An empirical investigation ~~of~~ into the arbitrage pricing ~~theory~~ theory of~~n~~ China A share market after financial crisis

Formatted: Bottom: 0.01"

Abstract

~~The purpose of t~~This paper ~~is to~~ examines the applicability of Arbitrage Pricing Theory APT ~~in~~ to China's A Share Market after the financial crisis~~,~~ using monthly data of 100 stocks from Shanghai Stock Exchange A ~~S~~share as proxy variable~~s~~ to the whole market. Above all, it~~this paper~~ obtains four priced common factors from ~~twenty-two~~22 macroeconomic variables by Exploratory Factor Analysis and builds relevant APT model for further empirical test. Then through multiple linear regressions, we test the explanatory power of APT ~~in China's A Share Market~~. ~~According to the~~Using empirical test results and the problem of poor goodness of fitting, we offer probable ~~give possible~~ reasons (for what?) and make ~~relevant~~some suggestions.

Comment [.1]: No abbreviations in abstract, and even in text only after the full form is fir introduced and abbreviation shown in brackets.

Comment [.2]: This sentence does not conve the intended meaning. Pleae rewrite.

Ⅰ.Introduction

Arbitrage Pricing Theory (APT), formulated by Ross (1976), is proposed to be an alternative to the mean variance Capital Asset Pricing Model (CAPM) which was introduced by Sharp (1969) and, Lintner (1965). In the analysis of APT, Ross (1976) suggested that the market portfolio does idn'ot play a specific role any more. It proved that the expected return to assets was systematically influenced by multiple exogenous macroeconomic factors. Depending on the assumption of a well-diversified, perfectly competitive and frictionless economy, APT specified the linear relationship between expected return to assets and a series of macroeconomic factors.

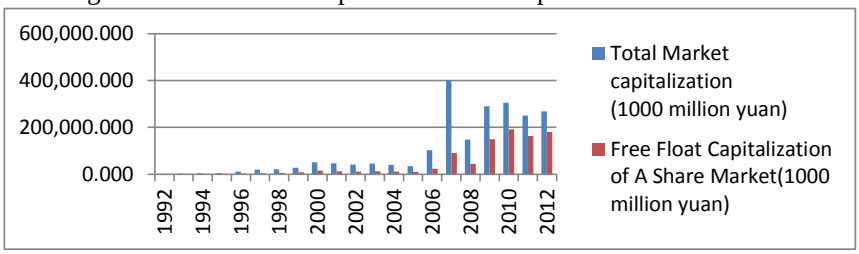
The paper of Chen, Roll and Ross (1986) selected the systematic forces that haved significant effect on the two determinants of stock returns: discount factors and expected cash flows. Through a vector autoregressive model, shock to real industrial production, changes in the risk premium, twist in the yield curve, unexpected inflation and change in expected inflation were explored to have systematic influences on stock market returns. In addition, by using a multi-factor model repeated in every subsample ofor each year, Chen, Roll and Ross obtained for each macroeconomic variable a time series of estimates of associated risk premium, which indicated the explanatory power to the multi-factor model. This kind of analytical method has been adopted till now for examining the validity of APT in the stock market.

During the recent decades, APT is considered to be the most significant analytic tool for to explaining the pricing behaviour in the assets market, mainly in some relatively developed markets which more than satisfy the assumptions of well-diversified, perfectly competitive and frictionless economy. In order tTo provide a better fitting tool to explain the market phenomenon, more mature commercial factor models like BIRR Model[1], RAM Model[2] are derived. While the emerging markets in developing countries under imperfect system are not mature and unhealthy, assumptions of APT cannot be satisfied completely. Pproper models for emerging markets are still being explored. China's stock market, the most typical emerging market is divided into three markets—A Share Market, B Share market and H Share market. Since foundationestablishment, A Share Market has always been the main driving force of China's stock market and guides the overall direction of stock the market, in both aspects of capitalization, and outstanding shares, which can be (shown in Figures 1 and Figure 2).

[1] Burmeister, Roll and Ross (2003) introduced BIRR Model in the paper "*Using Macroeconomic Factors to Control Portfolio Risk",* <http://www.ftse.com/Analytics/BIRR>

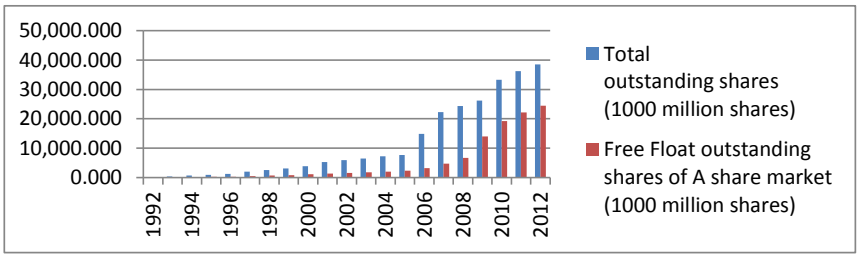
[2] RAM Model (Risk Attribution Model) was introduced by Salomon Brothers in 1986

Figure 1: Total market capitalization and capitalization in A Share Market.



Source: Wind

Figure 2: Total outstanding shares and outstanding shares in A Share market.



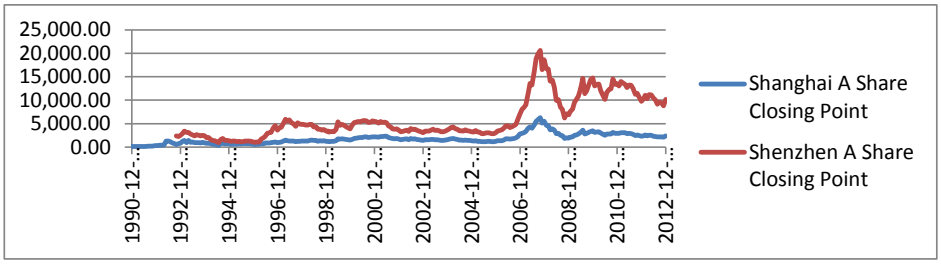
Source: Wind

A Share Market has ~~always~~ been developing during~~in~~ the unrest period and faces strict regulatory controls. ~~Not u~~Until 1996, ~~because of increasingly~~the ~~opened~~ capital market, the IPO system and reverse merger~~,~~ were not open ~~did~~ and when it did the number of listings by private enterprises increased rapidly. The year ~~of~~ 2007 ~~wa~~is considered ~~as~~ the most significant period~~plendid~~ ~~time~~ in the history of China's market development as also~~, the same~~ to the A Share Market. Reforms in ~~N~~non-tradable shares ~~reform~~ and ~~reform~~ in ~~the~~ Fund Industry pushed the A Share Market to ~~be~~ a bull run~~market~~. Both ~~of~~ the Shanghai A Share Index and the Shenzhen A Share Index reached their peaks, respectively, to ~~be~~ 6429.68[3] and 1629.7546[4], increasing by over twice the size ~~as~~of 2006. As shown in Figure 3, the closing point in both the exchanges ~~got to~~reached an unprecedented ~~peak~~level.

Comment [.3]: What's unrest?

Figure 3: A share closing point in Shanghai ~~Stock Exchange~~ and Shenzhen Stock Exchanges.

[3] Statistical yearbook 2012 of Shanghai Stock Exchange
[4] Statistical yearbook 2008 of Shenzhen Stock Exchange



Source: Wind

In 2008, the global financial crisis ~~spilled over to~~enveloped China too~~,~~ leading to a deep correction in China's equity market and a sharp contraction in terms of market value. China's equity market shrank dramatically, which ~~was embodied in the~~ depressed~~ion~~ ~~of~~the ~~–~~A share market, mainly because of the ~~P~~price ~~C~~comparison effect between the A ~~shares~~ and H shares. The year ~~of~~ 2008 ~~is~~ also saw ~~no doubt to be~~ another turning point in~~to~~ China's ~~in~~ history. After the crisis, the ~~on the one hand,~~ A share market remained grim~~–~~. The A Share Index swung back and forth around 2000 points during the recent five years and even~~r~~ reached 1806.79~~–~~[5] points in 2008. To strengthen~~reinforce~~ the A share market, RQFII mechanism[6] ~~–~~came into force in August ~~of~~ 2011, with the pumping of ~~starting from~~ RMB 20 billion ~~RMB~~. Even though the effect of RQFII has been hanging ~~i~~on fire~~doubt~~ so far, it is still ~~be~~ the focus of investors. On the other hand, ~~T~~the prices of shares on A ~~S~~share ~~M~~market are ~~has serious bubble displayed in the~~ overestimated ~~prices of A shares~~ without reference to ~~and~~ price~~–~~earnings ratio. Investors firmly believe that there ~~ex~~ist~~s~~ a great arbitrage opportunity~~chance~~ in the A ~~S~~share market.

Comment [.4]: The meaning is not clear.

The financial crisis forced the stock market to change, no matter whether in ~~terms of~~ structure or system. ~~The China's~~ A share market has become~~s~~ further open and mature. Economists wish~~insist~~ ~~on~~to examine~~ing~~ the validity of APT in China's market while the outcome~~result~~ is controversial. Opponents~~–~~ ~~believe~~consider that the ~~great~~ changes in China's stock market have not ~~didn't~~ improved the situation at all. ~~fitting degree of APT apparently.~~ Nevertheless, supporters feel ~~consider~~that investors can use the opportunity arising out of the crisis in the ~~combine the situation of~~ capital market and their own ~~people's~~ preference~~s,~~ and use~~ing~~ APT ~~in order~~ to make rational choice~~s in the crisis~~. From the perspective of the current market, macroeconomic factors like inflation, real~~–~~estate market and consumer price index play a dominant part.

The objective~~purpose~~ of this paper is to examine the performance of APT on China's stock market after the financial crisis. The paper is organized as follows. Section Ⅱ reviews the literature of APT. In Section Ⅲ，we briefly introduce the revised model

[5] Hexun A share index

[6] RQFII is the RMB qualified foreign institutional investor scheme. Launched in August 2011 it allowed a small number of Chinese financial firms to establish RMB-denominated funds in Hong Kong for investment in the mainland. The aim being to allow overseas investors to use offshore RMB deposits to invest in mainland securities markets.

based on China's physical situation and methodology in Exploratory Factor Analysis~~,~~ to select economic state variables which have systematic impact on returns from stock~~s return~~. Section Ⅳ introduces the data. Section Ⅴ describes the economic characteristics of the macroeconomic variables selected and conducts a cross-sectional regression analysis among ~~between~~ the factors and stock return to examine the validity of APT in China. SectionⅥ ~~provides the~~ concludes the work~~sion~~.

Ⅱ.Literature Review

Arbitrage means taking advantage of the price variance and trading among~~between~~ different markets to get~~obtain~~ riskless profit. ~~The a~~Arbitrage ~~behavior~~ is realized by buying high-return securities and selling low-return one~~securities~~ at the same time. As a consequence, the price of high-return securities will be pushed up while the rate of return will fall down and the low-return securities will be pulled down while the rate of return will go~~rise~~ up. This process finally brings equilibrium ~~makes~~ to the security market. ~~reach its equilibrium. The A~~~~rbitrage P~~~~ricing T~~~~heory~~ shows~~indicates~~, under unbalanced state, that there exists no arbitrage chance in the market. Moreover, the return to risky assets can be explained by multiple factors.

(1) Theoretical development of APT

Though ~~the Arbitrage Pricing Theory~~APT has become one of the mainstream models in the study of asset~~s~~ pricing issue, doubts are always heard, especially when compared with the Capital Assets Pricing Model. Roll and Ross (1980), Chen (1983), Chen, Roll and Ross (1986), Fama and French (1992), Groeneworld and Fraser (1997) are the most famous APT supporters. For one thing, ~~the~~ APT requires less but more realistic assumptions: a. multiple investment terms; b. taxes exist; c. investors cannot borrow money at the riskless rate. For another thing, APT shows better explanatory power than CAPM as it is a multi-factor model. Researchers like Dhrymes, Friends and Gultekin (1984) question~~ed~~ the testability both of APT and CAPM.

As mentioned in Roll and Ross (1980), "*APT has clear advantages over CAPM. APT is based on a linear return generating process and requires no utility assumptions beyond monotonicity and concavity.*" In addition, APT allows multiple factors and the equilibrium is characterized by ~~the~~ linear relationship between return to assets and the priced common factors. APT's empirical test indicate~~sd~~ its performance by cross-sectional regressions.

The idea of Chen, Roll and Ross (1986) has been discussed~~talked about~~ in the introduction. Chen (1983) using the daily data of stock return from 1963 to 1978~~,~~ compare~~d the evidence on~~ both APT and CAPM and test~~ed~~ if ~~whether~~ the APT can explain several empirically abnormal occurrence~~s~~ concerning ~~the~~ CAPM. Based on

~~the~~ empirical evidence,~~ the~~ APT cannot be disregarded ~~rejected~~ in favour of any other~~alternative~~ hypothesis and ~~the APT~~ performs better over CAPM in regard to ~~as implemented by~~ three market indices: S&P Index, Value Weighted Stock Index and Equally Weighted Stock Index. ~~The~~ APT ~~was proved to be~~is a better~~-~~fitting model ~~in~~to explain~~ing the~~ pricing behaviour in stock market.

Following Chen's methodology, Cagnetti (2002) tested and compared CAPM and APT in the Italian stock market with pre-set factors. All of the Davidson and Mackinnon Equation, the Posterior Odds Ratio and Residual Analysis strongly supported the performance of APT. In addition, through factor analysis, Cagnetti obtained~~got~~ eight common factors: market portfolio factor, fixed income securities, market index, foreign variables, production and monetary factor, inflation factor, industrial factor, people's expectation factor, which follow closely the logic of economic activity.

McElroy and Burmeister (1986) and Burmeister and McElroy (1988) recast the APT as a multivariate nonlinear regression model with across-equation restriction. They replaced the unknown random factors with observed macroeconomic variables. The nonlinear time~~-~~series methodology made the parametric APT testable rather than assuming that~~e~~ price~~ing~~ restrictions hold. Antoniou, Garrett and Priestley (1998) followed this method to investigate the performance of APT on London Stock Exchange. The paper randomly divided the securities into two subsamples to test whether ~~exist~~ the unique return~~-~~generating process across two subsamples exists as well as risk prices in different subsamples. Finally, Antoniou, Garrett and Priestley (1998) ~~got the~~ concluded~~sion~~ that to ~~get~~ develop a unique return~~-~~generating process, the model should be robust across different samples. ~~Moreover, t~~They got five common factors to price securities, among which are: unexpected inflation, money supply and excess returns on market portfolio carrying the same prices of risk in different samples, and the other two factors are: default risk and exchange rate only offerred as marginal improvement in the performance of APT. Azeez and Yonezawa (2006) adopted the same ~~way~~ method to examine the validity of APT in the Japanese stock market during the bubble period~~. They also~~and compared the robustness of priced factors ~~over~~ during the bubble period with pre- and post-bubble period data~~s~~. Four different types of risk factors—money supply, inflation, exchange rate and industrial production—were identified with having ~~that had~~ significant influence on the returns to stocks ~~turned out to be money supply, inflation, exchange rate and industrial production~~.

Fama and French (1992) compared the CAPM and APT from several aspects. Besides showing the weakness that had been found in Chen (1983) and Chen, Roll and Ross (1986), this paper also implied that CAPM overestimated the risk-free rate and underestimated the market risk premium. Moreover, the returns to high-beta assets will be overestimated and ~~the returns to low-beta assets~~ will be underestimated to low-beta assets. In ~~the aspect of~~ empirical test, Fama and French had different

ideas from traditional two-step method, which might have the errors-in-variables problem. Using the monthly data ~~for~~ of 19 sectors from the Australian Stock Exchange, Fama and French consider that~~ed~~ the beta of market cannot explain the variance among different returns ~~to~~ on stocks, while this can be interpreted by three classes of factors. Any changes in the real domestic activity, nominal domestic influences and foreign variables are supposed to change investors' expectations of future cash flows and afterwards influence current stock prices. T~~he t~~est results indicate~~d~~ that inflation rate ~~wa~~is more an empirical support than other macroeconomic factors.

In recent decades, with ~~the~~ improvement ~~of~~ in asset~~s~~ pricing theory, several commercial asset~~s~~ pricing model~~s~~ have been derived from the APT multi-factor model~~;~~, for instance, BIRR Model, RAM (Risk Attribution ~~–~~Model). BIRR Model was ~~raised~~ developed by Burmeister, Ibbotson, Roll and Ross (2003), who introduced five macroeconomic factors: confidence risk, time horizon risk, inflation risk, business cycle risk, and market timing risk. All the factors were divided into leading indicator, coincident indicator and lagging indicator and determined the order to build a better fitting model. As a consequence, four factors that indicate~~d~~ greater influence on return to stock~~, respectively~~are: ~~to be~~ macro-economic climate index，the CRB metal stock index，difference between the speed of deposit and loan，and one-year treasury yields. Among all the above, CRB metal stock index representing the commercial cycle had greater sensibility. RAM ~~(Risk Attribution Model)~~ considered ~~adopted~~ six factors to explain the pricing behaviour in the stock market: long-term economic growth expectation, short-term risk of the business cycle，long-term bond yields，short-term treasury yields, inflation and exchange rate.

(2) ~~The study results about the v~~Validity of APT in China's stock market

~~For~~ In the period of ~~unrest~~turmoil and rapid development ~~of~~in China's stock market, ~~plentiful~~ several economists have shown~~ed~~ great interest in the pricing issue of China's stock market, no matter before or after the financial crisis in 2008. As one of the most crucial assumptions of APT is a well-diversified, perfectly competitive and frictionless economy, most of the earlier studies ~~in~~ using it ~~APT before were to~~ examine~~d the~~ its performance ~~of APT~~ in developed stock market~~s~~. However, the rapid rise in emerging market~~s~~ or transition economy is obvious to all, especially in China, and hence~~.~~ Chinese economists devoted themselves to examine~~plain~~ the market behaviour in China with a revised model.

Li (2002) revised the multi-factor model in his paper *Experimental Test about Shenzhen Stock Market by APT* according to changes in market system. Price Limit System and T~~–~~+~~–~~1 Trading Rules were added ~~in~~to the regression model as two dummy variables. In the first step, by factor analysis, Li ~~got~~ listed four macroeconomic factors which~~which reflected~~ significantly impact ~~that~~ various industries listed ~~have~~ on the Shenzhen stock exchange~~market~~, as ~~well as~~also ~~the~~

Formatted: Font color: Red

Formatted: Font color: Red

Formatted: Font color: Red

Formatted: Font color: Red

Comment [.5]:

impact of earning performance and financial conditions status onf their own companies. In the second step, Li respectively examined the correctness accuracy of APT with the non-modified and modified models in Shenzhen market. The study test results showsed that both of the dummy variables had significant effect on returns to from stock, which means thatt unsystematic risk from policies had great effect on investors' behaviour and market direction. The physical significance of Li's suggests paper is that the most key point to promote the more radical marketing approach to improve China's stock market. is to speed up the marketization. A similar method will be adopted in this paper, – too but by adding a different dummy variable, RQFII System, will be added into the multi-factor model.

Liu and Qin (2004) adopted the test methodology ofin Brown and Weinstein (1984), respectively selecting two sample groups in two periods and examining if whether the generation process was the same among different subsample groups or as well as whether different sample groups faced the same factor structure. Before the test of APT, by Maximum Likelihood Method the authors proved there existenced of at least nine factors affecting the China's market and examined the relevancevalidity of the nine factor model to in China market. The test results suggested implied that, at that time, the APT model might not explain the economic behaviour in China's market. The pricing issue of China's stock market was influenced too much by policies and it was hard to considerform the real – value investment philosophy. However, this phenomenon coincided with China's emerging market or transition economy.

Tian (2006) followed the methodology of BIRR Mmodel to select macro risk factors in China's stock market. The procedure waiis as follows: firstly, divide all the alternative factors into leading, consistent and lagging factors through Variance Inflation Factor method and robust regression; secondly, build the best model with selected factors; last, to avoid the heteroscedasticity problems, use the GARCH Mmodel to obtainget risk sensitivity matrix. In this paper, factors including total energy production, 7-day interbank interest rates, M2, CPI, long-term loan, FDI, outstanding of deposits, were selected as macro risk factors.

After the financial crisis, China's stock market underwent several experienced great changes. Though badly affected by suffering a lot from the crisis, the Chinese market was opened by a relatively passively and slowly regained its became more normal tivestatus. Hu and Chen (2009) considered the at one assumption of APT was that risk-averse individual investor always preferred increasing income without exposing to higher risk, which could be achieved by arbitrage. This assumption waiis in accordance with the investors' psychology during the depression. Hence, APT model was well received infurther appreciated for China's stock market after the Ffinancial Ccrisis. Yin (2008) did a similar study on the validity of APT in China's stock market. The majority of preveious studies examined of the validity of APT in China and suggested indicated that policies had greater influence on the market than the

market force and too ~~over~~much regulation ~~supervision~~ might increase the risk.

III.The Model and Methodology

(1) ~~(1)~~The APT model starts with the following assumptions; ~~as follow:~~
Asset markets are perfectly competitive and frictionless. No arbitrage chance exists in the market. Individuals have homogeneous beliefs that asset~~s~~ returns are generated by a linear k-factor model;
The k factors that have effect on assets returns are stochastic and unexpected. Here we assume the mean of every factor to be~~is~~ zero;
All the investors are~~have the~~ risk--averse and monotone increasing utility function.
Under the above assumption~~s~~, Ross in 1976 generalized a multi-factor F_1, F_2, F_3……F_k systematically influencing the return~~s~~ from ~~on~~ assets

$$\mathbf{R_i = E(R_i) + b_{i1}F_1 + b_{i2}F_2 + \cdots + b_{ik}F_k + \varepsilon_i}, \text{ i=1, 2, 3…… m, m>k} \qquad (1)$$

~~W~~where R_i is random rate of return to asset I; F_j（j=1, 2…k）–is denoted as the systematic common factor, $E(F_iF_j)=0, i\neq j$, implie~~sying~~ that factors are independent of ~~between~~ each other; b_{ik} is the unique coefficient of sensitivity that asset *i* has on factor j, measuring the systematic risk–; ε_i is considered as a disturbance term, which measures the idiosyncratic and unsystematic risk, including all the unrelated information about other assets. $E(R_i)$– refers to the expected return from~~to~~ stock *i* when other risks equal ~~to~~ zero. We assume $E(\varepsilon_i) = 0$ and $E(\varepsilon_i\varepsilon_j) = 0, i\neq j$.

Associated with no risk arbitrage and sufficient portfolio diversification, the equilibrium relationship is as follows:

$$\mathbf{E(R_i) = \theta_0 + \theta_1 b_{i1} + \theta_2 b_{i2} + \cdots + \theta_k b_{ik}} \qquad (2)$$

where θ_0 is denoted as the expected return to riskless asset; θ_j can be interpreted as risk premiums corresponding to risk factor F_j(j=1,2…k).
In consideration of the RQFII, we intend to test whether the fluctuations in policy have significant influence on our regression, so we introduce this mechanism into APT model as dummy variable. The revised model is ~~shown~~ as follows:

$$\mathbf{R_i = E(R_i) + b_{i1}F_1 + b_{i2}F_2 + \cdots + b_{ik}F_k + \varepsilon_i + cD} \qquad (3)$$

where D_j is dummy variable , D = 0, before the RQFII
1, after the RQFII
c is the sensitivity of price to the dummy variable.
;

Formatted: Font: (Default) Times New Roman, 12 pt, Complex Script Font: Times New Roman, 12 pt

Formatted: List Paragraph, Numbered + Lev 1 + Numbering Style: 1, 2, 3, … + Start at: 1 Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"

Formatted: Font: (Default) Times New Roman, 12 pt, Complex Script Font: Times New Roman, 12 pt

Formatted: Font: (Default) Times New Roman, 12 pt, Complex Script Font: Times New Roman, 12 pt

Therefore, the model can be converted~~translated~~ to:

$$\mathbf{R_i} = \begin{cases} [\mathbf{E(R_i)} + \mathbf{c}] + \mathbf{b_{i1}F_1} + \mathbf{b_{i2}F_2} + \cdots + \mathbf{b_{ik}F_k} + \boldsymbol{\varepsilon_i} + \mathbf{cD}, & \mathbf{D=1} \\ \mathbf{E(R_i)} + \mathbf{b_{i1}F_1} + \mathbf{b_{i2}F_2} + \cdots + \mathbf{b_{ik}F_k} + \boldsymbol{\varepsilon_i}, & \mathbf{D=0} \end{cases}$$

(2) Exploratory Factor analysis

However, the APT model itself doesn't ~~have the~~ access ~~to~~ macroeconomic factors. In general, there exist two methods to determine the number of factors. One ~~way~~ is to run regressions with artificially stated factors on the rate of return to stock and to determine the final common factors~~,~~ based on test of goodness of fit. Yasushi (1988) adopted parallel data used in Chen, Roll and Ross (1986) and mixed with international trade factors to discuss the ~~different~~ significance~~s~~ of different pricing effects of similar macroeconomic factors in ~~a~~ parallel but with a distinguishing economy. According to recent studies, to support the APT model, the artificial~~ly~~ factors should be properly integrated. ~~combined.~~ Nevertheless, an agreement on how to correctly combine these factors which ~~can~~ persistently and completely explain the cross-sectional variation has not yet been arrived at. ~~cannot be reached an agreement.~~ In addition, artificial~~ly stated~~ factors may lead to information overlap of~~or the~~ multicollinearity between independent variables. ~~Simply~~ Any ~~to~~ reduction ~~e~~in the number of variables to solve the above problem may lead to information loss and lack of explanatory power ~~to model~~.

Another methodology is to pre-set several statistic variables as alternate common factors and use Principal Component Analysis or Maximum Likelihood Method to determine common factors. This procedure is called the Exploratory Factor Analysis (EFA), which is a common methodology of dimensionality reduction, making use of the correlation between original variables to best combine and simplify the multi-variable data. Interpreting most of the information of the original data can reduce difficulty in examining the validity of APT. The common factors after sifting mainly avoid linear correlation. The second method seems to provide more objective analysis and better fitness.

The starting point of EFA is to replace the major~~ity~~ information of original variables with fewer mutually independent factors. Assuming that there exist p variables, $x_i (i = 1, 2, \ldots p)$ are standardized variables; $F_j (j = 1, 2, \ldots m)$ are denoted as the systematic common factors, then,

$$\mathbf{x_1} = \boldsymbol{a_{11}F_1} + \boldsymbol{a_{12}F_2} + \cdots + \boldsymbol{a_{1m}F_m} + \boldsymbol{a_1\varepsilon_1}$$
$$\mathbf{x_2} = \boldsymbol{a_{21}F_1} + \boldsymbol{a_{22}F_2} + \cdots + \boldsymbol{a_{2m}F_m} + \boldsymbol{a_2\varepsilon_2}$$

……

$$\mathbf{x_{p-1}} = \boldsymbol{a_{(p-1)1}F_1} + \boldsymbol{a_{(p-1)2}F_2} + \cdots + \boldsymbol{a_{(p-1)m}F_m} + \boldsymbol{a_{p-1}\varepsilon_{p-1}}$$

$$\mathbf{x_p} = \boldsymbol{a_{p1}F_1} + \boldsymbol{a_{p2}F_2} + \cdots + \boldsymbol{a_{pm}F_m} + \boldsymbol{a_p\varepsilon_p}$$

This ~~Also~~ can also be expressed by the (m, p) matrix:

$$\mathbf{X} = \mathbf{AF} + \mathbf{a\varepsilon}$$

where A is denoted as factor loading matrix, and~~;~~ a_{ij} is called factor loading, representing correlation coefficient between original variable *i* and common factor *j*. The bigger the absolute value of a_{ij} is, the more significant variable *x* is to factor *j*;
The sum of squares of a_{ij} in line *i* in matrix A is defined as the ~~C~~communalities between original variables~~,~~ reflecting the contribution of common factors to original variables:

$$\mathbf{h}_i^2 = \sum_{j=1}^{m} \boldsymbol{a}_{ij}^2$$

Below the premise of standardized variables, the variance of x_i can be expressed as

$$\mathbf{h}_i^2 + \boldsymbol{\varepsilon}_i^2 = \mathbf{1}.$$~~;~~

The sum of squares of a_{ij} in column *j* is ~~V~~variance contribution:

$$\boldsymbol{S}_j = \sum_{i=1}^{p} \boldsymbol{a}_{ij}^2$$

The higher S_j is~~,~~ the more significant the common factor *j* is;
$\boldsymbol{\varepsilon}$ is regarded as the part that cannot be explained by common factors, equivalent to the residual term.

Common factors prove to have the following characteristics: a. the number of factors m is less than the number of original variables p; b. common factors are not simple reduction of original variables, but effective recombination; c. there doesn't exist correlation between different factors, which means one variable cannot be explained by multiple common factors; d. common factors have named explanatory according to the combined information.

(3) Procedure of EFA and cross-sectional regression
Basic steps of factor analysis focus on two central issues: how to construct, name and explain the common factors. Here, ~~This paper~~we will do the factor analysis as per~~ccording to~~ the following steps:

a. Judge the fitness of all the original variables to do EFA
The most crucial assumption of EFA is that there exist strong correlations between original variables. Therefore, correlation analysis will be carried out first~~ly~~. The simplest method is to conduct statistical test with the correlation coefficients matrix. If the majority of correlation coefficients is less than 0.3 and cannot pass the statistical test, factor analysis is inapplicable.

In this paper, two methodologies will be adopted. The first one is Kaiser-Meyer-Olkin test (KMO). KMO statistics is an index used ~~f~~to~~r~~ compare~~ing~~ the correlation coefficients and partial correlation coefficients.

$$\mathrm{KMO} = \frac{\sum\sum_{i\neq j} r_{ij}^2}{\sum\sum_{i\neq j} r_{ij}^2 + \sum\sum_{i\neq j} p_{ij}^2}.$$

r_{ij} is denoted as the simple correlation coefficient between variables i and j~~-~~, and p_{ij} is the partial correlation coefficient~~s~~.and KMO statistics is an index between 0 and 1. The closer the value is to 1, the more appropriate it is for~~to do~~ analysis. Kaiser even~~r~~ gave the measurement level: 0.9 means pretty appropriate; 0.7~~ to~~ 0.8 ~~means~~ suitable; results under 0.6 are not recommended for~~to do~~ factor analysis.

The second method is Bartlett Test of Sphericity, which starts from correlation coefficient~~s~~ matrix. The null hypothesis: correlation coefficient~~s~~ matrix is supposed to be an identity matrix. If the data~~statistic~~ is relatively large and p-value is less than significance level, then we reject the null hypothesis. This means the correlation coefficient~~s~~ matrix proves not to be an identity matrix which is appropriate for factor analysis; ~~otherwise, i~~f~~t is~~ not, it is not suitable~~ recommended~~ for~~to do~~ factor analysis.

b. Construct the common factors and calculate the factor~~ ~~loading matrix

After making sure the applicability of EFA, we have several traditional methodologies for constructing the common factors: Principal Component Analysis, Principal Axis Factoring, Maximum Likelihood Method and Least Square Method.

The key difference between ~~all~~the above method~~s~~ is the ~~as~~ following: Principal Component Analysis, just through rotating the coordinates, recombines the ordinary variables into a new linear combination, also called Principal Component; while Principal Axis Factoring, Maximum Likelihood Method and Least Square Method base on the Factor Model in which combination between potential variables and random effects variables represents the initial variables. This paper employs the most common method~~:~~. Principal Component Analysis. All the variables in the new combination are uncorrelated, and are expressed as follows:

$$y_1 = u_{11}x_1 + u_{21}x_2 + \cdots + u_{p1}x_p$$
$$y_2 = u_{12}x_1 + u_{22}x_2 + \cdots + u_{p2}x_p$$
……
$$y_p = u_{1p}x_1 + u_{2p}x_2 + \cdots + u_{pp}x_p$$

~~W~~where y_i is uncorrelated with y_j; the variances of $y_1, y_2, \ldots y_p$ as well as the ability to combine variables ~~are~~ decrease~~ing~~.

The coefficients u are required to be

$$u_{1k}^2 + u_{2k}^2 + \cdots + u_{pk}^2 = 1\ (k = 1,2, \ldots p).$$

Therefore, $y_1, y_2, \ldots y_p$ are referred to as the first, the second…and the last principal component of the original variables. In general, only several components with highest variance will be selected as principal component.

During the Principal Component Analysis, we obtain p eigenvalues and related eigenvectors used for calculating the factor loading matrix A. Based on the main purpose of Principal Component Analysis to decrease the number of variables, we usually select loading matrix of m common factors, m＜p. Eigenvalues $S_i(i = 1,2, \ldots p)$ of common factors are required to be larger than 1. What's more, the total cumulative variance contribution rate of these m factors $(S_1+S_2+\ldots+S_m) / \sum S_i$ should be larger than 80%. $S_i(i = 1,2, \ldots p)/\sum S_i$ is denoted as the percentage variance contribution of every factors.

c. Increase the interpretability of factors by rotation

The interpretation of common factors is another core issue of factor analysis. As eEvery common factor combines information from multiple variables;, in other words, one common factor only enables to explains small amount of information from every variable, a common factors is not a typically representative ofor any single variable. The situation above leads to cryptic interpretation of new factors. The ideal access to better interpreting the common factors is rotation. In this paper, we use Varimax Rotation Method to realisze a clearer factor interpretation, implying that the loading of a certain variable on one factor closes to 1 and loadings on other factors close to 0. One common factor turns to be a typical representative ofor one certain variable. But it is regrettable that the interpretations of factors isare quite subjective. We come across vVarious interpretations occur in various combinations. So far, there is no clearity criterionа for this.

d. Calculate the factor scores.

The final step is to calculate the factor scores according to coefficients of the score function by Regression Method.

$$F_j = \beta_{j1}x_1 + \beta_{j2}x_2 + \cdots + \beta_{jp}x_p \quad (j=1,2,\ldots\ldots m)$$

So far, we have examined the work of factor analysis is done. In the empirical part, we obtain the APT model by two-step multiple linear regressions. In the first part, the factor scores are used as independent variables and returns tofrom sample stocks in A Share market are used as dependent variables. In the second part, constant terms in the first regression representing the expected returns are used as dependent variables and coefficients are used as independent variables, getting the final APT model.

Ⅳ. Data

The data used are monthly closing prices ofor a randomly selected sample of 100

stocks in the Shanghai Stock Exchange A Share Index as proxy variables for China's A Share market from January 2008 to December 2012~~,~~ provided by Wind Information. To avoid ~~the~~ stock price distortion resulting from exclusion of~~ding~~ dividend and rights, all price parameters will ~~data~~ adopt~~s~~ backward adjusted price without missing data. For the monthly return to stocks cannot be obtained directly, we ~~use the~~ calculate~~ing method~~ as follows:

$$\frac{\text{Closing price in month T} - \text{Closing price in month (T} - 1)}{\text{Closing price in month (T} - 1)}$$

~~In terms of choosing the appropriate group of economic factors,~~ Since there is no formal theoretical guidance to choose the appropriate group of economic factors, we turn to~~.~~ Chen, Roll and Ross (1983) who gave a group of macroeconomic factors including *industrial production index*, *unexpected inflation*, *changes in default risk*, *unexpected change in term structure*.~~;~~ ~~b~~Based~~ing~~ on the conclusion of Chen, ~~Roll and Ross, Antoniou, Garrett and Priestley~~et al. (1998) added *real retail sales*, *commodity prices*, *exchange rate* also; Cagnetti (2002) obtained ~~8~~ eight factors by conducting factor analysis on ~~thirty three~~33 variables~~,~~~~,~~; these factors ~~respectively to be~~are: *market portfolio factor, fixed income securities factor, risk premium, foreign variables, production and monetary, inflation, industrial production and people's expectations*; Yasushi (1988) added international variables, especially oil price to examine the impact of international factors on the returns from~~to~~ stocks.

We have arrived at the following macroeconomic factors after ~~Having learnt~~ examining several previous works: ~~others' study results, the macroeconomic factors we use are~~ Currency in Circulation (M_0), Narrow money supply (M_1), Broad money supply (M_2), Consumer Price Index (CPI), Producer Price Index (PPI), Foreign Direct Investment (FDI), Net Trade (NT), Gold Reserve (GR), Foreign Currency Reserve (FCR), NYMEX Oil Future Price (OP), Exchange Rate by indirect quotation (ER), 3-month Treasury Security Yield (SB), 10-year Government Bond Yield (LB), Inflation Rate (IFR), Benchmark Interest Rate (BIR), Ratio of Exports to Imports for China (REI), Macro-economic Climate Index (MECI), Import Price Index (IPI), Consumer Confidence Index (CCI), Business Confidence Index (BCI), Financial Institution Short-term Loan (FSL), Financial Institution Long-term Loan (FLL). Unfortunately, monthly data of GDP and Industrial Production ~~a~~were not readily available.~~found.~~ The date of macroeconomic variables is~~are~~ ~~selected~~ based on monthly ~~level~~ data ~~during~~ of the same period ~~with data~~as of stock price and ~~there exist~~ no ~~missing~~ data is missing.

Comment [.6]: Concept not clear.

Table 1 : Correlation Matrix

Note: Circulation (M_0), Narrow money supply (M_1), Broad money supply (M_2), Consumer Price Index (CPI), Producer Price Index (PPI), Foreign Direct Investment (FDI), Net Trade (NT), Gold Reserve (GR), Foreign Currency Reserve (FCR), NYMEX Oil Future Price(OP), Exchange Rate by indirect quotation(ER), 3-month Treasury Security Yield(SB), 10-year Government Bond Yield (LB), Inflation Rate (IFR), Benchmark Interest Rate (BIR), Ratio of Exports to Imports for China (REI), Macro-economic Climate Index (MECI), Import Price Index (IPI), Consumer Confidence Index (CCI), Business Confidence Index (BCI), Financial Institution Short-term Loan (FSL), Financial Institution Long-term Loan (FLL).

Table 1 shows the correlation matrix of all the macroeconomic variables over 60 months from 2008 to 2012. Just taking several variables as ~~an~~ examples, the high correlations between (i~~1~~) money supply and CPI, (~~2~~ii) inflation and PPI, (~~3~~iii) net trade and ratio of exports to imports, and (iv~~4~~) foreign currency reserve and exchange rate ~~show the information of~~ related to variables can be explained by certain common factors. In addition, the low correlation between oil price and money supply implies the existence of no less than two common factors; a~~ny~~ single variable alone will ~~is~~ not be sufficient enough to explain the macro economy.

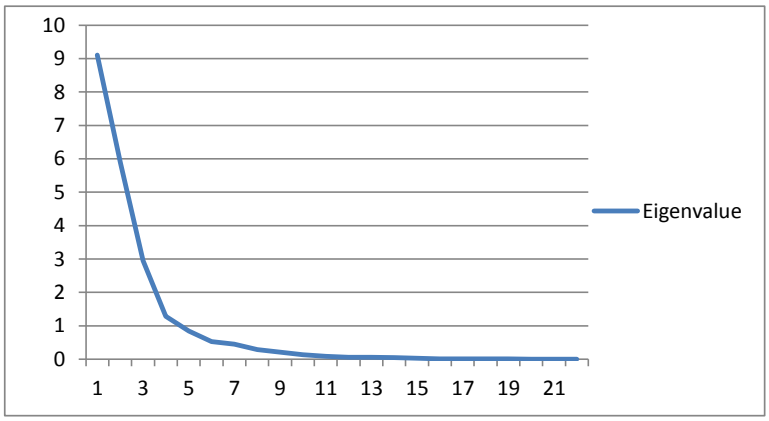
Ⅴ.Test~~ing~~ Results

(1) Exploratory Factor Analysis

The ~~statistic of~~ Kaiser-Meyer-Olkin test yields a factor of ~~is~~ 0.801; Bartlett Test of Sphericity gives Chi-square approximation of 2835.029 with degree of freedom of 231; the p-value is 0.000, less than the significance level 0.05. All results above prove that original variables are suitable for ~~to do~~ Exploratory Factor Analysis.

The following figure is the Screen Plot between common factors and eigenvalues, with horizontal axis representing the number of common factors and vertical axis ~~representing the eigenvalues~~ of the certain factor. ~~According to the Screen Plot,~~ ~~t~~The eigenvalues of first four factors show larger variations while ~~from~~ the fifth one ~~factor eigenvalues start to~~will level off. It implies that the first four factors have significant effect~~s~~ on ~~explaining~~ original variables. To increase the accuracy in the test of APT, we limit the study to ~~decide to keep~~ four common factors.

Figure 4: Screen Plot



As shown in ~~the Communalities~~ Table 2, the initial communalities of all the variables in column 2 are 1.000; the extraction communalities of 80% variables are more than 0.900, close to 1, which means the information can be explained by common factors ~~is~~by more than 90%, only with a small~~a little~~ information los~~t~~. Taking ~~Currency in Circulation (~~M_0~~)~~ as an example, 0.918 means 91.8% of variance can be illustrated by

Comment [.7]:

common factors.

Table 2: Communalities

Method: Principal Component Analysis

（3）Table 3 ~~i~~shows the Total Variance of ~~explained of~~ extraction factors as well as ~~that after~~ rotation. The second column in the table describes the variance contribution~~s~~ of all factors to initial variables in a decreasing order. ~~In t~~The third column~~, we get~~ shows the percentage of ~~variables~~ variance of variables explained by factors, also in ~~the~~ decreasing order ~~as well~~. The fourth column shows the accumulated percentages of variance explained by factors. The fifth to seventh columns ~~specially~~ show ~~the description of~~the four extracted factors. The variance contributions of factors are, respectively, ~~to be~~ 9.107, 5.898, 2.949, and 1.284, and~~.~~ ~~T~~the percentage of variance ~~explained for extracted factors~~ is ~~are separately to be~~ 41.396%，26.808%，13.404%，5.834%. For example, 41.396% equals ~~to~~ 20 divided by 9.107. The accumulated percentage~~s~~ of first four extracted factors are 41.396%, 68.205%, 81.609%, 87.443%. The information in table proves the ~~s~~results shown on ~~ame conclusion as~~ Screen Plot, recommending four common factors. The eighth to tenth columns show the results after rotation. The eigenvalues of four factors now are: ~~change to be~~ 8.911, 5.538, 2.658, 2.130, totally explaining 87.443% of the total variance of initial variables.

(3) Table 4 shows ~~contains~~ all the factor loadings before and after factor rotation~~,~~ showing ~~implying~~ the relation between initial variables and the common factors. We ~~give~~assign names with better interpretability to ~~the~~ common factors depending on the following factor loading matrix so that we can explain the structure and main characteristics further. As shown in ~~the t~~Table 4, before rotation, M_0, M_1, M_2, CPI, FDI, GR, FCR, EX, FSL, FLL have larger loadings on the first factor; PPI, OP, IFL, LB, MECI, IPI,CCI ~~have larger loadings~~ on the second factor~~,;~~ and NT, GR, BIR, REI, FCL ~~have large loadings~~ on the third factor~~?s~~. FDI, NT, GR have similar loadings on different factors. Apparently, there exists a common thread ~~coincident information in~~among all the four factors, which will ~~goes against to~~ explain the common factor and provide a~~nd find~~ proper economic direction. ~~implication.~~

We now employ ~~After using Varimax Method,~~the factor rotation method which makes factor loading matrix clear~~er~~ and gives better interpretation~~bility~~ than earlier~~before~~. Every variable can be explained by a single factor. The first factor mainly reflects the information concerning *Consumption and Investment Activity*

~~*Factor*,~~ including Currency in Circulation (M_0), Narrow money supply (M_1), Broad money supply (M_2), Consumer Price Index (CPI), Foreign Direct Investment (FDI), Gold Reserve (GR), Foreign Currency Reserve (FCR), Exchange Rate by indirect quotation (EX), Financial Institution Short-term Loan (FSL), Financial Institution Long-term Loan (FLL); the second factor covers the information in Producer Price Index (PPI), NYMEX Oil Future Price (OP), Inflation Rate (IFL), 3-month Treasury Security Yield (SB), 10-year Government Bond Yield (LB), Macro-economic Climate Index (MECI), Import Price Index (IPI), and Consumer Confidence Index (CCI). The ~~, we call the~~ second factor is the *Price Movement and Yield to Treasury Bonds Factor*; the third factor includes Net Trade (NT), Ratio of Exports to Imports for China (REI), Business Confidence Index (BCI), ~~is~~ named ~~as~~ *International Trade Factor*; then the fourth factor explains the left Benchmark Interest Rate (BIR), which can be defined as the *Benchmark Interest Rate Factor*. Every initial variable ~~has~~ belonged ~~to~~ a common factor.

Table 4: Factor Loadings

Formatted: Centered

Method: Orthogonal rotation method； Rotation converges after five~~5~~ iterations

Table 5 indicates the covariance matrix of common factors. Columns 2 ~~to~~ 5 express the covariance before rotation and ~~column~~ 6 ~~to~~ 9 ~~show the covariance~~ after rotation. According to the characteristics of factor analysis, factors are supposed to be orthogonal and unrelated to each other. Results in the covariance matrix after rotation prove this point, which is the appropriate ~~to be the~~ factor ~~in~~for APT.

Table 5: Component Score Covariance Matrix

(4) Table 6 shows us the coefficients of the score function by Regression Method.

Table 6: Component Score Coefficient Matrix

Extraction Method: Principal Component Analysis
Rotation Method: Varimax with Kaiser Normalizaion. Component Scores

Therefore, we can write the score function based on the following:~~is.~~

$$\begin{aligned} F_1 &= 0.112M_0 + 0.115M_1 + 0.113M_2 \ldots + 0.111\text{FSL} + 0.114\text{FLL} \\ F_2 &= 0.005M_0 + 0.031M_1 - 0.006M_2 \ldots - 0.026\text{FSL} + 0.025\text{FLL} \\ F_3 &= 0.040M_0 + 0.050M_1 + 0.026M_2 \ldots + 0.017\text{FSL} + 0.040\text{FLL} \\ F_4 &= -0.011M_0 - 0.108M_1 - 0.022M_2 \ldots + 0.057\text{FSL} - 0.085\text{FLL} \end{aligned}$$

Finally, we can calculate all the factor scores (see Appendix A). ~~which will be shown in the Appendix.~~

(2) Empirical Test of APT

Through the analysis above, we define four common factors: *Consumption and Investment Activity Factor*, *Price Movement and Yield to Treasury Bonds Factor*, *International Trade Factor*, *Benchmark Interest Rate Factor* and obtain~~get~~ the factor scores. To build the APT model, we take the monthly returns ~~t~~of the 100 sample stocks from 2008 to 2012 as explained variable and take the factor scores as explanatory variables for use in ~~to do~~ multiple linear regression. As a result, we obtain the constant terms as expected return ~~to~~from each stock and coefficients ~~to~~of four factors (see Appendix B). In all the results, about 35% of joint p-values are significant at the significance level of 5% and ~~there are~~ an additional~~further~~ ~~about~~ 20% ~~joint p-values are significant~~ at ~~the significance level of~~ 10% level. Then we take constant terms as explained variables and take coefficients as explanatory variables for ~~do~~ a second multiple linear regression. ~~again.~~

As mentioned in Section Ⅲ, we ~~this paper~~ add~~s~~ the RQFII as a dummy variable. Nevertheless, during all ~~the~~ regressions, RQFII is not significant ~~t~~for all the sample stock returns, and even lowers the whole significance of~~or~~ each regression. Hence, we will not perform the second regression with dummy variables. The test result~~s~~ suggests that the RQFII system hasn't played a significant role in improving China's A Share market so far.

Using the unrevised model, we get the regression equation as follows:

$$\begin{aligned} E(R_i) = -0.0069029 - 0.1520825\widehat{b_{i1}} - 0.2327547\widehat{b_{i2}} - 0.0027024\widehat{b_{i3}} \\ +0.012801\widehat{b_{i4}} \end{aligned}$$

We ~~can~~ notice the following points from the above regression equation: on ~~the~~ one hand, risky security *i* has negative sensitivity coefficients on the first three common factors, implying that the larger the risk premi~~ums~~ of *Consumption and Investment Activity Factor*, *Price Movement and Yield to Treasury Bonds Factor*, *International*

Trade Factor ~~are~~, the smaller is the expected return rate to the risky security *i* ~~is~~; on the other hand, risky security *i* has positive sensitivity coefficients on the last common factor, implying that~~e~~ larger the risk premium of *Benchmark Interest Rate Factor* is, the larger is the expected return rate to security *i* ~~is~~.

The p-value to the total test is 0.0395, smaller than the significance level $\alpha = 0.05$, which means that all the variables jointly have significant effect on expected return rates to stocks. However, the R-square which measures the goodness of fit is not satisfactory and there exists residual between the model and the real market. In this paper, we consider that this result may be caused by data deficiencies. Data of China's A Share Market has the problem of inadequate macroeconomic index and inconsistent statistical standard. Many macroeconomic factors only have annual data but not monthly data, such as *GDP*, *Industrial Production and Retail Sales*, which may cause the APT Model underspecified and further result in the low goodness of fit. Another reason is supposed to be that China's A Share market is not mature and capital market is not sufficiently opened. The macro-environment in China cannot satisfy the assumption of APT, especially the assumption of a well-diversified, perfectly competitive and frictionless economy. There may exist great arbitrage chances in China's A Share market, and the great bubble in stock market just proves this.

Ⅵ. Conclusion

In this paper, we obtain four common factors with clear economic interpretations by Exploratory Factor Analysis, separately to be *Consumption and Investment Activity Factor*, *Price Movement and Yield to Treasury Bonds Factor*, *International Trade Factor and Benchmark Interest Rate Factor*. All these factors are combined and simplified from 22~~twenty two~~ initial variables, such as Money supply, Net Trade, Interest Rate and Inflation Rate~~, etc~~. This paper takes the factors of international trade, monetary policy, macro-economic climate index, commodity price~~s~~ movements into account so that we can get significant macroeconomic factors as wide as possible and describe the current situation of the A Share market accurately. Data ~~A~~analysis ~~about statistics~~ shows the extraction effect~~s~~ is very good. The only pity and shortcoming is that data like GDP, Industrial Production which ~~have ever proved to perform significant roles in~~ affect~~ing~~ the expected returns significantly of~~to~~ risky securities cannot be provided, which may have influenced ~~on~~ the result of factor analysis.

In the empirical part, we take 100 sample stocks from Shanghai Stock Exchange A Share Index as ~~the~~ proxy variables to A Share market. Then this paper builds on the APT model based on four priced factors obtained by factor analysis and investigates the validity of the ~~APT~~ model in ~~the~~ China's A Share market. According to ~~the~~ test results, we found the negative effect of *Consumption and Investment Activity Factor*, *Price Movement and Yield to Treasury Bonds Factor*, *International Trade Factor* ~~have~~ on the expected return ~~f~~from risky securities and the positive effect of *Benchmark Interest Rate*. However, the APT model cannot be proved as a good fitting

model to current stock market, ~~likely because of~~for two reasons: first, underspecified model because of data deficiencies; second, the incomplete and relatively closed and controlled market environment. Current China's stock market is in ~~the~~ transition stage, behaving as not sufficiently open and imperfect market system. Many measures aiming at improving~~e~~ the A Share market are just testing the water, just like the RQFII system which hasn't had too much ~~positive~~ effect. ~~on the development of A Share market.~~ Further, ~~lagging of~~ policy lag~~ies~~ means that it will take time for policies to work. Keeping ~~Aiming at~~the above problem in mind~~s~~, this paper has several suggestions: for one thing, expand the data capacity for further study to increase the fitting degree of APT model; for another ~~thing~~, promote ~~the~~ marketization(???) process in China's A Share market to increase the sensitivity of market to changes in macroeconomic environment.