Herding evidence in Chinese stock market: a study of the relationship between stock price index and trading volume based on behavioral finance theory

Xiaofeng ZHAO
x4zhao@ln.edu.hk

Follow this and additional works at: https://commons.ln.edu.hk/ljbfe

Part of the Finance Commons, and the Finance and Financial Management Commons

Recommended Citation

This Article is brought to you for free and open access by the Department of Economics at Digital Commons @ Lingnan University. It has been accepted for inclusion in Lingnan Journal of Banking, Finance and Economics by an authorized editor of Digital Commons @ Lingnan University.

Xiaofeng ZHAO

Abstract
Over the last couple decades, more evidence has been found supporting the notion that investors are not always rational. Herding behaviors have been observed in both the stock market crash\(^1\) and financial bubbles\(^2\), which were beyond the understanding of modern finance theory. In this paper, the herding phenomenon was explored in the Chinese stock market by the study of the relationship between stock prices and trading volume over the past 7 years. It was found that the change of price is statistically significant to have caused the change of trading volume, but the reverse is not true. Theoretically, this identifies persistent herding phenomenon in the Chinese stock market. The findings provide useful investment guidance for investors and new considerations in financial reform for the government.

---

\(^1\)Stock Market Crash of 1929
October 24, 1929, known as the Black Thursday, marked the start of the crash of 1929. This was followed by Black Monday and Tuesday, the very next week. The Dow Jones Industrial Average fell more than 20% just on Monday and Tuesday alone. By mid-November, the index had dropped by around 40% compared to its peak just a couple of months ago. The index lost almost 90% of its value, before it began a gradual uptrend in mid 1932.

\(^2\) Stock Market Crash of 1987
This crash occurred on October 19, 1987, now called the Black Monday. On this day, the Dow Jones Industrial Average slipped more than 500 points, or 22% of its value. Other indices were also affected badly - the S&P500 lost 20% and the NASDAQ Composite fell 11%. Why NASDAQ didn't fall as much as the other two was probably because the underlying market making system failed and the system was totally deadlocked. So far, the 1987 stock market crash is the largest one-day decline in history.
Introduction
Modern finance theories have evolved based on the assumptions that people are unbiased in their predictions about the future, and rational in decision making. However, psychologists have known for a long time that those assumptions are inaccurate because people oftentimes make predictable errors in their forecasts, and act in a seemingly irrational manner. If many people suffer from irrationality simultaneously, then signs (e.g. herding) can be found in the markets. As a result, the market can sometimes be inefficient.

Behavioral finance proposes psychology based theories to explain stock market anomalies. These theories have two building blocks: cognitive psychology (how people think) and the limits to arbitrage (what people do, resulting in market inefficiency). These two blocks are closely related because how people respond/react to one thing is determined, to some extent, by the way they have interpreted that thing.

Cognitive bias: Cognitive psychologists have documented many patterns regarding how people behave. One important pattern is overconfidence which causes people to overestimate their knowledge, underestimate risks, and exaggerate their ability to control events. Considering venture capital as an example, starting a new business is risky. In fact, most new businesses fail. When 2994 new businesses owners were asked about their chances of success, they thought they had a chance of 70% to succeed, while only 39% thought that any other business like theirs is likely to be as successful as theirs (1988, Arnold C. Cooper, Carolyn Y. Woo). It’s overconfidence that causes people overestimate their skills. If later these businesses fail, most likely these businessmen will attempt to believe that it’s due to bad luck or the market, instead of any personal factors. This is called the self-attribution bias and this pattern also happens in the stock market and is the basic idea of this paper.

Limits to arbitrage: Misevaluation of financial assets is common in the market. Arbitrage opportunity, if present, will disappear quickly. However, under some circumstances, withdrawals of capital after a losing streak may actually result in selling pressure that intensifies the inefficiency. It is the same story when strong capital inflows happen. In both cases, people behave in a herding fashion. As someone learn what other people think about various stocks, the social consensus forms. One minute the herd is doing nothing, and the next minute the herd is galloping at full speed.

Emirer et al, indicates that the level of herding is higher in emerging market (12 Dec, Riza Emirer, Daigo Gubo and Ali M. Kutan). The Chinese stock market is known for being full of speculation; the lack of sophisticated investors and proper regulations intensifies the psychological bias, which leads to market inefficiency. In this paper we will use empirical method to explore whether the irrational pattern (herding) does really exist in the Chinese stock market. This will give us a better understanding of the Chinese stock market.

4 http://www.cryoblog.org/social-aspects-of-tavesting.html
1. **Story (theory blocks):**
Investors suffer from psychological and emotional bias in similar pattern, regardless of whether they are individual investors or financial institutions. Successful and more experienced forecasters have “learned to be overconfident”. It is not just individuals who are affected, markets also become overconfident when market returns have gets high (2005, Richard Deaves, Erik Lüders, Michael Schröder). As such, we will use index as the market proxy instead of exploring individual stock prices. And we will also assume that individual investors and financial institutions behave in indistinctive manners.

One contributing factor to the overconfident pattern is the self-attribution bias, which is the tendency of humans to attribute successful outcomes to their skill and to blame unsuccessful outcomes on their bad luck. Investors who experience a run of successful results start to develop an inflated opinion of their own skills, which may possibly result in both complacency and exaggerated risk taking\(^5\). Such cognitive psychology is dangerous as investors are in denial and therefore cannot successfully learn from his or her mistakes.

Just to put the idea in a simple example: assume you have just purchased one stock after a due diligent analysis. The price actually turned out to rise after purchasing. You will be happy and become more confident in yourself thinking, believing that it’s your distinctive analysis approach or your investment strategy that contributed to the success. You attribute the success to your good skills ignoring whether the market is still too good.

Warren Buffett once told a story: *There was a duck in the pond, after a torrential rainstorm, it began to float up, thinking that its paddling skills have caused it to rise instead of the rising water.*

This story tells the same thing. Once people’s concepts or beliefs have to be recognized, exaggerated risk are going to be taken to increase investment (chasing). Considering one after two investments in above example, the former one is decision after formal analysis (assuming it is rational), the later one is nothing but just chasing (as if purchasing without any brain). Unintentionally, the investment strategy has changed, and investors have not realized.

When many investors behave in this manner, trading volume tends to increase as a consequence. It would magnify the arbitrage opportunities if available and prevent the market from being efficient. This process is a phenomenon of herding. As such, we want to find the herding evidence by studying the relationship between index levels and trading volume.

2. **Assumptions and hypothesis**
The purpose of this paper is to explore whether irrational pattern (herding) does exist in the Chinese stock market. The researcher uses empirical method to study the study relationship between market index level and trading volume.

Three assumptions are made:

- In short term, arbitrage opportunity can exist in both rational and irrational markets
- Index is a good proxy for the market
- Individuals and financial institutions behave in indistinctive manners.

Hypotheses:

\( H_0 \): People are rational: they invest based on formal analysis, and not suffer from psychological and emotional factors, hence no herding

\( H_a \): People are not always rational: psychological and emotional factors involve in the investment decision making, hence there will be herding

3. Inferences procedures & estimation

If \( H_0 \) is correct, it’s reasonable to expect that trading will cause the change of price and the volume is therefore useful in the prediction of price. If people are rational, they should follow rational investment process. After formal analysis, if investors think some stocks are undervalued they will purchase immediately to catch the opportunity before it disappears. Increase in transactions (rising of trading volume) will drive the stock price up until they reach the efficient level. Because we assumed that, in short term, arbitrage opportunity can exist in both rational and irrational markets; stocks can be undervalued, even though \( H_0 \) is really true. So, if \( H_0 \) is true, trading being the cause of the change in stock price is expected.

If \( H_a \) turns out be true, it’s reasonable to expect that, change of stock price will cause trading, and that stock prices are useful in the prediction of volume. In other words, it is the change of the stock prices that gives investors the signal to buy or sell. If people are not rational, they would follow irrational investment process. Assume you already have some shares on hand, or made some forecast and have been waiting to buy; after watching price going up, and recognizing the superiority of personal skills and experience, you rarely hesitate to increase the investment to catch the herd. So if \( H_a \) is true, the change of the price causing trading is expected.

4. Data

We will use SSE Composite Index of Chinese stock market as the market index. It is an index of all stocks (A shares and B shares) that are traded on the Shanghai Stock Exchange. It is the most comprehensive measure and is widely used in the Chinese stock market. We will now introduce the notations of time series variables used in this paper.

\text{index}: \text{the level of SSE Composite Index}
\text{lnindex}: \text{log of index}
\text{dlnindex}: \text{log difference of index}
\text{volume}: \text{the trading volume of SSE Composite Index}
\text{lnvolume}: \text{log of volume}
\text{dlnvolume}: \text{log difference of volume}

The data of index and volume is daily basis ranging from, Jan 01 2004 to Nov 17 2010. Figure 1-1 shows lnindex and lnvolume. Involume has been minified 3.5 times.
so that two time series can compatibly put in one graph. The unit root test shows that lnindex is first-order stationary I(1), and lnvolume is stationary I(0), which will be discussed in next section (see appendix).

![Figure 1-1](C:\Users\User\Documents\Figure.png)

**Source:** CSMAR Database, Jan 01 2004—Nov 17 2010

### Modeling

1. **Basic model**
   From the hypothesis and the estimates we made, we try to explore one variable (e.g. index) at the angel of the other one (e.g. volume), and then turn around, which indicates the characteristics of being simultaneous. So, to make the model simply we will use VAR model. The basic format is:

   **Equation 1-1:**
   
   
   \[ y_t = \theta_0 + \theta_1 y_{t-1} + \phi_1 x_{t-1} + \varepsilon_t \]
   
   \[ x_t = \delta_0 + \delta_1 y_{t-1} + \alpha_1 x_{t-1} + u_t \]

2. **Unit root and cointegration test**
   All variables used in our model are time series data. Before we decide how to apply our model, we have to know whether those series are stationary or not. To do this and unit root tests are used and the results provided by E-views show that lnindex is first-order stationary I(1), and lnvolume is stationary I(0). (Details in appendix).

   Cointegration test between lnindex and lnvolume is not needed as lnvolume is stationary already. We have concluded that there should not be any long-term relationship between stock and trading volume even without cointegration test. It is because stock value is determined by the performance of the companies, not by how they are traded in the exchange and therefore, price and trading volume should not have any fundamental relationship in the long-term.

3. **VAR model**
   As lnindex is first-order stationary I (1), and lnvolume is stationary I (0), and cointegration is not available under this situation, we will use VAR (1) model instead of Vector Error Correction Model (VECM).

   We take difference of the lnindex and the lnvolume, to get dlnindex and dlnvolume respectively, which are then both stationary I (0). After replacing x and y with notated variables, we have:
**Equation 1-2:**

\[
\begin{align*}
dlnindex_t &= \theta_0 + \theta_1 dlnindex_{t-1} + \phi_1 dlnvolume_{t-1} + e_t \\
dlnvolume_t &= \delta_0 + \delta_1 dlnindex_{t-1} + \alpha_1 dlnvolume_{t-1} + u_t
\end{align*}
\]

However, there is also an economical meaning between log difference of index (dlnindex) and log of volume (dlnvolume). Empirical work based on time series data assumes that the underlying time series is stationary. This means running the Model (dlnindex, dlnvolume) also make sense. So instead, we will use:

**Equation 1-3:**

\[
\begin{align*}
dlnindex_t &= \theta_0 + \theta_1 dlnindex_{t-1} + \phi_1 dlnvolume_{t-1} + e_t \\
dlnvolume_t &= \delta_0 + \delta_1 dlnindex_{t-1} + \alpha_1 dlnvolume_{t-1} + u_t
\end{align*}
\]

4. **Optimal lag choosing**

Running the model (Equation 1-3), we have to choose the best lag first.

<table>
<thead>
<tr>
<th>Lag</th>
<th>FPE</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.000533</td>
<td>-1.860620</td>
</tr>
<tr>
<td>1</td>
<td>1.26e-05</td>
<td>-5.608320</td>
</tr>
<tr>
<td>2</td>
<td>1.16e-05</td>
<td>-5.689785</td>
</tr>
<tr>
<td>3</td>
<td>1.13e-05</td>
<td>-5.715787</td>
</tr>
<tr>
<td>4</td>
<td>1.10e-05</td>
<td>-5.738995</td>
</tr>
<tr>
<td>5</td>
<td>1.10e-05</td>
<td>-5.739409</td>
</tr>
<tr>
<td>6</td>
<td>1.10e-05</td>
<td>-5.740650</td>
</tr>
<tr>
<td>7</td>
<td>1.10e-05</td>
<td>-5.742040</td>
</tr>
<tr>
<td>8</td>
<td>1.10e-05*</td>
<td>-5.745179*</td>
</tr>
</tbody>
</table>

Based on the AIC criteria, the optimal lag is 8 because its value is the smallest one, which minimizes the sum of residual square and number of parameters.

5. **Granger Causality**

We will restructure the model, with the optimal lag of 8, and then we conduct the Granger causality tests to see whether price is useful in the prediction of volume, or the other way around.
Table 1-2:

Dependent variable: DLNINDEX

<table>
<thead>
<tr>
<th>Excluded</th>
<th>Chi-sq</th>
<th>df</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LNVOLUM</td>
<td>8.095391</td>
<td>8</td>
<td>0.4242</td>
</tr>
<tr>
<td>All</td>
<td>8.095391</td>
<td>8</td>
<td>0.4242</td>
</tr>
</tbody>
</table>

Dependent variable: LNVOLUME

<table>
<thead>
<tr>
<th>Excluded</th>
<th>Chi-sq</th>
<th>df</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLNINDEX</td>
<td>238.0195</td>
<td>8</td>
<td>0.0000</td>
</tr>
<tr>
<td>All</td>
<td>238.0195</td>
<td>8</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

At the 5% significant level, index is not significant to cause volume. Even at the 10% significant level, we still cannot reject the notion that index does not cause volume. However, volume is very significant to cause index. This result is good, as it supports our estimations and favors the alternative hypothesis.

6. Impulse Responses

To track the effect of a one-time shock to one of the innovations on the current and future values of the endogenous, we will make the impulse responses test:
Table 1-3:

<table>
<thead>
<tr>
<th>Response of DLNINDEX to DLNINDEX</th>
<th>Response of DLNINDEX to LNVOLUME</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Response of LNVOLUME to DLNINDEX</th>
<th>Response of LNVOLUME to LNVOLUME</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1-3 indicates that, index does not respond much to a one-time shock coming from volume (top right figure); however, trading volume responds explosively to index (left down figure), which means that the change of index level is influential to the trading volume, but not the other way around.

7. Variance Decomposition
To better understand and to determine how much of the forecast error variance of each of the variable can be explained by exogenous shocks to the other variables, we will make variance decomposition:
It indicates that, in forecasting the index, most or all error variance comes from itself and not from volume. Volume is useless in the prediction of index. On the other hand, in forecast of volume, about 75% error variance can be explained by itself, with the remaining by index. Index makes some contribution in the prediction of volume.

8. Implications
The results from the above model suggest that price makes contribution in the prediction of volume, and therefore, the change of price causes or influences the change of trading volume. However, this is not true for the reverse. It rejects the null hypothesis in favor of the alternative one. When people are not rational, psychological and emotional factors influence their investment decisions and they tend to behave in a herding manner. People trend to make transactions after observing the market move in the direction they expected, which typically follow the irrational investment process.
Limitations and Extensions

1. Limitations
The result seems perfect to reject $H_0$ and in favor of $H_a$, this implies that irrational pattern (herding) does exist in Chinese stock market. Such pattern suggests inefficiency of the market.

Even though herding exists, not all misvaluations are caused by psychological biases. Some are merely due to temporary supply and demand imbalances. For example, the tyranny of indexing can lead to demand shifts that are unrelated to the future cash flows of the firm. When Yahoo was added to the S&P 500 in December 1999, index fund managers had to buy the stock, even though it had a limited public float. This extra demand drove the price up by over 50% in a week, and over 100% in a month. Eighteen months later, the stock price was down by over 90% from where it was shortly after being added to the S&P.

The other limitation is that the assumption-- individuals and financial institutions behave in indistinctive manners-- may not be true in reality. Financial institutions are trained to be more professional; as such, they tend to invest based on unique strategies. They are subject to strict regulations, and therefore on a whole they are less likely to make emotional decisions when compared to individual investors.

2. Extensions
On one hand, one investor might profit from the market inefficiency if he or she is the only one person who understands the mechanism of the irrational patterns of the market. Just as Buffett said “Be fearful when others are greedy. Be greedy when others are fearful”. It can be a philosophy derived from behavioral study. On the other hand, the herding phenomenon is a big problem for the development of Chinese stock market. It makes the market very inefficient. Herding causes people to be extremely optimistic and thus market bubble. In 2007, The Shanghai Stock Exchange index rocketed from 2,700 points to 6,400 points, a 130 percent rise. However, the performance of those listed companies on which the growth of the domestic stock market depends has not made much noticeable progress although stock reforms have been carried out in majority of these firms.

To reduce herding activities, more financial education, especially behavioral finance, should be taught. Cognitive bias can be corrected through related education. Therefore as China accelerates the development of financial industry, she should also increase the financial education.

Conclusion
Seeking to verify whether irrational pattern (herding) does really exit in Chinese stock, we used empirical methodology based on behavioral finance theory blocks. Through the modeling of the market index and trading volume, we found that the change of index is statistically significant to have caused the change of trading volume and as such, the index level is useful in the prediction of trading volume but not the other way around. We reject the null hypothesis in favor of the alternative hypothesis that people are not always rational. Psychological and emotional factors are involved in their investment decision making, so there is herding behavior.
However, not all misevaluations are caused by psychological biases and individuals and financial investors might behave in different manners. Restructuring our model to take into consideration of the above limitations might strengthen the thesis.

**Reference**
Arnold C. Cooper, Carolyn Y. Woo, and William C. Dunkelberg, “Entrepreneurs’ Perceived Chances for Success”
Arnold C. Cooper, Carolyn Y. Woo, and William C. Dunkelberg, “Entrepreneurs’ Perceived Chances for Success”

**Appendix**

**Unit root test**

<table>
<thead>
<tr>
<th>Type of Test</th>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Dickey-Fuller test statistic on the level of Inindex with trend and intercept</td>
<td>-1.088691</td>
<td>0.9293</td>
</tr>
<tr>
<td>Test critical values:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1% level</td>
<td>-3.963581</td>
<td></td>
</tr>
<tr>
<td>5% level</td>
<td>-3.412519</td>
<td></td>
</tr>
<tr>
<td>10% level</td>
<td>-3.128214</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of Test</th>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Dickey-Fuller test statistic on 1st difference of Inindex with intercept</td>
<td>-40.44734</td>
<td>0.0000</td>
</tr>
<tr>
<td>Test critical values:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1% level</td>
<td>-3.434087</td>
<td></td>
</tr>
<tr>
<td>5% level</td>
<td>-2.863078</td>
<td></td>
</tr>
<tr>
<td>10% level</td>
<td>-2.567636</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of Test</th>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Dickey-Fuller test statistic on level of lnvolume with trend and intercept</td>
<td>-4.357155</td>
<td>0.0026</td>
</tr>
<tr>
<td>Test critical values:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1% level</td>
<td>-3.963591</td>
<td></td>
</tr>
<tr>
<td>5% level</td>
<td>-3.412523</td>
<td></td>
</tr>
<tr>
<td>10% level</td>
<td>-3.128217</td>
<td></td>
</tr>
</tbody>
</table>