Linking agent-based models and stochastic models of financial markets

Ling Feng[a,b,c,1], Baowen Li[a,b,1], Boris Podobnik[c,d,e,f], Tobias Preis[c,g,h], and H. Eugene Stanley[c,1]

[a]Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore 117456, Republic of Singapore; [b]Department of Physics and Centre for Computational Science and Engineering, National University of Singapore, Singapore 117542, Republic of Singapore; [c]Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215; [d]Faculty of Civil Engineering, University of Rijeka, 51000 Rijeka, Croatia; [e]Faculty of Economics, University of Ljubljana, 1000 Ljubljana, Slovenia; [f]Zagreb School of Economics and Management, 10000 Zagreb, Croatia; [g]Chair of Sociology, in particular of Modeling and Simulation, Eidgenössische Technische Hochschule Zurich, 8092 Zurich, Switzerland; and [h]Artemis Capital Asset Management GmbH, 65558 Holzheim, Germany

Contributed by H. Eugene Stanley, March 29, 2012 (sent for review December 6, 2011)

It is well-known that financial asset returns exhibit fat-tailed distributions and long-term memory. These empirical features are the main objectives of modeling efforts using (i) stochastic processes to quantitatively reproduce these features and (ii) agent-based simulations to understand the underlying microscopic interactions. After reviewing selected empirical and theoretical evidence documenting the behavior of traders, we construct an agent-based model to quantitatively demonstrate that "fat" tails in return distributions arise when traders share similar technical trading strategies and decisions. Extending our behavioral model to a stochastic model, we derive and explain a set of quantitative scaling relations of long-term memory from the empirical behavior of individual market participants. Our analysis provides a behavioral interpretation of the long-term memory of absolute and squared price returns: They are directly linked to the way investors evaluate their investments by applying technical strategies at different investment horizons, and this quantitative relationship is in agreement with empirical findings. Our approach provides a possible behavioral explanation for stochastic models for financial systems in general and provides a method to parameterize such models from market data rather than from statistical fitting.

complex systems | power law | scaling laws

Modeling price returns has become a central topic in the study of financial markets due to its key role in financial theory and its practical utility. Following models by Engle and Bollerslev (1, 2), many stochastic models have been proposed based on statistical studies of financial data to accurately reproduce price dynamics. In contrast to this stochastic approach, economists and physicists using the tools of statistical mechanics have adopted a bottom-up approach to simulate the same macroscopic regularity of price changes, with a focus on the behavior of individual market participants (3–10). Although the second so-called agent-based approach has provided a qualitative understanding of price mechanisms, it has not yet achieved sufficient quantitative accuracy to be widely accepted by practitioners.

Here, we combine the agent-based approach with the stochastic process approach and propose a model based on the empirically proven behavior of individual market participants that quantitatively reproduces fat-tailed return distributions and long-term memory properties (11–14).

Empirical and Theoretical Market Behaviors

We start by arguing that technical traders (usually agents seeking arbitrage opportunities and make their trading decisions based on price patterns) contribute much more to the dynamics of daily stock prices S_t (or log price $s_t \equiv \ln(S_t)$) than fundamentalists (who attempt to determine the fundamental values of stocks). Although fundamentalists hold a majority of the stocks, they trade infrequently (see *SI Appendix*, Fig. S6). In contrast, technical traders contribute most of the trading activities (15) by trading their minority holdings more frequently than fundamentalists. Market surveys (16–18) also provide clear evidence of the prevalence of technical analysis. We consider here only technical traders, assuming that fundamentalists contribute only to market noise. Our study is of the empirical data recorded prior to 2006 and ignores the effect of high frequency trading (HFT) that has become significant only in the past 5 y. We propose a behavioral agent-based model that is in agreement with the following empirical evidence:

i. Random trading decisions made by agents on a daily basis. n_0 technical traders use different trading strategies, hence their decisions to buy, sell, or hold a position appear to be random. A trading decision is made daily because empirical studies report the lack of intraday trading persistence in empirical trading data (19). Market survey (16) also shows that fund managers put very little emphasis on intraday tradings. We estimate the probability p of having daily trade empirically from trading volumes.
ii. Price returns. The price return $r_t \equiv s_t - s_{t-1}$ is controlled by the imbalance d_t between the demand and the supply of stocks—the difference in the number of buy and sell trades each day. The excess in total demand or supply moves the price up or down, where the largest r_t occurs when all traders act in unison, when they all either buy or sell their stocks. We assume this relationship between price change r_t and d_t to be linear each day, as supported by empirical findings (20, 21).
iii. Centralized interaction mechanism of returns on technical strategies. For technical traders, an important input parameter in their strategies is past price movement (22, 23). Consequently, prices and orders reflect a main interaction mechanism between agents. In many agent-based models, the interaction strength between agents need to be adjusted with agent population size (5, 24, 25) or interaction structure (26) to sustain "fat" tails in return distributions. Here we propose a centralized interaction mechanism (price change) among agents so that the strength of interaction grows with agent population and is unaffected by interaction structure.
iv. Opinion convergence due to price changes. This is the unique mechanism that distinguishes our model from other models. It specifies the collective behavior of technical traders. Duffy et al. (27) found that agents learn from each other and tend to adopt the strategy that gives the most payoff. Given the price patterns at any point in time, a few most profitable technical strategies dominate the market because every technical trader

Author contributions: L.F., B.L., B.P., T.P., and H.E.S. designed research; L.F., B.L., B.P., T.P., and H.E.S. performed research; L.F., B.L., B.P., T.P., and H.E.S. contributed new reagents/analytic tools; L.F., B.L., B.P., T.P., and H.E.S. analyzed data; and L.F., B.L., B.P., T.P., and H.E.S. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. E-mail: g0801896@nus.edu.sg, phylibw@nus.edu.sg, or hes@bu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1205013109/-/DCSupplemental.

wants to maximize his/her profit by using the most profitable strategy copied from each other (The most profitable strategy would become less profitable when most agents adopt it, and a new profitable strategy emerges from the new price trends; soon agents will flock to the new profitable strategy until it is no longer profitable. This is similar to the regime switching phenomenon in various agent-based models.). On the other hand, the individual strategies used by different technical traders differ in their parameterizations of the buy/sell time, amount of risk tolerated, or portfolio composition (15). So when the input signal—the previous price change r_{t-1}—is small, every agent acts independently. When the input signal is large, the agents act more in concert, irrespective of their differences in trading strategies. During the spreading panic of market crashes, for example, most agents sell their stocks (and market makers are likely to make losses in such circumstances). This supports the empirical finding that large price swings occur when the preponderance of trades have the same buy/sell decision indicated in the findings by Gabaix X. et al. (28).

Behavioral Model and Results

Based on the items of evidence (i)–(iv) listed above we construct a three-step behavioral model:

- Step 1. Based on evidence (i), we assume n_0 agents, each of equal size 1. Each day, a trading decision $\psi_i(t)$ is made by each agent i,

$$\psi_i(t) \equiv \begin{cases} 1 & \text{with probability } p \Rightarrow \text{buy;} \\ -1 & \text{with probability } p \Rightarrow \text{sell;} \\ 0 & \text{with probability } 1-2p \Rightarrow \text{hold.} \end{cases}$$

- Step 2. Based on evidence (ii), we define price change r_t to be proportional to the aggregate demand d_t, i.e., the difference in the number of agents willing to buy and sell,

$$r_t \equiv kd_t = k\sum_{i=1}^{n_0}\psi_i(t). \qquad [1]$$

Trading volume is equal to the total number of trades in this case because every agent has the same trading size of 1. Hence daily trading volume N_t is defined as

$$N_t \equiv \sum_{i=1}^{n_0}|\psi_i(t)| \qquad [2]$$

and k is the sensitivity of price change with respect to d_t. We set k to 1 because a choice for k does not affect the statistical properties. Note that, according to Step 1, maximum (minimum) d_t means that all agents are in collective mode when they behave the same, all of them either buy or sell stocks.

- Step 3. Based on evidence (iii) and (iv), at day $t+1$ each agent's opinion is randomly distributed into each of the c_{t+1} opinion groups where

$$c_{t+1} = (n_0/|r_t|)^{\omega}, \qquad [3]$$

in which all agents comprising the same opinion group execute the same action (buy, sell, or hold) with the same probability p as in Step 1. Because $1 \le c_{t+1} \le n_0$, Eq. **3** implies (a) when the previous return is maximum, $|r_t|_{\max} = n_0$ (everyone buys/sells), there is only one trading opinion among the agents; (b) when the previous return is minimum $|r_t|_{\min} = n_0^{(\omega-1)/\omega}$ (see *SI Appendix*, Section 5), there are n_0 opinions and every agent acts independently of one another. We use $\omega = 1$ in most of our simulations because it produces results that are numerically close to empirical findings. The case of $\omega = 1$ corresponds to $|r_t|_{\min} = 1$, and there is only one more buy/sell order than sell/buy order. The result with different ω is presented in the end of this section. When market noise is considered, $c_{t+1} \sim \mathcal{N}(n_0/|r_t|, \sigma_c^2)$, where $\mathcal{N}$ denotes the normal (Gaussian) distribution and $\sigma_c^2 = b \cdot n_0/|r_t|$ quantifies market noise due to external news events.

To obtain an empirical value for daily trading probability p, we choose 309 companies from the Standard and Poor's 500 index traded over the 10-y period 1997–2006. Only 309 of the 500 total stocks were consistently listed during the entire 10-y period, and they are the ones we choose. We define the trading velocity V of an agent as the total number of shares he/she trades in a year, divided by the number of shares, he/she owns on average over a year

$$V = \frac{\text{Number of trades}}{\text{Number of shares}}. \qquad [4]$$

From the Compustat database, we divide the total number of shares traded on the market by the number of outstanding shares of each stock and obtain the average yearly trading velocity, V, for these stocks through the 10-y period. Thus we obtain,

$$V = 1.64, \qquad [5]$$

where the value of V varies from 1.3 to 1.9 throughout the 10 y. Note that $V > 1$ means that, on average, each stock changes its owner more than once during a year. From the data documented by Yahoo! Finance, we assume that institutional owners are fundamentalists and that the rest of the traders are predominantly technical traders. The percentage of outstanding shares held by institutional owners is usually higher than 60%, and the average value is 83% (*SI Appendix*, Fig. S6A).

Assuming institutional owners are fundamentalists with investment (trading) horizons longer than 1 y (they trade less frequently than once a year), we calculate the value of p as follows. We let the average yearly trading velocity of fundamentalists be V_f, then the average yearly trading velocity of technical traders V_c is

$$V_c = (V - 0.83V_f)/(1 - 0.83). \qquad [6]$$

Because there are about 250 trading days in a year, we calculate the daily trading probability p

$$p = V_c/(250 \cdot 2). \qquad [7]$$

Because fundamentalists have an investment horizon longer than 1 y, $V_f < 1$ (see Eq. **4**). We arbitrarily set V_f to be 0.2, 0.4, 0.6, 0.8, and estimate the corresponding values of p as 0.0174, 0.0154, 0.0134, 0.0115, respectively. For $(V_f, p) = (0.4, 0.0154)$, we find that approximately 80% of the trading volume is contributed by technical traders.

We compare the simulation results from our behavioral model (Steps 1, 2, and 3) with those obtained for the empirical set of stocks comprising the S&P 500 index. For the cumulative distribution (CDF) of both absolute returns and number of trades. Fig. 1 shows that the model reproduces the empirical data in both the central region and the tails. In particular, both model and empirical distributions have power-law tails $(P(|r_t| > x) \sim x^{-\xi_r}, P(n_t > x) \sim x^{-\xi_n})$ (13, 29–31) where the power-law exponents ξ_r and ξ_n that we obtain for the empirical data by applying the Hill's estimator are in agreement with the ones obtained for empirical data (32).

It is worth noting that with different values of ω in Eq. **3**, the tail exponent ξ_r for the absolute return distribution varies as

APPLIED PHYSICAL SCIENCES

ECONOMIC SCIENCES

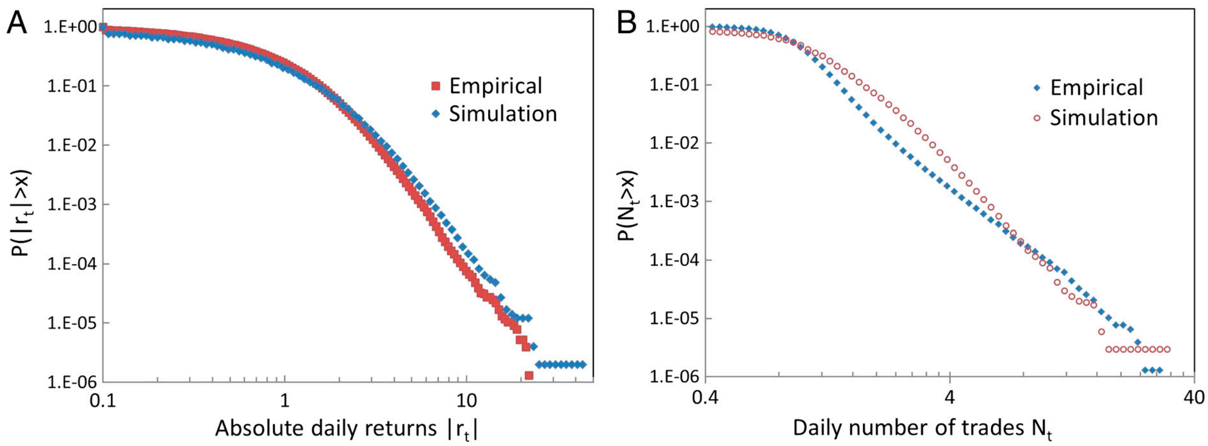

Fig. 1. Comparison of the distributional properties of simulations with empirical results for the parameter set $V_f = 0.4$, $n_0 = 2^{10}$, $b = 1.0$. V_f characterizes the how fast one agent's shares is traded. Empirical data are from the 309 stocks out of S&P 500 index components that have been consistently listed during the 10-y period 1997–2006. There are ≈800,000 data points in empirical data, and 1,000,000 sample points from simulation. (*A*) Cumulative distributions of daily returns, defined as $r_t \equiv \log[x_{t,\text{close}}] - \log[x_{t,\text{open}}]$, the difference between the daily opening and closing price of a stock on day *t*. Thus we ignore overnight returns arising, e.g., due to news events. The price for each stock is normalized to zero mean and unit variance before aggregation into a single distribution. The simulation results agree with the shape of the empirical distribution. The Hill estimator on 1% of tail region gives $\xi_{r,\text{S\&P500}} = 3.69 \pm 0.07$, $\xi_{r,\text{simulation}} = 4.08 \pm 0.08$. (*B*) The cumulative distribution of the same set of data but analyzed on the daily number of trades. Because the number of trades increases each year, we normalize the data by each stock mean number of trades on a yearly basis before aggregating them into the distributions. The simulation again reproduces the empirical results. The Hill estimator applied to 2.5% of the tail region gives $\xi_{n,\text{S\&P500}} = 4.02 \pm 0.07$, $\xi_{n,\text{simulation}} = 4.02 \pm 0.08$.

shown in Table 1. We find that $\omega = 1$ gives the tail exponent ξ_r that is closest to empirical finding (see *SI Appendix*, Section 5 for the implication of different values in ω).

Overall, we are able to generate power-law distributions in both returns and number of trades in agreement with empirical data. The fact that we are able to capture the trends in the two highly correlated quantities (33) implies a very plausible mechanism underlying our model. Furthermore, the tail exponent ξ_r is invariant with respect to different values of agent population size n_0 and fundamentalists' investment horizon V_f (Table 2 and *SI Appendix*, Section 1). In particular, the insensitivity of the result to the number of agents n_0 distinguishes this model to most other agent-based models (34, 35).

Extension to Stochastic Model

Next we study the underlying stochastic process of our behavioral model (Step 1 to Step 3). The mechanism of opinion convergence under price changes is incorporated into a mathematical form for an analytical understanding. Step 3 in our behavioral model indicates that the daily price change is not deterministic, where its variance $\sigma_t^2 \equiv E(r_t - E(r_t))^2$ is related to total number of opinion groups c_t, which is determined directly by previous return r_{t-1}. On average, each opinion group has $|r_{t-1}|$ agents, and we have

$$\begin{aligned}\sigma_t^2 &\equiv E(r_t^2|r_{t-1}) = c_t \cdot [p \cdot r_{t-1}^2 + p \cdot (-r_{t-1})^2 + (1-2p) \cdot 0] \\ \sigma_t^2 &= 2pn_0|r_{t-1}|.\end{aligned} \qquad [8]$$

Because σ_t presents the standard deviation of price change r_t, we can model price change as $r_t = \sigma_t \eta_t$, where η_{t+1} is a random variable with zero mean and unit variance, and its distribution is determined by Step 3. For simplicity, we use a normal distribution for η_t, although we find that other distributions, such as *t*-distribution, give similar results in our analysis (*SI Appendix*, Section 4).

Table 1 Tail exponents of absolute return distribution with different ω

ω	0.5	0.6	0.7	0.8	0.9	**1.0**	1.1	1.2	1.4	1.5	2.0
ξ_r	2.3	2.5	2.7	3.0	3.3	**3.8**	4.3	4.4	5.9	7.8	9.4

Again we note that the variance σ_t^2 in Eq. **8** is not constant but is time dependent, and time-dependent variances are commonly found in a variety of empirical outputs where phenomena are ranging from finance to physiology (2, 36). They are widely modeled with the Autoregressive Conditional Heteroskedasticity (ARCH) process (1). Here we provide a possible explanation for the ARCH effect found in financial data in terms of the behavior of technical traders—larger previous price change r_{t-1} brings traders' opinions closer to each other, resulting in large subsequent price fluctuations. Theoretical analysis on Eq. **8** leads to the fat-tailed return distributions (*SI Appendix*, Section 2) ubiquitously observed in real data, and in our case a power-law tail. Previous works have generated ARCH effect from other mechanisms that is different from our model (37).

The time-dependent variance of Eq. **8** defined by the most recent price change r_{t-1} is therefore based on short memory in the previous r_t. To this end, we further extend our behavioral model with two more items of empirical evidence:

v. Technical strategies are applied at different return intervals. In contrast to the simplistic realization in our behavioral model (Step 1 to Step 3) in which technical traders make their decisions only upon the most recent daily returns (see Eq. **8** and Step 3), market survey (16) indicates that technical strategies in practice are applied at different investment horizons ranging from 1 d to more than 1 y. This finding implies that agents calibrate their technical strategies based on returns of different time horizons, i.e., how stock preformed during the last day, week, month, up to year, or even longer. Most technical strategies are focused on short-term returns of a few days, and fewer are focused at yearly returns according to the survey (16).
vi. Increasing trading activity in volatile market conditions. Agents tend to trade more after large price movements. Different technical strategies set different thresholds on prices to trigger trading decisions (38), so large price fluctuations are likely to trigger more trades. Hence the probability of daily trade p is directly related to past returns. Because a large proportion of technical strategies is applied with short investment horizons, price changes from previous days have larger impact on p than price changes from past year.

Table 2 Tail exponents of absolute return distribution with different values of V_f and n_0

	$n_0 = 2^{10}$, $b = 1.0$				$V_f = 0.4$, $b = 1.0$			
Parameter value	$V_f = 0.2$	$V_f = 0.4$	$V_f = 0.6$	$V_f = 0.8$	$n_0 = 2^8$	$n_0 = 2^{10}$	$n_0 = 2^{12}$	$n_0 = 2^{14}$
Tail exponents ξ_r	3.69 ± 0.07	3.69 ± 0.07	3.74 ± 0.07	3.89 ± 0.07	3.86 ± 0.07	3.70 ± 0.07	3.66 ± 0.07	3.69 ± 0.07

The survey on US market (16) allows us to estimate the relative proportion of technical analysis applied at different investment horizons. Precisely, agents having investment horizon i days are affected by price change in the past i days, and the relative portion of such traders is characterized by α_i. Fig. 2 illustrates this plot, and we find that the curve follows a power-law decay with exponent $d = 1.12$, i.e., $\alpha_i \propto i^{-d}$. As different agents look at returns over time scales of differing lengths, their trading opinions are affected by the past returns of history longer than one day, so is the convergence of their collective opinions. Agents having investment horizon i days are affected by price difference in the past i days, and the relative portion of such traders are characterized by α_i. Hence the convergence of opinions is not only affected by the previous day's return but also the price difference of longer durations. Therefore Eq. **3** is better written as

$$c_{t+1} = \frac{n_0}{\sum_{i=1}^{M} \alpha_i |s_t - s_{t-i}|}, \qquad [9]$$

where M is the maximum investment horizon for which technical analysis is applied, and α_i is the proportion of agents focusing on investment horizon i and it decays with i as $\alpha_i \propto i^{-d}$ as in Fig. 2. When $M = 1$, Eq. **9** is identical to Eq. **3**, which is the scenario for homogeneous investment horizon of 1 d. $|s_t - s_{t-i}|$ is used as a simplified information content for technical traders with investment horizon i. Because there can be at least one opinion group, this boundary condition requires $\sum_{i=1}^{M} \alpha_i = 1$. From Eq. **9**, Eq. **8** is transformed into

$$\sigma_{t+1}^2 = 2pn_0 \sum_{i=1}^{M} \alpha_i |s_t - s_{t-i}| = 2pn_0 \alpha \Sigma_t, \qquad [10]$$

where $\alpha \equiv [\sum_{i=1}^{M} i^{-d}]^{-1}$ and $\Sigma_t \equiv \sum_{i=1}^{M} i^{-d} |s_t - s_{t-i}|$.

Note that Eq. **10** can be also intuitively derived. Because by α_i we denote the proportion of agents observing the price change in the past i days, the price change $|s_t - s_{t-i}|$ contributes to the overall opinion convergence with a weight α_i. Even without knowing the exact contribution of $|s_t - s_{t-i}|$ in the overall opinion convergence, by taking its first order effect with weight α_i, we would have the functional form of Eq. **10**.

To take into account the effect of increasing trading activity when market is more volatile, we assume that traders place technical thresholds based on returns of different horizons, and that the proportion of traders with different horizons is determined by the same survey results reported in ref. (16). Similar to the construction of cluster formation in Eq. **9**, we define the time-dependent probability of trading as $p_{t+1} \approx p_0 + \alpha' \Sigma_t$. By p_0 we denote the base trading activity (largely due to fundamentalists) and $\alpha' \Sigma_t$ is the additional trade due to threshold crossing by technical traders. Therefore we transform Eq. **10** to

$$\sigma_{t+1}^2 \approx 2p_{t+1} n_0 \alpha \Sigma_t = 2\alpha n_0 (p_0 + \alpha' \Sigma_t) \Sigma_t$$
$$= -\frac{\alpha n_0 p_0^2}{2\alpha'} + 2\alpha\alpha' n_0 \left(\Sigma_t + \frac{p_0}{2\alpha'}\right)^2.$$

The first term is very small compared to the second term because technical traders dominate trading activities. Hence, we can ignore the first constant term and focus on the second quadratic term. Therefore, the previous equation transforms to

$$\sigma_{t+1} \approx \sqrt{2\alpha\alpha'} \left(\Sigma_t + \frac{p_0}{2\alpha'}\right) = A + B\Sigma_t. \qquad [11]$$

Constant A is a scaling parameter defining the average size of returns, and B characterizes the relative portion of trades accomplished by technical traders (due to crossing the thresholds) vs. background trading activities (mostly by fundamentalists). Eq. **11** has similar functional form as σ_{t+1} in the fractional integrated

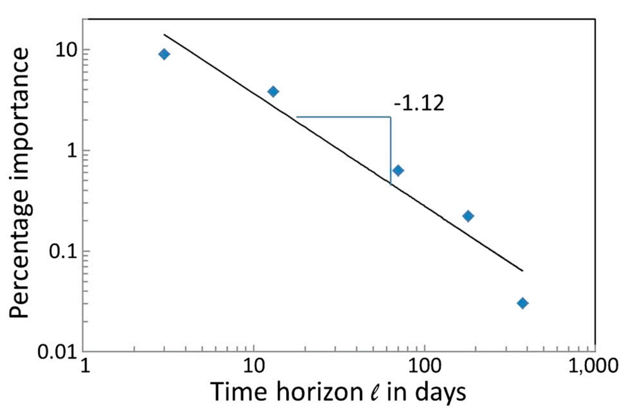

Fig. 2. Plot of survey result (16) on percent importance placed on technical analysis at different time horizons by U.S. fund managers. The plot shows a power-law function with exponent −1.12. We have combined the percent values for both flow and technical analysis at each time horizon, as flow analysis in a broad sense is one type of technical analysis. Technical analysis is heavily used at investment horizons of days to weeks and decays to close to zero at 500 d.

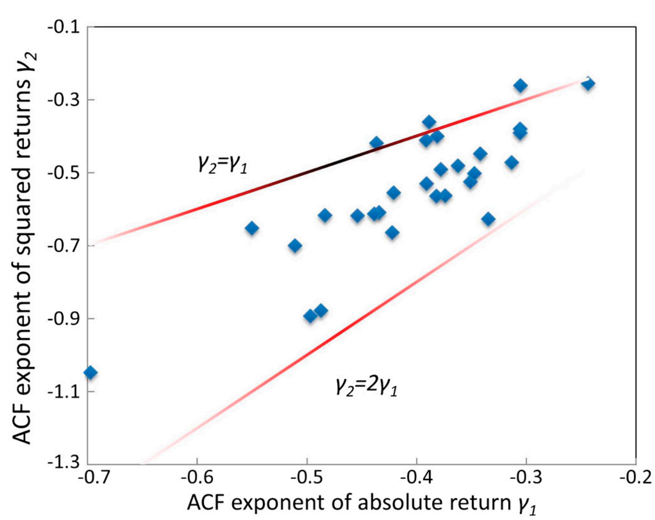

Fig. 3. Confirmation of the scaling relations Eq. **12**. Dependence of γ_1 against γ_2 for the 30 stocks in Dow Jones Indices components. Daily closing prices from 1971–2010 are used for each stock. The two exponents fulfill $\gamma_1 \le \gamma_2 \le 2\gamma_1$, so the scaling relation predicted by our stochastic model is consistent with these empirical data. The range of lag, $11 \le \ell \le 250$, is used because theoretical power-law decay of the ACF is valid for large ℓ, and for $\ell > 250$ the ACF values are close to the noise level.

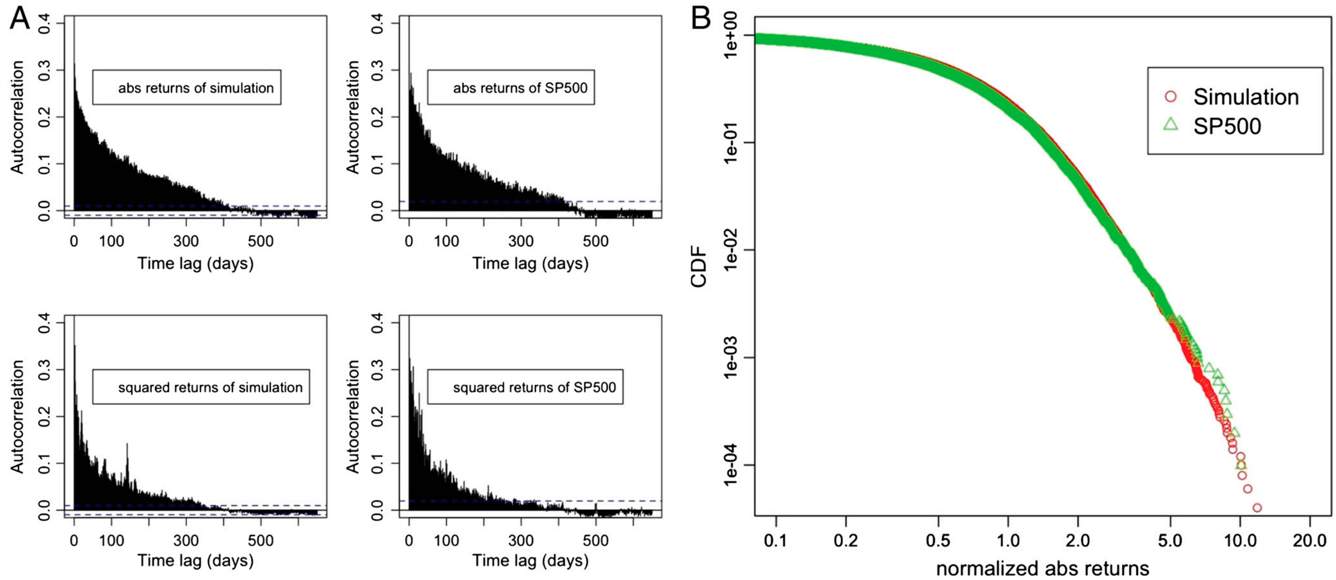

Fig. 4. Comparison between empirical data and simulation. The empirical data (about 10,000 data points) are from the S&P 500 index daily closing prices in the 40-y period 1971–2010. The simulation are carried out for 50,000 data points to obtain good convergence. Black Monday, October 19, 1989, is removed from the analysis of the ACF. (*A*) Autocorrelation of simulation vs. S&P 500 index. The simulation results show behavior similar to the S&P 500 for both absolute returns and squared returns. The rate of decay is also similar numerically. (*B*) The CDF of absolute returns for simulations and the S&P 500. The simulation reproduces the return distribution well.

processes (39). However, there are some differences in a way how past returns are used in σ_{t+1}—whereas in fractional integrated process σ_{t+1} depends on past daily returns, in Eq. **11** we use absolute values of past aggregate returns. Precisely, the absolute values of past daily, weekly, and monthly returns, to mention a few. Parke (40) has demonstrated how heterogeneous error durations among traders could result in fractional integration, and here we explicitly and quantitatively provide a behavioral interpretation of the long memory in absolute returns. The decaying dependence of standard deviation σ_{t+1} on past aggregate returns of different durations in Eq. **11** is a direct outcome of agents applying technical strategies at different investment horizons. We find this dependence to decay as a power law, and the power-law exponent d is calculated from empirical observations (16).

Analytical and Simulated Results

In general, the long memory in returns can be demonstrated by analyzing auto-correlation functions (ACF). We find that the ACFs of both absolute and squared returns (ρ_ℓ and ρ'_ℓ) decay as power laws with ℓ, i.e., $\rho_\ell(|r_t|, |r_{t-\ell}|) \propto \ell^{-\gamma_1}$ and $\rho'_\ell(r_t^2, r_{t-\ell}^2) \propto \ell^{-\gamma_2}$, where we obtain the scaling relations (*SI Appendix*, Section 2) similar to phase transitions in statistical mechanics,

$$\begin{cases} \gamma_1 = 2d - 2; \\ 2d - 2 \le \gamma_2 \le 4d - 4; \\ \gamma_1 \le \gamma_2 \le 2\gamma_1. \end{cases} \qquad [12]$$

Whereas Ding et al. have shown that correlations in $|r_t|$ decay more slowly than correlations in r_t^2 (41), we, instead, show quantitative scaling relations that have been derived and explained by the behavior of individual market participants (characterized by d). Empirical verification with the stock components of Dow Jones Indices in Fig. 3 confirms the validity of the scaling relation of Eq. **12**.

The behavioral understanding of our stochastic process Eq. **11** allows us to perform simulations in which every parameter is based on empirical calibration rather than on conventional statistical estimation. In Eq. **10** we replace the upper limit of M in the summation by 500 trading days, which corresponds to ≈2 years of investment horizon as implied by the survey (16). We set $d = 1.2$, in agreement with empirical finding (16). From the empirical trading volume fluctuations and unconditional variance of price changes, we estimate $A = 0.002$ and $B = 0.05$ (*SI Appendix*, Section 3). Therefore we obtain $\sigma_{t+1} \approx 0.002 + 0.05\sum_{i=1}^{500} i^{-1.2}|s_t - s_{t-i}|$. In Fig. 4 we compare the simulation results with the results for S&P 500 index. We demonstrate that the ACF of absolute and squared returns—for both simulations and the S&P 500 index—fulfill our scaling relations of Eq. **12** with $\gamma_{1,\text{simulation}} = 0.40$, $\gamma_{2,\text{simulation}} = 0.53$; $\gamma_{1,\text{S\&P500}} = 0.44$, and $\gamma_{2,\text{S\&P500}} = 0.7$.

Eq. **11** reflects the long-term memory of empirical data accurately while still being a stationary process (*SI Appendix*, Section 3) in contrast to many other power-law decaying processes (2, 39). In addition, the behavioral picture of Eq. **11** explains the origin of long-term memory through heterogeneous investment horizons of different technical traders based on empirical evidence.

Summary

Starting from the empirical behavior of agents, we construct an agent-based model that quantitatively explains the fat tails and long-term memory phenomena of financial time series without suffering from finite-size effects (35). The agent-based model and the derived stochastic model differ in construction but share the same mechanism of opinion convergence among technical traders. Whereas the agent-based model singles out the dominant market mechanism, the stochastic model allows a detailed analytical study. Both approaches allow their parameter values to be retrieved from market data with clear behavioral interpretations, thus allowing an in-depth study of this highly complex system of financial market.

The universality of various empirical features implies a dominant mechanism underlying market dynamics. Here we propose that this mechanism is driven by the use of technical analysis by market participants. In particular, past price fluctuations can directly induce convergence or divergence of agents' trading decisions, which in turn give rise to the "ARCH effect" in empirical findings. Additionally, the heterogeneity in agents' investment horizons gives rise to long-term memory in volatility.

ACKNOWLEDGMENTS We thank F. Pammolli, J. Cherian, and Y. Chen for discussions and suggestions. The work has been partially supported by the Grant R-144-000-285-646 from the National University of Singapore, the German Research Foundation through Grant PR 1305/1-1, the Ministry of Science of Croatia, the Keck Foundation Future Initiatives Program, Defense Threat Reduction Agency, and the National Science Foundation Grants PHY-0855453, CMMI-1125290, CHE-0911389, and CHE-0908218.

1. Engle RF (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50:987–1007.
2. Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. *J Econom* 31:307–327.
3. Lux T, Marchesi M (1999) Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature* 397:498–500.
4. Duffy J, Ünver MU (2006) Asset price bubbles and crashes with near-zero-intelligence traders. *Econ Theor* 27:537–563.
5. Cont R, Bouchaud JP (2000) Herd behavior and aggregate fluctuations in financial markets. *Macroecon Dyn* 4:170–199.
6. Preis T (2011) Econophysics-complex correlations and trend switchings in financial time series. *Eur Phys J-Spec Top* 194:5–86.
7. Preis T, et al. (2011) GPU-computing in econophysics and statistical physics. *Eur Phys J Spec Top* 194:87–119.
8. Preis T, Moat HS, Stanley HE, Bishop SR (2012) Quantifying the advantage of looking forward. *Sci Rep* 2:350.
9. Mike S, Farmer JD (2008) An empirical behavioral model of liquidity and volatility. *J Econ Dyn Control* 32:200–234.
10. Hommes CH (2002) Modeling the stylized facts in finance through simple nonlinear adaptive systems. *Proc Natl Acad Sci USA* 99:7221–7228.
11. Mandelbrot B (1963) The variation of certain speculative prices. *J Bus* 36:394–419.
12. Gopikrishnan P, Plerou V, Amaral LAN, Meyer m, Stanley HE (1999) Scaling of the distribution of fluctuations of financial market indices. *Phys Rev E* 60:5305–5316.
13. Plerou V, Gopikrishnan P, Amaral LAN, Gabaix X, Stanley HE (2000) Economic fluctuations and anomalous diffusion. *Phys Rev E* 62:R3023–R3026.
14. Maslov S, Mills M (2001) Price fluctuations from the order book perspective: Empirical facts and a simple model. *Physica A* 299:234–246.
15. Brown B (2010) Chasing the Same Signals: How Black Box Trading Influences Stock Markets from Wall Street to Shanghai. (Wiley, New York).
16. Menkhoff L (2010) The use of technical analysis by fund managers: International evidence. *J Bank Financ* 34:2573–2586.
17. Taylor MP, Allen H (1992) The use of technical analysis in the foreign exchange market. *J Int Money Financ* 11:304–314.
18. Cheung YW, Wong CYP (2000) A survey of market practitioners' views on exchange rate dynamics. *J Int Econ* 51:401–419.
19. Eisler Z, Kertèsz J (2007) Liquidity and the multiscaling properties of the volume traded on the stock market. *Europhys Lett* 77:28001.
20. Plerou V, Gopikrishnan P, Gabaix X, Stanley HE (2002) Quantifying stock-price response to demand fluctuations. *Phys Rev E* 66:027104.
21. Bouchaud JP, Farmer JD, Lillo F (2009) Handbook of Financial Markets: Dynamics and Evolution. (North-Holland, Amsterdam), pp 57–160.
22. Preis T, Paul W, Schneider JJ (2008) Fluctuation patterns in high-frequency financial asset returns. *Europhys Lett* 82:68005.
23. Preis T, Virnau P, Paul W, Schneider JJ (2009) Accelerated fluctuation analysis by graphic cards and complex pattern formation in financial markets. *New J Phys* 11:093024.
24. Egenter E, Lux T, Stauffer D (1999) Finite-size effects in Monte Carlo simulations of two stock market models. *Physica A* 268:250–256.
25. Lux T, Schornstein S (2005) Genetic learning as an explanation of stylized facts of foreign exchange markets. *J Math Econ* 41:169–196.
26. Lux T (2009) Handbook of Financial Markets: Dynamics and Evolution. (North-Holland, Amsterdam), pp 161–215.
27. Duffy J, Feltovich N (1999) Does observation of others affect learning in strategic environments? An experimental study. *Int J Game Theory* 28:131–152.
28. Gabaix X, Gopikrishnan P, Plerou V, Stanley HE (2003) A theory of power-law distributions in financial market fluctuations. *Nature* 423:267–270.
29. Plerou V, Stanley HE (2008) Stock return distributions: Tests of scaling and universality from three distinct stock markets. *Phys Rev E* 77:037101.
30. Podobnik B, Valentincic A, Horvatic D, Stanley HE (2011) Asymmetric levy flight in financial ratios. *Proc Natl Acad Sci USA* 108:17883–17888.
31. Podobnik B, Horvatic D, Petersen AM, Urosevic B, Stanley HE (2010) Bankruptcy risk model and empirical tests. *Proc Natl Acad Sci USA* 107:18325–18330.
32. Hill BM (1975) A simple general approach to inference about the tail of a distribution. *Ann Stat* 3:1163–1174.
33. Podobnik B, Horvatic D, Petersen AM, Stanley HE (2009) Cross-correlations between volume change and price change. *Proc Natl Acad Sci USA* 106:22079–22084.
34. Alfarano S, Lux T, Wagner F (2008) Time-variation of higher moments in financial markets with heterogeneous agents: An analytical approach. *J Econ Dyn Control* 32:101–136.
35. Irle A, Kauschke J (2011) Switching rates and the asymptotic behavior of herding models. *Adv Complex Syst* 14:359–376.
36. Ashkenazy Y, et al. (2001) Magnitude and sign correlations in heartbeat fluctuations. *Phys Rev Lett* 86:1900–1903.
37. Aoki M (2002) Open models of market shares with two dominant types of market participants. *J Econ Behav Organ* 49:199–216.
38. Cont R (2007) Long Memory in Economics. (Springer, Berlin), pp 289–309.
39. Baillie RT, Bollerslev T, Mikkelsen HO (1996) Fractionally integrated generalized autoregressive conditional heteroskedasticity. *J Econom* 74:3–30.
40. Parke WR (1999) What is fractional integration? *Rev Econ Stat* 81:632–638.
41. Ding Z, Granger CWJ, Engle RF (1993) A long memory property of stock market returns and a new model. *J Empir Financ* 1:83–106.