Testing for herding behaviour among energy sectors in Chinese stock exchange

Cuiting Shen

School of Management, Shanghai University, 99 Shangda Road, Baoshan District, Shanghai, 200444, China

1Email: Zoey_8023@hotmail.com

Abstract. Herding behaviour in stock markets leads a group of investors to imitate others and make the same economic decisions as other market participants, causing excess market volatility and price instability. This paper aims to test for the herding behaviour in different energy sectors of Chinese Stock Exchange. Firstly, Generalized Capital Asset Pricing Model (GCAPM) is employed to observe the stocks in energy sectors during financial crisis, which can test for the nonlinear relationship between return of particular portfolio and the average return of market. The empirical results indicate the presence of herding behaviour in energy sectors. Secondly, artificial neural networks are used to predict the herding behaviour in Chinese stock markets. It is advisable that during the periods when the stock market is unstable, investors should keep awake to avoid herding behaviour.

1. Introduction

The concept ‘herding behaviour’, regarded as the mutual simulation resulting in a convergence of action, has received considerable attention in researches examining the grouping behaviour of investors [1]. Such behaviour always appears in extreme financial markets, especially in financial crisis [2]. According to traditional Capital Asset Pricing Model (CAPM), Christie and Huang proposed the cross-sectional standard deviation (CSSD) of stock returns as an indicator to measure the presence of herding behaviour of market participants [3]. Subsequently, Chang put forward the cross-sectional absolute deviation (CSAD) to improve the measurement of herding behaviour [4]. Based on two above-mentioned studies, either developed financial markets, like American, Japan, Eurozone markets [5, 6], or emerging financial markets, such as China, Athens have been researched a lot [7-8].

This paper aims to investigate herding behaviour in energy sectors of Chinese Stock Exchange focusing on the extreme financial conditions happened in 2015. Because energy is core component in developing economy, eradicating poverty and improving living standards. It contributes directly to meeting both basic needs and more sophisticated human needs [9]. We apply Generalized Capital Asset Pricing Model (GCAPM) which is the extension of traditional CAPM to measure herding behaviours. Although, both cross-sectional standard deviation (CSSD) and cross-sectional absolute deviation (CSAD) focus on the dispersion of stock returns across the market, they can only detect the existence of herding behaviour in the entire market, instead of testing which asset portfolio investors decide to follow suit. Compared with CAPM, Dong and Han (2007) proposed a novel model according to the Generalized Capital Asset Pricing Model (GCAPM): if investors have herding behaviour on an asset portfolio, there will be a nonlinear relationship between the expected return rate of the portfolio and the average expected return rate of markets. Through the empirical experiments, it
was showed that 8 of the 10 tested stock sectors (New Energy, Wind Energy, Scarce Resources, Nuclear Power, Stored Energy, Biomass Energy et al.) in Chinese Stock Exchange have significant herding behaviour.

After observing the presence of herding behaviour in energy sectors of Chinese stock exchange, we intend to use artificial back-propagation neural network to predict herding behaviour. According to Behavioural Finance Theory, emotional contagion is thought to be one of the most crucial foundations of herding, causing consistent behaviours in animal groups, which can lead to the sharp rise or fall of stock trading volume. Therefore, in this paper, we take ‘trading volume’ as the representative of herding behaviour to carry out empirical experiments based on artificial neural networks.

The rest of paper is organized as follow: Section 2 presents the methodology; Section 3 demonstrates data employed and empirical outcomes and Section 4 concludes our findings.

2. Methodology

2.1. Generalized Capital Asset Pricing Model (GCAPM)

Christie and Huang (1995) proposed the cross sectional standard deviation (CSSD) to capture herding behaviour among investors. If the stock prices fluctuate sharply, investment behaviour among market participants will be market-oriented. As a result, the return on individual stocks and the return on stock market will converge, employing the cross sectional standard deviation of individual asset returns as follows:

\[ CSSD_t = \sqrt{\frac{\sum_{i=1}^{N}(R_{it} - R_{mt})^2}{N-1}} \]  

(1)

Where \( N \) is the number of all sample stocks in the stock portfolio, \( R_{it} \) is the return of stock \( i \) on day \( t \), and \( R_{mt} \) is the return of stock portfolio \( M \) on day \( t \). The calculation of cross sectional standard deviation (CSSD) apparently underestimates herding behaviour in the stock market, because the convergence of individual stock returns means that there must be serious herding behaviour among investors.

In order to deal with the drawbacks of Christie and Huang (1995), Chang et al. (2000) proposed cross-sectional absolute deviation (CSAD) to improve the sensitivity and accuracy of detecting herding behaviour, which is given by:

\[ CSAD_t = \frac{1}{N} \sum_{i=1}^{N} |R_{it} - R_{mt}| \]  

(2)

To conduct a test for detecting herding activity, previous studies have used the following model based on traditional CAPM model:

\[ CASD_t = \alpha + \gamma_1 |R_{mt}| + \gamma_2 R_{mt}^2 + \varepsilon_t \]  

(3)

Where \( CASD_t \) is the cross-sectional absolute deviation, \( R_{mt} \) is the return on the market index on day \( t \). If the regression coefficient of \( R_{mt}^2 \) is significantly 0, there is no herding behaviour among market investors. And when the regression coefficient of \( R_{mt}^2 \) is significantly negative, it indicates the serious presence of herding behaviour. While when the regression coefficient of \( R_{mt}^2 \) is positive, the herding behaviour in market is not obvious enough.

Nevertheless, whether CSSD or CSAD, both approaches are based on the traditional CAPM model which assumes that all investors will hold the same and effective market portfolio, which is not restrained in Generalized Capital Asset Pricing Model (GCAPM). The GCAPM considered that investors will have a small-scale investment portfolio due to the transaction cost and the indivisibility of investment. The number of stock in their own portfolio is lower than the number of stock in whole market portfolio. So, the average expected return of market is the weighted average of expected returns for each small-scale group. When there is herding behaviour in such small-scale group, the market share of this group will increase or decrease, which will impact the expected return of this portfolio as well as the average expected return of market. There are three important hypothesis in the model proposed by Dong and Han [10]:
Firstly, the assets in market are divided into two types: one risk-free asset and \( N \) risky assets. The market participants also have two styles: one is rational investors who will follow the assumption of CAPM and invest in an efficient market portfolio \( M \). The others are investors with herding behaviour who will only invest the portfolio \( H \) of \( n \) risky assets \((n<N)\). Although portfolio \( H \) does not include all the assets in the market, it is still an effective borderline formed by these \( n \) assets. If there is herding behaviour on portfolio \( H \), the proportion of assets invested in \( H \) will fluctuate with market conditions.

Secondly, it is supposed that the number of investors holding portfolio \( H \) is in a minority, while the number of investors holding portfolio \( M \) is in a majority. We take \( \theta \) as the ratio of investment to portfolio \( H \). That is to say, \((1-\theta)\) is the ratio of investment to portfolio \( M \). As a result, \( \theta \) can be assumed to be a number near zero.

Thirdly, as the ratio of investment, \( \theta \) will decrease or increase since investors will change their own follow-up investment behavior with the change of external factors. Although there are many factors that can affect herding behaviour, it is hypothesized that \( \theta \) is decided by the difference value between the expected return on market \((E(\bar{R}_m))\) and the expected return on portfolio \( H \) \((E(R_h))\). For example, when \((E(R_h) - E(\bar{R}_m))\) increase, the investors can observe this signal and then they will raise their investment in portfolio \( H \). Thus, the relationship between \( \theta \) and \((E(R_h) - E(\bar{R}_m))\) is given by:

\[
\theta = \lambda_0 + \lambda_1[E(R_h) - E(\bar{R}_m)]
\]  

Through the efficient market portfolio \( M \), we know that:

\[
E(R_h) = R_f + \beta_{mh}[E(R_m) - R_f]
\]  

Where \( E(R_h) \) is the expected return on portfolio \( H \), \( E(R_m) \) is the expected return on portfolio \( M \), \( R_f \) is the risk-free return rate, \( \beta_{mh} \) is the coefficient of portfolio \( H \).

The expected return of market is calculated by weighting:

\[
E(\bar{R}_m) = (1 - \theta)E(R_m) + \theta E(R_h)
\]  

To be noticed, in reality, the assets arbitrarily included in portfolio \( H \) may be also included in the effective portfolio \( M \). But the proportion of these assets in \( H \) is different from the proportion in \( M \). Therefore, we can regard portfolio \( H' \) as the same assets between \( H \) and \( M \). The coefficient of \( H' \) is \( \beta_{mh} \). The new portfolio \( G \) is then generated, which contains assets both in \( H \) and \( H' \). The expected return of portfolio \( G \) equals to the weighted average of \( H \) and \( H' \):

\[
E(R_G) = R_f + \frac{(1-\theta)\omega}{(1-\theta)\omega + \theta} \beta_{mh}'[E(R_m) - R_f] + \frac{\theta}{(1-\theta)\omega + \theta} \beta_{mh} \times [E(R_h) - R_f]
\]  

Where \( \omega \) is the proportion of investment \( H' \) in portfolio \( M \). From equation (5) and (7), we can get:

\[
E(R_G) = R_f + \left[\frac{(1-\theta)\omega}{(1-\theta)\omega + \theta}\beta_{mh}' + \frac{\theta}{(1-\theta)\omega + \theta}\beta_{mh}\right] \times [E(R_m) - R_f]
\]  

Then, the derivation of equation (6) is as follow:

\[
E(\bar{R}_m) = [E(R_m) - R_f](1 - \theta + \theta \beta_{mh}) + R_f
\]  

From equation (8) and (9), we can get the equation related to the expected return of \( G \) \((E(R_G))\) and the average expected return of market \((E(\bar{R}_m))\):

\[
E(R_G) = R_f + f(\theta)[E(\bar{R}_m) - R_f]
\]  

\[
f(\theta) = \left[\frac{(1-\theta)\omega}{(1-\theta)\omega + \theta}\beta_{mh}' + \frac{\theta}{(1-\theta)\omega + \theta}\beta_{mh}\right]^{\frac{1}{1-\theta + \theta \beta_{mh}}}
\]  

Where \( f(\theta) \) reflects the relationship between the expected return of \( G \) \((E(R_G))\) and the average expected return of market \((E(\bar{R}_m))\). We can know that there is a nonlinear relationship between \((E(R_G))\) and \((E(\bar{R}_m))\), instead of linear relationship in traditional CAPM.

Since \( \theta \) is very close to zero, \( f(\theta) \) is expanded at zero using the first-order Taylor formula:

\[
f(\theta) = f(0) + f'(0)\theta + o(\theta)
\]  

\[
f(\theta) = \left[\frac{\beta_{mh} - \beta_{mh}'}{\theta} + o(\theta)\right]^{\frac{1}{1-\theta + \theta \beta_{mh}}} + o(\theta)
\]  

From equation (4), (10) and (13), we can get:

\[
E(R_G) = R_f + \left[A_0 + A_1[E(R_h) - E(\bar{R}_m)]\right][E(\bar{R}_m) - R_f]
\]
Since $E(R_h)$ and $E(R_g)$ have high correlation (because portfolio H and G contain the same assets), so we can replace $E(R_h)$ with $E(R_g)$, so the equation (14) can be rewritten as:

$$E(R_g) - R_f = [A_0 + A_1[E(R_g) - E(R_m)]] [E(R_m) - R_f]$$ (15)

It can be seen from the above formula that as long as $A_1$ is significantly different from zero in the empirical analysis, the herding behaviour in the asset portfolio can be proved.

2.2. Artificial neural networks

The traditional models for stock analysis is artificial neural networks (ANNs), which can predict dynamic and non-linear stock prices in a robust way. The outcomes generated by ANNs are mostly the forecast about stock values. However, in this paper, we adopt the neural networks to analysing the herding behaviour of investors.

The artificial neural network used in this paper is Back Propagation (BP) neural network which is a kind of multi-layer feed-forward network. Neurons are arranged in layers and the model structure is divided into three layers: input layer, hidden layer and output layer. The signals propagate through input layers in succession, and generate the output information. The basic idea of BP learning algorithm is to modify the connection weights between nodes so that the global error of the network reaches a minimum value. The process of BP algorithm includes two stages: forward calculation and back propagation.

- **Forward calculation**

  The outcome of hidden layer:

  $$y^k_h = f \left( \sum_{i=1}^{N_1} \omega_i x^k_i + \theta_h \right)$$ (16)

  The outcome of output layer:

  $$z^k_j = f \left( \sum_{h=1}^{N_2} \omega_h y^k_h + \gamma_j \right)$$ (17)

  $$z^*_j = f \left[ \sum_{h=1}^{N_2} \omega_h [f \left( \sum_{i=1}^{N_1} \omega_i x^k_i + \theta_h \right) + \gamma_j] \right]$$ (18)

- **Back propagation**

  Error function:

  $$E = \frac{1}{2} \sum_{k,j}^P \left( T^k_j - z^k_j \right)^2$$ (19)

  Weight adjustment:

  $$\Delta \omega = -\eta \frac{\partial E}{\partial \omega}$$ (20)

  Weight correction:

  $$\omega = \omega + \Delta \omega = \omega - \eta \frac{\partial E}{\partial \omega}$$ (21)

  $X_i$, $Y_h$ relatively denotes the input node and output node of hidden layer. $Z_j$ represents the outcome of output layer which indicates that the target signal is the supervisor signal. The connection weight between input node $i$ and hidden node $h$ is $\omega_{ih}$ while the connection weight between hidden node $h$ and output node $j$ is $\omega_{hj}$. $N_1$, $N_2$, $N_3$ stand for the number of input layer, hidden layer and output layer. $\theta_h$, $\gamma_j$ relatively stands for the threshold of hidden layer node $h$ and output layer node $j$. The transfer function is $f$. $\eta$ is the learning step, and $P$ is the number of samples ($k = 1, 2 ... P$).

In this paper, we take stock trading volume as the output of BP neural network. According to Behavioural Finance, emotional contagion is thought to be one of the most crucial foundations of herding, causing consistent behaviour in animal groups, which can lead to the sharp rise or fall of stock trading volume. Once the stock value rises, the optimism of a few investors quickly begins to spread among traders and then infects other investors. Eventually, majority of investors will be affected to increase their subjective probability so that they seem anxious to buy in. On the contrary, when the stock value is barely satisfactory, pessimistic sentiment of investors will be generated. As a result, market participants may sell stocks in succession. These two kinds of trade (buy or sell) can be intuitively reflected by stock trading volume. Therefore, we can hypothesize that if there is herding behaviour occurred in stock markets, the volume will increase or decrease obviously. As for the input data of BP neural network, opening/closing prices of stocks, media attention, and technical indexes
will be appropriate choices because that there are various factors can lead to herding behaviour of market participants.

3. Empirical results

3.1. The presence of herding behavior
Since the Generalized Capital Asset Pricing Model (GCAPM) in Section 2 is applicable to any portfolio, we choose 10 energy sectors as the experiment objects, that is, we treat stocks of each sector as a specific asset portfolio and test whether the investors have herding behaviour. The data employed in this paper are daily close stock price and market values of energy sections in Chinese stock exchange, obtained from Resset Database. The energy sections include New Energy, Wind Energy, Scarce Resources, Nuclear Power, Stored Energy, Biomass Energy, Fuel Cell, Coal Chemical, Water and Power, Petroleum. Then we take national debt yields as risk-free return rate since the national debts are relatively safe and have steady profitability. The date range for the data used is 05/01/2015 to 31/12/2015, because the stock index in China experienced a sharp rise and fall in 2015, which is good for observing whether investors will generate herding behaviour due to changes in external factors and market conditions.

Table 1. Regression results for 10 energy sectors.

| Sector             | $A_0$  | $A_1$        | DW  |
|--------------------|--------|--------------|-----|
| New Energy         | 1.093  | -3.621       | 1.881 |
| Wind Energy        | 1.043  | -2.842       | 1.718 |
| Scarce Resources   | 1.111  | -5.439       | 2.211 |
| Nuclear Power      | 1.225  | -0.594       | 2.013 |
| Stored Energy      | 1.012  | -6.033       | 1.799 |
| Biomass Energy     | 1.151  | -4.548       | 1.461 |
| Fuel Cell          | 1.173  | -5.079       | 1.751 |
| Coal Chemical      | 1.215  | -3.732       | 1.858 |
| Water and Power    | 1.198  | 10.732       | 2.041 |
| Petroleum          | 0.873  | -9.027       | 2.052 |

Notes: The value in parentheses is t-Statistics.
*** Statistically significant at 1% level.
**  Statistically significant at 5% level.
*    Statistically significant at 10% level.

We select 10 energy sectors as the representatives of energy section in Chinese Stock Exchange. Table 1 indicates the presence of herding behavior occurred in 10 energy sectors. There are totally 8
sectors where market investors mimic others to decide their trading strategies. Among those sectors, Wind Energy is statistically significant at the 1% level, which presents that the herding behavior in this sector is the strongest. Besides, Scarce Resources, Stored Energy, Biomass Energy, Coal Chemical are statistically significant at the 5% level. Thus, the herding behavior in such four sectors have slightly lower strength. What is more, Fuel Cell, Water and Power, Petroleum are statistically significant at the 10% level, which indicates that the herding behavior in these sectors is weak. And there is no herding behavior in New Energy, Nuclear Power through empirical experiments.

From empirical results, we can know that herding behavior of investor is prevalent in the energy market in 2015. One reason for this phenomenon is the increasing attention from the public. After 2000, the global energy investment basically continued to increase except for 2009, which was affected by the Lehman crisis. Investments in the development and production of traditional energy sources such as coal, crude oil and natural gas have nearly tripled in 2014 and renewable energy has nearly quadrupled. With the economic growth in emerging market countries, the demand for energy has expanded and the investment for development has also increased. Take renewable energy (Photovoltaic and Wind Power) as example, their global investment in power generation equipment and technological innovations reached 313 billion U.S. dollars, accounting for nearly 20% of the total in 2015. Another reason why market participants simulated others in energy section is the supports from government policies. For example, the establishment of Shanghai Natural Gas Trading Centre. And constructing a global energy Internet which was proposed in 2015 can promote global demand for electricity in a clean and green manner. By 2050 the global energy Internet cumulative investment will exceed 100 trillion US dollars, as predicted by Chinese government. Other reasons that can cause herding behavior are frequent intervention of government, irregular information disclosure system or irrational investor structure and so on.

3.2. The prediction of herding behavior
Due to the limited space, we only choose Wind Energy as our experiment object in this section. Wind Energy has been proved to be statistically significant at the 1% level, which presents that the herding behaviour in this sector is completely strong. So, next step is to use BP neural network to observe whether there is a sharp rise or fall during the 05/01/2015 to 31/12/2015. The input information of BP model are opening/closing stock prices and 16 technical indicators (including MACD, KDJ, and CCI et al.).

![Figure 1. The training results of BP Neural network.](image1)

![Figure 2. The prediction results after training.](image2)

The figure 1 is the training results of BP neural network while the figure 2 is the prediction results after training. The symbol ‘+’ stands for the monitoring values of training set and the symbol ‘*’ represents the forecasting values of test set. In this neural network, there are two hidden layers which can be seen from figure 3. The training applied is the gradient-descent BP algorithm with adaptive learning rate and Mean Square Error (MSE) is the error indicator.
The predictive values in figure 2 shows that there is the trading volume in Wind Energy multiple staged rises and falls during 05/01/2015 to 31/12/2015 since the market condition in China was quite turbulent. It is advisable that during the periods when the stock market is unstable, investors should keep awake to avoid herding behaviour. We can find that although the predictive values are fairly close to monitoring values in training figure 1, the accuracy of figure 2 is not satisfactory. One of the main reason is that the BP neural networks adopted is original without any optimization algorithms. Although the BP algorithm has clear derivation, strong learning ability and easy to implement, there are still some defects that cannot be neglected to be improved, such as slow learning convergence. Nevertheless, the improvement of BP networks is not the emphasis of our paper, which can be retained to further discussion.

4. Conclusions
This study is in line with the models of Dong and Han (2007), providing further work into the herding behaviour of energy sectors in Chinese market from 05/01/2015 to 31/12/2015, i.e. in New Energy, Wind Energy, Scarce Resources, Nuclear Power, Stored Energy, Biomass Energy, Fuel Cell, Coal Chemical, Water and Power, Petroleum. In order to observe the simulation of market participants in specific stock section, we employ Generalized Capital Asset Pricing Model (GCAPM) to test for the nonlinear relationship between return of particular portfolio and average return of market. Then, we use BP neural network to predict whether there is herding behaviour or not.

From the regression results, it can be found that 8 of the 10 asset portfolios in energy section did not satisfy the linear relationship of the CAPM model. For the nonlinear relationship in CSAD, Chang (2000) did not give a strict mathematical derivation, but only to be logically explained. However, according to GCAPM model, we can easily find that the herding behaviour breaks the linear relationship of CAPM and naturally leads to the non-linear features in the CSAD method. In addition, through BP neural networks, we can find that although the accuracy of BP models was not satisfactory, the predictive value and monitoring value tend to the same tendency. Hence, the improvement of BP networks should have further discussion.

All in all, this paper aims to identify the herding behaviour in different portfolio or stock sections so as to prevent irrational behaviour of market participants. This needs joint efforts from market, government and investors.

References
[1] Chiang T C, Li J, Tan L and Nelling E 2013 Multinatnl. Finance. J. 17 165
[2] Chiang T C and Zheng D 2010 J. Bank. Financ. 34 1911
[3] Christie W G and Huang R D 1995 Financ. Anal. J. 51 31
[4] Chang E C, Cheng J W and Khorana A 2000 J. Bank. Financ. 26 1651
[5] Peltomäki J and Vähämäa E 2015 Finance. Res. Lett. 14 111
[6] Bensaâ Da A, Jlassi M and Litimi H 2015 Am. J. Financ. Acc. 4 50
[7] Economou F, Katsikas E and Vickers G 2016 Finance. Res. Lett. 14 334
[8] Lin T, Chiang T C, Mason J R and Nelling E 2007 Pacific-Basin. Finance. J. 16 61
[9] Batten J A, Kinateder H, Szilagyi P G and Wagner N F 2017 Energy. Econ. 66 559
[10] Zhiyong D. and Xu H. 2007 Finance Research. 5 108