

THE GREATER FOOLS THEORY AND THE BUBBLES IN CHINESE
STOCK MARKET: A BEHAVIORAL APPROACH

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Lin Sun

January 2011

© 2011 Lin Sun

ABSTRACT

The Chinese Stock Market encountered its largest bubble from early 2006 to late 2007, and then crashed to the ground in mid 2008. On contrary to the bubbles in developed countries, the sky-rocketing phenomenon of the Chinese stock index in a very short period could not be fully explained by the rational bubble theory. In this paper, I examine both theoretically and empirically the “Greater Fools Theory”, in the scope of irrational bubble theory. The result suggests that the Greater Fools Theory should be credited for this bubble in China. The theory part of this thesis presents an intuitive mathematical model to show the consistency of Greater Fools Theory to agents’ behavior in reality. The empirical part of this paper adopts a traditional factor-pricing model, the APT model, to show the explanation power of greater fools proxies.

BIOGRAPHICAL SKETCH

Lin Sun was born on November 3, 1985 in Tianjin, China to Shouping Sun and Jiakang Sun. He is the only son of the family.

He resided in Tianjin for 15 years and then moved to Beijing in 2001. He graduated from Peking University High School in 2004. Lin firstly matriculated into Tsinghua University, Beijing in 2004, majoring in Economics. One year later, Lin followed the family to move to Hong Kong. And hence he transferred into the University of Hong Kong, majoring in Economic and Finance.

Lin has been exchanged to Seifu High School, Osaka, Japan in 2003. Lin has also been exchanged out to University of Edinburgh, UK in 2007. Lin accumulated working experience from internships in ABN-AMRO, J.P. Morgan and ING Bank. He has visited many countries and has taken part in broader varieties of international meetings and events.

Lin began his studies at Cornell University in the pursuit of a MSc. degree in the fall of 2008. After successful receiving the MSc. degree, he continued to study toward a PhD. degree from the fall of 2010.

ACKNOWLEDGMENTS

Prof. Calum Turvey, my advisor, deserves many thanks for his constructive guidance and unreserved help throughout this study. Similarly, I am also thankful to my committee member, Dr. Vicki Bogan for her advices and helps.

I would also like to thank other professors in Dyson School of Applied Economics and Management, Johnson School of Management and the Department of Economics who have taught me during the past two years: Prof. Richard Boisvert, Prof. David Just, Prof. Ming Huang, Prof. David Easley, Prof. Lawrence Blume and Dr. Hazem Daouk.

Their instructions are inevitable in the growth of my knowledge and skills. I am also thankful to the graduate coordinator Linda Morehouse. She has provided countless helps since the first time I came to the department.

Finally, I would like to dedicate this thesis to my parents, Jiakang Sun and Shouping Sun.

TABLE OF CONTENTS

Biographical Sketch.....	iii
Acknowledgment.....	iv
Table of contents	v
List of figures	vii
List of tables	viii
Chapter 1 - Introduction	- 1 -
Chapter 2 - Background of Chinese Stock Market.....	- 5 -
2.1. Glorious Past	- 5 -
2.2. Fading Years.....	- 5 -
2.3. First Issuance and Establishment of Stock Exchanges.....	- 5 -
2.4. Market Development	- 7 -
2.5. “No” Dividends Policy	- 9 -
2.6. Banned Short Selling.....	- 10 -
2.7. Summary.....	- 11 -
Chapter 3 - Literature Review	- 12 -
3.1. Modern Assets Pricing Theory	- 12 -
3.1.1. Markowitz’s Portfolio Theory	- 12 -
3.1.2. Tobin’s Separation Theorem	- 13 -
3.1.3. Capital Assets Pricing Model	- 14 -
3.1.4. Arbitrage Pricing Theory Model	- 16 -
3.2. Explanations of the Stock Bubble	- 18 -
3.2.1. Rational Expectation Bubble Theory	- 19 -
3.2.2. Irrational Bubble Theory	- 23 -
3.2.2.1. Feedback Theory of Speculative Bubbles	- 23 -
3.2.2.2. Herding Behavior	- 25 -
3.2.2.3. Noise Trader Theory.....	- 26 -

Chapter 4 - Methodology.....	- 31 -
4.1. What is Greater Fools Theory?.....	- 31 -
4.2. A Simple Mathematical Proof	- 33 -
4.2.1. Optimal Duration to Hold A Stock.....	- 34 -
4.2.2. Net Expected Returns for Holding a Fixed Period.....	- 35 -
4.2.3. Net Expected Return and Convergence Condition.....	- 36 -
4.2.4. Heterogeneous Beliefs of Investors	- 38 -
4.2.5. Remarks	- 41 -
4.3. The Factor Model	- 42 -
Chapter 5 - Empirical Result	- 46 -
5.1. The Data Source	- 46 -
5.2. Summary Statistics	- 46 -
5.3. OLS Estimates	- 48 -
5.3.1. Autocorrelation and Heteroscedasticity Test.....	- 50 -
5.3.2. Stability Test.....	- 53 -
5.3.3. Non-linear Regression	- 55 -
5.3.4. Redundant Variable Test	- 57 -
5.3.5. OLS Remarks	- 58 -
5.4. AR-IGARCH-M Model.....	- 59 -
5.4.1. Stationarity and Unit Root Test.....	- 59 -
5.4.2. Model Construction	- 60 -
5.4.3. AR-IGARCH-M Estimates	- 65 -
5.5. Economics Explanation	- 68 -
Chapter 6 - Concluding Remarks	- 73 -
Bibliography	- 75 -

LIST OF FIGURES

Figure 1 <i>Chinese Stock Market Development</i>	- 9 -
Figure 2 <i>Sensitivity of the Convergence Condition</i>	- 38 -
Figure 3 <i>The Process of Greater Fools Theory</i>	- 42 -
Figure 4 <i>Actual, Fitted and Residual Graph</i>	- 49 -
Figure 5 <i>Serial Correlation Test I: Correlogram of Residuals and Q-Statistics</i>	- 51 -
Figure 6 <i>ARCH Test I: Correlogram of Residuals Squared and Q-Statistics</i>	- 52 -
Figure 7 <i>Stability Test I: CUSUM Test and CUSUM Squares Test</i>	- 54 -
Figure 8 <i>Non-Linear Model Test: Scatter Graphs and Nearest Neighbor Lines</i>	- 56 -
Figure 9 <i>Actual, Fitted and Residual Graph</i>	- 67 -
Figure 10 <i>Serial Correlation Test: Correlogram of Residuals and Q-Statistics</i>	- 67 -
Figure 11 <i>Heteroscedasticity Test I: Correlogram of Residuals Squared and Q-Statistics</i>	- 68 -

LIST OF TABLES

Table 1 <i>Summary Statistics of Raw Series</i>	- 47 -
Table 2 <i>Summary Statistics of Transformed Series</i>	- 48 -
Table 3 <i>OLS Estimates</i>	- 49 -
Table 4 <i>Serial Correlation Test II: Breusch-Godfrey</i>	- 51 -
Table 5 <i>Heteroscedasticity Test II: ARCH LM Test</i>	- 52 -
Table 6 <i>Stability Test II: Chow Breakpoint Test</i>	- 55 -
Table 7 <i>Redundant Variable Test</i>	- 58 -
Table 8 <i>Augmented Dickey-Fuller Unit Root Test</i>	- 60 -
Table 9 <i>Model Selection: AR(x) structures comparison</i>	- 63 -
Table 10 <i>Model Selection II: GARCH(q,p) structures comparison</i>	- 64 -
Table 11 <i>Model Selection III: GARCH-M structures comparison</i>	- 64 -
Table 12 <i>AR(1)-IGARCH(2,1)-M Estimates</i>	- 66 -
Table 13 <i>Heteroscedasticity Test II: ARCH LM Test</i>	- 68 -

CHAPTER 1

INTRODUCTION

The so-called bubble, a phenomenon that assets prices deviated from its intrinsic values based on market fundamentals, has intrigued economist for a long time. What makes the bubble and crash phenomenon so problematic is that this pattern in financial market prices is widely considered harmful for the allocation of investment capital. Market prices form the basis for the allocation of investment capital to its most efficient uses. When the market prices deviate from the right measure of underlying value, it will cause a misallocation of available resources for the economy. That is the very reason that we should study the bubble and crash phenomenon. And hopefully we can help to reduce the frequency of such phenomenon by gaining a better understanding of it.

Economists usually hold different views about stock bubbles. One group believes that the price of an asset should simply equal its discounted future cash flows, and reflect the market fundamentals of a company. These supporters assert the financial market is always in equilibrium and bubbles like Tulip Mania were described by Garber (1989) as unusual moves in the fundamental value. They argue any deviation from this fundamental value should be corrected by some sophisticated arbitrageurs in an efficient market. Gürkaynak (2008) reports that “for each paper that finds evidence of bubbles, there is another one that fits the data equally well without allowing for a bubble. We are still unable to distinguish bubbles from time-varying or regime-switching fundamentals, while many small sample econometrics problems of bubble tests remain unresolved.” Such statement is quite representative in this camp. Indeed, the supporters of this camp deny the existence of bubble, not only because

they are the believer of the fundamental value, but more importantly because they fear the difficulty of testing it.

A second group believes that the rationality of both behavior and future expectations could also imply a price deviation from its fundamental value. Such bubble can occur even when the traders act rationally and have rational expectations. The bubble today is simply a justification and realization of a higher bubble tomorrow. A group of macroeconomists, including Tirole (1982), Blanchard and Watson (1983), were the first to thoroughly formalize the possibility of the so called “Rational Bubble Theory”, finding that dynamic model of the price level could have indeterminate explosive solution even if the agents have rational expectations. Rational Bubble Theory makes a large step further towards the understanding of the bubble phenomenon, but it is not the end of the road. Rational bubble theory cannot survive in the case that agents have adaptive expectation or the case that some market limits, including limited life of assets, limited personal wealth and finite number of market participants, exists.

The last group believes that a market bubble is caused by some irrational investors in the market. Such irrationality is the consequences of agents’ psychological and behavioral factors. This group supports the so-called “Irrational Bubble Theory”. Influential works include Fashion Herding Theory by Shiller and Fisher (1984), Feedback Loop Theory by Shiller (1990) and Noise Trader Theory by Long, Shleifer et al. (1990). Their theories are shifting away from the rigorous rational bubble works of those macroeconomist, but are more in line with the cognitive behaviors of the investors in the market.

Due to the complex nature of the stock market bubble, it is not easy to simply attribute the bubble to any single side of the opinions. Hence, I vote to study bubble and crash in a more flexible context. By the term of “flexible”, I mean we should

really think out of the box and combine edges of different approaches. First, I will study the phenomenon in an equilibrium model, just like the Rational Bubble Theory. It is because the steady state market is the most representative and is much easier to characterize. Second, we have to find a theory that fits zealot investors' cognitive behavior during a bubble, just like the Irrational Bubble Theory. The Greater Fools Theory, an old but interesting concept, catches me by its rich intuition and less assumption constraints. It is an equilibrium model, but it doesn't require anything on the preference of agents. All it requires is the current fool holds a belief that next fool will buy the stock over at a higher price. It vividly describes a populous investors' trading strategy during a stock market bubble.

Now I come to the data problem. Bubble is still relatively rare event. They seldom recur in the same country or market sector within the same generation of participants, so we have to carefully select the data we will study with. In the bubble happened in which country and sector should I choose? Is the US stock market a suitable one? I don't think the answer is positive. First, US stock market could be largely explained by the Rational Bubble Theory as West (1987), Flood and Hodrick (1990) tested. Second, many trading activities in US nowadays are now conducted through computational trading program. So the market suffers far more less influence from one trading idea (the Greater Fool Theory) but more from different trading ideas that opposite each other. Third, US stock market has short-selling mechanism, so it is easier for arbitrageurs to short and trade against a bubble. On the contrary, Chinese stock market seems like natural laboratory for such study, because of the poorly educated investors, speculative atmosphere and banned short-selling environment. Besides, some researchers, like Chan, McQueen et al. (1998), has tested and generally rejected the validity of Rational Bubble Theory in Asian markets. Thus, as an alternative to the Rational Bubble Theory and a complement to the existing Irrational

Bubble Theory, this thesis investigates the Greater Fools Theory to study the bubble phenomenon in China.

To contribute to the line of literatures on irrational bubble theory, this paper innovates on the Greater Fools Theory in Chinese stock market on several fronts. First, a series of properties is derived from the basic mathematical expression of the Greater Fool Theory by Telser (2010). These properties are generally in line with observable investors' behavior. Second, unlike previous irrational bubble theory, this Greater Fools Theory is empirically tested, using greater fools proxies. Third, I apply a good econometric practice by using an AR-IGARCH-M hybrid structure to derive a very satisfactory result.

The rest of the thesis is organized as follows. Chapter II provides a background to the Chinese stock market. In Chapter III, I will review the main categories of literature that try to explain stock market and other asset bubbles and in this context provide an overview of assets pricing models. I will introduce the methodology and provide mathematics and test the model of the Greater Fools Theory in Chapter IV. Empirical result will be reported and interpreted in Chapter V. Chapter VI summarizes and concludes the thesis.

CHAPTER 2

BACKGROUND OF CHINESE STOCK MARKET

2.1. *Glorious Past*

Before 1949, Shanghai was a major banking and financial center in Far East with a stock market capitalization similar to that of Tokyo and considerably larger than that of Hong Kong. So powerful was the Shanghai Stock Exchange (“SSE”) that any fluctuation of it between 1919-1949 had a huge influence on other world-class financial markets. And in fact, Shanghai stock exchange was once believed to be the world’s third largest after New York and London, according to an article in Shanghai-Investment.com (2000).

2.2. *Fading Years*

After the People’s Republic of China (PRC) was established in 1949, the government strictly controlled virtually all channels of investment by the principle of communism. Prior to 1978, it was the state owned banks in China that controlled the entire financial system under social planner’s commands. At that time, financing were conducted through the direct grant of the state budgetary fund or the government allocated bank credits. The official process that allocated investment across regions and industrial sectors was often objective and bureaucratic. In many cases, the allocation caused inefficient use of resources. A stock market is simply not necessary to exist in such a centrally planned economy.

2.3. *First Issuance and Establishment of Stock Exchanges*

The Chinese government started its reform in 1978 after Chairman Mao died two years earlier. To release productivity from regulatory and political constraints and to provide incentives to the employees of State Owned Enterprises (SOEs), China launched its shareholding reforms in the early 1980s. Accompanying this reform, there

was an outbreak of share issuance by SOEs or collective companies to bring shareholdings to private investors and its employees. On July 8th, 1983, Shenzhen Bao'An Investment restructured itself into a shareholding company and issued stocks, the first stock after the founding of the People Republic of China, on a private placement basis to the public. According to Green (2004), those shares had been traded illegally on the factory floor or on the streets outside the factory, since a ban on share trading had not been lifted. The popularity of the street trading called for a formal platform for shareholders to trade. On September 26, 1986, the first over-the-counter (OTC) trading post opened in Shanghai to meet investors' demand of secondary market.

The inconveniency of OTC trading created an exploding demand for stock exchanges amongst investors. In response to meet this demand, the Shanghai Stock Exchange (SHSE) and Shenzhen Stock Exchange (SZSE) were established by the government on December 19 1990, and December 1 1990, respectively. Their establishment also replaced the paper based trading system with a computer aided trading system. On Nov 8th, 2009, a third stock exchange called ChiNext was launched in Shen Zhen.

ChiNext¹ offers a new capital platform tailor-made for the needs of enterprises engaged in independent innovation and other growing venture enterprises. It is often referred as the Chinese version of "NASDAQ". To be noted, all three stock exchanges operated under an auction market environment without a specialist or market maker. Thus, according to Hertz (1998), liquidity is not guaranteed, trading may be frequently ceased, and manipulation may be widely spread without the presence of a market maker in SSE.

¹ Source: <http://www.szse.cn/main/en/ChiNext/aboutchinext/>

2.4. Market Development

Since its establishment in 1990, the Chinese stock market has grown at a phenomenal pace: The number of listed stocks In Shanghai and Shenzhen Exchange increased from 13 in 1991 to 1,644 by the end of April, 2009, and the aggregate market capitalization rose from US\$0.85 billion to nearly US\$2,480 billion during the same period.² In terms of market capitalization, the Chinese stock markets are now the second largest in the Asia-Pacific region after Japan, and are ahead of such major markets as Australia, Hong Kong and Korea. Considering that China has one of the fastest growing economies in the world and, further, that the government is committed to continuous deregulation and liberalization, the rapid development of Chinese stock markets is likely to continue into the foreseeable future.

Before August 1992, there was no consensus within the central government to establish a regulator in the stock market. However, the riot in Shenzhen in August, 1992 triggered by an extremely over-subscribed IPO (people went on the street because they couldn't even buy one share of the issued stock) put the creation of a regulator on the agenda of the government. Hence two regulators were set up back in 1993: The China Securities Regulatory Commission (CSRC) and the State Council Securities Committee (SCSC). These are consolidated in 1998 to get the present day CSRC. In March 2000, the China Securities Depository and Clearing Company (CSDCC) was established as the central securities clearing company. Currently, the Chinese stock market consists of one regulator, three stock exchanges, one clearing company, and many listed companies.

By far, the Chinese stock market had experienced seven major bull/bear circles. The first bull market started in December 1990, ended in May 1992. This was widely regarded as the “Stock Fever” and such fever has been vividly described and

² Source: <http://www.sse.com.cn/> and <http://www.szse.cn/>

carefully analyzed by Hertz (1998). During this period, the number of listed companies increased from 8 to about 50, while the Shanghai composite index soured from 96.05 to 1,429, implying a 1380% increase. After May 1992, the index crashed to 386 in half a year, losing 73% of its value. The last bull/bear flip-over happened between Jan 2006 to May 2008. The Shanghai composite index sky-rocketed from 1,161 in Jan 2006 to 6,124 in October 2007, increasing for more than 400% in 22 months and then it dramatically fell to less than 3,000 in May 2008, which was far earlier than the Global Financial Crisis of 2008-2009.

The following graphs show the recent performance of Chinese Stock Market by examining its market capitalization and turnover.³ They indicate that Chinese stock markets expanded aggressively after 2006. From the graphs, we can observe that Chinese investors are very sentimental in their trading pattern. When the market is bullish, Chinese investors tend to trade more and their turnover is extremely high and vice versa when the market is bearish. In a mature market, the turnover of the investors should be quite stable, since their trading pattern is mostly determined by the institutional investors who hold the stocks for a longer time horizon.

³ Source: People Bank of China

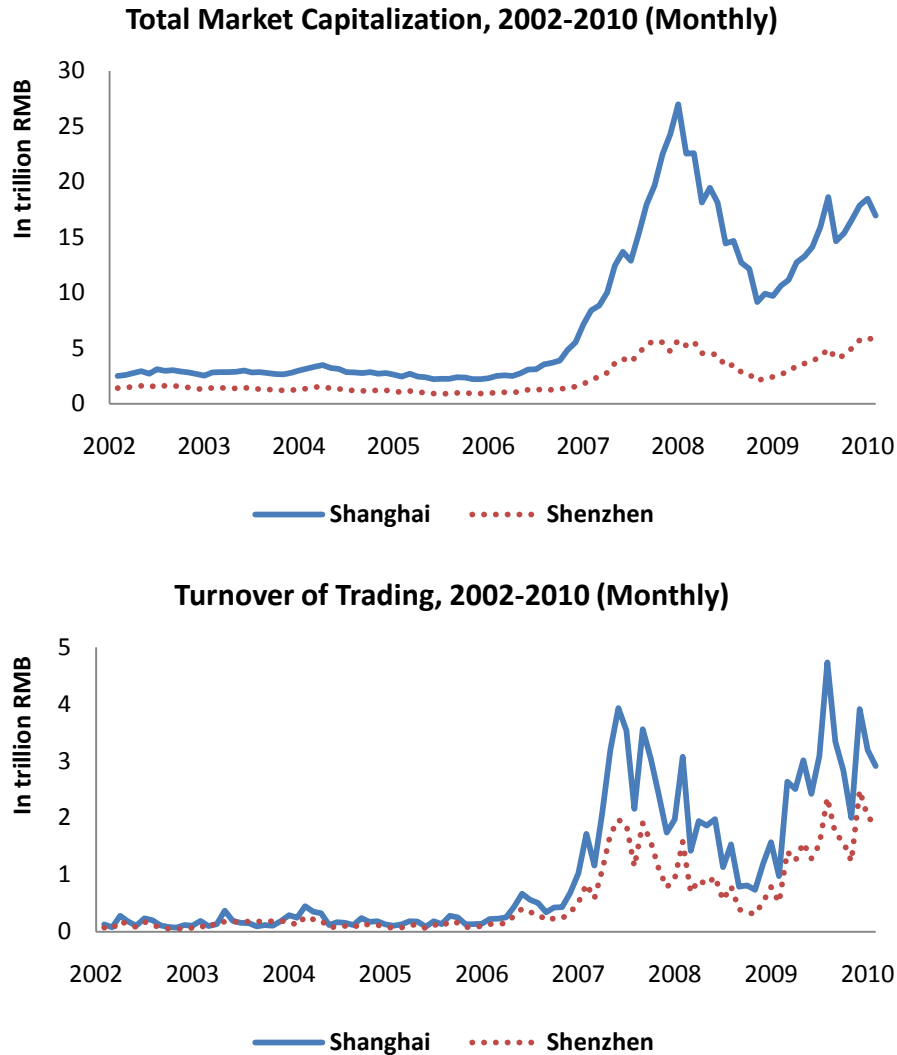


Figure 1 *Chinese Stock Market Development*

2.5. “No” Dividends Policy

Chinese companies, especially those non state-owned companies, are very unwilling to issue dividends to their investors, since their financing channel is very limited and self-financing is the last resort once their loan applications are rejected by the state owned banks. However, this habit of Chinese companies makes their stocks vulnerable to speculation. If a company issues affluent dividends, then value of its

stock would be easily calculated under the principle of discounting cash flow and all that investors care is the future profitability of one company. And such profitability could be examined and predicted by many financial analysts behind those securities. Although brokers' report is not impeccable, at least it would build the cornerstone of the stock price. On the contrary, if a company chose to ignore dividends issuances, investors' only return would be from the price appreciation, which heavily suffers from market sentiment, manipulation and investors' taste. In the case that non-dividend policy dominates the corporate world, stock market would become more volatile than ever, and the speculation power would find it easier for them to play around in such a market. Indeed, most Chinese investors have traded speculatively with very short holding periods. With round trip trading costs approaching 1% of the total transactions, the average annual turnover during the bull time roughly exceeded 12 times⁴!

2.6. *Banned Short Selling*

The Chinese stock market has no history of short sales. However, in 2007, the Chinese government, in an effort to increase the types of financial instruments available to market participants, considered introducing short selling to the market. As of 2009, the only short selling in the Chinese stock market occurs through the 11 brokerage firms that are part of the trial program. By far, there has not been any further announcement by the commission to make short selling a permanent feature of the Chinese stock market. And collectively, these 11 brokerage firms only contribute tiny trading volume to the market and hence their shorting activity could be easily ignored in front of any true market manipulators.

In a stock market that short selling is banned, fundamentalist who believed in the mean reversion of stock prices towards a fundamental value would find no place to

⁴ Source: Prof. Ming Huang's Lecture Notes

stay. And thus deviation from the fundamental value would be difficult to correct. This is the one of the very reasons that Greater Fools Theory seems quite fit for China.

2.7. *Summary*

This chapter has provided a brief overview of the historical development of Chinese stock exchange. The volatility of the market, set against huge swings in Bull-Bear market, wide seemingly inexplicable rises and falls, brings into question to market efficiency theory and rational bubble theory as reasonable explanations for market behavior. Against this backdrop, the next chapter reviews the historical development of portfolio choice and asset pricing theory that give rise to the current view of the existing bubble theories against which the existence of a Greater Fools Theory is tested.

CHAPTER 3

LITERATURE REVIEW

3.1. Modern Assets Pricing Theory

3.1.1. Markowitz's Portfolio Theory

Modern finance theory started from Markowitz's (1952) portfolio selection theory, which predicts how individual investors should allocate their assets by balancing the risk and return tradeoffs. His theory mainly assumes investors are risk averse and investors' risk preference can be described by a quadratic utility function, where only expected return and volatility matters. Investors are assumed to only care about the 1st moment of their utility – the mean and the 2nd moment of their utility – the variance. All higher moments, like skewness or kurtosis, are ignored. In his model, the expected return from the portfolio as a whole is: $E = \sum_{i=1}^N X_i u_i$ ⁵ and the variance of the portfolio is: $V = \sum_{j=1}^N \sum_{i=1}^N X_j X_i \sigma_{ji}$ ⁶. Markowitz notes that it is not a security's own risk that is important to an investor, but rather the contribution the security makes to the variance of his entire portfolio. And such contribution is determined by this security's covariance with all the other securities in his portfolio, which is σ_{ji} in his model. By combining different assets whose returns are not correlated, the total variance of the portfolio could be reduced. Markowitz also shows that if we have collected all the return, covariance information we needed, then for any level of risk, the efficient frontier identifies a point that is the highest returning portfolio in its risk class. By the same token, for any level of return, the frontier identifies the lowest risk portfolio in that return class. These portfolios form the efficient frontier, and

⁵ X_i is the percentage of investors' holding in i^{th} asset; u_i is the expected return of i^{th} asset.

⁶ σ_{ji} is the covariance between the j^{th} asset and the i^{th} asset.

Markowitz shows that for any investor that follows his assumption, it is economically efficient to limit his/her choices to portfolios that lie on this frontier.

3.1.2. *Tobin's Separation Theorem*

Tobin (1958) develops the modern portfolio theory one step further by showing how to identify which efficient portfolio should be held by an individual investor. The key “separation theorem” describes that in a world with one safe asset and a large number of risky assets, portfolio choice by any risk-averse investor can be described as a choice between the safe asset and a market portfolio of risky assets, which will be accepted by all investors. The proportion of the risky shares in this market portfolio of risky assets is the same for all risk-averse investors. The degree of risk aversion only determines the shares of the safe asset and the uniform portfolio of risky assets in the total portfolios for each investor. Mathematically, Tobin’s “Separation Theorem” can be concluded as below: The market portfolio of risky assets, which equals to $f = V^{-1}m/1^TV^{-1}m$ ⁷, could be found by maximizing the Sharpe ratio, which is $f^Tm/\sqrt{f^TVf}$. However, the fraction of wealth invested in the risky assets, which equals to $\alpha^{-1}V^{-1}m$, is individual specific, because of the risk aversion factor in the denominator. The fraction of wealth invested in the risk-free asset, which equals to $1 - 1^T\alpha^{-1}V^{-1}m$, is just the remaining portion of the total wealth. Graphically, Tobin indicates that a line drawing from the risk-free asset and is tangent with the Mean-Variance (MV) frontier is the Capital Market Line (CML). The tangent point lying on CML and MV frontier illustrates the market portfolio. Individual’s indifference curves that are tangent with the CML could decide the optimum proportions of the risk-free assets and the market portfolio.

⁷ “ m ” is the vector of mean excess returns on the risky assets and that “ V ” is the co-variance matrix. “ f ” denotes a portfolio of risky assets, in which “ f_i ” is the fraction of wealth invested in the i^{th} asset, normalized so that $f^T1 = 1$. “ α ” is the relative risk aversion.

3.1.3. Capital Assets Pricing Model

Building on the earlier work of Markowitz on diversification and Tobin's "Separation Theorem", Treynor (1961, 1962), Sharpe (1964), Lintner (1965) and Mossin (1966) developed the so called Capital Asset Pricing Model (CAPM). The CAPM mainly studies how investors' asset demand determines the relation between assets' risk and return in market equilibrium. CAPM assumes that all investors have the mean-variance preference, and more importantly they hold a homogeneous belief about the value of mean and variance. Based on Tobin's idea of market portfolio, CAPM concludes that any individual stock can be priced by the market portfolio as the formula below: $E(\tilde{R}_i) - R_f = \beta_{im}[E(\tilde{R}_m) - R_f] \forall i$ where $\beta_{im} = Cov(\tilde{R}_i, \tilde{R}_m)/Var(\tilde{R}_m)$. For the first time their theory clearly prescribes that it is the individual stocks' co-movements with the overall markets that determine stocks' expected returns (thus the stock prices). In other words, it is the systematic risk that matters in asset pricing. Unlike Sharpe and Lintner's CAPM, which assumes unlimited borrowing or lending at a risk-free rate, Black (1972) relax their assumption such that a riskless asset does not exist, but a substitute asset whose return are uncorrelated with the efficient portfolios can replace it. The investors behave slightly differently than before: they hold different efficient ("efficient" here means mean-variance efficient) portfolios, and the combination of their efficient portfolios creates the market portfolio and is also lying on the mean-variance frontier. The so called Zero Beta CAPM by Black (1972) can be described by the formula below: $E(\tilde{R}_i) - E(\tilde{R}_z) = \beta_{im}[E(\tilde{R}_m) - E(\tilde{R}_z)] \forall i$ where $\beta_{im} = Cov(\tilde{R}_i, \tilde{R}_m)/Var(\tilde{R}_m)$. $E(\tilde{R}_z)$ is called the zero-beta rate, which is the expected return on assets whose returns are uncorrelated with the market return. If CAPM is true, the model would have important implications for problems like capital budgeting, portfolio selection, cost-benefit analysis and any problem that requires the implementation of the relation

between risk and return. To justify the legitimacy of CAPM, many researchers started to test the model.

The initial test of the CAPM by Lintner (1965) was not successful. When average stock returns were plotted against betas of individual stocks, they found a lot of dispersion, and that the slope of the line was so flat that there was no realistic risk-free rate to support such model. Black & Myron (1973) and Fama & MacBeth (1973) tackled the problem by sorting stocks into portfolios based on betas. Black, Jensen, & Scholes (1972) test the CAPM by investigating whether the intercepts of the cross-sectional and time series regressions of excess return on market beta are zero. Fama & MacBeth (1973) followed another approach to examine the validity of CAPM model, by adding two additional explanatory variables to the cross-sectional regression: one is squared market beta and the other one is the residual variances from regressions of individual return on the market return. They used the squared beta to examine whether the relationship is linear. And they used residual variance to examine whether there are other measures of risk rather than the market beta that could help to explain the expected returns. Their result showed that both variables are not significant in explaining the returns. However, Roll (1977) challenged against the test of CAPM by making two statements. Firstly, based on pure mathematics, he showed that any mean-variance efficient portfolio \tilde{R}_p can satisfy the CAPM equation such that $E(\tilde{R}_i) - R_f = \beta_{ip}[E(\tilde{R}_p) - R_f]$. Secondly, a true market portfolio is unobservable; it contains all investment opportunities like bond, real estate, etc. Since previous tests use the stock market portfolio as a proxy for the true market portfolio, testing the CAPM equation is equivalent to testing mean-variance efficiency of stock market portfolio. Given that the stock market is assumed to be mean-variance efficient, testing the CAPM in this way is totally tautological.

3.1.4. Arbitrage Pricing Theory Model

The Arbitrage Pricing Theory (APT) Model is largely motivated by the empirical failure of the CAPM. It is a multi-factor model that uses factor loading to explain stock returns. The APT was developed primarily by Ross (1973). It is a one-period model in which every investor believes that the stochastic properties of returns of capital assets are consistent with a factor structure. Ross argues that if equilibrium prices offer no arbitrage opportunities over static portfolios of the assets, then the expected returns on the assets are approximately linearly related to the factor loadings. The factor loadings, or betas, are proportional to the returns' co-variances with the factors. Unlike CAPM, which is an equilibrium model and derived from individual portfolio optimization, APT is a statistical model, which tries to capture sources of systematic risk. A mathematical formula of APT could be described as below: $r_i - E(r_i) = b_{i1}F_1 + b_{i2}F_2 + \dots + b_{in}F_n + \epsilon_i \forall i$ ⁸. Almost from the inception of the APT, the choice of factors, number of factors and their interpretation has been hotly debated. One of the earliest empirical studies of the APT, by Roll & Ross (1980), uses factor analysis, a statistical technique that allows the researcher to infer the factors from the data on security returns. Their results indicate that there are four priced factors in the US stock market. The advantage of factor analytic techniques is that the factors determined from the data explain a large proportion of the risks in that particular dataset over the period under consideration. The drawback is that factors usually have no economic interpretation. An alternative to factor analytic techniques is to use observed macroeconomic variables as the risk factors. One of the first studies using observed factors was by Chen, Roll, & Ross (1986). Their argument is that at the most basic level we can imagine that some fundamental valuation model

⁸ $E(r_i)$ is the i th asset's expected return; $F_k, k = 1, \dots, n$ is the systematic factors; b_{ik} is the sensitivity of the i th asset to factor k ; and ϵ_i is the i th risky asset's idiosyncratic random shock

determines the prices of assets. That is, the price of a stock will be the correctly discounted expected future dividends. Therefore the choice of factors should include any systematic influences that impact future dividends, the way traders and investors form expectations and the rate at which investors discount future cash flows. They find that US stock prices are significantly related to (1) changes in industrial production (GNP), (2) the spread between the yield on short-term and long-term government bonds, (3) the spread between low- and high-grade bonds, (4) changes in expected inflation, and (5) changes in unexpected inflation. However, the APT still has the testability problem as the CAPM. Shanken (1982, 1985) asserts that for individual securities the approximation implied by Ross' APT is so imprecise that it makes it impossible ever to test whether the APT is true or false. Furthermore, Shanken argues that since the expected return for any security or portfolio is related only approximately to its factor sensitivities, to get an exact pricing relationship, additional assumptions are needed.

Although APT model has the similar testability problem as the CAPM model, I would still choose it as the factor assets pricing model in this paper, due to the following reasons. Firstly, APT is a multi-factor model, while CAPM is a single factor model. Confining systematic risk into a single “market portfolio” factor is not too compelling than decomposing the systematic risk into several meaningful factors. Secondly, APT requires no utility assumptions beyond monotonicity and concavity. On the contrast, CAPM requires a quadratic utility function, which is unlikely to be the case because of the higher order belief in greater fools’ decision making process and utility functions (please refer to the later part for the idea of higher order beliefs). Finally, because APT is based on a non-arbitrage condition it should hold for any subset of securities. On the contrast, CAPM has been criticized for a long time,

because of its dependence on a market portfolio of risky assets, which should be a world portfolio (that is not currently available) rather than the S&P 500.

3.2. *Explanations of the Stock Bubble*

What is a stock bubble? According to Kindleberger (2008) in the “The new Palgrave dictionary of economics”, he gave out a descriptive definition of bubble:

“Bubbles are typically associated with dramatic asset price increases followed by a collapse. Bubbles arise if the price exceeds the asset's fundamental value. This can occur if investors hold the asset because they believe that they can sell it at a higher price than some other investor even though the asset's price exceeds its fundamental value.”

People have studied bubbles for a long time. As early as 1841, the Scottish journalist Charles Mackay, in his book called “Extraordinary Popular Delusions and the Madness of Crowds”, proposed that crowds of people often behave irrationally to create bubbles. He mainly referred the Tulip Mania, along with the South Sea Bubble and the Mississippi Company scheme as his primary examples. Mackay's vivid book was popular among generations of economists and stock market participants.

However, at the early stage of modern finance study, researchers who believe market is always efficient typically chose to ignore the stock bubble rather than actually analyzing it. When Fama (1965) discussed the random-walk behavior of the stock market, he expressed an idea that stock price shouldn't deviated from its intrinsic value for two reason. First, he believed that some sophisticated traders can identify the situation when a price is running up its intrinsic value and prevent the bubbles from occurring by selling the stocks or shorting it. Second, once those astute chartists understood the nature of the dependencies in the series of successive price changes, they will be able to identify, statistically, situations where the price is

beginning to run up above the intrinsic value and thus erase the dependencies by selling the stocks or shorting it.

The real study of stock bubble began by the rational bubble theory in the 1980s and was complemented by the irrational bubble theories in the 1990s.

3.2.1. *Rational Expectation Bubble Theory*

By the definition given by Blanchard & Watson (1983), rational expectation bubble is the rational deviations of the price from their fundamental values. They believed that rationality of both behavior and of expectation often does not imply that the price of an asset be equal to its fundamental value.

They basically assume there are two types of assets in the market: one is the risk-free government bond; the other one is the risky stock. In the context of efficient stock market, investors (assumed to be risk neutral) would long stocks and short bonds, when the expected return of the stock exceeds the expected return of the bond. And hence the no arbitrage pricing condition should be satisfied and should be given by:

$$[E(P_{t+1}|I_t) - P_t + E(d_{t+1}|I_t)]/P_t = r_f^9 \quad (3.1)$$

Rearranging the above formula, we will get the famous linear rational expectation difference equation:

$$P_t = aE(P_{t+1}|I_t) + aE(d_{t+1}|I_t)^{10} \quad (3.2)$$

Given the assumption of rational expectation and that agents do not forget, the information set should have the following property that:

$$I_t \subset I_s, s \geq t \quad (3.3)$$

By the property of the conditional expectation, we will have:

⁹ “ P_t ” is the stock price at time t ; “ V_t ” is the true stock value at time t ; “ d_t ” is the dividends at time t ; “ r_f ” is the risk-free rate and it is constant over time.

¹⁰ $a = 1/(1 + r_f)$; $0 < a < 1$; “ I_t ” is the information set, assumed common to all agents

$$E(E(P_{s+1}|I_s)|I_t) = E(P_{s+1}|I_t) \text{ and } E(E(d_{s+1}|I_s)|I_t) = E(d_{s+1}|I_t) \quad (3.4)$$

Now taking expectation of (3.2) with respect to information set I_t , and incorporating the result of (3.4), we will have:

$$E(P_{t+k}|I_t) = aE(P_{t+k+1}|I_t) + aE(d_{t+k+1}|I_t), k = 1, 2, 3 \dots \quad (3.5)$$

Substitute (3.5) into formula (3.2) and iterated to time T , we can have:

$$P_t = a^T E(P_{t+T}|I_t) + \sum_{k=1}^T a^k E(d_{t+k}|I_t) \quad (3.6)$$

If we further assume the future stock price is bounded, we can have the transversality condition:

$$\lim_{T \rightarrow \infty} a^T E(P_{t+T}|I_t) = 0 \quad (3.7)$$

So the first term of (3.6) was eliminated and hence the only solution to (3.2) under transversality condition is:

$$P_t^* = V_t = \sum_{k=1}^T a^k E(d_{t+k}|I_t) \quad (3.8)$$

P_t^* is the present value of expected dividends and thus should be deemed as the fundamental value of the asset. However, P_t^* is not the only solution to (3.2). Without imposing the transversality condition, differential equation (3.2) could have multiple solutions, and they could be generalized as:

$$P_t = P_t^* + B_t^{11} \quad (3.9)$$

Substitute equation (3.9) back into (3.2), we will have:

$$\begin{aligned} P_t^* + B_t &= aE(P_{t+1}^* + B_{t+1}|I_t) + aE(d_{t+1}|I_t) \\ &= aE(P_{t+1}^*|I_t) + aE(B_{t+1}|I_t) + aE(d_{t+1}|I_t) \end{aligned} \quad (3.10)$$

Deduct the original equation (3.2) from (3.10), and we will get:

$$B_t = aE(B_{t+1}|I_t) \quad (3.11)$$

¹¹ “ B_t ” is the bubble component at time t .

So P_t in the form of (3.9) and (3.11) is the solution system to (3.2) without the transversality condition. (As we will discuss later, many follow-up works in the field of the rational bubble theory is to eliminate some implausible solutions by either mathematic condition or economic theories.) It tells us the current price of an asset reflects not only the fundamental value, but also the discounted price of the future bubble. And market participant would not earn excessive return, even with presence of a bubble.

Now we want to know what property could B_t have? Through iteration on (3.11), we will further get:

$$B_t = a^k E(B_{t+k}|I_t) \quad (3.12)$$

Equation (3.12) is the simplest form of a bubble, and it is called “deterministic bubble” by Blanchard & Watson (1983). In this case, from time t to time $t+k$, the bubble must grow at the rate of $1/a$, or equivalently $1 + r_f$, each period, in order to let the rational investors hold it. A further example given by Blanchard & Watson (1983) is that the “collapsing bubble” persists stochastically in each period only with probability of π and bursts with probability of $(1 - \pi)$. If the bubble continues, it has to grow in expectation at a rate of $(1 + r_f)/\pi$. This faster bubble growth rate (conditional on not bursting) is necessary to achieve an expected growth rate of r_f . The probability that this stochastic bubble lives for n periods converge to 0 as n grows large, so bubble can exist even though rational investors know they will eventually burst.

Now back to the “deterministic bubble”, since $0 < a < 1$, when $k \rightarrow \infty$, taking limit towards (3.12) and we will get:

$$\lim_{k \rightarrow \infty} E(B_{t+k}|I_t) = \lim_{k \rightarrow \infty} B_t/a^k = \begin{cases} +\infty, & \text{when } B_t > 0 \\ -\infty, & \text{when } B_t < 0 \end{cases} \quad (3.13)$$

What is the implication of equation (3.13)? We can see a negative bubble cannot emerge since (3.13) implies that the asset price has to become negative in

expectation at some point in time (like time $t+k$). Once investors anticipate a negative price in the future (like time $t+k$), they would not short the assets and trigger the negative bubble in the first place (like time t).

Blanchard and Watson's theory formulate a basic picture of what a rational bubble looks like. It grows at a constant rate (maybe an increasing rate, if agents are risk averse) and a negative bubble would never happen. On the one hand, inspired by their pioneering works, many researchers have proposed new types of rational bubbles like: Intrinsic Bubble by Froot & Obstfeld (1991), Periodically Collapsing Bubble by Evans (1991), etc. On the other hand, many researchers have put sufficient energy to reduce the size of the solution sets to the difference equation. Tirole (1982) showed that a positive rational bubble in equation (3.13) would not exist with constant number of asset holders and infinite time horizons; and Tirole (1985) shows that a rational bubble can arise in asset prices in a model with an infinite succession of overlapping generations of asset holders with finite planning horizons, as long as the growth rate of the economy is greater than or equal to the required rate of return. Diba & Grossman (1988) argue that negative rational bubbles cannot exist because it would imply that investors expect that the price of the asset will become negative at a finite future date.

The rational bubble theory only describes the condition the bubble relies on to sustain itself and the scenario of how asset prices grow. It is incapable of capturing the subtle interaction between the bubble and the market participants. This contributes a huge disadvantage for it. Furthermore, rational bubble theory cannot survive in the case that agents have adaptive expectation, which is very likely to be true in China, or the case that some market limits, including limited life of assets, limited personal wealth and finite number of market participants, exists. Beside of these theoretical constraints, rational bubble theory in empirical test cannot coincide with rapid asset appreciation in a very short period, which is just what has happened in China. Thus, I

believe rational bubble theory would not be a plausible explanation to the stock bubble phenomenon in China.

3.2.2. Irrational Bubble Theory

When the price appreciation in the stock market is so intensive that it exceeds the reasonable explanatory range of the rational bubble theory, we should consider irrational bubble theory. Irrational bubble theory emphasizes the irrationality of the market participants and there are basically three influential and popular explanations in the irrational camp.

3.2.2.1. Feedback Theory of Speculative Bubbles

In the book of “Irrational Exuberance” by Shiller (2000), he provided a summary of the feedback loop theory, to investigate the mechanism of a speculative bubble. Feedback loop theory describes a phenomenon that initial price increase could lead to higher sequential prices, by feed backing the current price into the future prices through the assets demand function of the next period.

There are two popular versions of the feedback theory. The first one assumes agents behave with adaptive expectation. In psychology, representativeness bias helps to explain why many investors seem to extrapolate price movements: investors see an investment with recent price increases as representative of longer-term successful investments. So if the current price increases, they will set current price as a reference point, and hence increase their expectations of future prices relative to it. The above example is a simple price-to-price feedback: investors modify their expectation of future price based on the current price. There are also lots of deviated forms of feedback theory stemming from adaptive expectations. For example, Akerlof (2009) mentioned the price-to-GDP-to-price feedback which conjectures that as the price of the stock market increase, investors start to feel they get wealthier and hence consume more than before; their consumptions impact positively on the GDP; investors

interpret the advance of the GDP as an evidence of improved economy rather than the consequence of a bubble, and hence they adjust upwards their expectation of price, which finally caused them to bid the price up further more. According to Akerlof (2009), the price-to-corporate earning-to-price is also a form of the adaptive expectation version. The only difference is that it traces the corporate earnings, instead of GDP, as an improving factor in the economy. Another popular version of the feedback theory assumes that a price increase could produce a positive feedback on investors' confidence, and thus push them to heighten their demands for future stocks. This version is believed to be more fitting to the consistent price increase pattern, rather than a sudden price increase. The rationale behind this is that people build up their confidence slowly, and the confidence can only be established by a continuous sequence of successes.

There is empirical evidence to support the feedback theory. For example, De Bondt (1993) studied 38,000 forecasts of stock prices and exchange rates by surveying peoples. He found that non-experts would most likely to expect past trends in prices to continue. It could be observed that they are optimistic in bull markets and pessimistic in bear markets.

Feedback fits the reality better than the rational bubble theory. Firstly, it doesn't limit itself to a constant bubble growth rate, but could accommodate a huge price increase in a short time, as long as the "enthusiasm-ness" in investors' expectation is high. Secondly, unlike the rational bubble theory, feedback theory could predict a negative bubble: when people observed the negative price movement, they will adjust downwards their future expectations and hence push the stock market to a lower position. It is just a reverse direction of the standard statement. However, feedback theory didn't explain the burst of the stock bubble very well. In Shiller's book, he use the argument that "According to the adaptive expectation version of

feedback theory, we would expect a bursting of the bubble, since investors no longer think prices will continue to rise and therefore no longer see a good reason to hold the stocks.” Such argument is really weak and it didn’t give a plausible reason to us why a stock market could crash suddenly, like in one day, since investors adjust their expectation slowly.

3.2.2.2. *Herding Behavior*

From the prospective of human psychology, individuals tend to do what the group would do. The reason of this herding behavior has been given by psychologist from several angels: people may feel the social pressure of expressing an opposite opinion; or they trust the authority of some experts in the group, or they simply feel that when a large group of people reach agreement on a question, they are almost certainly right.

As Shiller (2000) said in his book: “The (herding) behavior, although individually rational, produces group behavior that is, in a well-defined sense, irrational. This herdlike behavior is said to arise from an *information cascade*.” What Shiller said implies that imitating the behavior of other investors and disregarding his own information may be the best strategy for one investor. However, if everyone holds this kind of idea and reply on others’ decision, the group decision made by them may not be the most desirable one. It is because in this case they cannot make use of each other’s information, since they do not reveal their own information to others when they merely follow them.

It should be noted that not only the weakly informed individual investors would suffer from the herding behavior, but also the professional investors would conduct similar behaviors. Their herding behavior doesn’t follow the similar psychological biases as the individual would do, but originated from a series of job security consideration and performance consideration. If a fund manager is losing

money while the others are making money, the fund manager's job may be in danger. If a fund manager is losing money while the others are also losing money, there is more job security. So it could be beneficial for the job-concerned managers to follow the group behavior, not to beat the market, but merely to keep their jobs. Another herding scenario is that fund managers have to participate in the stock bubble and follow the others, since there would be tremendous performance pressure if they didn't. Most funds set the average fund return as a benchmark of their manager's performance. So if a fund manager ignores the opportunity of bubbles, his performance will be lagging behind the others in short run, and he will very likely be fired, (although he is surely right in the long run).

3.2.2.3. *Noise Trader Theory*

Long, et al. (1990) has shown a noise trader model, to explain the pricing bubble with the presence of noise traders and sophisticated investors in the market. They defined the noise trader as the type of trader that falsely believes that they learned some pseudo-signals of a stock and irrationally trade stocks based on such pseudo-signals. They adopted the overlapping generations model of Samuelson (1958) with the above two types of agents that each lives for two periods. In period " t ", there would be no consumption and no bequest for both agents, the only thing they are allowed to do is to choose an optimal portfolio to hold. In period $t+1$, their portfolios could be sold at a certain price and the wealth would be transformed into consumption in this period.

There are two asset classes that both pay r dividends at time $t+1$. The risk-free asset called s is convertible to equivalent units of goods at any time. S is in perfectly elastic supply and its price level is fixed at 1. The risky asset called u is only tradable between traders at p_t in time t . It can only be redeemed into consumption at p_{t+1} in time $t+1$. U is not in perfectly elastic supply and thus its quality is normalized to one

unit. The noise trader is denoted by n and their proportion in the market is μ ; the sophisticated trader is denoted by i and their proportion in the market is $1 - \mu$. These two types of traders will hold different beliefs on the distribution of p_{t+1} , and select their portfolios accordingly based on the principle of utility maximizing. They believe the misperception of the noise traders should be market-wide rather than idiosyncratic. Otherwise, arbitrageurs would easily eliminate such mispricing. Thus, they assumed noise trader's misperception at time t about the expected price of the risky asset at time $t+1$ should be represented by an i.i.d.¹² normal random variable $\pi_t: \pi_t \sim N(\pi^*, \sigma_\pi^2)$. π^* reflects the average bullishness (or bearishness) of the noise traders, and σ_π^2 is the variance of noise traders' misperceptions.

They assumed the agents measure their utility by a constant absolute risk aversion function of wealth at time $t+1$, which is:

$$U = -e^{-2\gamma w} \quad ^{13} \quad (3.14)$$

Maximizing equation (3.14) is equivalent to maximizing the following equation, since the return is assumed to be normally distributed:

$$\bar{w} - \gamma \sigma_w^2 \quad ^{14} \quad (3.15)$$

Understanding this, we can write down the respective expected utilities that the sophisticated traders and the noise traders try to maximize below:

$$E_t(U^i) = C_0 + \lambda_t^i [r + E_t(p_{t+1}) - p_t(1 + r)] - \gamma (\lambda_t^i)^2 [E_t(\sigma_{p_{t+1}}^2)] \quad ^{15} \quad (3.16)$$

¹² independent and identically distributed

¹³ γ is the coefficient of absolute risk aversion. w is the wealth at time $t+1$

¹⁴ \bar{w} is the expected wealth at time $t+1$, and σ_w^2 is the one period ahead variance of wealth.

¹⁵ C_0 is the fixed income at time t . $E_t(\sigma_{p_{t+1}}^2) = E_t\{[p_{t+1} - E_t(p_{t+1})]^2\}$ is the variance of the price at time $t+1$. λ_t^i and λ_t^n are the holdings of risky assets for sophisticated traders and noise traders, respectively.

$$E_t(U^n) = C_0 + \lambda_t^n [r + E_t(p_{t+1}) - p_t(1 + r)] - \gamma(\lambda_t^n)^2 [E_t(\sigma_{p_{t+1}}^2)] + \lambda_t^n(\pi_t) \quad (3.17)$$

As we can see, the only different between equation (3.16) and (3.17) is noise traders' misperception term about the risky assets' return. Taking first order derivatives of equations (3.16) and (3.17), we can get the optimal quantities of risky assets that each trader should purchase as below:

$$\lambda_t^i = \frac{r + E_t(p_{t+1}) - p_t(1 + r)}{2\gamma E_t(\sigma_{p_{t+1}}^2)} \quad (3.18)$$

$$\lambda_t^n = \frac{r + E_t(p_{t+1}) - p_t(1 + r)}{2\gamma E_t(\sigma_{p_{t+1}}^2)} + \frac{\pi_t}{2\gamma E_t(\sigma_{p_{t+1}}^2)} \quad (3.19)$$

It should be noted that λ_t^i and λ_t^n can be negative, which means arbitrage is allowed in this model. Since the risky asset is not in elastic supply and its supply has been normalized to 1, we can set up the market clearing condition and plug equation (3.18) and (3.19) into it:

$$\begin{aligned} & (1 - \mu)\lambda_t^i + \mu\lambda_t^n \\ &= (1 - \mu) \frac{r + E_t(p_{t+1}) - p_t(1 + r)}{2\gamma E_t(\sigma_{p_{t+1}}^2)} \\ &+ \mu \left[\frac{r + E_t(p_{t+1}) - p_t(1 + r)}{2\gamma E_t(\sigma_{p_{t+1}}^2)} + \frac{\pi_t}{2\gamma E_t(\sigma_{p_{t+1}}^2)} \right] \\ &= 1 \end{aligned} \quad (3.20)$$

Equation (3.20) contains two unknowns (p_t and $E_t(p_{t+1})$), and thus we can solve the equilibrium price of the risky asset, p_t , in terms of p_{t+1} :

$$p_t = \frac{1}{1 + r} [r + E_t(p_{t+1}) - 2\gamma E_t(\sigma_{p_{t+1}}^2) + \mu\pi_t] \quad (3.21)$$

Equation (3.21) can be solved recursively and we will have the equilibrium price in the steady state:

$$p_t = 1 + \frac{\mu(\pi_t - \pi^*)}{1 + r} + \frac{\mu\pi^*}{r} - \frac{2\gamma}{r} [E_t(\sigma_{p_{t+1}}^2)] \quad (3.22)$$

Since the variance of the price is solely caused by noise traders' misperception, we can express $E_t(\sigma_{p_{t+1}}^2)$ in terms of:

$$E_t(\sigma_{p_{t+1}}^2) = \mu^2 \sigma_\pi^2 / (1 + r)^2 \quad (3.23)$$

Substitute equation (3.23) back into (3.22), we will get the final solution of risky asset's price:

$$p_t = 1 + \frac{\mu(\pi_t - \pi^*)}{1 + r} + \frac{\mu\pi^*}{r} - \frac{2\gamma}{r} \left[\frac{\mu^2 \sigma_\pi^2}{(1 + r)^2} \right] \quad (3.24)$$

In final equation (3.24) derived by Long, et al. (1990), “ I ” is the fundamental value of the risky asset. The second term in (3.24) implies that the price of the risky assets could be affected by the degree of variation of noise traders' misperception away from its mean value. The third term in (3.24) represents the price fluctuation due to the average noise traders' misperception. And the last term in (3.24) represents the uncertainty that drives the price down and its return up. This uncertainty stems from the belief of the noise traders in the next period. And the lower price and higher return driven by the last term is the compensation to the risk created by noise traders' speculation.

The noise trader model gives us pretty good intuition of why a bubble cannot burst with the presence of arbitrageurs (i.e. sophisticated traders). It is because the belief of noise traders is unpredictable in the future and thus arbitrage need to bear the risk that misperceptions become even more extreme tomorrow than today. In short, stock bubble could sustain, because of the limit of arbitrage. Considering the additional risk created by the noise traders, arbitrageurs would not bet too much to correct price way back to fundamental value. The noise trader model is a good model that captures the true subtle interaction between the arbitrageurs and the retail investors on Wall Street. It portraits the true reality that a large portion of the trading activities conducted by the professional arbitrageurs can be seen as a response to noise trading rather than as trading independently on fundamentals. However, it is may not

be wise to attribute Chinese stock bubble solely to the limit of arbitrage. First, noise trader theory assumes that noise traders hold misperceptions. This may not be the case in reality. Retail investors are not as simple and naïve as they are modeled. As I will discuss later in the Greater Fools Theory, only the last fool who found no other fools to resell his shares are the true fools. Otherwise, the players in the early stage of stock bubble are pretty smart. Second, noise trader theory describes a steady-state equilibrium stock price. It is not appropriate to apply it to a booming stock bubble, which is definitely not in steady state.

CHAPTER 4

METHODOLOGY

4.1. What is Greater Fools Theory?

The intuition of the “Greater Fools” theory originated from the famous Keynesian “Beauty Contest” theory in his General Theory. In the following famous passage, Keynes (1936) used metaphor to portrait the investment activities in the stock market by a beauty contest in which the winners are those who anticipate the average opinion. This is the first attempt made by economist to link the higher orders beliefs of human being to the stock market investment.

“Professional investment may be likened to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole: so that each competitor has to pick, not those faces which he himself finds prettiest, but those which he thinks likeliest to catch the fancy of the other competitors, all of whom are looking at the problem from the same point of view. It is not a case of choosing those which, to the best of one’s judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practice the fourth, fifth and higher degrees.”¹⁶

Keynes’s insights opened a brand new window for us to observe the financial market. Traditionally, many financial researchers hold a view that the price of an asset should be equal to the current expectations of its future pay-offs by a representative

¹⁶ The General Theory of Employment, Interest and Money, Page 140

agent. This statement is largely based on the property of a martingale process and the existence of a representative agent. For a martingale process, it has been shown that higher order expectations are redundant in the process of calculating the asset price by Tirole (1982), and thus expectation could be iterated over time in a martingale process. Iterated expectation implies one investors' current expectation of future expectation of future pay-offs should be equal to his current expectation of future pay-offs. Also, in a typical assets pricing literature, a financial researcher would create the so called "representative agent", since they believe a large group of investors on average can be represented by a uniform agent that holds a certain belief and preference. By assuming the existence of representative agent, one can conclude the average current expectation of the future pay-offs in the market should be in line with the representative agent's current expectation of the future pay-offs. Nevertheless, this simplification ignores the interaction between the investors when they are forming their expectations. We have to concede the settings of martingale process and representative agents are less prudent in a reality, if we consider the inspiration from Keynes' passage more carefully. Indeed, in Keynes' beauty contest example, he revealed the importance of higher order belief in people's decision-making process. By applying the same rationale in the assets pricing world, we can easily see that the calculation of asset price not only requires the understanding of investors' beliefs about future cash flows, but also of investors' beliefs about other investors' beliefs, and higher order beliefs. Hence, Keynes' beauty contest idea actually built the corner stone of the greater fool theory by reminding us the importance of higher order beliefs.

At the basis of the beauty contest, the Greater Fools Theory states that it does not matter if the price paid for an asset is higher than the fundamental value, as long as someone (the greater fool) is willing to pay an even higher price. The anticipation of other greater fools' move is a typical higher order belief. When acting in accordance

with the Greater Fools Theory, an investor buys securities without any regard to their qualities on his own belief, but with regard to their qualities on others' beliefs.

4.2. A Simple Mathematical Proof

According to the Greater Fools Theory, investors who participate in the stock bubble are still rational: they weigh the probability of further rises against the probability of falls. And we know that probability of further rise is associated with the supply of greater fools: if the supply of fools plummets, stock market are more likely to fall; if the supply of fools sours, stock market are more likely to rise. So it would be rational for an investor to buy shares, knowing that they are overvalued, as long as the probability weighted expectation of gain exceeds the probability-weighted expectation of loss. To maximize their return, they should hold overvalued stocks until the expectation of the stock return peaks, or in other words, until the supply of greater fools starts to decline. Selling too soon will cause a loss of potential profits for them.

Telser (2010) presents a model in scratch about “Greater Fools” to characterize several behaviors of speculators. I will use his proof as a basis and add some of my modification and understanding to show this theory in mathematics is consistent with what we observed in the market. First, let us define some notations in the model and make some assumptions to facilitate the discussion.

r = rate of return for each period; $0 < r < 1$

p_t = probability of finding a greater fool buyer in each period; $0 < p_t < 1$

$q_t = 1 - p_t$ = probability of not finding a greater fool buyer in each period; $0 < q_t < 1$

Assumption 1A: Assume finding a greater fool is a random event, and is independently and identically distributed (“I.I.D.”). Therefore, $p_t \equiv p, q_t = 1 - p, \forall t$. If the current stockholder is able to find a great fool and sell the stock to him, the

bubble will continue. Vice versa, the bubble will burst. Hence, the probability of a stock bubble that lasts for $t + 1$ period is $p^t q$.¹⁷

Assumption 1B: Expected return of holding the stock forever must be convergent; otherwise nobody would consider an exit strategy and they are all happy to hold it forever.

Based on the above two assumptions, we can write the expected return of holding a stock forever is structured as follow:

$$E(r, t) = \sum_{t=0}^{\infty} (1 + r)^t p^t q \quad (4.1)$$

The necessary condition for the expected return to be convergent is:

$$(1 + r)p \leq 1 \quad (4.2)$$

Using the expected return equation, we can infer certain behaviors of the investors from this simple model and check whether such behaviors are consistent with the reality.

4.2.1. Optimal Duration to Hold A Stock

$$E(T) = \sum_{t=0}^{\infty} (t + 1) p^t q \quad (4.3)$$

The expected optimal duration to hold a stock, “ $E(T)$ ”, is simply the sum of the probability weighted time in the right hand side. “ $t+1$ ” measure the time of a bubble cycle and “ $p^t q$ ” is the probability of a bubble lasts for “ $t+1$ ”. A good example of this formula is: when $p=1/2$, the bubble burst at time 1 with probability of $1/2$, at time 2 with probability of $1/4$ and so on. After applying the weighting of the probability, the optimal duration, “ $E(T)$ ”, equals $1/2 * 1 + 1/4 * 2 + 1/4 * 3 +$

¹⁷ An I.I.D. distribution is quite unreliable in reality, but purely for the matter of simplicity.

$1/16 * 4 + \dots = 2$.¹⁸ Hence in this example, the investor should hold the underlying assets for only 2 time periods.

The exogenous variable in this formula is “ p ”, which implies the higher probability one can find a greater fool in each period; the longer one should hold the stock in average. This is consistent with the observation that investors participate in the early stage of a bubble hold the stock longer, since the probability of finding a greater fool is higher at the initial point. We should also note that the duration is not related with the return (“ r ”) at all. This may suggest that in a bubble, investors are not referring the price and return information to optimize their holding methods. This finding is at least partly true in some cases. In the DotCom bubble, many investors still long the overvalued DotCom stocks one night before the crash. They totally ignore the warning sent by the high price level. Why? At least one reason is they believe more people will join and continue this game.

4.2.2. Net Expected Returns for Holding a Fixed Period

We already know the optimal duration one should hold the stock for. But what is the optimal time point to enter and exit the stock market? Let’s calculate the net expected return for an investor who was supposed to enter the market at time “ t ” and exit it at time “ $t+s$ ”.

$$\begin{aligned}
E(\text{Net Return}) &= E(\text{Return} - \text{Cost}) \\
&= E(\text{Return} | \text{Bubble not bursts before time } "t + s") \\
&\quad + E(\text{Return} | \text{Bubble bursts before time } "t + s") \\
&\quad - E(\text{Cost at time } "t") \\
&= p^s(1+r)^{t+s} + \sum_{i=1}^{s-1} p^i q^{s-i} * 0 - (1+r)^t \\
&= (1+r)^t [(p^s(1+r)^s - 1)]
\end{aligned} \tag{4.4}$$

¹⁸ Let $\frac{n}{2^n} = s(n)$. $(1/2)s(n) = n/2^{(n+1)} = (n-1)/2^n$. Hence $(1/2)s(n) = s(n) - (1/2)s(n) = 1/2^n$. Since $\sum_1^\infty 1/2^n = 1$, we can have $\sum_1^\infty s(n) = 2$.

From the previous convergence condition (4.2), we know $(1 + r)p \leq 1$. Now let us have a look at this final product of the net return, $(1 + r)t[(ps1+rs-1)]$. Since the second term is negative, the net return will in best case only hit a zero profit, which indicates that investing based on “Greater Fools” theory is not a profitable strategy. How come it just achieves a negative or zero profit? Is there anything wrong with the model setting? Is it because of the I.I.D. distribution of the probability? Let’s hold it for a moment, and I will discuss it in the next section immediately.

We also found that the net return are only related to “ t ” and “ s ”, the larger “ t ” and “ s ” is, the larger one’s loss. It means the later one enters the market and the longer one holds this stock, the fewer return one will get. It is consistent with our empirical observation that only early birds buy in earlier and hold for relatively shorter period could win in a stock bubble. Buying too late and holding too long will make one suffer from the crash. It is also consistent with the demographical logics that the supply of greater fools is limited and cannot be sustained forever.

4.2.3. *Net Expected Return and Convergence Condition*

In the last section, we derived the result that the net expected return is negative, which is not quite consistent with the reality. Is it due to our simplified assumption that investors receive nothing when the market crashes? We know that in the real world, even when the investors cannot find a greater fool and market collapses, they could still recover a proportion of their wealth by cutting their losses. So let us further assume:

$\varphi = \text{the proportion of investment recovered in the event of a bubble burst}$

After injecting this parameter, let us consider the net expected return in a two periods’ case to illustrate its effect:

$$E(Return) = p(1 + r) + \varphi * q - 1 \quad (4.5)$$

In period 1, investors invest one dollar in a certain asset. In period 2, they could earn $1 + r$ dollars with probability of p , or they could recover φ dollar of their investment with probability of q . So it seems that the new parameter, φ , has the power to flip the expected return back to positive, and to satisfy the convergence condition $(1 + r)p \leq 1$ at the same time, when:

$$p(1 + r) + \varphi * q > 1 \quad (4.6)$$

However, if we incorporate the concept of “loss cutting”, the convergence condition would not be the same as “ $(1 + r)p \leq 1$ ” anymore. The new convergence condition should be exactly the same as “ $p(1 + r) + \varphi * q \leq 1$ ”, if we think over the intuition of the convergence condition more carefully. From the two periods’ case, we observe that the expected return is mathematically equivalent to the convergence condition. Thus, we can conclude the negative expected return in the previous section is purely caused by strictly application of the convergence condition in all periods, rather than the lack of “loss cutting” behavior in the model formation. If we relax the convergence condition in some periods but strictly obey it when the time approaches infinity, the expected return would firstly increase in the positive range and then decrease to the negative range, which fits the reality very well.

But is it true that the convergence condition could be initially not bounded, and eventually bounded with the only varying exogenous parameter “ p ”? The answer is “yes”! As long as the probability increases over a threshold level, the convergence condition could be relaxed and become greater than 1 at some time. The expected return would be positive if the convergence condition exceeds 1 and be negative if it is not. The following graph shows the sensitivity of the convergence condition. In the graph, “ X ” axis describes the probability of “ p ”; “ Y ” axis describes the loss-cutting

ratio “ φ ” and “ Z ” axis describes the value of the convergence condition. For a given value of “ φ ”, we could find plenty of “ p ” to support the value of the convergence condition to be greater than 1. We can also see the net convergence condition, as well as the net expected return, will increase in “ p ” and “ φ ”. This graph gives us a very intuitive way to see the linkage between the Greater Fools Theory and the return.

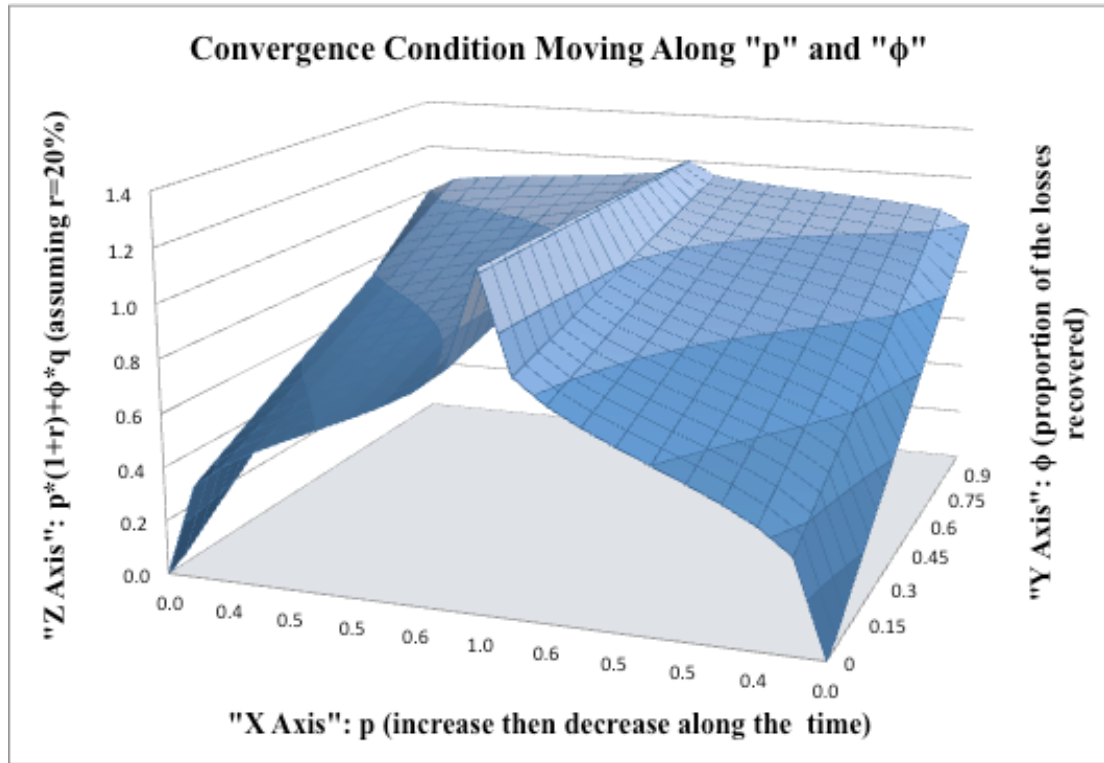


Figure 2 *Sensitivity of the Convergence Condition*

4.2.4. *Heterogeneous Beliefs of Investors*

In our basic model, we treat all investors equivalently as fools and assume all investors behave in the same way. Besides, our model involves no uncertainty and expectation in the parameters. Now we want to relax the assumptions on parameters to accommodate some heterogeneous beliefs of investors.

Let's assume all investors are facing a decision-making problem at time $t-1$: whether or not to stay the market in the next period. Investors purchased the stock at "1". In contrast to the constant rate of return we used before, we allow time varying returns (r_{t-1}, r_t, etc) and we also allow the investors to form different expectations about returns at time t , based on the observable return at time $t-1$ now. Investor will earn a payoff of $1 + E_{t-1}(r_t)$ when they successfully sell it to the next fool. Loss cutting ratio will be ignored, instead, we will assume investors could have earned an alternative return, " $1 + E_{t-1}(r'_t)$ ", when they cannot find a greater fool with a probability of " q ". Since the investors make the decision at time $t-1$, they will try to maximize the following equation.

$$E_{t-1}(Return) = p_t[1 + E_{t-1}(r_t)] + q_t[1 + E_{t-1}(r'_t)] - 1 \quad (4.7)$$

It is a two periods' case. The return at time t was formed into expectation, ex ante, on time $t-1$. Without loss of generality, we assume there are two types of investors. One type is fundamental investor. They believe in the fundamental value of a stock estimated by future cash flows. They hold this belief, maybe because they are long-term investors in the market and they trust the mean reversion in the long run. They believe stock price will eventually come back to its value, which is also assumed to be "1" in this case. The other type is the momentum investor. They don't buy the idea of fundamental value. They believe market is efficient, and the latest price already reflects the value of the stock. According to the above intuition, we can make the following assumptions about these two types of investors.

Assumption 2A: Fundamental investors believe they can only sell the stock at its value ("1"), when they cannot find a fool to resell it. Hence their expectation of r'_t will be equation (4.8).

$$E_{t-1}^F(r'_t) = f(r_{t-1}) = \frac{1 - (1 + r_{t-1})}{1 + r_{t-1}} = \frac{-r_{t-1}}{1 + r_{t-1}} < 0 \quad (4.8)$$

Clearly, the derivative of $E_{t-1}^F(r'_t)$ with respect to r_{t-1} is:

$$\frac{d[E_{t-1}^F(r'_t)]}{d(r_{t-1})} = \frac{-1}{(1+r_{t-1})^2} < 0 \quad (4.9)$$

It implies the fundamentalists believe “higher last period’s return is, lower the alternative return will be in this term.”

Assumption 2B: Momentum investors believe they can sell the stock at least at the last period’s price (“ $1 + r_{t-1}$ ”), even when they cannot find a fool to sell it. It may be because they believe last period’s price already reflected the latest value of the stock. Hence their expectation of r'_t will be greater than 0 as equation (4.10) indicates.

$$E_{t-1}^M(r'_t) = f(r_{t-1}) \geq \frac{(1+r_{t-1}) - (1+r_{t-1})}{1+r_{t-1}} = 0 \quad (4.10)$$

Since they believe in the momentum of stock investment, it is naturally to assume the derivative of $E_{t-1}^M(r'_t)$ with respect to r_{t-1} is also greater than 0:

$$\frac{d[E_{t-1}^M(r'_t)]}{d(r_{t-1})} \geq 0 \quad (4.11)$$

It implies the momentum investors believe “higher last period’s return is, higher the momentum will be resided in this term’s alternative return.”

Assumption 2C: Both types of investors are earning the same return, when they can sell the stock to the next fool, since we believe the fools can also reach a consensus about the price they offered to buy stocks. Therefore:

$$E_{t-1}^F(r_t) = E_{t-1}^M(r_t) = E_{t-1}(r_t) \quad (4.12)$$

Now from equation (4.7), we can solve for a threshold probability for both types of investors to stay in the market. This “ p_t^* ” is the probability that makes the investors breakeven.

$$p_t^* = \frac{E_{t-1}(r'_t)}{E_{t-1}(r'_t) - E_{t-1}(r_t)} \quad (4.13)$$

The derivative of p_t^* with respect to r_{t-1} can be easily shown by chain rule for both types of investors:

$$p_t^F = \frac{E_{t-1}^F(r'_t)}{E_{t-1}^F(r'_t) - E_{t-1}(r_t)} \text{ and } \frac{d[p_t^F]}{d(r_{t-1})} > 0 \quad (4.14)$$

$$p_t^M = \frac{E_{t-1}^M(r_t)}{E_{t-1}^M(r_t) - E_{t-1}(r_t)} \text{ and } \frac{d[p_t^M]}{d(r_{t-1})} \leq 0 \quad (4.15)$$

From (4.14), we can see fundamentalist will have a higher threshold probability (of finding a greater fool) to stay when they saw a higher r_{t-1} , while the momentum guys will have a lower threshold probability instead. Therefore, our model will accommodate two types of investors who hold heterogeneous beliefs. Fundamentalist will be more likely to leave the market, while the observed return (" r_{t-1} ") is rising, because the threshold probability for them to stay is getting tougher as a bubble thrives. In contrast, the momentum guys will be more likely to join the market, while the observed return (" r_{t-1} ") is rising, because the threshold probability for them to stay is getting easier as a bubble thrives.

4.2.5. *Remarks*

On the first hand, after discussing some implications from this mathematical model, we found that the probability of finding a greater fool, or " p_t ", is important as it determines both the net expected return and the optimal holding period for investors in a stock bubble. And because of demographic reasons, the probability of finding a greater fool is directly linked with the supply of greater fools. So if we could use supply of greater fools as a proxy to examine the relationship of " p_t " and return " r_t ", the mathematical proof above will be empirically tested. The null hypothesis for this is "as the probability of finding a fool increases, the return should also increase."

On the other hand, we also found in section 4.2.4 that the observed return " r_{t-1} " and threshold probabilities " p_t " of heterogeneous investors are correlated. For fundamentalist, as r_{t-1} goes up, p_t will goes down. For momentum guys, as r_{t-1} goes up, p_t will also up. It implies the fundamentalist could work as a stabilizing force against the greater fools. Therefore, remaining hypotheses are "as the observed return of last period increases, the probability of finding a fundamentalist will decrease" and "as the observed return of last period increases, the probability of finding momentum

investors will increase. These hypotheses will not be officially tested, but will be casually discussed in section 5.5.

In summary, the general process of a bubble under a Greater Fools Theory would most likely follow the processes below:

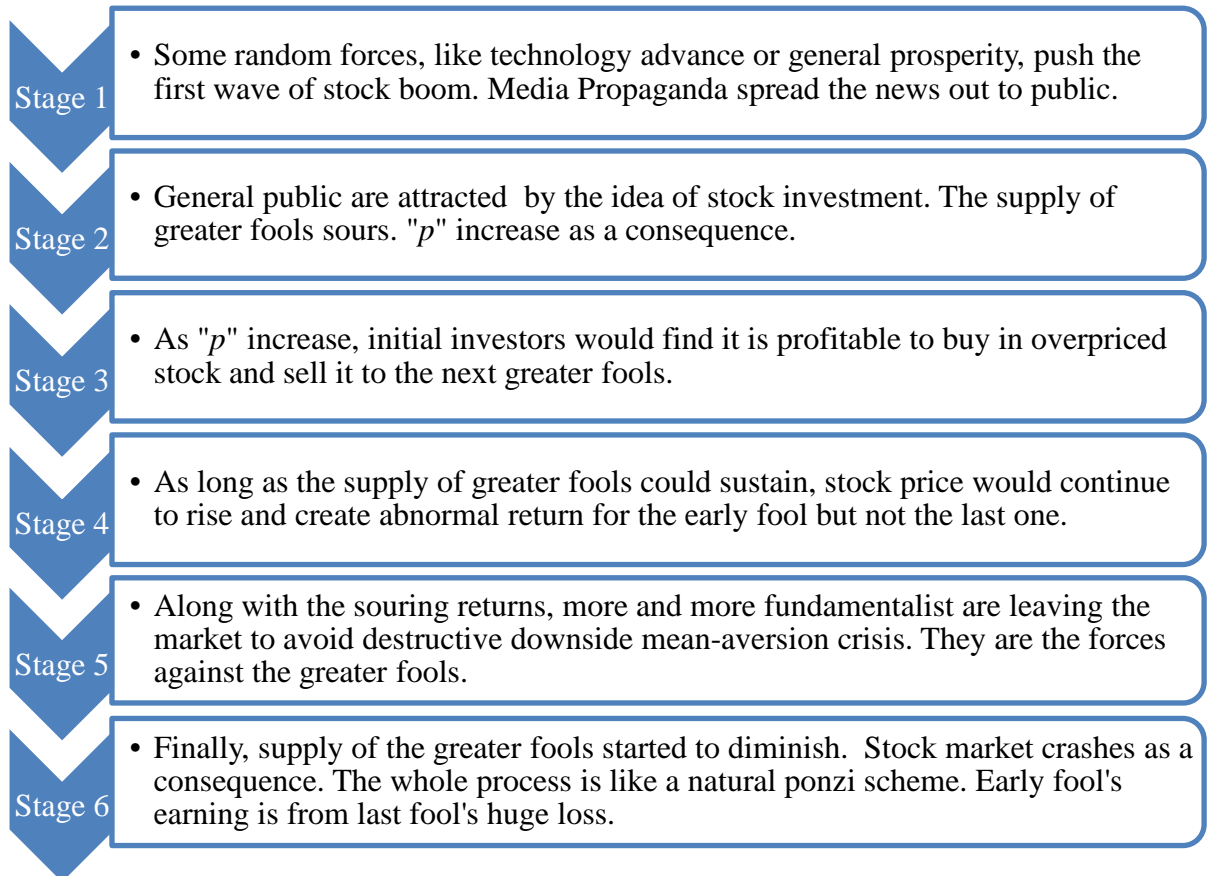


Figure 3 *The Process of Greater Fools Theory*

4.3. The Factor Model

The assets pricing model I chose to test the theory is the APT model. According to my suggestion above, the explained variable will be the return. And the explanatory variable will include the supply of greater fools and some others.

APT model contains the following major assumptions, according to Ross (1973):

- i. Investors are risk averse and seek to maximize their terminal wealth. In other terms, preferences are monotonic and convex.
- ii. Investors can borrow and lend at the risk free rate.
- iii. There are no market frictions (transaction costs, taxes, or restrictions on short selling).
- iv. Investors agree on the number and identity of the factors that are important systematically in pricing assets.
- v. There is no riskless arbitrage profit opportunities left in the market.

Although (iii) and (v) are probably not true in Chinese market, it should be clear that APT model requires fewer assumptions than any other assets pricing models. Thus it may not be the best model to fit the reality, but must not be the worst to do so.

APT model is basically a multi-factor model. However, as we have mentioned in the literature review session, the specific risk factors have not been identified or discovered in the APT model. It requires practitioner to employ some ad-hoc specifications of factors, in order to capture the systematic risks involved in the investments. Such theoretical structure creates tremendous convenience for me. I can add a variable that proxies the supply of greater fool freely and safely into the model as a risk factor. However, I have to concede that market may be mainly lead by the greater fool factors but it could also be affected by some other factors. These factors should be macroeconomics related, since they capture variations in the underlying reasons why an asset's pay-offs and cash flows change over time. Indeed, risk factors can also be identified at a microeconomic level by focusing on relevant characteristics of the companies themselves, such like the size of the firm and the value of the firm. This approach is just what Fama & French (1992) did in their famous three-factor model. However, as I am studying the index as a whole, firm level risk should not be considered here.

In Equation (4.16), r_t is the observed asset return, r_f is the risk free rate and r_t^* is the risk premium. To facilitate the expression of my methodology, I created two new expressions, “equilibrium return” and “bubble return” in the right hand side of Equation (4.16). Through some simple algebra, the excess return ($r_t - r_f$) can be discomposed into two parts: the equilibrium return ($r_t^* - r_f$) and the bubble return ($r_t - r_t^*$).

$$r_t - r_f = (r_t - r_t^*) + (r_t^* - r_f) \quad (4.16)$$

The dependent variable on the left should be the excess return for the month.

ER = excess return of stock index over the risk free rate

In Equation (4.16), equilibrium return is the return that investors earn on the economic fundamentals. Any variation of macroeconomics conditions should give a change to the equilibrium return. Bubble return is the return that investors earn on the event of successfully finding a greater fool and selling stocks to him. The more supply of greater fools, the higher chance to find one and sell the stocks and hence the higher return. Thus any variation of the greater fools supply should give a change to the bubble return.

Inspired by the empirical work of (Chen, et al., 1986), I choose the following monthly macroeconomics variables in China for the inputs of my APT model to explain the equilibrium returns ($r_t^* - r_f$):

DIP = the growth rate in industrial production

DI = the change in inflation, measured by CPI

DL = the growth rate in aggregate liquidity, measured by M1 supply

DE = the growth rate in exports

DMMR = the change in money market rate

Thus the equilibrium return should be governed by a set of broad economic influences in the following fashion:

$$r_t^* - r_f = a_t + [b_{1t}DIP_t + b_{2t}DI_t + b_{3t}DL_t + b_{4t}DE_t + b_{5t}DMMR_t] + \varepsilon_t \quad (4.17)$$

To mimic the supply of greater fools, I choose the following two factors to explain the bubble returns ($r_t^* - r_f$).

DRI = the growth rate in retail investors

, measured by the "newly registered retail investor stock accounts"

DII = the growth rate in institutional investors

, measured by the "newly registered institutional investor stock accounts"

Based on these new factors, the bubble return should be governed by a set of variables that proxies the supply of greater fools in the following way:

$$r_t - r_t^* = a_t + [b_{1t}DRI_t + b_{2t}DII_t] + \varepsilon_t \quad (4.18)$$

Consolidating Equation (4.16), (4.17) and (4.18), we can build up the model for empirical tests as:

$$r_t - r_{ft} = a_t + [b_{1t}DIP_t + b_{2t}DI_t + b_{3t}DL_t + b_{4t}DE_t + b_{5t}DMMR_t + b_{6t}DRI_t + b_{7t}DII_t] + \varepsilon_t \quad (4.19)$$

where I will regress the monthly excess return against all of my monthly macroeconomic factors and greater fool proxies.

CHAPTER 5

EMPIRICAL RESULT

5.1. *The Data Source*

The data sources of this study are the most reliable and have been widely used by researchers. Some of the macroeconomic variables, including CPI, Export, Money Market Rates and Stock of M1, were retrieved from the EIU, “Economic Intelligence Unit”.¹⁹ The Industrial Production Index was retrieved from IMF, “International Monetary Fund”.²⁰ The Shanghai Stock Index information was retrieved from the CSRC, “China Securities Regulatory Commission”.²¹ The newly discovered data series, “Monthly Registered Stock Accounts”, was retrieved from CSD&C, “China Securities Depository and Clearing Corporation”.²² Since the public available information of the CSD&C only starts from January 2005, my data series is limited to start from the same time. The ending date is December 2009. Therefore, a total of 60 months’ observations have been included in my study.

5.2. *Summary Statistics*

In this section, I will try to describe and explain the raw and the transformed data series I am using. Shanghai Stock Index is the monthly average number, rather than a simple monthly close figure. Money Market Rate is serving both as the risk free rate and the basic lending rate in the model. The month-on-month industrial production growth figures in January 2007, 2008 and 2009 are missing from the data series. Several sources (like University of Michigan China Data Center, Chinese

¹⁹ www.eiu.com

²⁰ www.imf.org/external/data.htm

²¹ www.csrc.org.cn

²² www.chinaclear.cn

Statistic Bureau, etc.) contain such missing value as well and there is clearly no way to avoid them. Summary statistics of the raw data are provided in Table 1.

Table 1 *Summary Statistics of Raw Series*

Series Name	Obs.	Mean	S.D.	Minimum	Maximum
Shanghai Index	60	2556.04	1287.42	1042.18	5765.03
Retail Investors (000')	60	1211.243	1290.89	49.66	5600.00
Institutional Investors	60	4399.55	3957.75	304.00	18104.00
Money Market Rate (%)	60	3.00	1.03	4.77	1.38
CPI (indexed to 100)	60	106.69	5.40	99.00	114.50
Export (\$bn)	60	93.01	23.17	44.28	136.72
Industrial Production Growth (%)	60	15.20	3.83	5.40	22.95
Stock of M1 Money (¥bn)	60	13907.73	3432.18	9269.77	22000.20

APT model specifies the independent variables to be the incremental part of a macroeconomics variable. Since some of our raw data series, like the “Industrial Production Growth”, is incremental value, so it will fit with the model specification out of the box. While the remaining data are not aligned with the model specification, we need to take some transformation of them. The procedures are described below.

$$DIP_t = \text{Industrial Production Growth}_t$$

$$DI_t = \log(CPI_t) - \log(CPI_{t-1})$$

$$DL_t = \log(M1_t) - \log(M1_{t-1})$$

$$DE_t = \log(\text{Export}_t) - \log(\text{Export}_{t-1})$$

$$DMMR_t = \text{Money Market Rate}_t - \text{Money Market Rate}_{t-1}$$

$$DRI_t = \log(\text{Retail Investors}_t) - \log(\text{Retail Investors}_{t-1})$$

$$DII = \log(\text{Institutional Investors}_t) - \log(\text{Institutional Investors}_{t-1})$$

$$ER = \log(\text{Shanghai Index}_t) - \log(\text{Shanghai Index}_{t-1}) - \text{Money Market Rate}_t$$

Money Market Rate is measured in percentage terms, so it is only taken by the first order difference. All the rest data series were taken first order logarithm

difference to capture the percentage changes on the original data. After the transformation, the summary statistics of the transformed data are provided in Table 2.

Table 2 *Summary Statistics of Transformed Series*

Series Name	Obs.	Mean	S.D.	Minimum	Maximum
DIP (%)	59	15.10	3.79	5.40	22.95
DI (%)	59	0.22	0.67	-0.89	2.57
DL (%)	59	1.39	1.91	-4.50	6.07
DE (%)	59	1.61	12.26	-33.26	36.59
DMMR (%)	59	-0.02	0.53	-2.11	1.75
DRI (%)	59	5.77	49.30	-103.35	154.12
DII (%)	59	4.43	48.84	-149.73	139.24
ER (%)	59	-1.37	9.18	-21.54	16.53

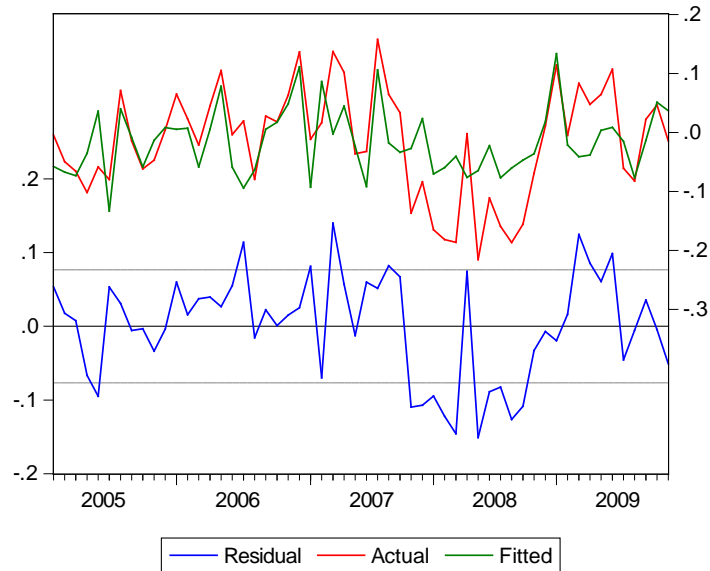
5.3. OLS Estimates

Table 3 shows the regression results estimated via OLS. Among the independent variables, it is found that the intercept, the changes in retail investor (DRI) and the liquidity (DL) are significant. And the R-squared amounts to 0.380, which should be considered as an acceptable result, since all the variables are already 1st order difference and explaining stock return usually doesn't get back a high R-squared. However, we cannot give the estimates any economic explanation before we conduct a series of tests against the classic assumptions of OLS. Most importantly, we should test the assumption of "no serial correlation and no heteroscedasticity". It is because if the classical assumptions don't hold, the OLS will not be BLUE and lose its efficiency even asymptotically.

Table 3 OLS Estimates

R-squared	0.380	Mean dependent var	-0.016
Adjusted R-squared	0.289	S.D. dependent var	0.091
S.E. of regression	0.077	Akaike info criterion	-2.168
Sum squared resid	0.282	Schwarz criterion	-1.878
Log likelihood	68.694	Hannan-Quinn criter.	-2.055
F-statistic	4.197	Durbin-Watson stat	1.262
Prob(F-statistic)	0.001		

Variable	Estimate	Std. Error	t-Statistic	Prob.> t
Intercept	-0.078	0.045	-1.736	0.089
DRI	0.127	0.033	3.831	0.000
DII	-0.061	0.040	-1.518	0.136
DE	-0.111	0.110	-1.015	0.315
DI	2.435	1.878	1.297	0.201
DIP	0.186	0.289	0.645	0.522
DL	2.036	0.700	2.908	0.006
DMMR	-0.162	2.063	-0.078	0.938

**Figure 4 Actual, Fitted and Residual Graph**

5.3.1. *Autocorrelation and Heteroscedasticity Test*

With time series data, the ordinary regression residuals are usually autocorrelated over time. In such a case, it will not be desirable to simply use ordinary regression analysis since the assumption that the errors are independent of each other will be violated. Therefore, a correlogram graph and Q-statistics are deployed to graphically detect if autocorrelation issue exists. From the length of the bar in Figure 5, we can roughly guess the residuals are not a white noise process. The residual is autocorrelated with its own lagged term with up to 5 periods and the significant Q-statistics and the lower p -values confirm such finding. Additionally, a Breusch-Godfrey Lagrange Multiplier test is also used to test for autocorrelation. Breusch-Godfrey test set the residuals as the dependent variable and regress it against all the original explanatory variables and lagged residuals. The null hypothesis is the residuals contain no serial correlation. After the regression, it employs the F-statistics and the chi-square-statistics for inference. In Table 4, the result of the Breusch-Godfrey test is presented with lag term of 1 to 5. If we set the critical interval as 5%, we will see residuals are autocorrelated with the lagged residuals up to three continuous periods. Furthermore, if we extend the critical interval to 10%, we will see the residuals are autocorrelated up to four periods. It is a very significant result. This confirms the residuals are not white noise, and thus an autoregressive model needs to be considered to improve the quality of the regression.

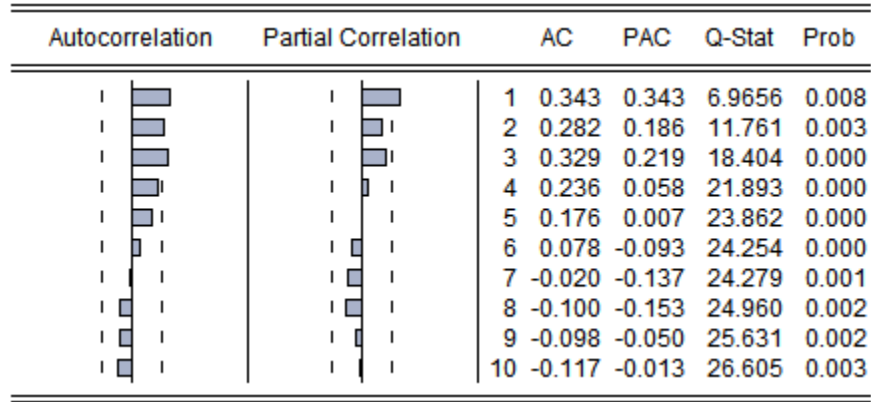


Figure 5 Serial Correlation Test I: Correlogram of Residuals and Q-Statistics

Table 4 Serial Correlation Test II: Breusch-Godfrey

Breusch-Godfrey LM Test	Lag=	1	2	3	4	5
F-statistic		8.178	5.942	4.800	3.919	3.027
Prob. F		0.006	0.005	0.006	0.008	0.020
Obs*R-squared		8.300	11.497	13.576	14.711	14.579
Prob. Chi-Square(2)		0.004	0.003	0.004	0.005	0.012

In a time-series financial data, the heteroscedasticity issue often appears in the autoregressive conditional form, which is widely referred as the “ARCH”. ARCH describes a phenomenon that fluctuations in volatility tend to be grouped into clusters when viewed over time: high volatility is most likely to be followed by high volatilities; low volatility is most likely to be followed by low volatilities. The intrinsic force to push this phenomenon could be market sentiment, rumor, common economic conditions, etc. In terms of statistics, we can see the variance of the error term at time t is autocorrelated with the squares of the previous error terms. Or in terms of mathematics, we can express it as below.

$$\text{Var}(u_t) = \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_n u_{t-n}^2 \quad (5.1)$$

To examine if ARCH exists in the residual of our model, we can initially check the correlogram of residual squared graph as shown in Figure 6. From the length of the

bars, we can see the residuals squared are at least correlated with up to three continuous terms. And the significant Q-statistics and the low p -value also depict the severity of the ARCH problem. Alternatively, we can also run a heteroscedasticity test for ARCH. The regression function of this test is just based on the equation above, regressing the residual squared terms against the lagged residual squared terms. I try different lag setting and store the results in Table 5. Once again, I confirm the fact ARCH phenomenon is persistent in the residuals squared. And it is significant at 5% level for four lagged terms. So this result reminds us the existence of the ARCH problem and it calls for an ARCH or a GARCH (“Generalized ARCH”) model to complement my study.

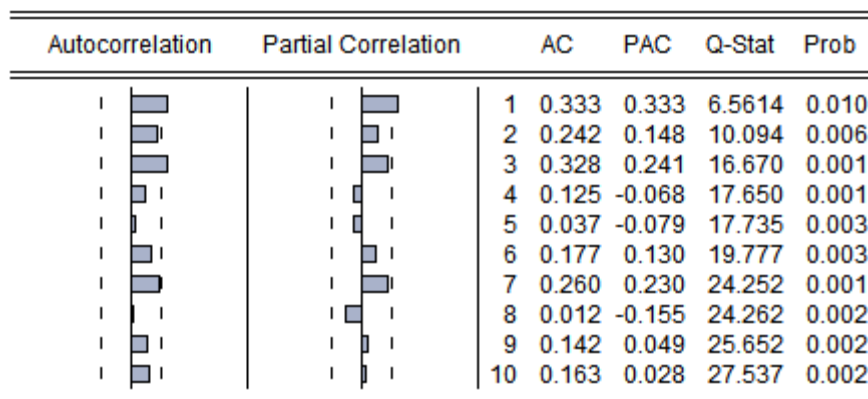


Figure 6 ARCH Test I: Correlogram of Residuals Squared and Q-Statistics

Table 5 Heteroscedasticity Test II: ARCH LM Test

ARCH LM Test	Lag=	1	2	3	4	5
F-statistic		6.775	3.310	3.741	3.055	2.192
Prob. F		0.012	0.046	0.019	0.029	0.082
Obs*R-squared		6.205	6.156	9.640	10.352	9.632
Prob. Chi-Square		0.013	0.046	0.022	0.035	0.086

5.3.2. *Stability Test*

Stability test is useful and necessary among time series studies. Apart from the test against some important classic assumption, we should also consider whether the coefficient we estimated is stable over time. If it is not, it means the model may need some modification or some dummy variable to be added, in order to accommodate such instability. The first testing method is the RLS (“Recursive Least Square”) estimation. Suppose there are “k” explanatory variables in the model, RLS will use the first k sample to derive a set of coefficient estimates, and then iterate such set of coefficients into the fitted value calculation for the k+1 sample. RLS will derive error terms for all the remaining n-k samples recursively and use such special error terms for inference. The CUSUM test is just a test against the special error terms of the RLS. In the following graph, the CUSUM test and the CUSUM Squares Test will draw the boundaries of the 5% critical interval, if the graph of the error terms go beyond the boundary, it will indicate a coefficient instability at such time point. From Figure 7, we can see that the CUSUM test does not indicate any model instability at 5% significance level. However, the CUSUM Squares test does find a potential model break point starting from January 2007 to February 2008. The error terms slightly draw outside the boundary, and we’d better use another test to further test whether our model is correct during this period. A Chow Breakpoint Test has been introduced to test the potential breakpoint starting from January 2007. From Table 6, the null hypothesis cannot be rejected at 10% significance level, which means the coefficients are consistent and stable across time. Therefore, we can conclude it is not necessary to make structural change to our model, like adding dummy variables or splitting the sample into sub samples according to certain events.

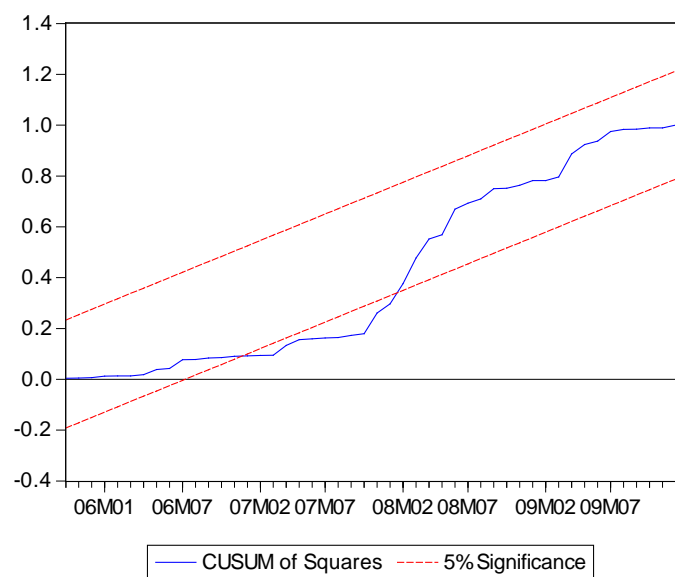
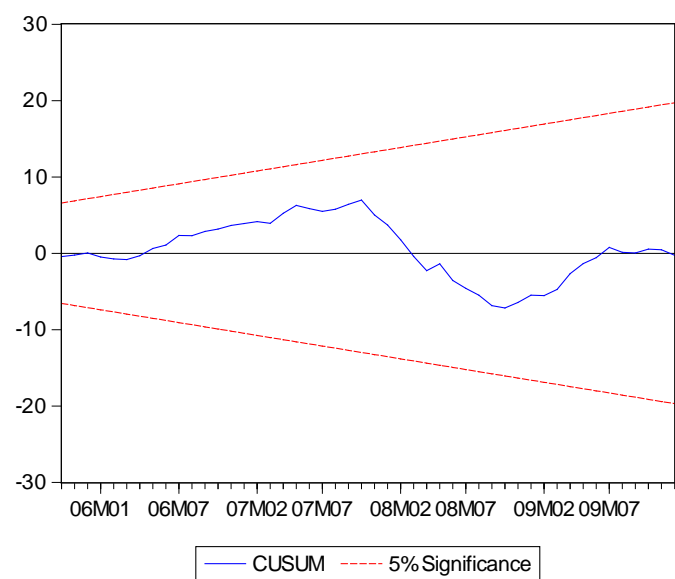


Figure 7 *Stability Test I: CUSUM Test and CUSUM Squares Test*

Table 6 *Stability Test II: Chow Breakpoint Test*

Chow Breakpoint Test:	2007M01		
Null Hypothesis:	No breaks at specified breakpoints		
Varying regressors:	All equation variables		
Equation Sample:	2005M02-2009M12		
F-statistic	1.200955	Prob. F(8,40)	0.3231
Log likelihood ratio	12.05486	Prob. Chi-Square(8)	0.1488
Wald Statistic	9.607638	Prob. Chi-Square(8)	0.2937

5.3.3. *Non-linear Regression*

Sometimes, the dependent variable may show some non-linear relationship with the independent variables. In a non-linear model, the estimated coefficients are still linear, but the variables are in the form of a non-linear function (exponential, log, inverse, square, etc.) of the originals. In graph, we can observe a curved regression line, rather than a straight one with a non-linear regression. Therefore, the quickest way to determine whether a non-linear model should be adopted is to observe the scatter graph. If the dependent variable and the independent variable show some special curvatures, we should be cautious and try some non-linear forms to fit the data. A series of scatter graphs and nearest neighbor fit lines have been drawn to gauge the curvatures between variables. Nearest neighbor fit is a nonparametric regression method that fit local polynomials, so it should help us to find the non-linear relationships between variables. From Figure 8, we can see all explanatory variables have no obvious non-linear relationship with the excess return. Although all the lines show intensive volatility, but not a single one shows any shape of rectangular hyperbola (inverse form in variables), parabola (quadratic form in variables), exponential curve or logarithm curve. Thus we believe there is no non-linearity between the variables and it is not necessary to conduct non-linear regressions. Furthermore, from Figure 5 I observe some scatter graphs are in the shape of a sphere,

and thus I suspect a linear relationship does not even exist in some graphs. So a redundant variable test should be conducted to eliminate those potentially insignificant variables.

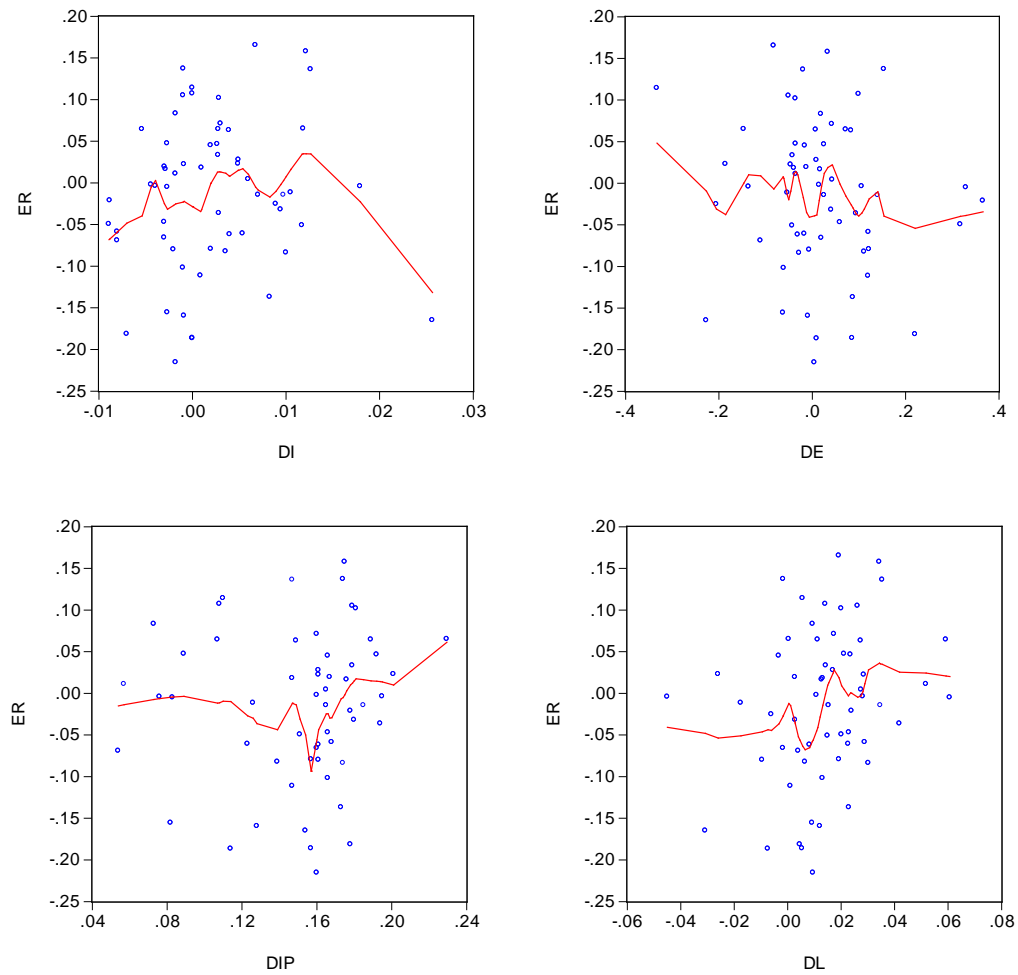
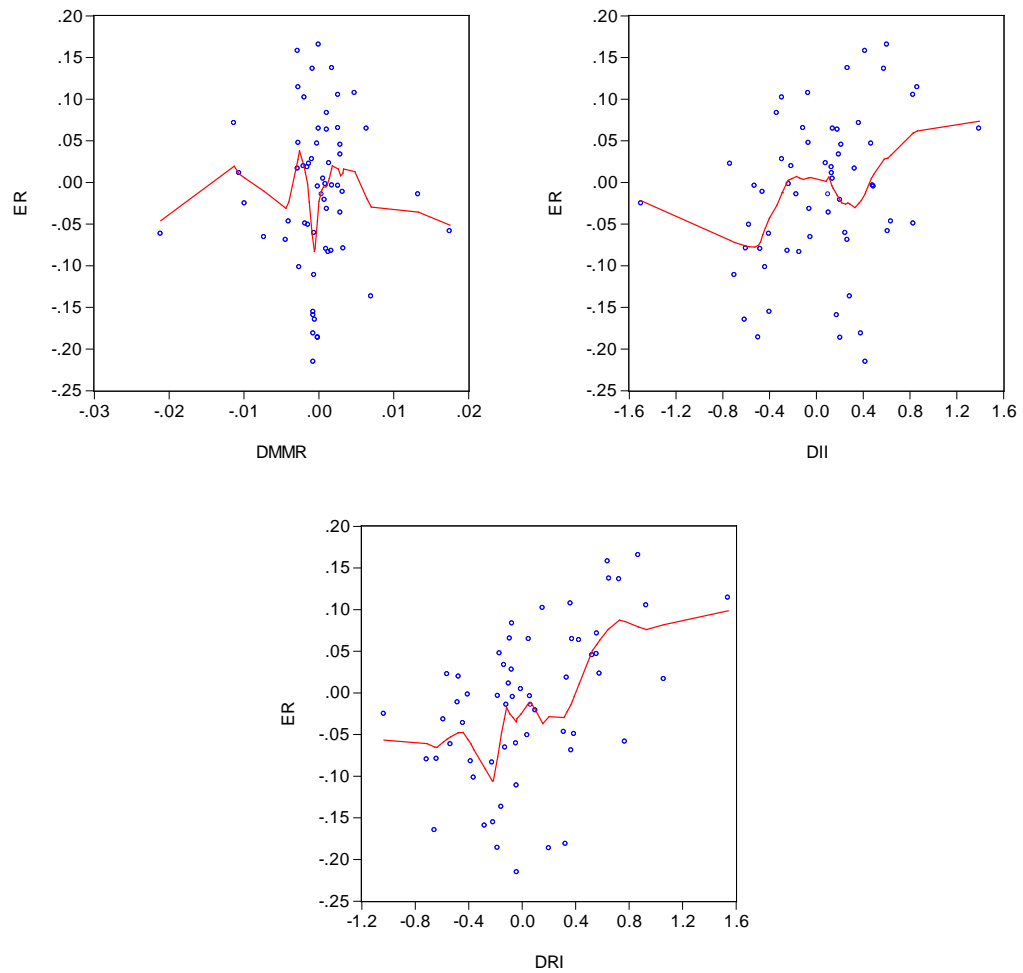


Figure 8 *Non-Linear Model Test: Scatter Graphs and Nearest Neighbor Lines*

Figure 8 (Continued)



5.3.4. Redundant Variable Test

The redundant variables test allows us to test for the statistical significance of a subset of included variables. More formally, the test is for whether a subset of variables in an equation has zero coefficients and might thus be deleted from the equation. The test statistics are the F-statistic and the Log likelihood ratio. The F-statistic is more reliable under finite sample, and the LR test is an asymptotic test. Given the limited number of observations in our test, we should rely more on the F-statistic. In Table 7, we can see only two variables DL and DRI are significant at even

1% level. DII is significant at the 15% level, but not at the 10% level. And DI is significant around the 20% level, but not at the 15% level. To be careful, I will keep DII and DI to see their importance in my following models. The other variables of DE, DIP and DMMR are not significant at an even higher level. To be more precise, I also test whether they three are collectively significant and the result is not desirable as well. Thus I will eliminate these three macroeconomic variables from my model.

Table 7 *Redundant Variable Test*

Redundant Variable Test	Variable=	DE	DI	DIP	DL	DMMR	DRI	DII
F-statistic		1.030	1.682	0.416	8.454	0.006	14.673	2.305
Prob. F(1,48)		0.315	0.201	0.522	0.006	0.938	0.000	0.136
Log likelihood ratio		1.189	1.928	0.484	9.085	0.007	14.937	2.626
Prob. Chi- Square(1)		0.276	0.165	0.487	0.003	0.933	0.000	0.105

5.3.5. *OLS Remarks*

Although we cannot draw too many economic explanations from the flawed OLS regression, we still learn a lot from it about what we should do next. First, an autoregressive model should be considered, which means an AR(X) structure should be added into the repressors. Second, a GARCH model should be adopted because of the autoregressive conditional heteroscedasticity issue. Third, the model is stable over the time; there is no need to add any dummy variable or conduct any sub-sample regressions. Fourth, there isn't any obvious non-linearity between the variables, so we don't need to transform any variable into a non-linear function of itself. Fifth, Three out of seven independent variables are significantly redundant and I will not include them in my following model constructions. And I can conclude from here that the change in export, industrial production and lending rate contribute very few to the change of excess return in Chinese stock market.

5.4. *AR-IGARCH-M Model*

Most of the time series statistical tools are designed to model the conditional mean of a random variable. On the contrary, GARCH will not only model the conditional mean but also the conditional variance, or volatility, of a variable. There are several reasons that we may wish to use GARCH model and thus forecast volatility. First, we may need to analyze the risk of holding an asset or the value of an option. Second, forecast confidence intervals may be time varying, so that more accurate intervals can be obtained by modeling the variance of the errors. Third, more efficient estimators can be obtained if heteroscedasticity in the errors is handled properly.

After several theoretic checks and empirical trials, I finally decided an AR(1)-IGARCH-M model should be used to derive the right estimates.

5.4.1. *Stationarity and Unit Root Test*

Before the injection of any AR(x) terms into the model, it should be clear that the theory behind ARMA estimation is based on stationary time series. A series is said to be stationary if the mean and auto-covariance of the series do not depend on time. For example, a most common non-stationary series is a random walk ($y_t = y_{t-1} + \varepsilon_t$). However, the first order difference of such random walk is stationary ($\Delta y_t = y_t - y_{t-1} + \varepsilon_t$). Since a random walk series could become stationary after a first order difference, it could be called $I(1)$, “integrated of order 1”. Also, we can call the random walk series contain one unit root.

Standard inference procedures do not apply to regressions that contain an integrated dependent variable or integrated regressors. Therefore, it is important to check whether a series is stationary or not before using it in a regression. The formal method to test the stationarity of a series is the unit root test.

The unit root test generally contains one test equation, which is in the form of “ $\Delta y_t = \eta y_{t-1} + \alpha + \delta t + \sum_{i=1}^p \beta_i \Delta y_{t-i} + u_t$ ”. The null hypothesis is that the time series “ y_t ” has one unit root, or equivalently “ $H_0: \eta = 0; H_1: \eta < 0$ ”. There are several options for the unit root test. First, we can add some lagged difference terms (Δy_{t-i}) based on some automatic selection criteria (AIC, SIC, etc.). Second, we can add an intercept term and a trend term (t) into the test equation. The most widely accepted unit root test is the ADF “Augmented Dickey-Fuller” test. I will present my test statistics in Table 7 based on this method. It can be shown that all the variables are stationary after the first order difference originated from the APT theory. Therefore, any regression including ARMA structures should be safe now, and no spurious regression will be created as a result.

Table 8 *Augmented Dickey-Fuller Unit Root Test*

Variable	Exogenous	Lagged Δy_{t-i} terms (AIC)	ADF statistic	5% level	Prob.
ER	none	0	-4.101	-1.947	0.000
DRI	none	0	-8.527	-1.947	0.000
DL	none	1	-3.048	-1.947	0.003
DII	none	1	-7.145	-1.947	0.000
DI	none	0	-5.416	-1.946	0.000

5.4.2. Model Construction

To cure the autocorrelation, an AR(x) structure has to be added into the mean equation. The original AR(x) structure started with equation (5.2) and (5.3). Especially, (5.3) describe the situation where current residual is correlated with last period’s residual.

$$\text{Var}(u_t) = \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_n u_{t-n}^2 \quad (5.2)$$

$$\epsilon_t = \phi \epsilon_{t-1} + u_t \quad (5.3)$$

Through some substitution and iteration, we can transformed (5.2) and (5.3) to the following form as (5.4) shown.

$$Y_t = \phi Y_{t-1} + X_t \cdot \vec{\theta} + \epsilon_t \quad (5.4)$$

Thus, it is clear that we can simply add some lagged dependent variables into the right hand side of the mean equation, to accommodate the AR(x) structure.

Besides of a time varying conditional mean of financial time series, most of them also exhibit changes in volatility over time. While modeling such time series, we cannot use homoscedastic models. The simplest way to allow volatility to vary is to model conditional variance using a simple autoregressive process. Basic ARCH models were introduced by Engle (1982) and generalized as GARCH by Bollerslev (1986) and Taylor (1986). GARCH model is consistent with the some stylized facts that have been uncovered by researchers, including volatility clustering, fat tail and volatility mean reversion. A simple GARCH(1,1) model that should be constructed by three parts, the conditional mean equation, the conditional normality assumption and the conditional variance equation.

$$Y_t = X_t \cdot \vec{\theta} + \epsilon_t \quad (5.5)$$

$$(\epsilon_t | I_{t-1}) \sim N(0, \sigma_t^2) \quad (5.6)$$

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2, 0 \leq \alpha, \beta < 1 \quad (5.7)$$

In the mean equation given above, conditional mean (Y_t) is related with an array of exogenous variables (X_t). The conditional normality assumption implies the error term (ϵ_t) should follow a normal distribution with the forecasted variance (σ_t^2). The left hand side of the variance equation (σ_t^2) is one-period ahead forecasted variance based on past information. The right hand side of the variance equation contains three parts: a constant term (ω), news about volatility from the previous period, measured as the lag of the squared residual from the mean equation (ϵ_{t-1}^2), and last period's forecast variance (σ_{t-1}^2). Furthermore, ϵ_{t-1}^2 and σ_{t-1}^2 terms are often

referred as the ARCH term and the GARCH term, respectively. We should notice that a covariance stationary conditions ($0 \leq \alpha, \beta < 1$) should apply for the GARCH model; otherwise, the variance will continue to increase over time and eventually explode. The above model is a typical GARCH(1,1) model. We can easily extend it into a GARCH(q,p) model, which contains ARCH terms lagged up to q periods and GARCH terms lagged up to p periods in their variance equation. Due to the heteroscedasticity issue I spotted above, it is quite consistent to use GARCH as a curing regression method.

GARCH-in-Mean (GARCH-M) model is a modification on the mean equation of GARCH by Engle, Lilien, & Robins (1987). In financial investment, high risk is often expected to lead to high returns. GARCH-M extends the basic GARCH model so that the conditional volatility can generate a risk premium that is part of the expected returns. GARCH-M introduces an arbitrary function of volatility (σ_t) into the mean equation as shown in (5.8).

$$Y_t = X_t \cdot \theta + \lambda f(\sigma_t) + \epsilon_t \quad (5.8)$$

Since my study is also using also using stock returns data as the dependent variable, it should be logic and necessary to incorporate GARCH-M into my model, in order to capture the risk premium generated by high volatility in the market.

IGARCH (Integrated-GARCH) model added a coefficient restriction to the existing GARCH model. In an IGARCH model, the sum of the coefficients (α and β) of the ARCH and GARCH terms should always equal to one. The conditional variance of the IGARCH model is clearly non-stationary and integrated. This has important implications for interpreting the volatility of such a time series: the volatility of the model is not mean reverting; any shock to the volatility is persistent. The rationale behind IGARCH fits the reality in China very well. For example: when financial crisis happened, the volatility in China just jumps to the roof and stay there for a long time.

So it is better to assume persistence of volatility, rather than to assume the mean-reversion of it.

After discussing the concept and rationale of some chosen testing structures, it is time to do the real practice to select the best-fitted one from varieties of them.

From Table 9, we can see that an IGARCH-M model with an AR(1) structure has the highest AIC figure and the second highest R-squared figure. Besides, both AR(2) and AR(3) structure are not significant in the regression. Thus only an AR(1) term will be incorporated into the model.

Table 9 *Model Selection: AR(x) structures comparison*

AR(X)=	1	2	3
R-squared	0.609	0.608	0.613
AIC	-2.656	-2.623	-2.607
Prob. of AR(1)	0.000	0.001	0.000
Prob. of AR(2)		0.662	0.931
Prob. of AR(3)			0.742

I also tried different GARCH(q,p) setting as shown in Table 10. At the first glance, we can see all ARCH and GARCH terms in GARCH(1,1) and GARCH(2,1) are significant at a 10% level. In contrast, some terms in GARCH(1,2) and GARCH(2,2) are not significant at all. Now, among GARCH(1,1) and GARCH(2,1), we found the latter has the lowest AIC and SIC figures, indicating the best model fitness. Thus we should adopt GARCH(2,1) structure rather than the more common GARCH(1,1) structure.

Table 10 *Model Selection II: GARCH(q,p) structures comparison*

	GARCH(1,1)	GARCH(2,1)	GARCH(1,2)	GARCH(2,2)
R-squared	0.641	0.609	0.615	0.606
AIC	-2.636	-2.656	-2.628	-2.630
SIC	-2.280	-2.336	-2.308	-2.274
Prob. of ARCH(1)	0.001	0.002	0.019	0.002
Prob. of ARCH(2)	NA	0.054	NA	0.227
Prob. of GARCH(1)	0.000	0.000	0.597	0.294
Prob. of GARCH(1)	NA	NA	0.305	0.681

For the GARCH-M structure, I tried standard deviation (σ_t), normal variance (σ_t^2) and logarithm of variance ($\log [\sigma_t^2]$) in the mean equation. All of them are significant. σ_t derives the best R-squared, and $\log [\sigma_t^2]$ derives the best AIC and SIC figures. To balance these two benchmarks, I chose to use σ_t , a variable gives satisfactory scores in both side.

Table 11 *Model Selection III: GARCH-M structures comparison*

	σ_t	σ_t^2	$\log [\sigma_t^2]$
R-squared	0.609	0.621	0.527
AIC	-2.656	-2.652	-2.739
SIC	-2.336	-2.332	-2.420
Prob. of σ_t	0.000	NA	NA
Prob. of σ_t^2	NA	0.000	NA
Prob. of $\log [\sigma_t^2]$	NA	NA	0.000

Therefore, I finally decided to choose an AR(1)-IGARCH(2,1)-M(σ_t) model as equations (5.9), (5.10) and (5.11) shown below:

$$ER_t = [ER_{t-1}, DI_t, DII_t, DRI_t, DL_t] \cdot \vec{\theta} + \lambda \sigma_t + \epsilon_t \quad (5.9)$$

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \epsilon_{t-2}^2 + \gamma \sigma_{t-1}^2 \quad (5.10)$$

$$\alpha + \beta + \gamma = 1 \quad (5.11)$$

5.4.3. *AR-IGARCH-M Estimates*

Table 12 shows the regression results estimated via AR(1)-IGARCH(2,1)-M model. Compared with the OLS method, I have to mention several improvements from this new method. First, the R-squared increases from 0.38 in OLS to 0.61 in the refined model, indicating a larger proportion of the variance is explained in the new model. Second, F-statistics that test the overall significance of the regression model increase from 4.20 to 9.53, implying the new regression method justifies a closer relationship between the dependent variable and the independent ones. Third, as the model selection criterion AIC decrease from -2.17 to -2.656 and SIC decreases from -1.88 to -2.34, we can see the new model is superior to the old one in the closeness of fit. Fourth, the Durban-Watson statistic rises from 1.26 to 1.70. Since small values of the Durbin-Watson statistic indicate the presence of autocorrelation, we can see the autocorrelation problem has been largely corrected in the new model.

Table 12 *AR(1)-IGARCH(2,1)-M Estimates*

R-squared	0.609	Mean dependent var	-0.014
Adjusted R-squared	0.545	S.D. dependent var	0.093
S.E. of regression	0.062	Akaike info criterion	-2.656
Sum squared resid	0.191	Schwarz criterion	-2.336
Log likelihood	86.014	Hannan-Quinn criter.	-2.531
F-statistic	9.526	Durbin-Watson stat	1.704
Prob(F-statistic)	0.000		

<u>Mean Equation</u>				
Variable	Estimate	Std. Error	t-Statistic	Prob.> t
Intercept	0.044	0.011	3.847	0.000
SQRT(GARCH)	-1.327	0.274	-4.842	0.000
ER(-1)	0.371	0.073	5.058	0.000
DRI	0.119	0.016	7.635	0.000
DII	-0.077	0.021	-3.690	0.000
DI	1.221	0.830	1.471	0.141
DL	0.747	0.357	2.091	0.037

<u>Variance Equation</u>				
Variable	Estimate	Std. Error	t-Statistic	Prob.> t
RESID(-1)^2	0.552	0.175	3.163	0.002
RESID(-2)^2	-0.373	0.193	-1.929	0.054
GARCH(-1)	0.820	0.153	5.348	0.000

(5.12) and (5.13) are the estimation equations with substituted coefficients:

$$ER_t = 0.04 - 1.33 * \sigma_t + 0.37 * ER_{t-1} + 1.22 * DI_t + 0.12 * DRI_t - 0.08 * DII_t + 0.75 * DL_t \quad (5.12)$$

$$\sigma_t^2 = 0.55 * \epsilon_{t-1}^2 - 0.37 * \epsilon_{t-2}^2 + 0.82 * \sigma_{t-1}^2 \quad (5.13)$$

From the ACF graph, we can also easily observe the improvement. OLS did very poorly in predicting the excess return from September 2007 to October 2008. In OLS, the fitted line is deviated below from the actual line as we can see from Figure 1. On the contrary, the new model plots a better graph where fitted line are leveled up and moving along closer with the actual line in Figure 9.

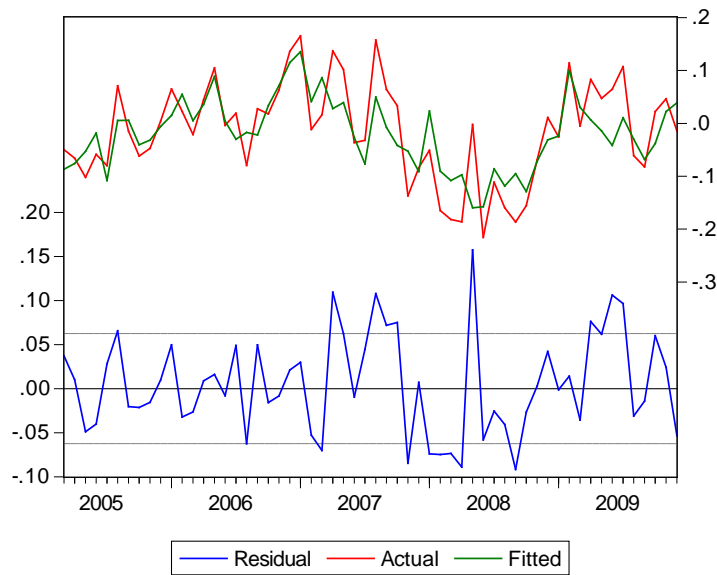


Figure 9 *Actual, Fitted and Residual Graph*

Other typical tests have been repeated under the new model as well. For the autocorrelation problem, we found the bars in the correlogram are no longer lengthy and the Q-statistics are not significant any more as shown in Figure 10. So we can conclude that the autocorrelation problem has been successfully cured by the AR(1) structure.





















Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.039	0.039	0.0942	0.759
		2	-0.047	-0.049	0.2317	0.891
		3	-0.040	-0.036	0.3327	0.954
		4	0.040	0.041	0.4336	0.980
		5	-0.008	-0.015	0.4383	0.994
		6	-0.017	-0.014	0.4579	0.998
		7	-0.119	-0.116	1.4192	0.985
		8	-0.154	-0.152	3.0800	0.929
		9	0.048	0.048	3.2412	0.954
		10	-0.111	-0.142	4.1374	0.941

Figure 10 *Serial Correlation Test: Correlogram of Residuals and Q-Statistics*

For the heteroscedasticity problem, the correlogram and the Q-statistics shows it has be cleaned away as shown in Figure 11. An additional ARCH LM test has been also conducted to further qualify the issue. From Table 13, we can judge that even at the 10% critical level, not a single lagged term has the ARCH problem. So the GARCH structure in the new model successfully cured the problem as well.

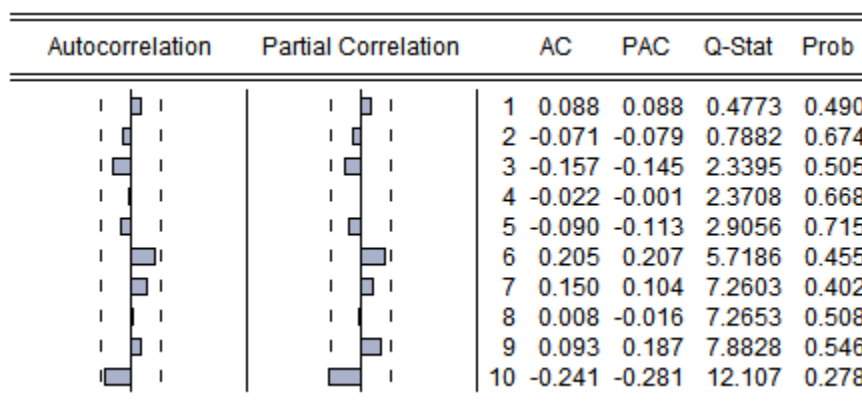


Figure 11 *Heteroscedasticity Test I: Correlogram of Residuals Squared and Q-Statistics*

Table 13 *Heteroscedasticity Test II: ARCH LM Test*

ARCH LM Test	Lag=	1	2	3	4	5
F-statistic		0.439	0.335	0.597	0.483	0.574
Prob. F		0.510	0.717	0.620	0.748	0.720
Obs*R-squared		0.451	0.700	1.865	2.048	3.050
Prob. Chi-Squared		0.502	0.705	0.601	0.727	0.692

5.5. *Economics Explanation*

The results shown above in Table 12 deliver the following implications to us. Some of them are consistent with simple intuition and experience of other countries; some of them are new and unique in China.

- 1) There is ARCH effect in the excess return rate of Chinese stock market. It means volatility clustering phenomenon does exist in China. I think one possible

explanation of such phenomenon could be developed through our heterogeneous investors model, if we further assume our two groups of investors could switch behaviors among different periods. In this new setting, a significant proportion of fundamental investors may switch to the momentum type, when the threshold probability for momentum investors to stay is low enough. And these newly switched (from fundamental to) momentum investors will create extra buy/sell activities in an upward/downward trend of the market. Finally, these extra buy/sell activities trigger the outbreak of volatilities and constitute the elevated volatility interspersed among more tranquil periods. This finding is consistent with the basic stylized fact of financial series that is observed elsewhere.

2) One term lagged volatility and forecasted volatility (ϵ_{t-1}^2 and σ_{t-1}^2) affect the current forecasted volatility positively and significantly at a 0.000 level. Two term lagged volatility (ϵ_{t-2}^2) effect the current forecasted volatility negatively and significantly at a 0.054 level. The sign was flipped from two periods lagged ARCH term to one period lagged ARCH term, proving that the volatility clustering in China happens in a very short horizon. Today's high volatility may indicate a likely increase of volatility in the next period, but the volatility after the next period is most likely to be adjusted down. Such finding is consistent with the nature and environment of a speculative stock market.

3) The IGARCH restriction was proved as legitimate by the result, showing that any shock to the volatility is persistent.

4) The GARCH-M structure in the mean equation affects the excess return negatively and significantly at a 0.000 level, which is a very interesting result. It means excess return doesn't compensate for higher risk (or expected volatility). Such fact could be explained by a series of plausible frictions in reality. First, it could be the case that investors in the Chinese stock market may not realize volatility is the source

of risk. This explanation relies on the poorly educated investors. Second, it could be the case that investors may favor volatility rather than avoiding it, since volatility provides the context for them to gamble and speculate. This explanation involves the concept of naïve investor. Third, it could be the case that investor are exploiting price movement too much. When they see a price increase/decrease (and consequentially the high volatility), they will overbuy/oversell the stocks and results a lower return. Fourth, it could most likely be the case that some predatory investors rely on volatility to trade against retail investors. When the former predicts a high volatility, they will inject their capitals into the market to trade and gain from the latter, because the former are better trained than the latter to trade under a high volatility environment. Their general capital injection could lead to a higher price and lower return of stocks.

5) The percentage change of retail investors loads positively and significantly against the excess return at a 0.000 level. This is consistent with the Greater Fools Theory. More retailed investors rushed in the market, higher the chance a greater fool could be found and higher the price and the return. This result also indirectly proved the greater fools in the simple could be deemed as momentum investors in the heterogeneous belief model. For momentum investors, a higher observed return in last period will make the threshold probability for them to stay lower. Hence, as the bubble thrives and the return sours, momentum investor are more likely to join the market, implying the number of them will increase. I didn't directly testing this positive relationship. But through the positive relationship we tested between " ER_{t-1} " and " ER_t ", and the positive relationship we tested between " ER_t " and " DRI_t ", I casually established a npositive relationship between " ER_{t-1} " and " DRI_t ", and hence indirectly confirmed the second hypothesis in section 4.2.5. Thus, I believe retail investors could be classified not only as greater fools, but also as momentum investors. Although the coefficient is merely 0.119, we have to notice the

standard deviation of DRI is almost 49.3%, which means if DRI moves up/down by one S.D., excess return will increase/decrease by 5.87%! It is a very bold effect on the stock market.

6) The percentage change of institutional investors loads negatively and significantly against the excess return at a 0.000 level. This result shows the institutional investors are not qualified as a greater fool. Alternatively, they could be deemed as the fundamentalist discussed in section 4.2.4. For fundamentalist, a higher observed return in last period will make the threshold probability for them to stay tougher. Hence, as the bubble thrives and the return sours, fundamentalist are more likely to leave or not join the market, implying the number of them will decline. I didn't directly testing this negative relationship. But through the positive relationship we tested between " ER_{t-1} " and " ER_t ", and the negative relationship we tested between " ER_t " and " DII_t ", I casually established a negative relationship between " ER_{t-1} " and " DII_t ", and hence indirectly confirmed the second hypothesis in section 4.2.5. Thus, I believe institutional investors couldn't be classified as a part of the greater fools. They as a whole behave like a fundamentalist and hold different belief from the retail investors.

7) The percentage change of liquidity affects the excess return positively and significantly at a 0.037 level. This result is consistent with the empirical result found in elsewhere. Liquidity could push the stock market in many means. I will list the most intuitive two. In the real side, more liquidity could increase the lending to the firm, make the firm committed to more investment and hence make their stocks more attractive and higher priced. In the nominal side, more liquidity could increase everyone's nominal income, thus people will invest more nominal money into the stock than before, which will push the stock price to a higher level. The coefficient of DL is 0.747, and one standard deviation of DL is 1.91%, which means if DL moves

up/down by one S.D., excess return will increase /decrease by 1.43%. So liquidity is the third largest factor that explains the variance of the excess return.

8) The percentage change of inflation affects the excess return positively at a 0.141 level. This variable is not very significant. I believe it is because the liquidity variable DL already works as a substitute of inflation. Inflation happens when the extra liquidity reflected on the commodities. So liquidity and inflation is related. Besides, since the formation of the basket is arbitrary and the menu cost prevents immediate reaction, inflation will always lag behind the liquidity. Hence, it is really no surprise to see inflation is not as significant as liquidity.

9) As we have shown before in the redundant variable test, all the other macroeconomic variables, including export, industrial production and interest rate, are not significant in the study. This contradicts the finding in the developed countries. I believe this interesting result could be supported by the Chinese stock market background I introduced earlier and the Greater Fools Theory. First, Chinese market is most likely not an efficient market driven by public information and data. Second, since the Chinese economy is still under the intervention of the central government, investor concern more about the incoming government policies rather than the incoming economic indicators. Third, following the Greater Fools Theory, as long as the investors believes more greater fools will enter the arena, they will buy the stocks anyway regardless with any indicators. Fourth, due to the lack of shorting mechanism, positive indicators might be reflected in the market, but negative ones may not be immediately reflected. Thus, the general explanatory power of economic indicators would be weakened.

CHAPTER 6

CONCLUDING REMARKS

My study tried to characterize the Greater Fools Theory, test whether it is empirically valid and statistically significant. The objective of this study has been fulfilled with a satisfactory result. Properties derived from the basic mathematic formula are quite consistent with people's irrational behaviors during a bubble. The empirical test finds out the supply of greater fools is the single largest force that drives the stock market return into an abnormal level. And institutional investors are the second largest force that stands against the influence of those greater fools. Liquidity affects the stock market in a third place by exerting influence on real side and nominal side of the economy. The statistical practice is also very successful. The AR-IGARCH-M model corrected all common problems for a time series regression. Thus the result should be reliable and trustworthy.

However, some further improvement could be done on this study. First, we could further incorporate more features into the model of the Greater Fools Theory, like trying non I.I.D. distributions, different expectation and even some utility functions. Second, we could include a wider range of testing periods, if data in the earlier time has been published. Third, we could test the Greater Fools Theory in a firm level, by using some panel data. The factor model should also be adjusted to the Fama-French form, and also include the greater fools proxies. Fourth, some Vector Autoregressive Regression method could be discussed to see its fitness with the data.

By revealing the significance of the Greater Fools Theory to public, my research has deep implications either to the market participant or the policy maker in China. First, it reminds the investor that the greater fools' thinking dominates in Chinese stock market. If most Chinese investors become more aware of the Greater Fools Theory, the stock markets will be less volatile, avoiding the huge and abrupt

boom and bust in the last few years. By that time, the stock markets will become a more reliable venue for investment, supporting China to become a more developed economy. Second, the policy maker can play a more critical role in preventing the birth and burst of bubbles, should they understand the relationship between the Greater Fools Theory and Chinese stock bubbles. They could provide more information about the fools in the market, to avoid exaggerated expectation of next incoming fools. They should allow the short-selling of stocks in order to create more arbitrageurs than fools in the market. And they should also encourage dividends issuance to create more fundamentalist than fools in the market.

BIBLIOGRAPHY

- Akerlof, G. A., & Shiller, R. J. (2009). *Animal spirits : how human psychology drives the economy, and why it matters for global capitalism*. Princeton: Princeton University Press.
- Black, F. (1972). Capital Market Equilibrium with Restricted Borrowing. *The Journal of Business*, 45(3), 444-455.
- Black , F., Jensen, M. C., & Scholes, M. S. (1972). The Capital Asset Pricing Model: Some Empirical Tests. *Michael C. Jensen, STUDIES IN THE THEORY OF CAPITAL MARKETS*, Praeger Publishers Inc., 1972.
- Black, F., & Myron, S. (1973). The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy*, 81(3), 637-654.
- Blanchard, O. J., & Watson, M. W. (1983). Bubbles, Rational Expectations and Financial Markets. *SSRN eLibrary*.
- Bollerslev, T. (1986). *Generalized autoregressive conditional heteroskedasticity with applications in finance*.
- Chan, K., McQueen, G., & Thorley, S. (1998). Are there rational speculative bubbles in Asian stock markets? *PACIFIC BASIN FINANCE JOURNAL*, 6(1-2), 125-151.
- Chen, N.-F., Roll, R., & Ross, S. A. (1986). Economic Forces and the Stock Market. *The Journal of Business*, 59(3), 383-403.
- De Bondt, W. P. M. (1993). Betting on trends: Intuitive forecasts of financial risk and return. [doi: DOI: 10.1016/0169-2070(93)90030-Q]. *International Journal of Forecasting*, 9(3), 355-371.
- Diba, B. T., & Grossman, H. I. (1988). The Theory of Rational Bubbles in Stock Prices. *The Economic Journal*, 98(392), 746-754.

- Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4), 987-1007.
- Engle, R. F., Lilien, D. M., & Robins, R. P. (1987). Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model. *Econometrica: Journal of the Econometric Society*, 55(2), 391-407.
- Evans, G. W. (1991). Pitfalls in Testing for Explosive Bubbles in Asset Prices. *The American Economic Review*, 81(4).
- Fama, E. F. (1965). The Behavior of Stock-Market Prices. *The Journal of Business*, 38(1), 34-105.
- Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47(2), 427-465.
- Fama, E. F., & MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *The Journal of Political Economy*, 81(3), 607-636.
- Flood, R. P., & Hodrick, R. J. (1990). On Testing for Speculative Bubbles. *The Journal of Economic Perspectives*, 4(2), 85-101.
- Froot, K. A., & Obstfeld, M. (1991). Intrinsic Bubbles: The Case of Stock Prices. *The American Economic Review*, 81(5), 1189-1214.
- Garber, P. M. (1989). Tulipmania. *Journal of Political Economy*, 97(3), 535-560.
- Green, S. (2004). *The development of China's stock market, 1984-2002 : equity politics and market institutions*. London; New York: Routledge.
- Gürkaynak, R. S. (2008). ECONOMETRIC TESTS OF ASSET PRICE BUBBLES: TAKING STOCK. *Journal of Economic Surveys*, 22(1), 166-186.
- Hertz, E. (1998). *The trading crowd : an ethnography of the Shanghai stock market* (Vol. 108). Cambridge [England] ;New York: Cambridge University Press.
- Keynes, J. M. (1936). *The general theory of employment, interest and money*. New York: Harcourt, Brace.

- Kindleberger, C. P. (2008). bubbles in history. In S. N. Durlauf & L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics*. Basingstoke: Palgrave Macmillan.
- Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, 47(1), 13-37.
- Long, J. B. D., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise Trader Risk in Financial Markets. *The Journal of Political Economy*, 98(4), 703-738.
- Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1), 77-91.
- Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, 34(4), 768-783.
- Roll, R. (1977). A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory. [doi: DOI: 10.1016/0304-405X(77)90009-5]. *Journal of Financial Economics*, 4(2), 129-176.
- Roll, R., & Ross, S. A. (1980). An Empirical Investigation of the Arbitrage Pricing Theory. *Journal of Finance*, 35(5), 1073-1103.
- Ross, S. A. (1973). *The arbitrage theory of capital asset pricing*. Philadelphia: Rodney L. White Center for Financial Research, University of Pennsylvania, The Wharton School.
- Samuelson, P. A. (1958). An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money. *The Journal of Political Economy*, 66(6), 467-482.
- Shanghai-Investment.com. (2000). China's stock market to become largest in Asia by 2010 China's stock market. from <http://www.shanghai-investment.com/archives/000724.html>

- Shanken, J. (1982). The Arbitrage Pricing Theory: Is it Testable? *Journal of Finance*, 37(5), 1129-1140.
- Shanken, J. (1985). Multi-Beta CAPM or Equilibrium-APT?: A Reply. *Journal of Finance*, 40(4).
- Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, 19(3), 425-442.
- Shiller, R. J. (1990). Market Volatility and Investor Behavior. *The American Economic Review*, 80(2), 58-62.
- Shiller, R. J. (2000). *Irrational exuberance*. Princeton, NJ: Princeton University Press.
- Shiller, R. J., Fischer, S., & Friedman, B. M. (1984). Stock Prices and Social Dynamics. *Brookings Papers on Economic Activity*, 1984(2), 457-510.
- Taylor, S. (1986). *Modelling financial time series*. Chichester [West Sussex]; New York: Wiley.
- Telser, L. G. (2010, 1/14/2010). Are Speculators Foolish in Keynes' Greater Fool Model. *The Selected Works of Lester G. Telser*, from http://works.bepress.com/lester_telser/16
- Tirole, J. (1982). On the Possibility of Speculation under Rational Expectations. *Econometrica*, 50(5), 1163-1181.
- Tirole, J. (1985). Asset Bubbles and Overlapping Generations. *Econometrica*, 53(6), 1499-1528.
- Tobin, J. (1958). Liquidity Preference as Behavior Towards Risk. *The Review of Economic Studies*, 25(2), 65-86.
- Treynor, J. L. (1961). Market Value, Time, and Risk. *Unpublished manuscript*.
- Treynor, J. L. (1962). Toward a theory of market value of risky assets. In R. A. Korajczyk (Ed.), *Asset pricing and portfolio performance : models, strategy, and performance metrics*. London: Risk Books.

West, K. D. (1987). A Specification Test for Speculative Bubbles. *The Quarterly Journal of Economics*, 102(3), 553-580.