1

# How volatilities nonlocal in time affect the price dynamics in complex financial systems

Lei Tan[1,2], Bo Zheng[1,2,*], Jun-Jie Chen[1,2], Xiong-Fei Jiang[1],

**1 Department of Physics, Zhejiang University, Hangzhou 310027, China**
**2 Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing 210093, China**
**\* E-mail: zhengbo@zju.edu.cn**

## Abstract

What is the dominating mechanism of the price dynamics in financial systems is of great interest to scientists. The problem whether and how volatilities affect the price movement draws much attention. Although many efforts have been made, it remains challenging. Physicists usually apply the concepts and methods in statistical physics, such as temporal correlation functions, to study financial dynamics. However, the usual volatility-return correlation function, which is local in time, typically fluctuates around zero. Here we construct dynamic observables nonlocal in time to explore the volatility-return correlation, based on the empirical data of hundreds of individual stocks and 25 stock market indices in different countries. Strikingly, the correlation is discovered to be non-zero, with an amplitude of a few percent and a duration of over two weeks. This result provides compelling evidence that past volatilities nonlocal in time affect future returns. Further, we introduce an agent-based model with a novel mechanism, that is, the asymmetric trading preference in volatile and stable markets, to understand the microscopic origin of the volatility-return correlation nonlocal in time.

## Author Summary

## Introduction

Financial markets, as a kind of typical complex systems with many-body interactions, have drawn much attention of scientists. In recent years, for example, various concepts and methods in statistical physics have been applied and much progress has been achieved [1–22]. Following the trend towards quantitative analysis in finance, the efforts of scientists in different fields promote each other and deepen our understanding of financial systems [1, 6, 17, 23–37].

From the perspective of physicists, a financial market is regarded as a dynamic system, and the price dynamics, i.e. the time evolution of stock prices, can be characterized by temporal correlation functions, which describe how one variable statistically changes with another. It is well-known that the price volatilities are long-range correlated in time, which is called volatility clustering. Many activities have been devoted to the study of the collective behaviors related to volatility clustering in stock markets [3, 5, 6, 38, 39]. However, our understanding on the movement of the price return itself is very much limited. The autocorrelating time of returns is extremely short, that is, on the order of minutes [3, 38]. As to higher-order time correlations, it is discovered that the return-volatility correlation is negative — in other words, past negative returns enhance future volatilities [4, 9, 23, 40, 41]. This is the so-called leverage effect. As far as we know, all stock markets in the world exhibit the leverage effect except for the Chinese stock market, which unexpectedly shows an anti-leverage effect, i.e., the correlation between past returns and future volatilities is positive [9, 41]. Returns represent the price changes, and volatilities measure the fluctuations of the price movement. The leverage and anti-leverage effects characterize how price changes induce fluctuations. At this stage, one may ask what affects the return itself. It has been discovered that future returns can be predicted by the dividend-price ratio [42, 43], which is corroborated by subsequent studies. However, the predictive power of the dividend-price ratio is sensitive to the selection of the

sample period [44, 45]. Recently, price extrema are found to be linked with peaks in the volume time series [13]. Moreover, it is reported that massive data sources, such as Google Trends and Wikipedia, contain early signs of market moves. The argument is that these "big data" capture investors' attempts to gather information before decisions are made [35, 46, 47]. These researches provide insight into the price dynamics.

What is the dominating mechanism of the price dynamics is highly complicated. The problem how volatilities affect the price dynamics has drawn much attention. Although many efforts have been made, it remains enormously challenging. According to a hypothesis known as the volatility feedback effect, an anticipated increase in volatility would raise the required return in the future. To allow for higher future returns, the current stock price decreases [24, 25]. Based on this hypothesis, various models, such as Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) model [48] and Exponential GARCH (EGARCH) model [49] have been applied to examine the correlation between past volatilities and future returns, and the results are controversial. The correlation is discovered to be positive in some researches [24, 25], while negative in others [49–51]. Often the coefficient linking past volatilities to future returns is statistically insignificant [26]. On the other hand, the volatility-return correlation function can be used to characterize the correlation between past volatilities and future returns. If the hypothesis of the volatility feedback effect is valid, the volatility-return correlation function should be non-zero. However, it typically fluctuates around zero [4, 9]. Such a volatility-return correlation function can only characterize the correlation local in time. In fact, the scenario in financial markets may be more complicated. Interactions, and thus correlations could be nonlocal in time.

In this study, we construct a class of dynamic observables nonlocal in time to explore the volatility-return correlation, based on the empirical data of hundreds of individual stocks in the New York and Shanghai stock exchanges, as well as 25 stock market indices in different countries. Strikingly, the correlation is discovered to be non-zero, with an amplitude of a few percent and a duration over two weeks. This result provides compelling evidence that past volatilities nonlocal in time affect future returns. Further, we introduce an agent-based model with a novel mechanism, that is, the asymmetric trading preference in volatile and stable markets, to understand the microscopic origin of the volatility-return correlation nonlocal in time.

## Materials

We collect the daily closing prices of 200 individual stocks in the New York Stock Exchange (NYSE), 200 individual stocks in the Shanghai Stock Exchange (SSE) and 25 stock market indices in different countries. The time periods of the individual stocks and stock market indices are presented in Table 1. All these data are obtained from Yahoo! Finance (finance.yahoo.com). To keep the time periods for all stocks exactly the same and as long as possible, we select 200 stocks in the SSE, most of which are large-cap stocks. For comparison, 200 stocks in the NYSE are collected.

**Table 1. The time period, effective pair of time windows and maximum $AP_0$.** The time period, effective pair of $T_1$ and $T_2$, and maximum $AP_0$ for the individual stocks in the NYSE and SSE, as well as 18 stock indices. The volatility-return correlation nonlocal in time is positive for all these indices and stocks, except for the Australia and Japan indices, which exhibit a negative volatility-return correlation. For other 7 indices, nonzero $\Delta P(t)$ could not be detected for almost all pairs of $T_1$ and $T_2$. These indices include MERV (Argentina 1996-2012), BSESN (India 1997-2012), KLSE (Malaysia 1993-2012), KJSE (Indonesia 1997-2011), OMXC20.CO (Denmark 2000-2012), OSEAX (Norway 2001-2012) and FTSE (England 1984-2012), which are not listed in this table.

| Index | Period | Effective $T_1$ | Effective $T_2$ | max $AP_0$ |
|---|---|---|---|---|
| 200 stocks in the NYSE | 1990-2006 | 6-36 | 45-250 | 0.006 |
| 200 stocks in the SSE | 1997-2007 | 4-44 | 95-250 | 0.032 |
| BVSP (Brazil) | 1993-2012 | 26-44 | 60-105 | 0.027 |
| GSPTSE (Canada) | 1977-2012 | 19-40 | 190-240 | 0.029 |
| IPSA (Chile) | 2003-2012 | 31-44 | 190-220 | 0.032 |
| Shanghai Index (China) | 1990-2009 | 27-44 | 80-105 | 0.036 |
| S&P 500 (America) | 1950-2011 | 29-36 | 185-225 | 0.012 |
| DAX (German) | 1959-2009 | 28-44 | 85-250 | 0.015 |
| KOSPI (Korea) | 1997-2012 | 26-44 | 90-115 | 0.027 |
| MXX (Mexico) | 1991-2012 | 6-19 | 65-140 | 0.029 |
| NZ50 (New Zealand) | 2004-2012 | 27-32 | 100-120 | 0.033 |
| IBEX (Spanish) | 1993-2012 | 11-25 | 105-225 | 0.026 |
| OMX (Sweden) | 1998-2012 | 7-17 | 55-160 | 0.043 |
| SSMI (Switzerland) | 1990-2012 | 27-44 | 80-110 | 0.023 |
| FCHI (France) | 1990-2012 | 18-19 | 135-145 | 0.021 |
| AEX (Holland) | 1982-2012 | 21-23 | 50-65 | 0.023 |
| Shenzhen (China) | 1991-2009 | 3-31 | 90-250 | 0.045 |
| DJI (America) | 1928-2011 | 39-41 | 200-225 | 0.006 |
| AORD (Australia) | 1984-2012 | 11-44 | 45-250 | -0.032 |
| N225 (Japan) | 1984-2011 | 34-44 | 200-250 | -0.017 |

# Methods and Results

## Asymmetric conditional probability in volatile and stable markets

To explore the volatility-return correlation in stock markets, we construct a class of observables, including conditional probabilities and correlation functions. We first discuss the conditional probabilities.

The price of a financial index or individual stock at time $t'$ is denoted by $Y(t')$, and the logarithmic return is defined as $R(t') \equiv \ln Y(t') - \ln Y(t'-1)$. For comparison of different indices or stocks, we introduce the normalized return

$$r(t') \equiv [R(t') - \langle R(t') \rangle]/\sigma. \tag{1}$$

Here $\langle \cdots \rangle$ represents the average over time $t'$. In other words, $\langle R(t') \rangle = [\sum_{i=1}^{n} R(i)]/n$ is the average of the time series $R(t')$, where $n$ denotes the total number of the data points of $R(t')$, and $\sigma = [\langle R^2 \rangle - \langle R \rangle^2]^{1/2}$ is the standard deviation of $R(t')$. There may be various definitions of volatility, a simplified one is

$$v(t') = |r(t')|, \tag{2}$$

which measures the magnitude of the price fluctuation.

One may compute temporal correlation functions to investigate the dynamic correlations. The usual volatility-return correlation function is defined as $f(t) = \langle v(t') \cdot r(t'+t) \rangle$ with $t > 0$, and it characterizes

how the volatility at $t'$ influences the return at $t' + t$. However, this correlation function fluctuates around zero [4,9]. It is noteworthy that such a kind of $f(t)$ is local in time, while interactions such as information exchanges in financial markets may be more complicated, leading to correlations nonlocal in time.

To explore the correlations nonlocal in time, we first define an average volatility at $t'$ over a past period of time $T$,

$$\langle v(t') \rangle_T = \frac{1}{T} \sum_{i=1}^{T} v(t' - i + 1). \tag{3}$$

To evaluate whether the average fluctuation in a short time period $T_1$ is strong or weak, we compare it with a background fluctuation, which is defined over a much longer period of time $T_2$ in the past. Therefore, we introduce the difference of the average volatilities in two different time windows,

$$\Delta v(t') = \langle v(t') \rangle_{T_1} - \langle v(t') \rangle_{T_2}, \tag{4}$$

with $T_2 \gg T_1$. $T_1$ and $T_2$ are called the short window and long window, respectively. When $\Delta v(t') > 0$, the stock market in the time window $T_1$ is volatile; otherwise, it is relatively stable.

Next, we compute the conditional probability $P^+(t)|_{\Delta v(t')>0}$, which is the probability of $r(t' + t) > 0$ on the condition of $\Delta v(t') > 0$. Here we consider only $t > 0$. Correspondingly, the conditional probability $P^+(t)|_{\Delta v(t')<0}$ is the probability of $r(t' + t) > 0$ for $\Delta v(t') < 0$. We do not observe any $r(t' + t)$ equal to 0 in the normalized return series. Thus, the conditional probability of $r(t' + t) < 0$ is $1 - P^+(t)|_{\Delta v(t')>0}$ and $1 - P^+(t)|_{\Delta v(t')<0}$, respectively. In a time series of returns, the total number of positive returns is generally different from that of negative ones. Let us denote the unconditional probability that the return is positive by $P_0(t)$, which is the percentage of the positive returns in all returns without any condition.

The specific calculations for $P^+(t)|_{\Delta v(t')>0}$, $P^+(t)|_{\Delta v(t')<0}$ and $P_0(t)$ are described in S1 Appendix. If past volatilities and future returns do not correlate with each other, both $P^+(t)|_{\Delta v(t')>0}$ and $P^+(t)|_{\Delta v(t')<0}$ should be equal to $P_0(t)$. In other words, if $P^+(t)|_{\Delta v(t')>0}$ and $P^+(t)|_{\Delta v(t')<0}$ are different from $P_0(t)$, i.e., if the conditional probability of returns is asymmetric in volatile and stable markets, there exists a non-zero volatility-return correlation and such a correlation is nonlocal in time. In this case, it can be proven that if $P^+(t)|_{\Delta v(t')>0} > P_0(t)$, we have $P^+(t)|_{\Delta v(t')<0} < P_0(t)$, otherwise, we have $P^+(t)|_{\Delta v(t')<0} > P_0(t)$ (see S1 Appendix). To describe the asymmetric conditional probability in volatile and stable markets, we introduce

$$\Delta P(t) = P^+(t)|_{\Delta v(t')>0} - P^+(t)|_{\Delta v(t')<0}. \tag{5}$$

It is important that the probability difference $\Delta P(t)$ relies on $\Delta v(t')$, thereby depending on the time windows $T_1$ and $T_2$. Even though the volatility-return correlation function local in time is zero, the nonlocal observable $\Delta P(t)$ can be non-zero. We call a pair of $T_1$ and $T_2$ at which $\Delta P(t)$ is non-zero an effective pair.

At the time windows $T_1 = 24$ and $T_2 = 205$, for instance, we compute $\Delta P(t)$ for 200 stocks in the NYSE and take an average over these stocks. As displayed in Fig. 1(a), the average $\Delta P(t)$ remains positive for over 20 days with an amplitude of 1 percent. The result indicates that the past volatilities nonlocal in time enhance the positive returns in the future. For comparison, three curves for $\Delta P(t)$ averaged over 150, 100 and 50 randomly chosen stocks are also displayed. Within fluctuations, these three curves are consistent with that for $\Delta P(t)$ averaged over 200 stocks. We take the average over many stocks for the purpose of exploring the collective behavior of stocks. For a single stock, the price dynamics is much more complicated, and $\Delta P(t)$ fluctuates more strongly. Then we perform the same computation for 200 stocks in the SSE at the time windows $T_1 = 10$ and $T_2 = 210$. As displayed in Fig. 1(b), $\Delta P(t)$ remains positive for about 40 days and the amplitude is about 5 percent. Compared with the results for the NYSE, the amplitude and duration of $\Delta P(t)$ for the SSE are respectively much larger and longer. The reason may be that the US stock market is highly developed, with large market size and diversified investment philosophies, while the Chinese stock market is emerging and of small market size, in which the investment philosophies of investors resemble each other.
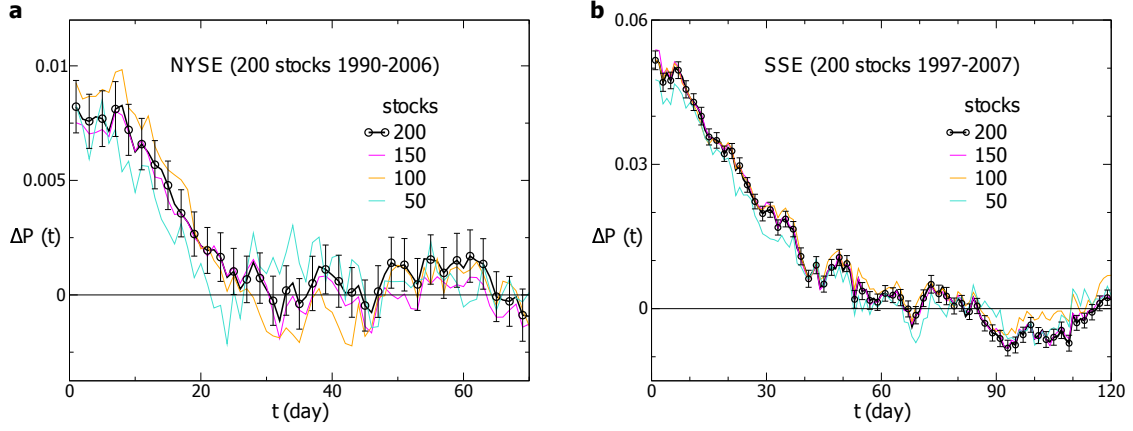
**Figure 1. The probability difference for the individual stocks**. The probability difference $\Delta P(t)$ for (**a**) 200 individual stocks in the SSE and (**b**) 200 individual stocks in the NYSE. The black line shows $\Delta P(t)$ averaged over 200 stocks with error bars. The other lines represent $\Delta P(t)$ averaged over 150, 100, and 50 randomly chosen stocks. The time windows are $T_1 = 24$ and $T_2 = 205$ for the NYSE, and $T_1 = 10$ and $T_2 = 210$ for the SSE.

For the validation of our methods, each point of $\Delta P(t)$ in Fig. 1 is analyzed by performing Student's $t$-test. In general, a $p$-value less than 0.01 is considered statistically significant. For the NYSE, the smallest $p$-value is in the order of $10^{-12}$, and all the $p$-values for $1 \leqslant t \leqslant 19$ are less than 0.01. For the SSE, the $p$-values are even smaller, and less than 0.01 for $1 \leqslant t \leqslant 52$.

Actually the definition of volatility in Eq. (2) is a simplified one. A more standard definition of volatility at $t'$ is

$$v_1 \left(t'\right) = \left[\frac{1}{m} \sum_{i=1}^{m} r^2 \left(t' - i + 1\right)\right]^{1/2}, \tag{6}$$

where $m$ represents a relatively small time window, which may be set to be 5 days, i.e., the number of the trading days in a week. Given that these two definitions $v(t')$ and $v_1(t')$ may lead to different results in extreme volatility regimes, we consider both of them in our calculations. For $v_1\left(t'\right)$, the average volatility at $t'$ over a past period of time $T$ is $\langle v_1(t')\rangle_T = [\sum_{i=1}^{T-m+1} v_1\left(t' - i + 1\right)]/\left(T - m + 1\right)$, with $T \geqslant m$. Thus, the difference of the average volatilities in two different time windows is $\Delta v_1\left(t'\right) = \langle v_1\left(t'\right)\rangle_{T_1} - \langle v_1\left(t'\right)\rangle_{T_2}$.

For further comparison, one may also define the average volatility at $t'$ over a past period of time $T$ as $\langle v_2(t')\rangle_T = [1/T \cdot \sum_{i=1}^{T} r^2(t' - i + 1)]^{1/2}$. Thus the difference of the average volatilities in two different time windows is $\Delta v_2\left(t'\right) = \langle v_2\left(t'\right)\rangle_{T_1} - \langle v_2\left(t'\right)\rangle_{T_2}$. For $\Delta v_1\left(t'\right)$ and $\Delta v_2\left(t'\right)$ respectively, the probability difference is

$$\Delta P_1(t) = P^+(t)|_{\Delta v_1(t')>0} - P^+(t)|_{\Delta v_1(t')<0} \tag{7}$$

and

$$\Delta P_2(t) = P^+(t)|_{\Delta v_2(t')>0} - P^+(t)|_{\Delta v_2(t')<0}. \tag{8}$$

For the NYSE and SSE respectively, we compute $\Delta P_1(t)$ and $\Delta P_2(t)$, and take an average over individual stocks. The time windows are the same as those for $\Delta P(t)$ in Fig. 1. As displayed in Fig. 2, the curves for $\Delta P(t)$, $\Delta P_1(t)$ and $\Delta P_2(t)$ overlap each other within fluctuations. In the following calculations, we mainly consider $\Delta P(t)$ and $\Delta P_1(t)$.
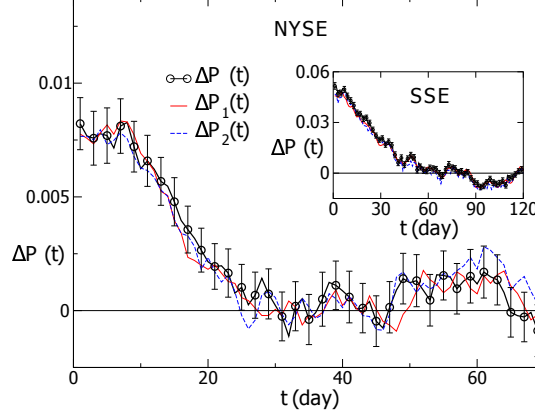
**Figure 2. The probability differences for three definitions of volatility.** The results are averaged over individual stocks. For the NYSE, the time windows are $T_1 = 24$ and $T_2 = 205$. The result for the SSE is displayed in the inset, with the time windows $T_1 = 10$ and $T_2 = 210$.

## Effective pairs of $T_1$ and $T_2$

In the calculation of $\Delta P(t)$, the time windows $T_1$ and $T_2$ are crucial. $T_1$ represents the recent period of time, and investors measure the current fluctuation of prices according to the volatility averaged over $T_1$. Thus, $T_1$ should be relatively small. In our calculations, $T_1$ ranges from 1 to 44 days. Here 44 is the number of trading days in two months. $T_2$ stands for the period of time in which one estimates the background of volatilities in the past. Theoretically, $T_2$ should be much larger than $T_1$. On the other hand, $T_2$ should not be arbitrarily large either: firstly, the memory of investors may not last very long; secondly, maybe more importantly, $T_2$ reflects the long-term fluctuation of stock markets, which should be reasonably fixed. In our calculations, $T_2$ ranges from 45 to 250 days. Here 250 is the number of trading days in a year. In fact, $T_2$ is more crucial to $\Delta P(t)$ than $T_1$. If $T_2$ were equal to the total length of the volatility series, $\langle v(t') \rangle_{T_2}$ would be a constant, and $\Delta P(t)$ would become a local observable, which is just a volatility-return correlation function local in time but averaged over a $T_1$-day moving window.

In Fig. 1(a) and (b), we display $\Delta P(t)$ computed with a specific effective pair of $T_1$ and $T_2$. Actually, the effective pair of $T_1$ and $T_2$ is not unique, and exists in a particular region. Therefore, we compute $\Delta P(t)$ with each pair of $T_1$ and $T_2$, and identify the effective pairs at which $\Delta P(t)$ is significantly non-zero. Since $\Delta P(t)$ needs to be computed in a large region of $T_1$ and $T_2$, it is inefficient to observe the behavior of $\Delta P(t)$ by eyes. Besides, due to the fluctuation of $\Delta P(t)$, the visual observation could be difficult in some cases. Therefore, we propose technical criteria to efficiently discriminate the non-zero $\Delta P(t)$.

The schematic diagram of the criteria is displayed in Fig. 3. The criteria comprise four steps:
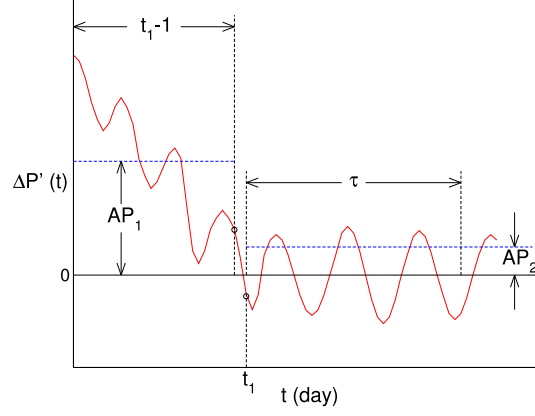
**Figure 3. A schematic graph of the criteria for <mark>identifying non-zero $\Delta P$</mark>.** The red line represents $\Delta P'(t)$, which is the 3-point smoothed $\Delta P(t)$. $t_1$ is the day when the sign of $\Delta P'(t)$ changes for the first time. We define $\Delta P'(t)$ in the range of $1 \leq t \leq t_1 - 1$ as the first part, and that in the range of $t_1 \leq t \leq t_1 + \tau - 1$ as the second part. The average absolute values for the first and second parts are denoted by $AP_1$ and $AP_2$, respectively.

(1) $\Delta P(t)$ is smoothed with a 3-day moving window and the result is denoted by $\Delta P'(t)$.

(2) Supposing $\Delta P'(t)$ changes sign for the first time at $t_1$, we define $\Delta P'(t)$ in the range of $1 \leq t \leq t_1 - 1$ as the first part, and that in the range of $t_1 \leq t \leq t_1 - 1 + \tau$ as the second part. A non-zero $\Delta P'(t)$ would remain positive or negative in the first part, while fluctuate around zero in the second part. We set $\tau$ to be 44, i.e., the number of the trading days in two months, which is long enough to confirm whether the second part of $\Delta P'(t)$ fluctuates around zero.

(3) we calculate the average absolute values for the first and second parts of $\Delta P'(t)$, denoted by $AP_1$ and $AP_2$ respectively.

(4) For a non-zero $\Delta P(t)$, it has to be satisfied that (i) $\Delta P'(t) > AP_2$ for $1 \leqslant t \leqslant t_0$, and $t_0 > 10$; (ii) each value of $\Delta P'(t)$ in the second part is smaller than $AP_1$. With these conditions, we sift out the non-zero $\Delta P(t)$ preliminarily. To measure how significantly $\Delta P(t)$ differs from zero, we calculate the average value of $\Delta P'(t)$ for $1 \leqslant t \leqslant t_0$, which is denoted by $AP_0$. Actually, the larger $|AP_0|$ is, the more significantly $\Delta P(t)$ differs from zero. The average of non-zero $AP_0$ over different pairs of $T_1$ and $T_2$ is denoted by $\overline{AP_0}$. To consolidate our results, we identify those non-zero $\Delta P(t)$, which meet an additional requirement: (iii) $|AP_0| > |\overline{AP_0}|$. $AP_0$ is set to 0 unless $\Delta P(t)$ satisfies all the requirements above.

Now we compute $\Delta P(t)$ for the individual stocks in the NYSE with each pair of $T_1$ and $T_2$. $\Delta P(t)$ is averaged over 200 stocks, and the corresponding $AP_0$ is calculated. The landscape of $AP_0$ is displayed in Fig. 4(a). The result indicates that the effective pairs of $T_1$ and $T_2$ do exist in a particular region, and both $T_1$ and $T_2$ are characteristics of the stock markets. In Fig. 4(a), the effective pairs of $T_1$ and $T_2$ are basically adjacent to each other, suggesting that $\Delta P(t)$ locally is not very sensitive to $T_1$ and $T_2$. This is somehow expected, since $\Delta P(t)$ is computed from the volatilities averaged over $T_1$ and $T_2$, and a little alteration in $T_1$ or $T_2$ would not dramatically change $\Delta P(t)$. From this perspective, the gaps between the disconnected regions in Fig. 4(a) probably result from the fluctuations, especially taking into account the relatively small amplitude of non-zero $\Delta P(t)$ for the NYSE.
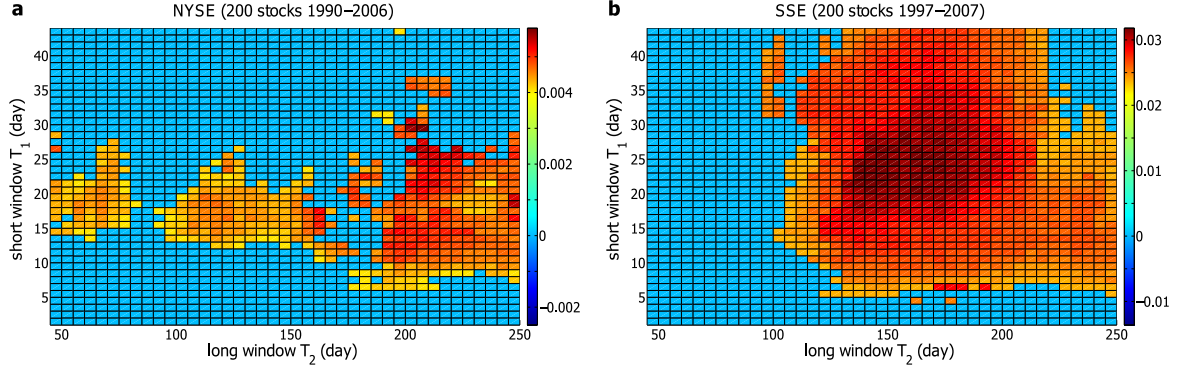
**Figure 4. The landscape for the amplitude of $\Delta P(t)$.** The amplitude $AP_0$ of $\Delta P(t)$ at different time windows $T_1$ and $T_2$ for (**a**) 200 individual stocks in the NYSE and (**b**) 200 individual stocks in the SSE. $T_1$ ranges from 1 day to 44 days, and the increment is 1 day. $T_2$ is from 45 to 250 days, with an increment of 5 days. The larger $AP_0$ is, the more significantly $\Delta P(t)$ differs from zero. For $\Delta P(t)$ fluctuating around zero, $AP_0 = 0$.

Next, we perform a parallel analysis on the 200 stocks in the SSE, and the landscape for the amplitude of $\Delta P(t)$ is displayed in Fig. 4(b). Similar with the result for the NYSE, a large region of non-zero $\Delta P(t)$ is observed for the SSE. At a single pair of $T_1$ and $T_2$, $\Delta P(t)$ averaged over 200 stocks would generally be non-zero, if $\Delta P(t)$ of some stocks is non-zero. Moreover, the region of non-zero $\Delta P(t)$ varies from one stock to another. Therefore, the average $\Delta P(t)$ of the individual stocks is non-zero in a relatively large region for both the NYSE and SSE. Additionally, as displayed in Fig. 4(b), there exists only one connected region of non-zero $\Delta P(t)$ for the SSE, with the amplitude dwindling from the center to the edge. Compared with the result for the NYSE in Fig. 4(a), the region of non-zero $AP_0$ in Fig. 4(b) is broader, without gaps, and the value of $AP_0$ is almost an order of magnitude larger. The reason may be traced back to the fact that the Chinese stock market is emerging, and less efficient than the US stock market. To further validate our methods, we perform Student's $t$-test on each point of non-zero $\Delta P(t)$ in Fig. 4. A $p$-value less than 0.01 is considered statistically significant. At an effective pair of $T_1$ and $T_2$, $\Delta P(t)$ is confirmed to be non-zero, if all the $p$-values are less than 0.01 for $1 \leqslant t \leqslant 10$. All non-zero $\Delta P(t)$ are confirmed except for a few ones at very small $T_1$.

We also compute $\Delta P_1(t)$ with different pairs of $T_1$ and $T_2$ for the NYSE and SSE. Since $m$ in Eq. (6) is set to be 5, $T_1$ should not be smaller than 5. The landscapes for the amplitude of $\Delta P_1(t)$ are almost the same as those for the amplitude of $\Delta P(t)$.

Further, we compute $\Delta P(t)$ for the 25 stock market indices in different countries. The volatility-return correlation is positive for 16 indices, and the corresponding effective pairs of $T_1$ and $T_2$, as well as the maximum $AP_0$, are given in Table 1. For most of these indices, the maximum $AP_0$ is over 2 percent, indicating that the correlation is rather prominent. In Fig. 5(a), we display the regions of effective pairs of $T_1$ and $T_2$ for 5 representative indices including the Brazil, Shanghai, Mexico, Spanish and S&P 500 indices. For other 7 indices, nonzero $\Delta P(t)$ could not be detected for almost all pairs of $T_1$ and $T_2$. Exceptionally, the Australia and Japan indices exhibit a negative volatility-return correlation, i.e., the volatilities in a past period of time enhance the negative returns in future times. The effective pairs of $T_1$ and $T_2$, as well as the maximum $AP_0$ for these two indices, are also presented in Table 1. We also compute $\Delta P_1(t)$ for the 5 representative indices, and the regions of effective pairs of $T_1$ and $T_2$ are shown in Fig. 5(b). Compared with Fig. 5(a), the regions of the effective pairs of $T_1$ and $T_2$ in Fig. 5(b) change slightly. The reason may be that the fluctuation of $\Delta P(t)$ and $\Delta P_1(t)$ for indices is stronger than that for the individual stocks.
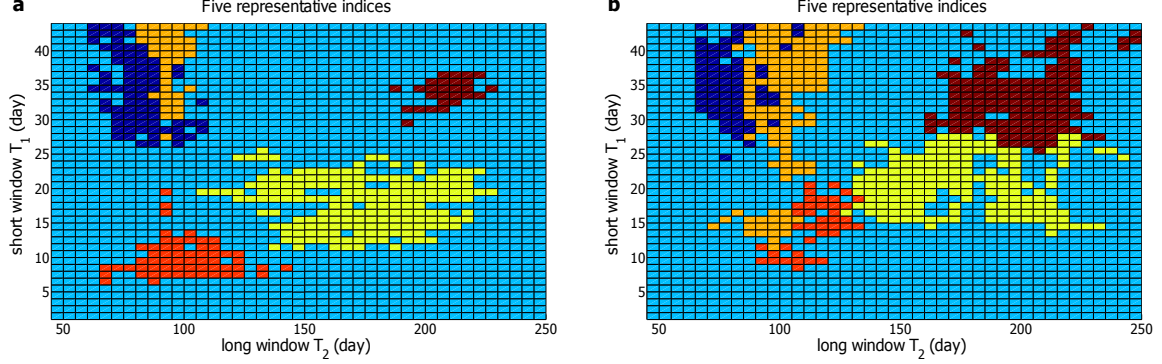
**Figure 5. The effective pairs of time windows for five representative indices.** The probability difference (**a**) $\Delta P(t)$ and (**b**) $\Delta P_1(t)$ for five stock market indices. Different colors represent the regions of effective pairs of $T_1$ and $T_2$ for different indices. Specifically, navy stands for the Brazil Index, orange stands for the Shanghai Index, red stands for the Mexico Index, yellow stands for the Spanish Index and crimson stands for the S&P 500 Index. For clarity, we display only one color at the overlapping regions, given that these regions are small. Some scattered points are also omitted.

To confirm that the nonlocal volatility-return correlation is indeed a nontrivial dynamic property of the stock markets, we randomly shuffle the time series of returns, i.e., randomize the time order of the returns, and perform the same calculation. In this case, $\Delta P(t)$ just fluctuates around zero. The result provides evidence that the correlation does originate from the interactions between past volatilities and future returns.

## Volatility-return correlation functions nonlocal in time

Up to now, we have only concerned with the signs of $\Delta v(t')$ and $r(t'+t)$ in computing $\Delta P(t)$. Actually, the magnitudes of $\Delta v(t')$ and $r(t'+t)$ should also be important to both theoretical analysis and practical applications. Taking into account the magnitudes of $\Delta v(t')$ and $r(t'+t)$, we may explicitly construct a correlation function nonlocal in time to describe the volatility-return correlations,

$$F(t) = \langle \Delta v(t') \cdot r(t'+t) \rangle. \tag{9}$$

Both $\Delta P(t)$ and $F(t)$ reflect the asymmetric behavior of $r(t'+t)$ in volatile and stable markets, but $\Delta P(t)$ should be more fundamental. When $\Delta P(t)$ is non-zero, $F(t)$ would be zero only if the contributions of $r(t'+t) > 0$ and $r(t'+t) < 0$ happen to cancel each other.

We compute $F(t)$ with different pairs of $T_1$ and $T_2$ for the 200 stocks in the NYSE and SSE respectively, and identify the non-zero ones with the same criteria for the non-zero $\Delta P(t)$. We also introduce $AF_0$ to describe how significantly $F(t)$ differs from zero, of which the definition is the same as $AP_0$ for $\Delta P(t)$. $F(t)$ is averaged over 200 stocks, and the landscape of the corresponding $AF_0$ is displayed in Fig. 6. The dynamic behavior of $F(t)$ is qualitatively the same as that of $\Delta P(t)$ but quantitatively different. Both the amplitude of $F(t)$ and the region of effective pairs of $T_1$ and $T_2$ are smaller than those of $\Delta P(t)$. The fluctuation of $F(t)$ is also somewhat stronger. Student's $t$-test is performed on the non-zero $F(t)$ and almost all of them are confirmed to be non-zero. We also compute $F_1(t)$, which is defined as $F_1(t) = \langle \Delta v_1(t') \cdot r(t'+t) \rangle$, with each pair of $T_1$ and $T_2$ for the NYSE and SSE, and the results are almost the same as those for $F(t)$.
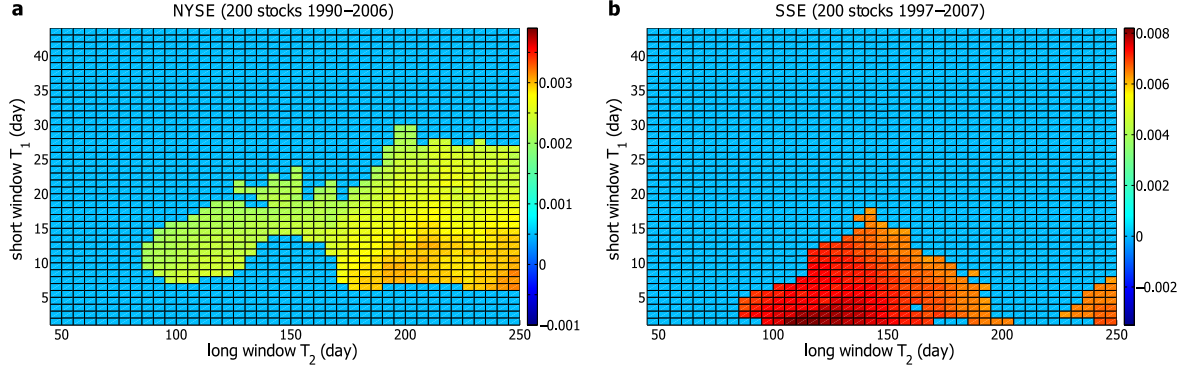
**Figure 6. The landscape for the amplitude of $F(t)$.** The amplitude $AF_0$ of $F(t)$ at different time windows $T_1$ and $T_2$ for (**a**) 200 individual stocks in the NYSE and (**b**) 200 individual stocks in the SSE.

In fact, $\Delta P(t)$ can be expressed as the correlation function $G(t) = \langle sgn(\Delta v(t')) \cdot sgn(r(t'+t)) \rangle$. Here $sgn(x)$ represents the sign of $x$. $G(t)$ behaves almost the same way as $\Delta P(t)$ does. Additionally, one may also define another volatility-return correlation function $H(t) = \langle sgn(\Delta v(t')) \cdot r(t'+t) \rangle$. Since only the magnitude of $r(t'+t)$ is taken into consideration, $H(t)$ is less fluctuating than $F(t)$, whereas the result looks qualitatively similar.

There have been many researches with different methods focusing on how volatilities affect returns in financial markets. A direct way is to calculate the usual volatility-return correlation function, which is defined as $f(t) = \langle v(t') \cdot r(t'+t) \rangle$ with $t > 0$. However, the result fluctuates around zero [4,9]. In the past years, various GARCH-like models are applied to investigate the correlation between past volatilities and future returns. In these models, the future returns are assumed to be correlated with the past volatilities, and there are coefficients quantifying the correlation. The results are controversial. The correlation is discovered to be positive in some researches [24, 25], but negative in others [49–51]. More often, the coefficient linking past volatilities and future returns is statistically insignificant [26]. From our perspective, these studies only characterize the volatility-return correlation local in time. In our work, however, both $\Delta P(t)$ and $F(t)$ are nonlocal in time, which are constructed based on the difference between the average volatilities in two different time windows. The correlation characterized by $\Delta P(t)$ and $F(t)$ is more complicated and of higher-order.

## Agent-based model with asymmetric trading preference

We construct an agent-based model to investigate the microscopic origin of the nonlocal volatility-return correlation. Agent-based modeling is a promising approach in complex systems, and has been applied successfully to study the fundamental properties in financial markets, such as the fat-tail distribution of returns, the long-range temporal correlation of volatilities, and the leverage and anti-leverage effects [15, 18, 27, 39, 52–57].

The basic structure of our model is borrowed from the models in refs. [15, 18], which is built on agents' daily trading, i.e., buying, selling and holding stocks. Since the information for investors is highly incomplete, an agent's decision of buying, selling or holding is assumed to be random. Due to the lack of persistent intraday trading in the empirical trading data, we consider that only one trading decision is made by each agent in a single day. In our model, there are $N$ agents and each agent only operates one share of stock each day. On day $t$, we denote the trading decision of agent $i$ by

$$S_i(t) = \begin{cases} 1 & \text{buy} \\ -1 & \text{sell} \\ 0 & \text{hold} \end{cases}. \tag{10}$$

The probability of buying, selling and holding decisions are denoted by $P_{buy}$, $P_{sell}$ and $P_{hold}$, respectively. Assuming that the price is determined by the difference between the demand and supply of the stock, we define the return $R(t)$ as

$$R(t) = \sum_{i=1}^{N} S_i(t). \tag{11}$$

Next, we introduce the investment horizon based on the fact that investors make decisions according to the previous market performance of different time horizons. It is found that the relative portion $\gamma_i$ of investors with $i$ days investment horizon follows a power-law decay $\gamma_i \propto i^{-\eta}$ with $\eta = 1.12$. With the condition of $\sum_{i=1}^{M} \gamma_i = 1$, $\gamma_i$ is normalized to be $\gamma_i = i^{-\eta} / \sum_{i=1}^{M} i^{-\eta}$, where $M$ is the maximum investment horizon. Considering different investment horizons of various agents, we introduce a weighted average return $R'(t)$ to describe the integrated investment basis of all agents. Specifically, $R'(t)$ is defined as

$$R'(t) = k \sum_{i=1}^{M} \left[ \gamma_i \sum_{j=0}^{i-1} R(t-j) \right], \tag{12}$$

where $k$ is a proportional coefficient. We set $k = 1/(\sum_{i=1}^{M} \sum_{j=i}^{M} \gamma_j)$, so that $|R'(t)|_{max} = N = |R(t)|_{max}$. According to ref. [18], the maximum investment horizon $M$ is set to 150.

Herding behavior is an important collective behavior in financial markets. We define a herding degree $D(t)$ to describe the clustering degree of the herding behavior,

$$D(t+1) = |R'(t)|/N. \tag{13}$$

On day $t+1$, the average number of agents in each group is $N \cdot D(t+1)$, and therefore we divide all agents into $1/D(t+1)$ groups. The agents in a same group make a same trading decision with the same trading probability. In ref. [15], it is assumed that the probabilities of buying and selling are equal, i.e., $P_{buy} = P_{sell} = p$, and $p$ is a constant estimated to be 0.0154. Therefore the trading probability is $P_{trade} = P_{buy} + P_{sell} = 2p$ and the holding probability is $P_{hold} = 1 - 2p$. In our model, the trading probability is also kept to be $2p$ and remains constant during the dynamic evolution.

Now we introduce a novel mechanism in our model, that is, the asymmetric trading preference in volatile and stable markets. In financial markets, the market behaviors of buying and selling are not always in balance [58]. Hence, $P_{buy}$ and $P_{sell}$ are not always equal to each other. They are affected by previous volatilities, and the more volatile the market is, the more $P_{buy}$ differs from $P_{sell}$.

For an agent with $i$ days investment horizon, the average volatility over previous $i$ days is taken into account, which is defined as

$$V_i(t) = \frac{1}{i} \sum_{j=1}^{i} V(t-j+1). \tag{14}$$

Then we define the background volatility as $V_M(t)$, where $M$ is the maximum investment horizon. On day $t$, the agent with $i$ days investment horizon estimates the volatility of the market by comparing $V_i(t)$ with $V_M(t)$. Therefore, the integrated perspective of all agents on the recent market volatility is defined as

$$\xi(t) = \left[ \sum_{i=1}^{M} \gamma_i V_i(t) \right] / V_M(t). \tag{15}$$

Thus, we define the probabilities of buying and selling as

$$\begin{cases} P_{buy}(t+1) = p[c \cdot \xi(t) + (1-c)] \\ P_{sell}(t+1) = 2p - P_{buy}(t+1) \end{cases}. \tag{16}$$

Here the parameter $c$ measures the degree of agents' asymmetric trading preference in volatile and stable markets. Compared with the model in ref. [18], $c$ is the only new parameter added in our model. We speculate that $c$ can be determined from the trade and quote data of stock markets. Unfortunately, the data are currently not available to us.

To judge from the amplitude of the volatility-return correlation, $c$ should be a small number. Let us set $c$ to be 1/80. The total number of the agents, $N$, is 10000. The returns of the initial 150 time steps are set to be random values following a standard Gaussian distribution. On day $t$, we randomly divide $N$ agents into $1/D(t)$ groups. The agents in a same group make a same trading decision with the same probability. After each agent makes his decision, the return $R(t)$ can be computed. Repeating the procedure we produce 20000 data points of $R(t)$ in each simulation, and abandon the first 15000 data points for equilibration. Thus we obtain a sample with 5000 data points.

After the time series $R(t)$ generated from our model is normalized to $r(t)$, we compute $\Delta P(t)$ with the time windows $T_1 = 3$ and $T_2 = 150$. The result is averaged over 100 samples and displayed in Fig. 7(a). $\Delta P(t)$ is significantly non-zero with an amplitude of 3 percent, lasting for about 20 days. For comparison, three curves for $\Delta P(t)$ averaged over 75, 50 and 25 randomly chosen samples are also displayed. Within fluctuations, these four curves are consistent with each other and in agreement with the empirical results.
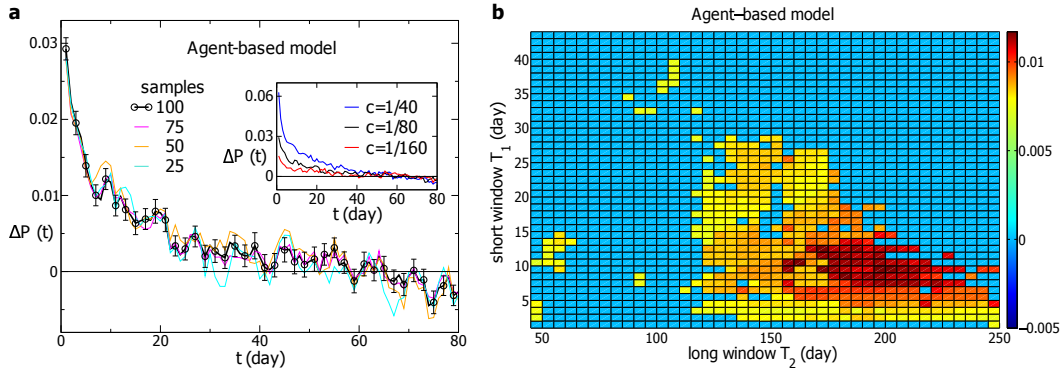


**Figure 7. The simulation results of the agent-based model.** (**a**) The probability difference $\Delta P(t)$ computed with the simulated returns at the time windows $T_1 = 3$ and $T_2 = 150$. The parameter $c$ is 1/80. The black line represents $\Delta P(t)$ averaged over 100 samples with error bars and the other lines stand for $\Delta P(t)$ averaged over 75, 50 and 25 randomly chosen samples. $\Delta P(t)$ for different values of $c$ are displayed in the inset, with $T_1 = 3$ and $T_2 = 150$. (**b**) The amplitude $AP_0$ of $\Delta P(t)$ at different time windows $T_1$ and $T_2$. Each $\Delta P(t)$ is averaged over 100 samples. The larger $AP_0$ is, the more significantly $\Delta P(t)$ differs from zero. For $\Delta P(t)$ fluctuating around zero, $AP_0 = 0$.

We also perform the simulation with $c = 1/40$ and $c = 1/160$, respectively, to investigate the dependence of $\Delta P(t)$ on $c$. As displayed in the inset of Fig. 7(a), the amplitude of $\Delta P(t)$ increases with $c$, i.e., the magnitude of $c$ determines the amplitude of the volatility-return correlation. For $c = 1/40$, the amplitude of $\Delta P(t)$ is about 6 percent, which is in the order of that for the SSE and other markets with a strong volatility-return correlation. For $c = 1/160$, the amplitude of $\Delta P(t)$ is close to that for the S&P500 index, of which the volatility-return correlation is relatively weak. Therefore, with $c$ ranging from 1/160 to 1/40, our model produces the volatility-return correlation consistent with the empirical results. Additionally, if $c$ is negative, the volatility-return correlation will be negative, i.e., the sign of $c$ fixes the correlation to be positive or negative.

Next, we compute $\Delta P(t)$ with different pairs of $T_1$ and $T_2$, and determine the region of effective pairs of $T_1$ and $T_2$. $\Delta P(t)$ is averaged over 100 samples, and the landscape of $AP_0$ is shown in Fig. 7(b). A single region with non-zero $\Delta P(t)$ is observed. For $T_2$ smaller than 120, for example, $\Delta P(t)$ is almost

zero. In other words, the effective pairs of $T_1$ and $T_2$ exist in a particular region, which is consistent with the empirical results.

# Discussion

We construct a class of dynamic observables nonlocal in time to explore the correlation between past volatilities and future returns in stock markets. Strikingly, the volatility-return correlation is discovered to be non-zero, with an amplitude of a few percent and a duration of over two weeks. The result indicates that past volatilities nonlocal in time affect future returns. Both the nonlocal dynamic observables $\Delta P(t)$ and $F(t)$ rely on two time windows $T_1$ and $T_2$. The effective pairs of $T_1$ and $T_2$ exist in a particular region, suggesting that both $T_1$ and $T_2$ are the characteristics of the stock markets.

Our results are robust for not only individual stocks but also stock market indices. The volatility-return correlation nonlocal in time is detected to be positive for individual stocks in the New York and Shanghai stock exchanges, as well as 16 stock indices. For other 7 indices, $\Delta P(t)$ fluctuates around zero. However, we suppose there may exist some higher-order correlations between volatilities and returns for these indices, which could be described by more complicated nonlocal observables. Exceptionally, other 2 indices exhibit a negative volatility-return correlation.

To investigate the microscopic origin of the volatility-return correlation, we construct an agent-based model with a novel mechanism, that is, the asymmetric trading preference in volatile and stable markets. Accordingly, a parameter $c$ is introduced to describe the degree of the asymmetric trading preference. The simulation results exhibit a positive correlation which is in agreement with the empirical ones. More importantly, the effective pairs of $T_1$ and $T_2$ for simulation results exist in a particular region, which is also consistent with the empirical ones. Actually, our model can also produce a negative correlation by changing the sign of $c$. The results reveal that both the positive and negative correlations arise from the asymmetric trading preference in volatile and stable markets. In our model, the nonlocality arises from the interaction between the integrated perspective on the recent market volatility and the probabilities of buying and selling.

Our results provide new insight into the price dynamics. Contrary to the assumptions in various models, the rise and fall of prices turn out to be far from random. To the best of our knowledge, the volatility-return correlation nonlocal in time is the only property concerning the control of the price dynamics, given that the autocorrelating time of returns is extremely short. This non-zero volatility-return correlation implies that there may exist higher-order correlations of returns, which deserves further investigation in the future, especially for those 7 indices with $\Delta P(t)$ fluctuating around zero. Furthermore, our results indicate that nonlocality is an intrinsic characteristic in the financial markets, which is more important than we thought before. Besides the volatility-return correlation in the stock markets, many other nonlocal correlations in financial systems are to be explored, which serves as our future agenda.

# Supporting Information

## S1 Appendix

**Calculation for $P^+(t)|_{\Delta v(t')>0}$, $P^+(t)|_{\Delta v(t')<0}$ and $P_0(t)$, and relation among them.**

## Acknowledgments

## References

1. Mantegna RN, Stanley HE. Scaling behavior in the dynamics of an economic index. Nature. 1995;376: 46.

2. Plerou V, Gopikrishnan P, Rosenow B, Amaral LAN, Stanley HE. Universal and nonuniversal properties of cross correlations in financial time series. Phys Rev Lett. 1999;83: 1471.

3. Gopikrishnan P, Plerou V, Amaral LAN, Meyer M, Stanley HE. Scaling of the distribution of fluctuations of financial market indices. Phys Rev E. 1999;60: 5305.

4. Bouchaud JP, Matacz A, Potters M. Leverage effect in financial markets: The retarded volatility model. Phys Rev Lett. 2001;87: 228701.

5. Krawiecki A, Hołyst JA, Helbing D. Volatility clustering and scaling for financial time series due to attractor bubbling. Phys Rev Lett. 2002;89: 158701.

6. Gabaix X, Gopikrishnan P, Plerou V, Stanley HE. A theory of power-law distributions in financial market fluctuations. Nature. 2003;423: 267.

7. Onnela JP, Chakraborti A, Kaski K, Kertesz J, Kanto A. Dynamics of market correlations: taxonomy and portfolio analysis. Phys Rev E. 2003;68: 056110.

8. Sornette D. Critical market crashes. Phys Rep. 2003;378: 1–98.

9. Qiu T, Zheng B, Ren F, Trimper S. Return-volatility correlation in financial dynamics. Phys Rev E. 2006;73: 065103.

10. Podobnik B, Horvatic D, Petersen AM, Stanley HE. Cross-correlations between volume change and price change. Proc Natl Acad Sci USA. 2009;106: 22079.

11. Shen J, Zheng B. Cross-correlation in financial dynamics. Europhys Lett. 2009;86: 48005.

12. Kenett DY, Shapira Y, Madi A, Bransburg-Zabary S, Gur-Gershgoren G, Ben-Jacob E. Index cohesive force analysis reveals that the us market became prone to systemic collapses since 2002. PLoS One. 2011;6: e19378.

13. Preis T, Schneider JJ, Stanley HE. Switching processes in financial markets. Proc Natl Acad Sci USA. 2011;108: 7674–7678.

14. Shapira Y, Kenett DY, Raviv O, Ben-Jacob E. Hidden temporal order unveiled in stock market volatility variance. AIP Advances. 2011;1: 022127.

15. Feng L, Li BW, Podobnik B, Preis T, Stanley HE. Linking agent-based models and stochastic models of financial markets. Proc Natl Acad Sci USA. 2012;109: 8388–8393.

16. Kenett DY, Preis T, Gur-Gershgoren G, Ben-Jacob E. Quantifying meta-correlations in financial markets. Europhys Lett. 2012;99: 38001.

17. Preis T, Kenett DY, Stanley HE, Helbing D, Ben-Jacob E. Quantifying the behavior of stock correlations under market stress. Sci Rep. 2012;2: 752.

18. Chen JJ, Zheng B, Tan L. Agent-based model with asymmetric trading and herding for complex financial systems. PLoS One. 2013;8: e79531.

19. Kenett DY, Ben-Jacob E, Stanley HE, Gur-Gershgoren G. How high frequency trading affects a market index. Sci Rep. 2013;3: 2110.

20. Jiang XF, Chen TT, Zheng B. Structure of local interactions in complex financial dynamics. Sci Rep. 2014;4: 5321.

21. Majdandzic A, Podobnik B, Buldyrev SV, Kenett DY, Havlin S, Stanley HE. Spontaneous recovery in dynamical networks. Nat Phys. 2014;10: 34–38.

22. Yura Y, Takayasu H, Sornette D, Takayasu M. Financial brownian particle in the layered order-book fluid and fluctuation-dissipation relations. Phys Rev Lett. 2014;112: 098703.

23. Black F. Studies of stock price volatility changes. Alexandria: Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economical Statistics Section. 1976;177–181.

24. French KR, Schwert GW, Stambaugh RF. Expected stock returns and volatility. J financ econ. 1987;19: 3–29.

25. Campbell JY, Hentschel L. No news is good news: An asymmetric model of changing volatility in stock returns. J financ econ. 1992;31: 281–318.

26. Bekaert G, Wu G. Asymmetric volatility and risk in equity markets. Rev Financ Stud. 2000;13: 1–42.

27. Cont R, Bouchaud JP. Herd behavior and aggregate fluctuations in financial markets. Macroeconomic Dyn. 2000;4: 170.

28. Yamasaki K, Muchnik L, Havlin S, Bunde A, Stanley HE. Scaling and memory in volatility return intervals in financial markets. Proc Natl Acad Sci USA. 2005;102: 9424–9428.

29. Bollerslev T, Litvinova J, Tauchen G. Leverage and volatility feedback effects in highfrequency data. J financ econ. 2006;4: 353–384.

30. Osipov GV, Kurths J, Zhou C. Synchronization in oscillatory networks. 1st ed. Berlin: Springer; 2007.

31. Shapira Y, Kenett DY, Ben-Jacob E. The index cohesive effect on stock market correlations. Eur Phys J B. 2009;72: 657–669.

32. Kenett DY, Shapira Y, Madi A, Bransburg-Zabary S, Gur-Gershgoren G, Ben-Jacob E. Dynamics of stock market correlations. AUCO Czech Economic Review. 2010;4: 330–341.

33. Kenett DY, Tumminello M, Madi A, Gur-Gershgoren G, Mantegna RN, Ben-Jacob E. Dominating clasp of the financial sector revealed by partial correlation analysis of the stock market. PLoS One. 2010;5: e15032.

34. Ren F, Zhou WX. Recurrence interval analysis of high-frequency financial returns and its application to risk estimation. New J Phys. 2010;12: 075030.

35. Da Z, Engelberg J, Gao PJ. In search of attention. J Finance. 2011;66: 1461–1499.

36. Zhao L, Yang G, Wang W, Chen Y, Huang JP, Ohashi H, et al. Herd behavior in a complex adaptive system. Proc Natl Acad Sci USA. 2011;108: 15058 –15063.

37. Jiang XF, Zheng B. Anti-correlation and subsector structure in financial systems. Europhys Lett. 2012;97: 48006.

38. Liu Y, Gopikrishnan P, Cizeau P, Meyer M, Peng CK, Stanley HE. Statistical properties of the volatility of price fluctuations. Phys Rev E. 1999;60: 1390.

39. Eguiluz VM, Zimmermann MG. Transmission of information and herd behavior: An application to financial markets. Phys Rev Lett. 2000;85: 5659.

40. Cox JC, Ross SA. The valuation of options for alternative stochastic processes. J financ econ. 1976;3: 145.

41. Shen J, Zheng B. On return-volatility correlation in financial dynamics. Europhys Lett. 2009;88: 28003.

42. Campbell JY, Shiller RJ. The dividend-price ratio and expectations of future dividends and discount factors. Rev Financ Stud. 1988;1: 195–228.

43. Fama EF, French KR. Dividend yields and expected stock returns. J financ econ. 1988;22: 3–25.

44. Valkanov R. Long-horizon regressions: theoretical results and applications. J financ econ. 2003;68: 201–232.

45. Boudoukh J, Michaely R, Richardson M, Roberts MR. On the importance of measuring payout yield: Implications for empirical asset pricing. J Finance. 2007;62: 877–915.

46. Moat HS, Curme C, Avakian A, Kenett DY, Stanley HE, Preis T. Quantifying wikipedia usage patterns before stock market moves. Sci Rep. 2013;3: 1801.

47. Preis T, Moat HS, Stanley HE. Quantifying trading behavior in financial markets using google trends. Sci Rep. 2013;3: 1684.

48. Bollerslev T. Generalized autoregressive conditional heteroskedasticity. J Econom. 1986;31: 307–327.

49. Nelson DB. Conditional heteroskedasticity in asset returns: A new approach. Econometrica. 1991;59: 347–370.

50. Turner CM, Startz R, Nelson CR. A markov model of heteroskedasticity, risk, and learning in the stock market. J financ econ. 1989;25: 3–22.

51. Glosten LR, Jaganathan R, Runkle DE. On the relation between the expected value and the volatility of the nominal excess return on stocks. J Finance. 1993;48: 1779–1801.

52. Lux T, Marchesi M. Scaling and criticality in a stochastic multi-agent model of a financial market. Nature. 1999;397: 498.

53. Bonabeau E. Agent-based modeling: Methods and techniques for simulating human systems. Proc Natl Acad Sci USA. 2002;99: 7280–7287.

54. Hommes CH. Modeling the stylized facts in finance through simple nonlinear adaptive systems. Proc Natl Acad Sci USA. 2002;99: 7221–7228.

55. Samanidou E, Zschischang E, Stauffer D, Lux T. Agent-based models of financial markets. Rep Prog Phys. 2007;70: 409.

56. Farmer JD, Foley D. The economy needs agent-based modelling. Nature. 2009;460: 685–686.

57. Sornette D. Physics and financial economics (1776–2014): puzzles, ising and agent-based models. Rep Prog Phys. 2014;77: 062001.

58. Plerou V, Gopikrishnan P, Stanley HE. Econophysics: Two-phase behaviour of financial markets. Nature. 2003;421: 130.