Research Article

Crosscorrelation Analysis between P2P Lending Market and Stock Market in China

Bin Wang, Zhonghui Ding, Xiang Wang, and Kai Shi

Correspondence should be addressed to Kai Shi; mail_shikai@foxmail.com

Received 31 March 2020; Revised 21 May 2020; Accepted 22 May 2020; Published 24 June 2020

Academic Editor: Emilio Gómez-Déniz

Copyright © 2020 Bin Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We first use the Hurst index and the $V_n$ statistic to study the respective characteristics of the P2P lending market and the stock market by rescaled range analysis. In terms of fluctuations, there is an antipersistence in the P2P lending market and long-term persistence in the stock market. Then, we studied the crosscorrelation between the daily logarithmic return series of P2P lending market and stock market and found that the crosscorrelation was multifractal and antipersistent.

1. Introduction

P2P (Peer to Peer) lending refers to the process of matching financing and investment by borrowers and lenders without the participation of traditional financial intermediaries. It is a new way of borrowing. The world’s first P2P lending company, ZOPA, was established in 2005, marking the birth of the P2P lending market. Subsequently, the United States, Germany, Japan, and other developed countries have successively established P2P lending platforms. Recently, P2P lending market has become a hot topic of concern in China. The rise of the P2P lending market has, on the one hand, solved the financing difficulties and tedious procedures, saving a lot of transaction costs. On the other hand, it also stimulates the enthusiasm of entrepreneurs and investors. In order to learn more about the many changes that this emerging market has brought to society, more and more scholars have begun to study this emerging market from different aspects such as the success rate of loan, the risk, the relations to other markets, and so on. From the aspect of success rate of loan, Herzenstein et al. [1] studied the loan data of the American P2P lending company Prosper and found the borrower’s credit situation, personal background characteristics, and the description of the loan application have effect on the success rate. And Pope and Sydnor [2] found the evidence of obvious ethnic differences in the United States. Black people have less possibility of obtaining funds and higher interest rates under the premise of obtaining loans. Duarte et al. [3] obtained that borrowers who appear more trustworthy have higher success probabilities. Besides the personal characteristics affected the success, social networks also did. Freedman and Jin [4] found that social relationship information has varying degrees of influence on the success rate in the USA. Lin et al. [5] analyzed the P2P lending platform, Prosper.com, and found that the online friendships of borrowers would affect the probability of successful funding, lower interest rates on funded loans. Barasinska and Schäfer [6] used a multiple regression analysis model to analyze the German P2P lending platform Samava and found that gender had no effect on the success rate of the P2P lending market which is different from traditional banking. From the viewpoint of risk, Emekter et al. [7] considered factors such as the success rate and risk of lending and provided reliable suggestions for the sustainable development of the P2P lending market. Shi [8] analyzed the influencing factors of lending yield and the volume of transaction by studying the existing P2P lending models in China and put forward suggestions on how to properly avoid and control the risk of P2P lending. Hu and Song [9] obtained that investors would show the pursuit of returns and the avoidance of risks, and when both returns and risks exist, there is an interest rate level with the
strongest investor preference. Jiang and Li [10] built the P2P network platform credit risk evaluation model by BP neural network and obtained that it is capable of using P2P network platform on personal credit risk evaluation scientifically. Li et al. [11] modeled and analyzed the systemic risk in P2P lending market in the perspective of complex network. Shi [8] deeply analyzed the influencing factors and risk control of our P2P lending market and obtained that the government should the establishment of a regulatory system gradually to resolve the impulsive need for different financial institutions to make profits using P2P lending. Lee and Lee [12] obtained that there is a lender herding behavior in P2P lending market in Korea. While from the perspective of the relation to other markets, Huo [13] analyzed that P2P and lending market in Korea. While from the perspective of the stock market have similar investment and financing relation to other markets, Huo [13] found that there is a “seesaw” effect between P2P lending market and other markets. This paper uses the multifractal detrending crosscorrelation analysis (MF-DCCA) method to analyze the crosscorrelation between the return rate of the P2P lending market and the return rate of the stock market, hoping to provide decision-making basis for China’s P2P lending market.

2. Methodology

2.1. Rescaled Range Analysis. Rescaled range analysis (R/S analysis) is a statistical technique designed to assess the nature of variability in data over time. It has been used to detect and evaluate the amount of persistence and randomness in financial market time series data. The basic principle is as follows:

(1) For a given time series \( x_t, t = 1, \ldots, N \), where \( N \) is the length of the time series, divide it into multiple adjacent subintervals of length \( n \), which means the number of such intervals is \( A = N/n \). Denote its subintervals as \( I_a \), \( a = 1, \ldots, A \). And denote the elements in \( I_a \) as \( x_{ak} \), \( k = 1, \ldots, n \) and \( a = 1, \ldots, A \). The mean of the \( a \)th subinterval is \( E_a = \sum_{k=1}^{n} x_{ak} \).

(2) For each \( I_a \), define \( Y_{ak} = \sum_{i=1}^{k} (x_{ai} - E_a) \), and let \( R_{Ia} = \max [Y_{ak}] - \min [Y_{ak}], 1 \leq k \leq n \). Its standard deviation is \( S_{Ia} = \left( \sum_{k=1}^{n} (x_{ak} - E_a)^2 / n \right)^{1/2} \).

(3) The rescaled range for each range in the time series is defined as

\[
\left( \frac{R_{Ia}}{S_{Ia}} \right)_{n} = \left( \frac{1}{A} \sum_{a=1}^{A} \frac{R_{Ia}}{S_{Ia}} \right)
\]

(1)

(4) Repeat the above procedure until \( n = N \). In this way, we get \( \{(R/S)_{n}\}_{n=1}^{N} \).

(5) Calculate the slope of the data to find the Hurst exponent \( H \). Here, \( H \) is the slope of the plot of each range’s \( \ln (R/S) \) versus each range’s \( \ln n \).

The Hurst index \( H \) is one of the indicators used to determine whether a sequence has random fluctuations. There are three general forms. (1) When \( H \in (0, 0.5) \) indicates that the sequence has antipersistence, a large value in a variable may be followed by a small value and a small value in a variable may be followed by a larger amount. The closer \( H \) is to 0, the stronger the antipersistence. (2) When \( H = 0.5 \) indicates that the sequence is a random sequence, there is no crosscorrelation between the sequence samples. (3) When \( H \in (0.5, 1) \) represents a long-term persistent relationship, there is a certain regularity in the change of sequence samples. The large value in one variable may be followed by the large value in another variable. The closer \( H \) is to 1, the more persistent and stable the trend.

This paper further tests the stability of \( R/S \) through the extended \( V_n \) statistic of Peters [18]. When the \( R/S \) statistic is noisy, the statistic \( V_n \) is more excellent in judging the cycle and estimated length. The formula for calculating statistics is \( V_n = (R/S)_n / \sqrt{n} \). If the time series is random, that is, \( H = 0.5 \), the curve of \( V_n = f (\ln n) \) should be a horizontal
line; if antipersistent, that is, $H \in (0, 0.5)$, it should slope downward; if persistent, that is, $H \in (0.5, 1)$, it should slope upward. When the curve shows a sharp turn, it indicates that the influence of the historical state on the future state disappears. The corresponding time length $n$ at this time is the average cycle length of the sequence.

2.2. Multifractal Detrended Crosscorrelation Analysis. Fractal methods are widely applied in economics and finance, such as the detrending crosscorrelation analysis and the detrending moving averages, see [19–29] and the references therein. We will use the method of multifractal detrending crosscorrelation analysis (MF-DCCA), which is proposed by Podobnik and Stanley [27] and improved by Zhou [29], to calculate the crosscorrelation index to quantify the fractal features of the crosscorrelation between logarithmic returns of the P2P lending market and the stock market. We take modulus of the detrended covariance function in order to get rid of its negative signs. See [15–17, 30] etc., for such treatment.

Assume that there are two time series $x(t)$ and $y(t)$, $t = 1, 2, \ldots, N$, where $N$ is the length of the time series.

First, construct two new time series by finding the cumulative difference between the two sets of series for the mean:

$$
\bar{X}(t) = \sum_{k=1}^{t} (x(k) - \bar{x}),
$$

$$
\bar{Y}(t) = \sum_{k=1}^{t} (y(k) - \bar{y}),
$$

where $\bar{x} = (1/N) \sum_{i=1}^{N} x(i)$ and $\bar{y} = (1/N) \sum_{i=1}^{N} y(i)$.

Then, divide the constructed two new time series into subsequences of equal length that do not overlap each other, where each subsequence contains $s$ data. That is, the number of nonoverlapping segments is $N_s = \lfloor N/s \rfloor$. Because the total length $N$ of the original sequence may not be an integral multiple of the length $s$ of the subsequence, there will be residual data at the end of the original sequence. In order not to discard these data, the procedure is also repeated for the reverse sequence of the original sequence. In this way, $2N_s$ segments can be obtained.

Thirdly, fit each segment by the least squares method to get the local trends $X'(t)$ and $Y'(t)$, $t = 1, \ldots, s$ and $v = 1, \ldots, 2N_s$. Here, $X'(t)$ and $Y'(t)$ are the fitting polynomials with order $m$ in segment $v$ (in this paper, we take $m = 3$). Then, the detrended covariance is determined by

$$
F^2(s, v) = \frac{1}{s} \sum_{t=1}^{s} \left| \bar{X}((v-1)s + t) - X'(t) \right| \cdot \left| \bar{Y}((v-1)s + t) - Y'(t) \right|.
$$

for $v = 1, \ldots, N_s$; while for $v = N_s + 1, \ldots, 2N_s$,

$$
F^2(s, v) = \frac{1}{s} \sum_{t=1}^{s} \left| \bar{X}((v-N_s)s + t) - X'(t) \right| \cdot \left| \bar{Y}((v-N_s)s + t) - Y'(t) \right|.
$$

Fourthly, define the $q$th-order ($q \neq 0$) the fluctuation function as follows:

$$
F_q(s) = \left\{ \frac{1}{2N_s} \sum_{v=1}^{2N_s} \left[ F^2(s, v) \right]^{q/2} \right\}^{1/q}.
$$

If $q = 0$, then

$$
F_0(s) = \exp \left\{ \frac{1}{4N_s} \sum_{v=1}^{2N_s} \ln \left[ F^2(s, v) \right] \right\}.
$$

At last, analyze the scaling behavior of the fluctuations by observing log-log plots $F_q(s)$ vs. $s$ for each fixed $q$. If the two series are long-range crosscorrelated, then we can obtain a power-law expression $F_q(s) \sim s^{-\nu(q)}$. That is,

$$
\ln F_q(s) = H_{xy}(q) \cdot \ln s + C.
$$

It is well known that $H_{xy}(q)$ is the generalized cross-correlation exponent, called generalized Hurst exponent, and represents the degree of the crosscorrelation between the two time series. When $H_{xy}(q)$ depends on $q$, the crosscorrelation of two time series is multifractal, otherwise it is monofractal. When $H_{xy}(q) = 0.5$, there are no crosscorrelations between the two series. When $H_{xy}(q) > 0.5$, the crosscorrelations are long-range persistent; while when $H_{xy}(q) < 0.5$, the crosscorrelations are antipersistent. Especially, the bivariate Hurst exponent $H_{xy}(2)$ has similar properties and interpretation as a univariate Hurst exponent. If $H_{xy}(2) > 0.5$, then the crosscorrelations between two time series are long-range persistent, otherwise crosscorrelations are antipersistent. And $\Delta H_{xy} = \max H_{xy} - \min H_{xy}$ is used to measure the time-varying degree of multifractality.

3. Empirical Results

3.1. Sample Selection and Descriptive Statistics. Due to that the data missed so much on the P2P lending market data before January 1, 2015, the P2P lending market is based on Interest-rate Working-day Composite Index (according to the definition from the webpage of Shenzhen Qiancheng Internet Finance Research Institute, the Interest-rate Working-day Composite Index comprehensively characterizes the following variables of P2P lending market in China: the interest, the transaction volume, the number of P2P platforms, and various incentive measures from the P2P platforms) published by Shenzhen Qiancheng Internet Finance Research Institute (http://www.qqcs.com.cn/index/finance/dsj.htm) from January 1, 2015, to January 17, 2020 (a total of 1249 working days). While the Composite Index is usually used to characterize the stock market and
the method MF-DCCA, we used requires two time series data to be synchronized. Therefore, we utilize the daily data of the Shenzhen Composite Index from (http://www.szse.cn/market/trend/index.html?code=399106) January 1, 2015, to January 17, 2020, as the research object of the stock market. For the missing values of unequal time in the two markets, we have separately established ARMA models to complement them. These two indices are very typical in China for each market. Let \( P_t \) be the composite index of each market at day \( t \). To eliminate the possible heteroscedasticity of time series, we take the daily logarithmic return \( r_t \), which is \( r_t = \log P_t - \log P_{t-1} \). The daily logarithmic returns of these two markets are illustrated in Figure 1 and the descriptive statistics are presented in Table 1.

It is shown that the standard deviation of the logarithmic return of the P2P lending market is larger than that of stock market. The skewness of the research object shows that the returns of the P2P lending market and the stock market both show a leftward bias. The kurtosis coefficient of the logarithmic return of the P2P lending market is 22.7437 and the kurtosis coefficient of the stock market is 7.2573, indicating that both are sharp peaks and fat tails which are common in financial data. The results of the Jarque-Bera statistical test show that the logarithmic return series are stationary series.

3.2. R/S Analysis of P2P Lending Market and Stock Market.

The Hurst index is calculated and analyzed for the logarithmic return data of the P2P lending market and the stock market by R/S analysis, which is shown in Table 2.

We can see that the value of the Hurst index in both markets is not 0.5, indicating that both sequences are not random. The Hurst index of the P2P lending market is 0.37, which belongs to \((0, 0.5)\), indicating that the P2P lending market has antipersistence, that is, if the P2P lending market’s index was up last period, then it is more likely to be down next period. The Hurst index of the stock market is 0.60 > 0.5, which indicates that the composite index of the stock market has a long-term persistent relationship.

Furthermore, from the relationship between \( V_n \) and \( \ln n \) in Figure 2, we can see that the P2P lending market on the left shows a downward trend, indicating that the sequence is antipersistent. In recent years, in order to rectify the chaos in P2P lending, the Chinese government has implemented strict policy controls on the P2P lending market. As a result, many unqualified platforms and nonprofit platforms have closed down successively, and the P2P lending market has shown a downward trend. And \( V_n \) statistics are reaching a regional extreme at the point \( \ln n = 3.3672 \), corresponding to \( n = 28 \), so the aperiodic cycle of the P2P lending market is about 28 days. The stock market on the right shows an overall upward trend, indicating that the stock market has long-term persistence; \( V_n \) statistic reached a regional extreme at the point \( \ln n = 2.9444 \) which means \( n = 18 \), indicating the aperiodic cycle of the stock market is about 18 days.

3.3. Crosscorrelation Test of P2P Lending Market and Stock Market.

In order to perform a certain preliminary analysis on the two time series and verify whether there is a crosscorrelation between the two time series, we adopt the qualitative method proposed by Podobnik et al. [31]. For the two time series, \( x(t) \) and \( y(t), t = 1, 2, \ldots , N \), where \( N \) is the length of the time series, set the test statistic

\[
Q_{cc}(m) = N^2 \sum_{i=1}^{m} \frac{c_i^2}{N-i},
\]

where \( c_i \) is the crosscorrelation function which is defined as

\[
c_i = \frac{\sum_{k=1}^{N} x(k)y(k-i)}{\sqrt{\left( \sum_{k=1}^{N} x^2(k) \right) \left( \sum_{k=1}^{N} y^2(k) \right)}}.
\]

The crosscorrelation statistic \( Q_{cc}(m) \) is approximately distributed \( \chi^2(m) \) with the degree of freedom \( m \). It can be used to test the null hypothesis that none of the first \( m \) crosscorrelation coefficient is different from zero. When \( Q_{cc}(m) \) is different from the critical value, it indicates that there is a crosscorrelation between the two time series significantly at a certain significance level. Moreover, if \( Q_{cc}(m) \) is larger than the critical values of \( \chi^2(m) \), we claim that there are not only crosscorrelations but also long-range crosscorrelations, see [31].

The crosscorrelation relationship between the logarithmic return series of the two markets was verified according to the construction of the crosscorrelation statistics described above. Figure 3 shows the statistics of the crosscorrelation coefficient of the logarithmic returns of the P2P lending market and the stock market and the critical value of the chi-square distribution. According to Figure 3, \( Q_{cc}(m) \) is generally larger than the critical value of the chi-square distribution at the 5% level of significance when \( m \) is larger than 880. Thus, we can overall reject the null hypothesis of no crosscorrelations. In other words, the long-range crosscorrelations existed between the returns of the P2P lending market and the stock market. As the crosscorrelation test based on the statistic \( Q_{cc}(m) \) can only test the existence of crosscorrelation between two time series qualitatively, in order to quantitatively study the crosscorrelations between these return series, we will use MF-DCCA in Section 3.4.

3.4. MF-DCCA.

In this section, we use the method of MF-DCCA to quantify the fractal features of the crosscorrelation between logarithmic return of the P2P lending market and the stock market (it was performed using R. The code is available at https://github.com/skjq/MFDCCA2020). The subinterval length is based on the recommendation of Peng et al. [32] with \( s_{\min} = 5 \) and \( s_{\max} = N/5 \). Peng et al. pointed out that when the value of \( s \) is less than \( s_{\min} = 5 \), the function \( F_q(s) \) will be affected by the sample, resulting in a deterministic component. If the value of \( s \) is greater than \( s_{\max} = N/5 \), \( F_q(s) \) will exhibit scale behavior that does not obey the power law. Therefore, conservatively, we take \( s \) in \([40, 50, 60, \ldots , 200]\) in this paper. At the same time, in order to analyze the small and large fluctuations in the time series, the selection of \([40, 50, 60, \ldots , 200]\) needs to be composed of positive \( q \) and negative \( q \). When \( q \) is positive, \( \mathcal{H}_{xx}(q) \) describes the behavior...
Table 1: Descriptive statistics of the daily logarithmic returns of P2P and stock market.

<table>
<thead>
<tr>
<th>Item</th>
<th>P2P</th>
<th>Stock market</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum</td>
<td>0.1632</td>
<td>0.0274</td>
</tr>
<tr>
<td>Minimum</td>
<td>-0.1719</td>
<td>-0.0374</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.0002</td>
<td>0.00006</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.0181</td>
<td>0.0076</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.5528</td>
<td>-0.9741</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>22.7437</td>
<td>7.2573</td>
</tr>
<tr>
<td>Jarque-Bera statistics**</td>
<td>20350.1700</td>
<td>1141.6440</td>
</tr>
<tr>
<td>ADF statistics**</td>
<td>-19.5900</td>
<td>-32.9750</td>
</tr>
<tr>
<td>Prob.</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Observations</td>
<td>1249</td>
<td>1249</td>
</tr>
</tbody>
</table>

Significance level is 5%.

Table 2: Hurst index for markets.

<table>
<thead>
<tr>
<th>Market</th>
<th>P2P</th>
<th>Stock</th>
</tr>
</thead>
<tbody>
<tr>
<td>(H)</td>
<td>0.37</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Figure 1: The daily logarithmic return of (a) P2P lending market and (b) stock market.

Figure 2: \(V_n\) statistics for (a) P2P lending market and (b) stock market.
of the smaller fluctuations, while when \( q \) takes a negative value, \( H_{xy}(q) \) describes the behavior of the larger fluctuations. In this way, different \( q \) can describe the effect of different fluctuations on \( F_q(s) \). Therefore, the value of \( q \) in this paper ranges from \(-10\) to \(10\), with an interval of 1.

It can be seen from Table 3 or Figure 4 that the crosscorrelation index \( H_{xy}(q) \) decreases with the increase of the order \( q \) from \(-10\) to 10 and is not a constant constant. This can explain that the crosscorrelation between the logarithmic return of the P2P lending market and the stock market has multifractal features. It shows once again that there is a cross-correlation relationship between the two markets and that the crosscorrelation relationship has multifractal features instead of monofractal. The multifractal properties of time series are also studied by multifractal spectrum analysis on multifractal scale index \( \tau(q) \) which is defined as \( \tau(q) = qH_{xy}(q) - 1 \). If \( \tau(q) \) is a nonlinear function of \( q \), the time series has multifractal natures. It can be seen from Figure 6 that \( \tau(q) \) is nonlinear function of \( q \). This again shows that there are multifractal features between the P2P lending market and the stock market.

To further study the crosscorrelation between the two markets, based on the relationship between the fluctuations function \( F_q(s) \) and the time scale \( s \), we obtain Figure 5 for different \( q \) values and find that there is a power law relationship. This shows once again that there is a cross-correlation relationship between the two markets and that the crosscorrelation relationship has multifractal features instead of monofractal. The multifractal properties of time series are also studied by multifractal spectrum analysis on multifractal scale index \( \tau(q) \) which is defined as \( \tau(q) = qH_{xy}(q) - 1 \). If \( \tau(q) \) is a nonlinear function of \( q \), the time series has multifractal natures. It can be seen from Figure 6 that \( \tau(q) \) is nonlinear function of \( q \). This again shows that there are multifractal features between the P2P lending market and the stock market.

In order to test its fluctuation, we analyze the multifractal spectrum. The singularity exponent \( \alpha \) and its multifractal spectrum \( f(\alpha) \) can be obtained through the following transform:

\[
\alpha = H_{xy}(q) + qH'_{xy}(q),
\]

\[
f(\alpha) = q[\alpha - H_{xy}(q)] + 1.
\]  

(10)

If it is multifractal, the graph of \( f(\alpha) \sim \alpha \) is the shape of a single peak bell. The multifractal spectrum width
\[ \Delta \alpha = \alpha_{\text{max}} - \alpha_{\text{min}} \] is usually used to quantify the multifractality degree of time series. The larger the \( \Delta \alpha \), the more uneven the distribution of time series and thus the stronger the multifractality [15]. Figure 7 shows that it is indeed a single peak bell with left hook (i.e., \( \Delta f = f(\alpha_{\text{min}}) - f(\alpha_{\text{max}}) > 0 \)) which means it is of multifractal nature, and the multifractal spectrum width is 0.3079 < 0.5, indicating that the multifractal strength of the crosscorrelation coefficient between P2P lending market and stock market is weak. It also shows that the singular spectral width of the stock market is 0.5740 which is smaller than the singular spectral width of the P2P lending market 1.098, indicating that the multifractal strength of the stock market is less than the multifractal strength of the P2P lending market, which means that the fluctuation of the P2P lending market is greater than the stock market fluctuations.

### 4. Conclusion

This paper first performs a descriptive statistical analysis and stationarity test on the logarithmic returns of the P2P lending market and the stock market. It is concluded that the logarithmic return series of both markets show peaks and fat tails and are stable. Then, the research found that the P2P lending market and the stock market have different fractal characteristics. Compared with the P2P lending market, the stock market has a long-term persistent development. This is also the reason why the development status of the two industries is greatly different. At last, the analysis of the crosscorrelation between the P2P lending market and the stock market confirmed the crosscorrelation relationship between the two markets is nonlinear and has multifractal characteristics. Based on this, in order to drive the development of today’s P2P lending market and to promote financial institution reform and improve the financial system, the Chinese government should take the stock market and P2P lending market as a whole and learn from the management mechanism of the stock market to strengthen effective supervision of P2P lending market. For example, similar to the stock market, the government can establish a unified national regulatory standard, a complete P2P lending market trading system, and implement the qualifications of
margin and shareholder. When investors make investment decisions on P2P, they need to consider the two markets comprehensively and analyze the actual impact of the current fluctuations of the two markets on future returns in order to better achieve the effect of risk aversion.

**Data Availability**

The data used to support the findings of this study are included within the article.

**Conflicts of Interest**

The authors declare that they have no conflicts of interest.

**Authors’ Contributions**

All authors contributed equally to this work.

**Acknowledgments**

We would like to thank Professor Zhou Weixing and Wang Fang for helpful discussion. The work was supported partially by the NSFC (11961015 and 11401127), GXNSF (2018GXNSFAA050031), and startup Grant in Guilin University of Technology.

**References**


