A: Statistical Mechanics and its Applications*, 271, 496–506. [1367]

Vives, Xavier (2008), *Information and Learning in Markets: The Impact of Market Microstructure*. Princeton University Press, Princeton. [1367, 1375]

Vives, Xavier (2011), "Strategic supply function competition with private information." *Econometrica*, 79, 1919–1966. [1367]

Co-editor Florian Scheuer handled this manuscript.

Manuscript received 21 November, 2018; final version accepted 23 January, 2020; available online 29 January, 2020.