西南财经大学

Southwestern University of Finance and Economics

2016 届

本科毕业论文（设计）

论文题目： Multiple-Factor Risk Model

--Application in China’s A-share Market

学生姓名： 汪嘉骏

所在学院： 经济与管理研究院

专　　业： 金融学（经济与管理国际化创新人才班）

学　　号： 41209009

指导教师： 黄霖

成　　绩：

2016 年 4 月

西南财经大学

本科毕业论文（设计）原创性及知识产权声明

本人郑重声明：所呈交的毕业论文（设计）是本人在导师的指导下取得的成果。对本论文（设计）的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。因本毕业论文（设计）引起的法律结果完全由本人承担。

本毕业论文（设计）成果归西南财经大学所有。

特此声明

毕业论文（设计）作者签名：

作者专业：金融学

作者学号：41209009

年　月　日

西南财经大学本科学生毕业论文（设计）开题报告表

论文（设计）名称	结构化多因子风险模型在中国 A 股中的应用				
论文（设计）来源	自选	论文（设计）类型	B	导 师	黄霖
学生姓名	汪嘉骏	学 号	41209009	专 业	经济与管理研究院

开题报告内容：（设计目的、要求、思路与预期成果；任务完成的阶段内容及时间安排；资料收集计划；完成论文（设计）所具备的条件因素；写作的基本思路等。）

一、设计目的：

本论文旨在探究结构化多因子模型在中国股票市场中的应用，将个股的超额收益率进行因子分解，分解成不同因子的因子收益率与特异收益率（残差项）之和。在得到模型可以解释的个股收益率之后在一段时间窗口下动态计算该时期的投资组合总风险（以标准差来衡量），以组合总风险形成的时间序列进行对于总风险偏高时期的判断，最后完成风险归因，包括因子归因和资产归因，即找寻出使得整个组合风险偏大的原因或具体仓位，进行相应的减仓建议。

二、设计要求与思路

（一）因子寻找及收益率分解

在结构化风险模型中，收益率将被以以下形式进行分解：

$$\mathbf{r} = \mathbf{X} \cdot \mathbf{b} + \mathbf{u}$$

上式中，r 为个股超额收益率向量，X 为因子暴露矩阵，b 为各因子的因子收益率向量，剩下的 u 为特异收益率残差项的向量。X 记录的是每一个个股在各个因子上面的因子暴露，b 中的因子收益率是对每个个股都相同的收益率。对于模型构建需要达到的效果是 u 向量中的个股特异收益率在一段时间中的序列满足个股与个股之间相关性几乎为零，说明收益率分解比较完美，即公共的收益率都被在因子上面暴露乘以公共的因子收益率所解释，只剩下个股特异的收益率残差项，如果互不相关即说明因子找寻有效。

对于模型构建的思路，首先是因子的寻找，分为主要四种因子：技术指标因子（波动率，MACD，成交量持仓量，KDJ，动量指标、市场 beta 等）、行业因子（初步选为 0-1 变量作为因子暴露）、财务因子（市净率市盈率、总市值、财务杠杆等）以及宏观因子（国际油价、CPI、PMI 等）。数据将采用月度数据，因为时间长度够长且混合频率数据的处理比较容易。构建思路如下：

（1） 得到个股每个月的技术指标，运用 Fama-MacBeth 两步回归的思路，先针对每只个股进行时间序列回归得到个股在各个因子上面的因子暴露 X，进行第二步的横截面回归得到每个公共因子每天的因子收益率 b；

（2） 对于行业因子，如果采用 0-1 变量则其本身就是 X 因子暴露矩阵了，所以直接横截面回归得到因子收益率 b；

（3） 宏观因子在本文中将被考虑作为 b 因子收益率序列，所以进行时序回归得到因子暴露 X；

（4） 财务因子可能会采用差值拟合的方法升高频率为月度，然后类似技术指标因子的处理方式进行两步回归。

（二） 风险模型

上述四步将会得到收益率分解的模型，之后完成风险模型的构建：

$$\mathbf{V} = \mathbf{X} \cdot \mathbf{F} \cdot \boldsymbol{X}^{\boldsymbol{T}} + \Delta$$

V 为个股的超额收益率方差协方差矩阵，X 为因子暴露矩阵，F 为因子收益率矩阵，$\boldsymbol{X}^{\boldsymbol{T}}$为因子暴露矩阵的转置，它们三个相乘即为模型解释的收益率之方差协方差矩阵，加上最后的特异收益率矩阵Δ。模型想要达到的效果是特异收益率矩阵Δ几乎是一个对角矩阵，即个股之间的特意收益率不存在明显的相关性。利用所得的解释风险进行加入持仓权重，个股被解释收益的方差协方差矩阵加入持仓比例向量之后会得到整个投资组合的总风险（这里将用标准差来进行度量），为月度，在一段时间内即可以得到组合的风险时间序列，利用类似布林线的方法在其序列均值上下 2 或 3 个标准差作出上下通道，如果风险序列上传上方通道说明组合总体风险过高，则进行风险归因。这部分方法还有待考究，BARRA 提供了因子归因和资产归因两种方法，即先寻找到对组合风险贡献较大的因子，然后再找出因子中出问题的持仓资产，找寻到适当量化方法后可以提出相应的减仓建议，即对哪只个股减仓以及减仓多少。

三、阶段内容即时间安排

时间	需完成的任务
2016.1-2 月	继续查阅相关文献和研报，定出需要计算的因子，给出因子 list
2016.3	下载数据开始数据预处理，计算出因子暴露和因子收益率，完成收益率分解模型
2016.4	选定滚动窗口，完成风险模型，解释风险归因的理由和方法，完成风险归因
2016.5	完成测试，准备最后的答辩

四、资料收集和完成所需条件

本论文所需资料将以 BARRA 提供的研报为主，将他们研究的美股市场转移到中国会遇上很多不一样的情况，所以对于因子的选定会有较大适应性改变，对于其模型思路要进行理解后的提升和改善，使得更加适应研究中国股市所需的情景。另外，对于文献的查阅将集中搜索多因子模型的应用，不局限与股市比如在期货市场的多因子模型也将被参考，确保建模过程的细节考虑更加完善。

所需条件，数据将主要来自于万得数据库，因为需要处理混频数据，财务数据和部分宏观数据是季度而模型视角为月度，所以寻找到适当的差值方法是论文完成的一个必要条件。另外，对于模型的修正需要参考大量的研报，每一个细节都需要做得较为精致，才能使得对收益率分解完成得干净，特异收益率呈不存在明显相关性。

五、写作基本思路

依照时间主线，按照模型的简历顺序进行书写，大体分为：

Introduction--Literature Review—Data Pre-progressing—Factor Model—Risk Model--Conclusions

其中的具体步骤将会依照实施过程进行适当调整，最终形成完整论文。

指导教师签名：　　　　　　　　　　　　　　日期：

论文（设计）类型：A—理论研究；B—应用研究；C—软件设计等；

Multiple-Factor Risk Model

--Application in China's A-share Market

Abstract

This paper refers to the BARRA's Multiple-Factor Model (MFM). According to the research ideas of constructing the MFM, in total 48 factors from the respective 5 aspects including technical indices, fundamental economy, market access return, industry allocation as well as firm characteristic factors are used to divide the individual stock abnormal return. These chosen factors are responsible for measuring various relevant situations of the risks and benefits of the individual stocks in China's A-share market. This kind of modelling puts itself in favorable position in dividing the individual return into factor returns and specific returns cleanly.

After rolling the factor return estimation window during every 120 months into a dynamic model, every month this paper calculates the total risk of a portfolio given a certain weight of positions. Total risk here is measured by the annual standard deviation estimated from the variance-covariance matrix of the predicted individual abnormal return by MFM. With this tool the portfolio manager can firstly observe the trend and the current total risk he is bearing.

The last section researches into the way of attributing the risks to every asset in the portfolio. By taking the partial derivative of the total risk with respect to the weights of stocks, the marginal contribution of each asset to the risk can be acquired. This serves as an application for the investor to make corresponding adjustment in his positions in the whole portfolio according to their marginal contribution on total risk.

Key words: BARRA Multiple-Factor Model, Risk Analysis

Contents

1. Introduction

In correspondence with the development of modern financial theory as well as the demand of market investment in reality, mathematical models are playing an increasingly important role in the construction of investment portfolio. Other than pursuing a higher active return based on the management of the portfolio, modelling the risk of a certain investment method selecting the target financial assets serves as another significant topic people are concerned about.

Among the various ways of calculating the total risk of a portfolio, the multiple-factor model (MFM) proposed by *BARRA* (1998) [1] has an expansionary application in the real practitioner's world. By using the multiple-factor model one can calculate the reasonable explained risk of a portfolio. It is based on the following specific thoughts of construction:

The return of the stocks can be explained by a group of common factors and an individual-stock-related specific return. We can regard these common factors as a force impacting the movement of the stock prices. This kind of decomposition can provide us with a better way of mainly tracing the specific returns of the stocks after dividing the return from the contribution of the common part. The specific returns refer to those which only vary with the stocks. And modelling this part of returns can give us a better understanding of the specific movements of their prices after subtracting the common forces offered by the common factors we choose. As for the common factors, their part can be explained by their loadings on the factors as well as the factor returns of these factors. This will be illustrated in the following section of the model construction.

2. Literature Review

In 1935, *Hicks John Richard* [2] proposed in his research that people's demand of currency arises from the pursuit of high return in line with low risk. And he also thinks that the index of risk has to be considered in the macro analysis in order to conduct researches about the returns in the asset market.

Hicks [3] also issued the concept of risk premium in his theory of general equilibrium in 1936. In their view risks can be dispersed in some appropriate ways. And it was cited by *Jacob Marschak (1959)* [4] of the phenomenon where people tend to capture high return while burdening low risk. Till 1952, the Portfolio Selection by *Markowitz* [5] serves as a foundation of the modern active portfolio management together with its fresh concept of market efficiency. Another important point for this paper to be referred is that it proposes the way of using standard deviation as the risk indicator, which is also the target of our model in the final part.

In 1964, CAPM arose from the paper of *William Sharpe* [6]. It proposed that the market return is not only affected by the return of risk-free investment tools but also the return from the risk premium of the market investment tools and the risk-free assets. They claim that personal investors are faced with two kinds of risks. Systematic risk cannot be diversified by adjusting the investment portfolio or increasing investment assets. And generally it refers to those factors not determined by the individual units but by some whole economic environment. Another one is unsystematic risk, which is totally decided by individual assets and can be diversified to zero by enhancing the kinds of investment tools.

It was claimed by *Rosenberg (1974)* [7] that in the short run, macro factors have obviously weaker explanatory power on returns and the movement of risks in contrast with technical factor and industry factor. However, in the long run macro factors seem to be more convincing than other factors in decomposing the market returns.

In 1976, *Stephen Rose* [8] proposed the theory of Arbitrage Pricing Theory (APT) in his paper published in Journal of Economic Theory. This is actually a development of CAPM theory. But the difference lies in the fact that it researches into the basis of one dependent variable with multiple independent variables. And in fact this is another direction of modern asset return research.

Fama and Frence issued their famous Fama-French 3-factor model in 1993 [9]. They discovered the significant relationship of book-to-market ratio and size towards the individual returns in the market, which utilized the information from the side of firm characteristic factors.

And *Grinnold and Ronald* (2000) [10] found that compared with other factor models, BARRA model can have a better result in one-side test. Besides, BARRA MFM can make use of more factors and thus more information from more aspects to analyze the stock returns in a wider range of sources.

More recently, Thierry and Pim (2004) [11] proposed their result of testing that multiple factors indeed contribute more explanatory impacts on value and momentum portfolios in time periods after 1963 using the USA stock market data.

In addition, in 2014 Lubos, Robert and Lucian [12] found that active funds' potential to earn more than passive benchmarks declines as the active industry increases, which offers the evidence of the significant power of industry factors. This gives rise to the thought of putting industry factors into the model of stock returns.

As a disagreement, Charoula and Alexandros (2014) [13] doubted the existence of common factors in dividing the individual returns in the commodity market. They issued their result that none of the multiple-factor models can be useful in modelling the futures returns. Therefore, after construction MFM for China's stock market, we plan to conduct the related researches in China's futures market, too.

3. BARRA Model Construction

3.1 Overview and Model Assumptions

This part will be responsible for the introduction of BARRA MFM.

3.1.1 Model Outline

Structured risk model firstly decomposes the stock return linearly. There are four major parts during the process of decomposition: abnormal return of individual stocks, factor loading or factor exposure of stocks, factor return and the stocks' specific return. Formula of Equation 3.1.1 below demonstrates the four parts respectively.

$$r_i - r_f = \sum_{k=1}^{K} X_{i,k} \cdot b_k + \epsilon_i \tag{3.1.1}$$

Among them their meanings are as follows.

$r_i - r_f$: individual return of stock *i* minus the risk-free return, which is the abnormal return of a certain stock;

$X_{i,k}$: loading of stock *i* on the factor *k*. And within these loadings the industry loading will be dummy variables, which will be shown in the following parts;

b_k : factor return of factor *k*. In this model every month's factor return will be estimated by regressing the individual abnormal return on the loadings;

ϵ_i : individual stock's specific return, this will be estimated as the residual terms in the final cross sectional regression.

And if we denote our model with the vector and matrix specification, it will be as in Equation 3.1.2 below.

$$\boldsymbol{r} = \boldsymbol{X} \cdot \boldsymbol{b} + \boldsymbol{\epsilon} \tag{3.1.2}$$

Here $\boldsymbol{r}$ is the N by 1 abnormal return vector. $\boldsymbol{X}$ is the factor loading matrix. $\boldsymbol{b}$ is the K by 1 factor return vector as a time series and $\boldsymbol{\epsilon}$ stands for N by one vector of the stocks' specific return.

Then under the assumption that specific returns of all the stocks are uncorrelated with each other, which means that we decompose the stocks' return cleanly, we can estimate the risk structure of the portfolio as is shown in Equation 3.1.3.

$$V_{i,j} = \sum_{k_1,k_2=1}^{K} X_{i,k_1} \cdot F_{k_1,k_2} \cdot X_{j,k_2} + \Delta_{i,j} \tag{3.1.3}$$

Among them,

$V_{i,j}$: total covariance of stock *i* and stock *j* and specially when *i*=*j* it is the total variance of stock *i*;

X_{i,k_1} : loading of stock *i* on the factor k_1;

F_{k_1,k_2} : variance-covariance of the factor returns k_1 and k_2;

$\Delta_{i,j}$: covariance of stock *i* and stock *j*'s specific return.

3.1.2 Construction Procedure

This process of estimating the covariance of the assets within our portfolio and finally getting the total portfolio variance as the indicator of the total risk will serve as the target of the risk model. Figure 3.1 summarizes the above steps in the next page.

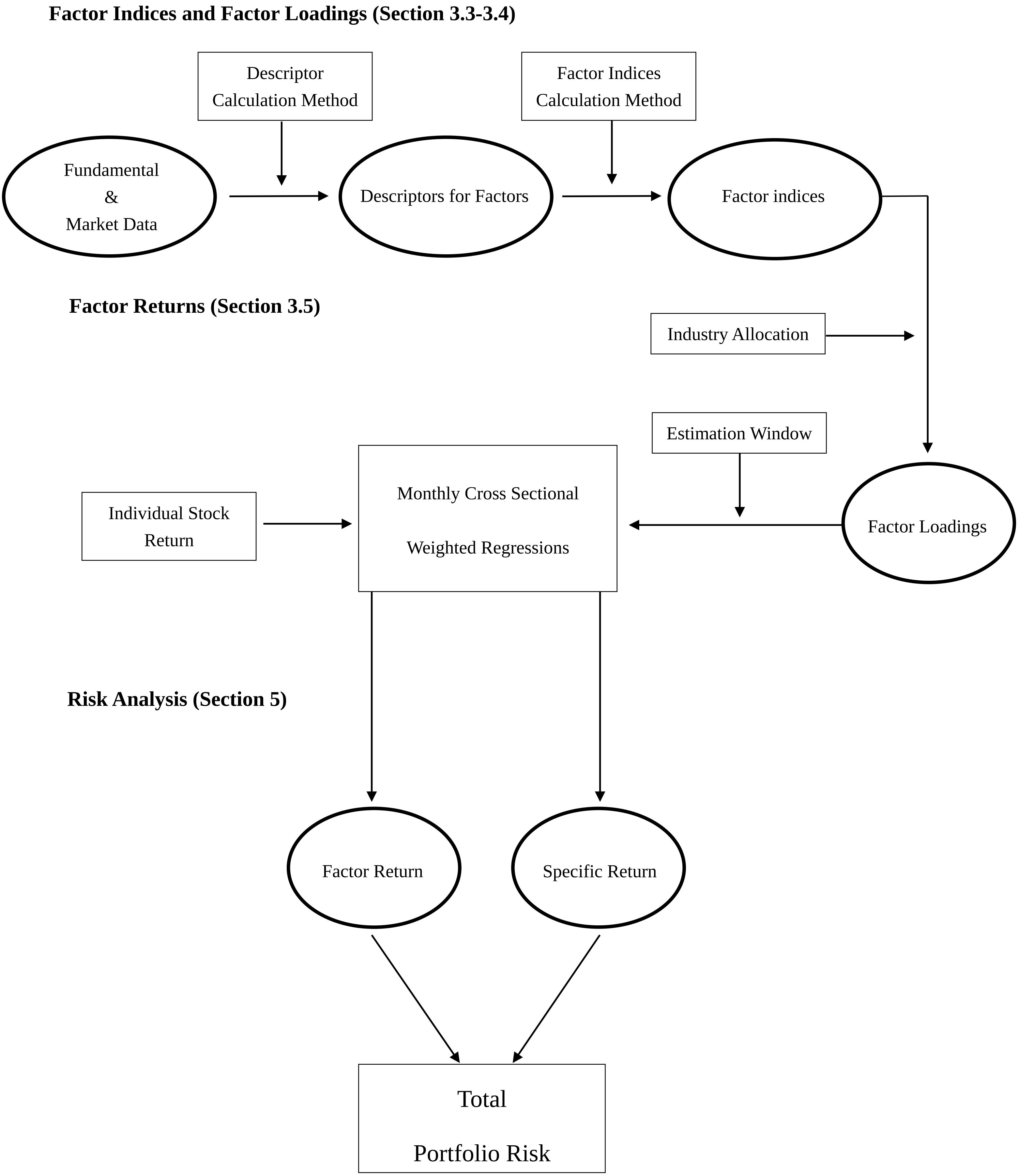

Factor Indices and Factor Loadings (Section 3.3-3.4)
Descriptor Calculation Method
Factor Indices Calculation Method
Fundamental & Market Data
Descriptors for Factors
Factor indices
Factor Returns (Section 3.5)
Industry Allocation
Estimation Window
Individual Stock Return
Monthly Cross Sectional Weighted Regressions
Factor Loadings
Risk Analysis (Section 5)
Factor Return
Specific Return
Total Portfolio Risk

Figure 3.1

3.1.3 Model Assumptions

In order to simplify the MFM this paper makes the following assumptions:

(1) According to many papers including Fama-French 3-factor mode, all of the financial market descriptors are mainly the individual returns, as well as trading volumes, and all of the technical indices calculated from them are stationary;

(2) Every company selected in our model will not change the industry they are in;

(3) The stock return can be divided into factor return and specific return linearly;

(4) If a certain factor is not significant in the time series regression of the individual return, it is assumed that this stock has no significant exposure on that factor but it does not mean that this factor has no influence on the stock.

3.2 Data Acquisition

The main data source for market data is WIND. This paper selects the stocks within Hushen 300 index. And after data cleaning process dealing with some missing values of the stocks data, the monthly raw market data are those of the 107 stocks in Hushen 300 index from January 2000 to December 2015, in total of 107 stocks with the whole 192 months.

In addition, macro data and firm characteristic data are captured from CSMAR. All macro data are at monthly frequency and firm characteristic data are at quarterly frequency. Data of exchange rate are cited from Federal Reserve Economic Data.

3.3 Risk Factor Indices

The most important step of the whole process is the selection of common factors. There are 5 types of factors in this model: technical index factors, market factor, industry factors, macro factors and firm characteristic factors. All of them will be introduced and calculated in detail in the part below.

3.3.1 Descriptor Standardization

The risk indices consist of descriptors designed to capture all the relevant risk features of a certain list company. Before we begin to handle them in the regressions they are standardized with respect of the estimation universe first in order to get rid of the effect of the different scales. The normalization process is summarized by the following relation:

$$\text{normalized descriptor} = \frac{raw\ descriptor - mean}{standard\ deviation} \tag{3.3.1}$$

3.3.2 Technical Index Factors

Technical indices are those which can most accurately and directly reflect the fluctuation situations in the financial market. Moreover, they have been used for many practitioners in the real world of stock investment. Based on this, technical indices gradually lose their power in predicting the movement of stocks in the next period but in the other way around, they can serve as very good descriptors for market expectations and moods. This is exactly going in line with the target of risk modelling instead of investment forecasting. In this paper particularly, there are in total 10 priori-factor of the field of technical indices. All of them will serve as descriptors first and will be used for acquiring the loadings. Attention has to be paid to the fact that the original technical indices are calculated on the day-count basis and here this paper just switches it into a month-count basis directly for simplicity. After testing, step by step some of them will be dropped due to insignificant impact of their factor loading on the returns.

(1) Volatility (5 months)

This paper chooses the standard deviation within the past 5 months almost half a year as the volatility. This is a comparatively long period and can thus contribute the general movement intense of the stocks.

(2) KDJ Index (9 months)

This index was proposed by *Doctor George Lane* (1950) and was firstly put into use in the futures market. As a stochastic index, KDJ measures the relative trend force of the buy side and the sell side. K stands for the fast confirmation line, D stands for the slow main line while J means direction sensitivity line. The calculation process is shown from Equation 3.3.2 to Equation 3.3.5. The first step of calculating the Raw Stochastic Value (RSV) is the most import concept in KDJ analysis: whether the market is in favor of the buy side or the sell side.

$$RSV_9 = \frac{\mathrm{close}_t - minimum\ price_{t-9,t}}{maximum\ price_{t-9,t} - minimum\ price_{t-9,t}} \cdot 100 \tag{3.3.2}$$

$$K_3 = \frac{RSV_t + K_{t-1}}{3} \tag{3.3.3}$$

$$D_3 = \frac{K_t + D_{t-1}}{3} \tag{3.3.4}$$

$$J_t = 3 \cdot K_t - 2 \cdot D_t \tag{3.3.5}$$

And the J-value is the descriptor of KDJ in each month.

(3) MACD (Moving Average Convergence and Divergence)

In the analysis of MACD this paper chooses the descriptor of EMA (Exponential Moving Average) as the MACD value to be used in the model construction. When faster moving average (MA) line crosses the slower MA line upwards, it signs a point where the prices of the stock is increasing at a growing speed and symbol an increasing trend of the prices. The calculation method is presented from Equation 3.3.6 to Equation 3.5.8.

$$\mathrm{DIF} = MA_{12}(close\ price) - MA_{26}(close\ price) \tag{3.3.6}$$

$$\mathrm{DEA} = MA_{12}(DIF) \tag{3.3.7}$$

$$\mathrm{EMA(MACD)} = DIF - DEA \tag{3.3.8}$$

Then EMA value is the descriptor of MACD in each month.

(4) Trading Volume

This is the direct factor as a measurement of people's thought and psychological mood towards the trend in financial market.

(5) Buy Signal

The *Granvi's 8 rules of trade (1960)* [14] using the MA line is becoming more and more popular in the trading operation methods for the investors in stock market. With a combination of MACD and the trading volume, we can construct another signal for capturing the long-side market. Under the condition that faster MA line crosses the slower MA line upwards and with an increasing trading volume, investors will regard it as a point when they can buy in the asset. The dummy variable of buy signal can be gained in the following method:

$$\text{Buy Signal} = \begin{cases} 1, if\ MACD > 0\ and\ trade\ volume\ increases \\ 0, otherwise \end{cases} \tag{3.3.9}$$

In this way, every month will have a dummy variable for buy signal of each stock. Figure 3.3.1 is an example (data source: Sina Finance) of a point with buy signal which is a time when people may consider to be a buying point (Stock: 002673 Xibu Security [15]).

Figure 3.3.1

(6) Sell Signal

Similarly, the dummy variable of sell signal can be constructed in the following way. When EMA is negative it means that the market is expecting a downward trend of the asset price. Together with a decrease of the trading volume, investors are mainly withdrawing their money from the market. At this time, we can assume that this is a sign of the short-side market.

$$\text{Sell Signal} = \begin{cases} 1, if\ MACD < 0\ and\ trade\ volume\ decreases \\ 0, otherwise \end{cases}$$

(3.3.10)

In this way, every month will have a dummy variable for buy signal of each stock. Figure 3.3.2 is an example of a point with buy signal which is a time when people may consider to be a buying point (Stock: 000025 Teli A [16]).

Figure 3.3.3

(7) RSI (Relative Intense Index, 3 months)

This index measures at a certain time period, the percentage of the fluctuation brought about by the price increase with respect of the whole fluctuation. Its formula is in Equation 3.3.11.

$$\mathrm{RSI}(3) = \frac{\text{Total price increase perentage}}{\text{Total price increase perentage} + \text{Total price decrease perentage}}$$

(3.3.11)

As is often the case, when RSI in the period is below 30 people will think that they will present a weak direction of buying. If it is above 75 it will illustrate a market with relative strong buying force.

(8) Momentum (9 months)

The simple calculation of momentum is subtracting the close price 9 months ago from the current close price. This is frequently used in trend analysis. Many stock analysts have the thought that the moving trend of the stock prices will proceed for a while. Therefore, the speed of the increase or decrease will approximately maintain the same. Momentum index will measure the lasting trend of the assets' movement.

$$\text{Momentum} = Close\ Price_t - Close\ Price_{t-9} \tag{3.3.12}$$

With this descriptor, this paper gets an indicator of the trend process within the past 9-month period.

(9) William Index (14 months)

Formula in Equation 3.3.13 introduces the calculation of this descriptor.

$$William(14) = \frac{Highest\ Price_{t-14,t} - Close\ Price_t}{Highest\ Price_{t-14,t} - Lowest\ Price_{t-14,t}} \tag{3.3.13}$$

This is a measure of the proportion of the force from the short-side with respect of the total force. The descriptor gained will indicate the relative force relationship in a different way other than RSI.

(10) Market Value of the Flowing Shares

According to the research result of Fama Frence 3-factor model, the market value of a certain firm (size effect) will have a negative impact on its stock return. In addition, there is the phenomenon that stocks in China are not completely flowing in the market with some part of shares locked there. Therefore, this paper uses the monthly market value of only the flowing shares as the indicator of size effect since only this part engages in the market trading operations directly.

3.3.3 Market Factor

Based on the construction of CAPM, which is a special case of APT model this paper involves the market factor—β. This factor will serve as the loading in the later part, but in itself market return is the original descriptor. And this paper calculates it by first capturing the return on market from Shanghai Composite Index monthly data. Then subtracting the risk-free rate from it will present us the market access return. In the whole 192 months from 2000 to 2015 for our model construction the market access return can be shown as Figure 3.3.4.

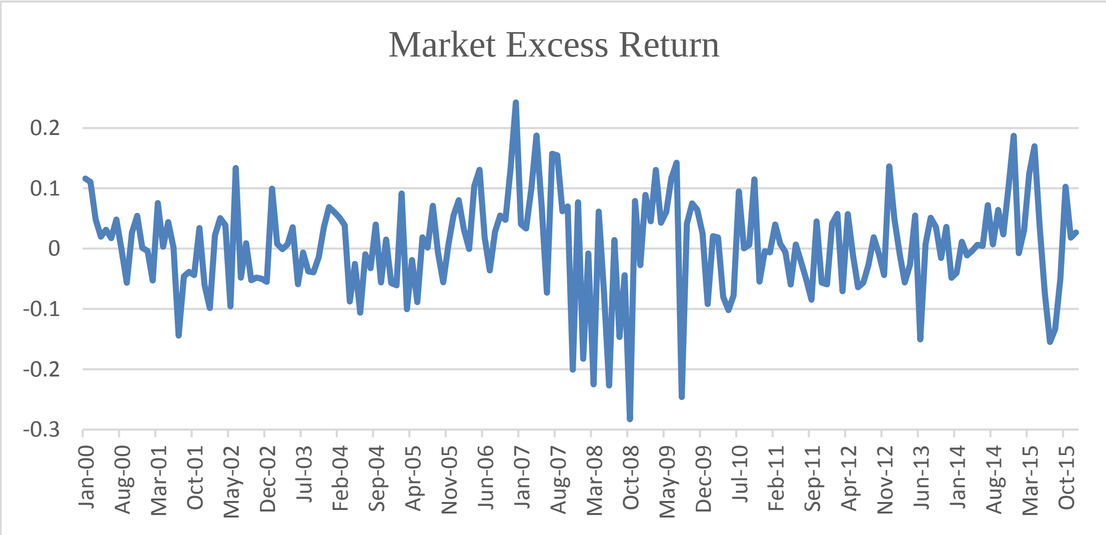


Figure 3.3.4

3.3.4 Macro Factors

In order to take the information from the fundamental economy into consideration, 7 macro indices are chosen to be the descriptors for the loading calculation. They reflect enough movement of the macro situations.

(1) Industrial Growth

This index is a time series similar to GDP and presenting the whole picture of the macro economy growth. The reason of using this is that it is monthly instead of quarterly which GDP is. As for the method of adjusting the frequency of GDP data some interpolation solution such as sample function method, Kriging interpolation and inverse-distance weighing method are used in some researches [17]. But considering that these methods of interpolation will add some noises into the raw information the GDP series can present, this paper gives up the thought of interpolation and uses industrial growth as a proxy of GDP. Every month this indicator will demonstrate how the macro economy is functioning. Stocks related to production field will be strongly affected by the macro growth situation. In the window of this paper it can be shown in Figure 3.3.5.

Figure 3.3.5

(2) CPI

CPI is the best indicator of the price level of a Country. Almost every stock will be influenced by the prices of goods in reality. It has a trend pictured in Figure 3.3.6.

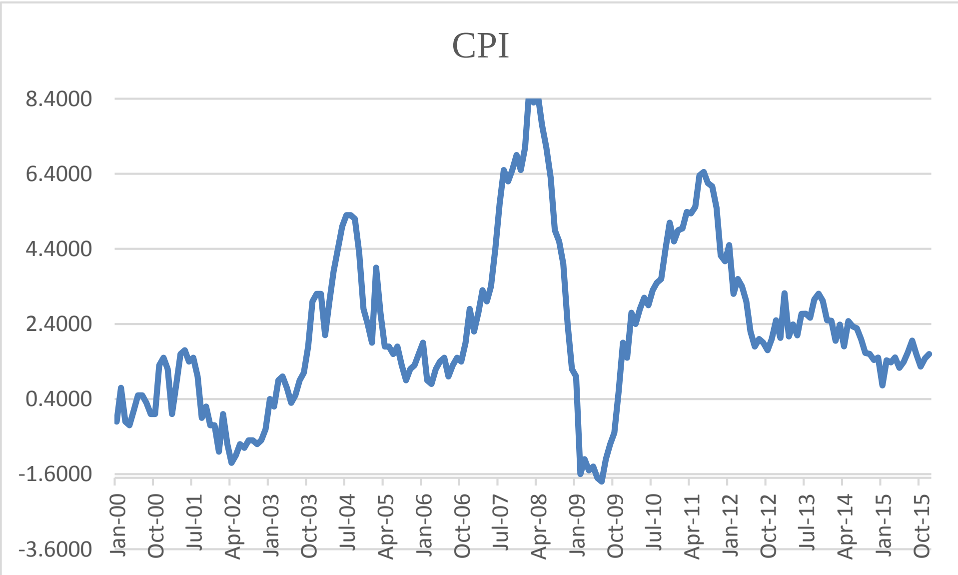


Figure 3.3.6

(3) Trade Surplus

As an index measuring the condition of exports and imports, trade surplus will affect the stocks whose underlying companies have oversea business or have to import raw materials from abroad. Its look is shown in Figure 3.3.7. It seems that

China’s trade account is acquiring more and more surplus and the exporting increase every year plays an important role to this change gradually.

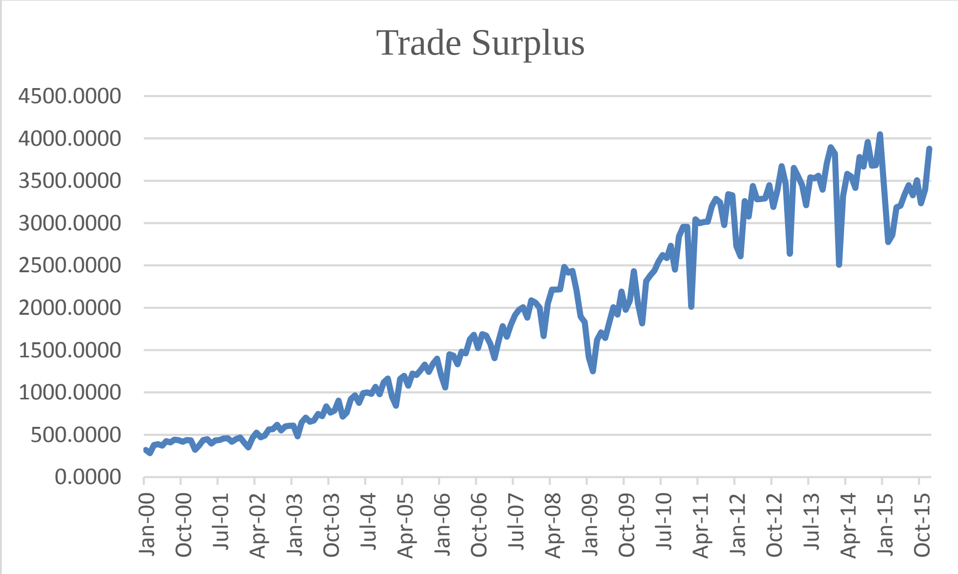


Figure 3.3.7

(4) Fiscal Deficit

From the perspective of the government, fiscal deficit especially its major factor of government spending will have an impact on stocks in construction, manufacturing and environmental sectors and so on. In Figure 3.3.8 is its time series every month.

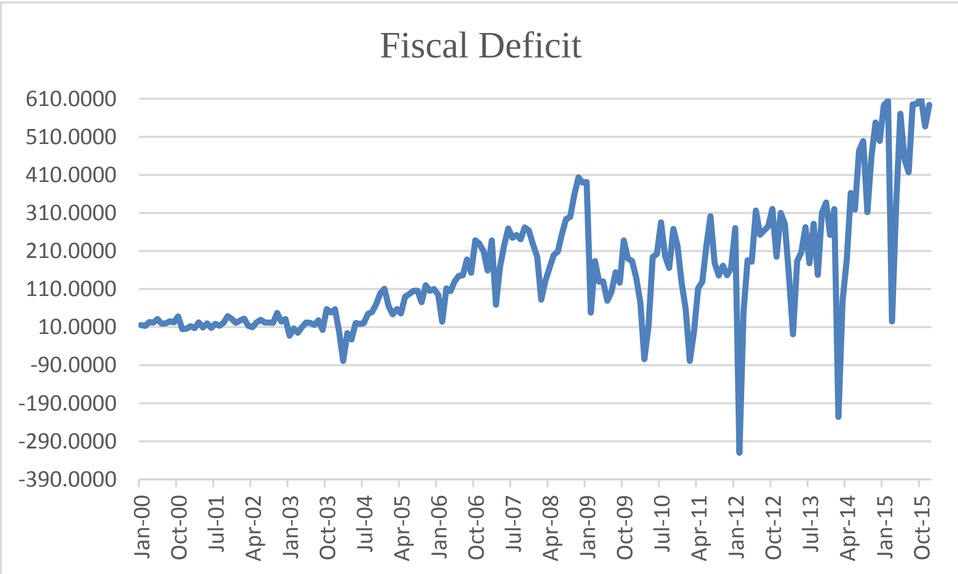


Figure 3.3.8

(5) M2 Money Stock

M2 serves as an important factor of the money supply, which is also a significant indicator of the macro monetary level. It is shown in Figure 3.3.9.

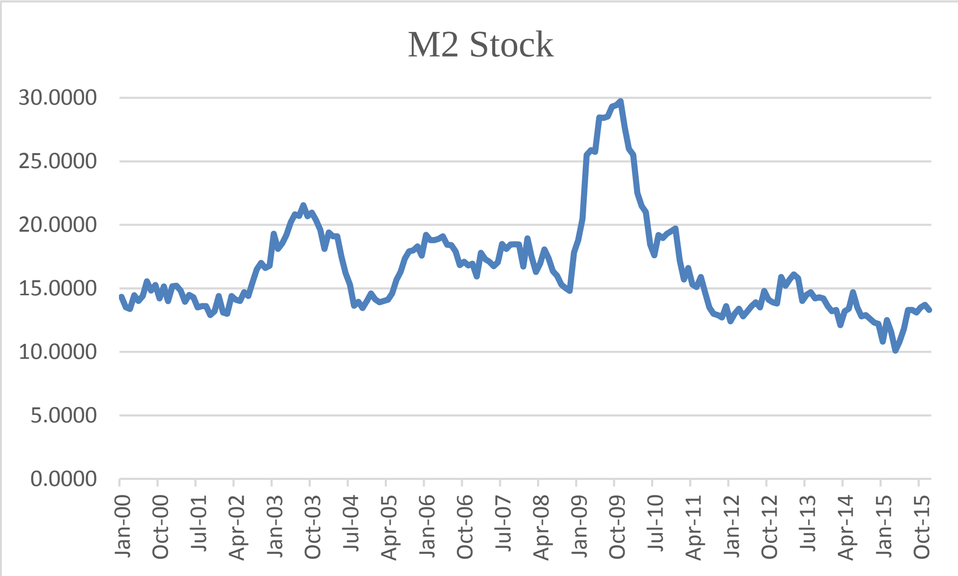


Figure 3.3.9

(6) Consumer Confidence Index (CCI)

In the data library of macro confidence, this index is a general sign of the mood from the consumers towards the development of the economy. Many papers have researched into the relationships between the stock return and the investors' mood and it turns out that there is significant impact form the investors' thought on the stocks. Figure 3.3.10 is a combination of CCI and market index.

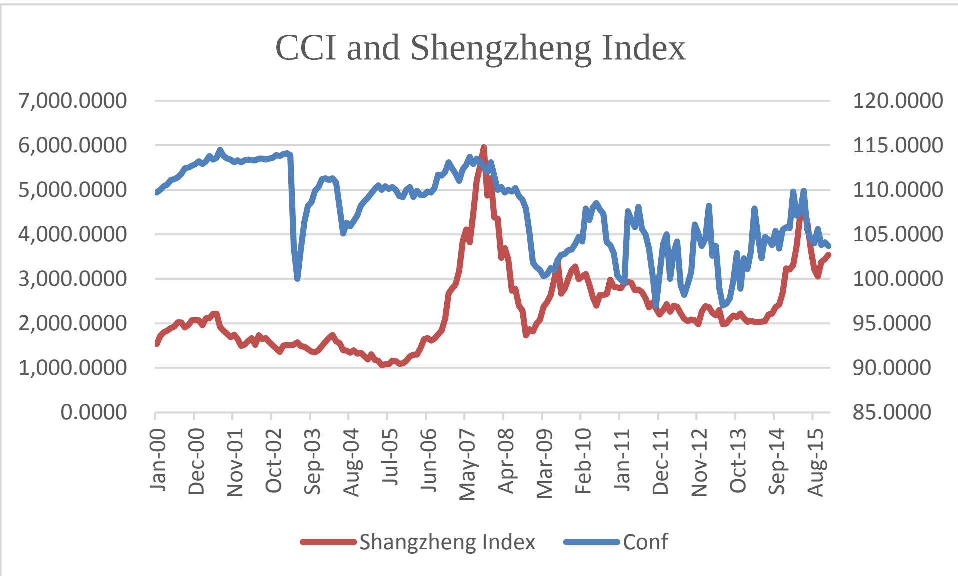


Figure 3.3.10

We can see that CCI move with the Shanghai Composite Index generally.

(7) RMB Effective Exchange Rate

Exchange rate is an import indicator related to imports and exports, which will influence many list companies in the A-share market if they have frequent commercial communication with foreign countries. In this way, the exchange rate (in Figure 3.3.11) will symbol the purchase power of Chinese yuan. (Data source: Federal Reserve Economic Data [18])

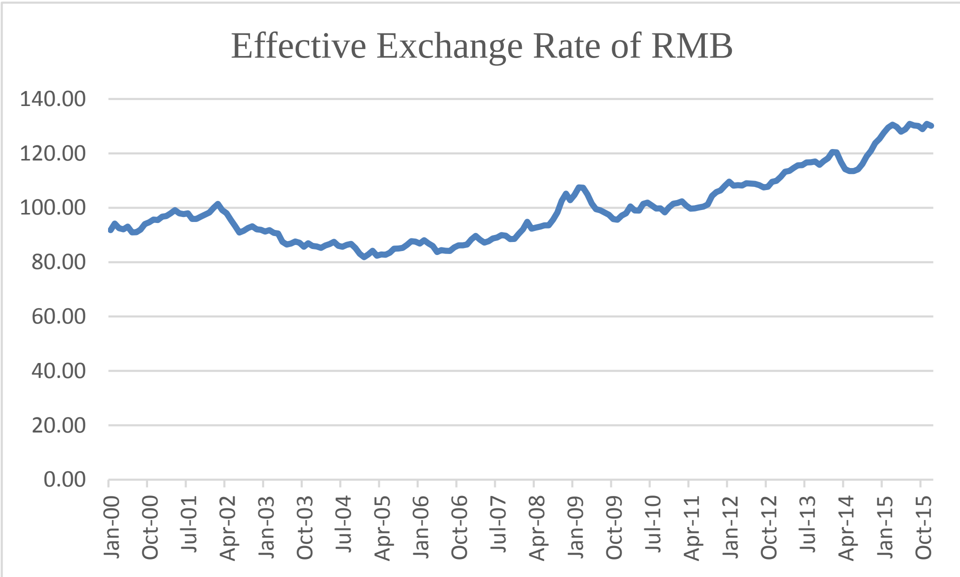


Figure 3.3.11

Now 7 macro factors have been selected as the original descriptors of the risk indices, the next step is filtering the insignificant indices.

3.3.5 Firm Characteristic Factors

In order to consider the impacts from the firm characteristic side, 8 factors are collected here to capture information from the companies themselves. They are: financial leverage, cash flow, dividend yield, EPS, operating expense, profitability, integrated leverage and operating leverage. All of these data were collected from CSMAR with a lot of missing values. Therefore, they are not used as descriptor but directly used as loadings.

3.3.6 Industry Factors

This type of factor is essential in explaining the individual movement of stock prices which are in the same sector. By referring to Shenwan Level 1 industry classification [19] we can construct 26 industry dummy variables. In fact there are 27 industries in the Shenwan Level 1 industry classification but we have to avoid the dummy variable trap so only 26 dummies are used. The formula of dummy variable for industry i is in Equation 3.3.14.

$$\text{industry}_i = \begin{cases} 1, \text{if the stock is in this industry} \\ 0, otherwise \end{cases} \tag{3.3.14}$$

In this way, we have 26 loadings of the stocks which will not change during different times. Here this paper assumes that the sector a stock is in will not vary, which means cases like Xiang'eqing who once changed its major business of restaurant into real estate will be excluded. All of the stocks will be in an invariant industry.

Now, after selecting all of the factors to be used in our model, we begin to test whether some of them will be excluded as factors in MFM.

3.4 Factor Loadings

In this section, time series regression will be utilized to calculate the loadings of each stock selected on the every factor. In total there are 10 technical index factors, 1 market factor, 7 macro factors, 8 firm characteristic factors and 26 industry factors to be summed as 52 factors in the basic factor library for us to choose. And then this paper tests the significance of the factor loadings of the former 3 types of factors by the method of Fama-Macbeth (1973) regression, while leaving the others as they are because the latter 2 types of factors will not be regressed to get the factor loadings.

In order to test whether the factor chosen is significant enough to be included in MFM, this paper establishes an original way of hypothesis testing.

After we get the loading of a certain factor by doing 107 time series regressions, every month we have 107 individual abnormal returns and 107 factor loadings. By doing cross sectional regressions every month under the significance level of 10% we can get the significance results month by month. Then, the significance level at 10% indicates that in every independent cross sectional regression, the probability of getting the wrong test result is 10%.

Therefore, all of the 192 months of cross sectional regressions consist of a new random event: whether the significant factor is invalid. And it is assumed that this event (the significant regression result is wrong) as a random variable follows a binomial distribution of probability 10%. Its CDF (cumulative distribution function) can be drawn as Figure 3.4.1.

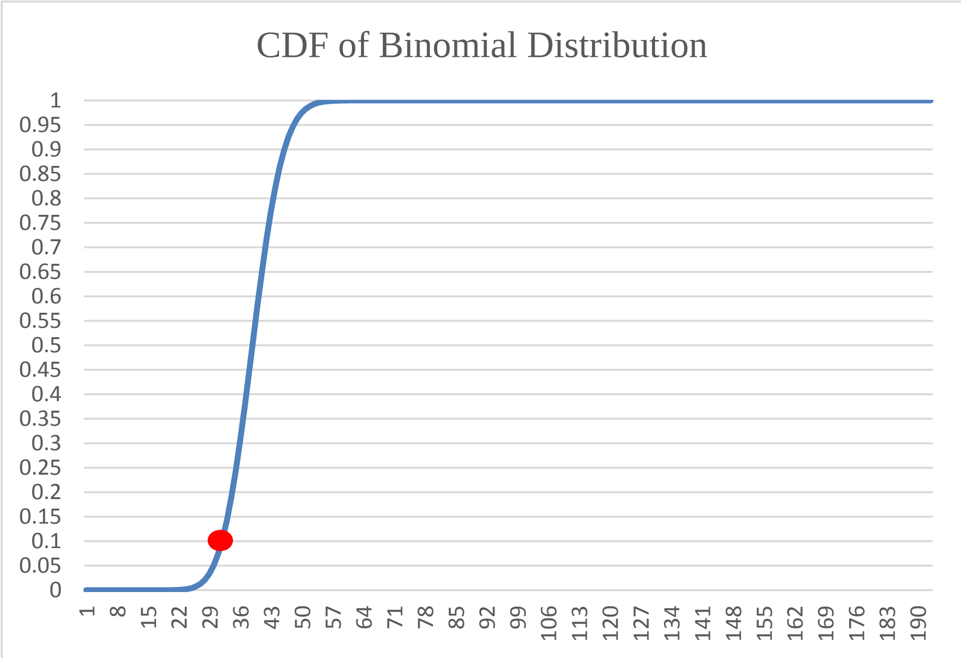


Figure 3.4.1

It can be seen in the figure and actually by reporting the vector value in MATLAB, after 32 months this random reaches the cumulative probability of 10% which is the level of significance. This implies that, it needs at least 32 months of the whole 192 months to claim that the factor after tested can be used in MFM.

3.4.1 Technical Index Factor Loading

(1) Step One—Get Loadings

The first step of the Fama-Macbeth regression is the respective time series regression of one certain factor descriptor on the individual return series. In order to test whether our model is feasible first we use the whole time period as the estimation window, which is the whole 192 months as the universe.

The regression model of a certain factor k is as in Equation 3.4.1.

$$\text{stock abnormal return}_t = \text{loading}_k \cdot descriptor_t + u_t \tag{3.4.1}$$

For example the first time series should regress the abnormal return of Pingan Bank (000001) in 192 months on its volatility in these 192 months, in total 192 observations. And the coefficient will serve as the factor loading of the factor volatility on the first stock. The specific regression process of one stock, such as the ith stock, in vector form is as Equation 3.4.2.

$$\begin{bmatrix} r_{stock_i,t_1} \\ r_{stock_i,t_2} \\ r_{stock_i,t_3} \\ \dots \\ r_{stock_i,t_{192}} \end{bmatrix} = \text{loading}_{stock_i,k} \cdot \begin{bmatrix} descriptor_{stock_i,t_1} \\ descriptor_{stock_i,t_2} \\ descriptor_{stock_i,t_3} \\ \dots \\ descriptor_{stock_i,t_{192}} \end{bmatrix} + \begin{bmatrix} u_{stock_i,t_1} \\ u_{stock_i,t_2} \\ u_{stock_i,t_3} \\ \dots \\ u_{stock_i,t_{192}} \end{bmatrix}$$

(3.4.2)

Since we have assumed that all the time series data are stationary themselves, we conduct the time series regression directly without doing the co-integration test. And in this step we do not care that much the significance of each time-series regression as is assumed in the assumptions in 3.1.3. The significance level of regression is 10%.

(2) Step Two--Test

For each factor, after we get the time series regression coefficients as the factor loadings on every stock, we can conduct the second step for Fama-Macbeth method. This is a cross sectional regression procedure. The vector form of regression for one factor, such as the *k*th factor is in Equation 3.4.3 which is the cross sectional regression at time t. The significance level of regression is also 10%.

$$\begin{bmatrix} r_{stock_1,t} \\ r_{stock_2,t} \\ r_{stock_3,t} \\ \dots \\ r_{stock_{107},t} \end{bmatrix} = \text{factor return}_{k,t} \cdot \begin{bmatrix} loading_{stock_1,k} \\ loading_{stock_2,k} \\ loading_{stock_3,k} \\ \dots \\ loading_{stock_{107},k} \end{bmatrix} + \begin{bmatrix} e_{stock_1,t} \\ e_{stock_2,t} \\ e_{stock_3,t} \\ \dots \\ e_{stock_{107},t} \end{bmatrix}$$

(3.4.3)

Because every month we have a cross sectional regression, every month we have a coefficient and its *p-value*. In this step we will test whether this factor makes sense in explaining the abnormal return of the stocks. By collecting the results of all the regressions in the whole 192 months we can gain the following results (Table 3.4.1) of the number of significant impacts of factor loadings on the return.

Technical Indices	Buy Signal	Sell Signal	William Index	**MACD**	RSI
Months of Significance	50	49	82	**29**	72
Technical Indices	**Trading volume**	Flowing Market Value	KDJ	Volatility	Momentum
Months of Significance	**26**	40	71	60	53

Table 3.4.1

It can be concluded in the table that the majority of the index descriptors are significant in at least 32 months over the total 192 months except MACD and trading volume. Therefore, we get rid the factors of MACD and trading volume out of our model leaving the rest 8 technical factors.

It is suspected that the explanation power of MACD and trading volume has been diminished by the frequent use of some integrated signal based on them in the practitioner's world. The research report of the Orient Securities named *Efficiency Hypothesis Testing of MACD Index in China's A-share Market* [20] has demonstrated that in many times the power of MACD alone is declining as more and more people can observe the operation points. This paper concludes that using either MACD or trading volume alone is not enough to capture the operation signals for investors. But the simplest integrated signal as constructed by buy-signal and sell-signal can contribute a lot of explanation power by extracting that of the MACD and trading volume to the movement of the stock returns. This presents the reason why MACD and trading volume did not pass the significance test of this paper as an indicator to be used in the MFM.

Then, the 8 technical loadings gained in the Step One formed as a 107×8 matrix serves as the technical index factor loadings (in picture 3.4.1).

$$\text{107 Individual Stocks}\left\{\overbrace{\begin{bmatrix} Buy\,Signal_{Stock_1} & \cdots & Volatility_{Stock_1} \\ \vdots & \ddots & \vdots \\ Buy\,Signal_{Stock_{107}} & \cdots & Volatility_{Stock_{107}} \end{bmatrix}}^{\text{8 Technical Factors}}\right.$$

Picture 3.4.1

3.4.2 Market Factor Loading

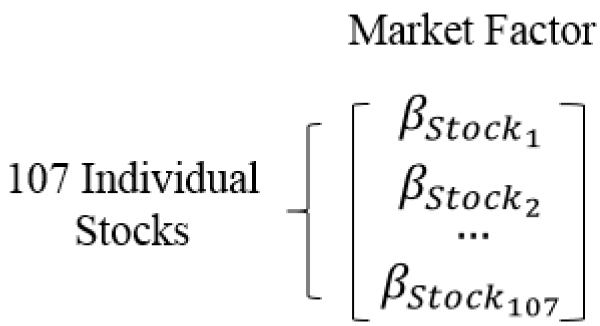
The stocks' β towards the market return premium will serve as factor loading in this part. By imitating the procedure of CAPM, using the similar method as it is conducted for the technical factors we first regress the individual abnormal returns of all the stocks selected on market premium. Then regressing the abnormal returns in each month on the coefficients of the former regressions of each stock. We can get the testing result for market factor as in Table 3.4.2 in the following part.

Market Factor	Market β
Months of Significance	52

Table 3.4.2

In total there are 52 months of the whole 192 months when the market loading is significant, it also passes the testing method constructed in 3.4.1. As a result, the market factor will be included in the MFM here.

Then, the market loadings gained in the Step One formed as a 107×1 vector serves as the market factor loadings (in picture 3.4.2).



Picture 3.4.2

3.4.3 Macro Factor Loading

In our model, macro indices play the similar role of the raw descriptors. And in the same way we did in 3.4.1 and 3.4.3 we can test their significance using the method of two-step regression with the results in Table 3.4.3.

Macro Factor	Industrial Growth	CPI	Trade Surplus	**Fiscal Deficit**
Months of Significance	42	42	62	**32**
Macro Factor	M2 Stock	CCI	Exchange Rate	
Months of Significance	40	53	53	

Table 3.4.2

By testing the macro factors, only fiscal deficit has to be deleted from the MFM. Then, the 6 macro loadings gained in the Step One formed as a 107×6 matrix serves as the macro factor loadings (in Picture 3.4.3).

$$\text{107 Individual Stocks}\;\left\{\;\overbrace{\begin{bmatrix} Industrial\ Growth_{Stock_1} & \cdots & Exchange\ Rate_{Stock_1} \\ \vdots & \ddots & \vdots \\ Industrial\ Growth_{Stock_{107}} & \cdots & Exchange\ Rate_{Stock_{107}} \end{bmatrix}}^{\text{6 Macro Factor}}\right.$$

Picture 3.4.3

So far, in total 15 factors containing technical factors, market factor as well as macro factors have been selected in the MFM. Next step is to merge them with industry factors and firm characteristic factors.

3.4.4 Industry Factor Loading

In this paper, the above three kinds of factors were first preprocessed as raw descriptors and by doing time series regressions as the first step we get the 15 factor loadings for them. But industry factor is a different story.

As has been described in 3.3.6, the industry factors are dummy variables indicating whether a certain stock belongs to a certain industry or not. So, without the time series regressions we can directly get a 107×26 matrix serving as the industry factor loadings (in picture 3.4.4).

$$\text{107 Individual Stocks}\;\left\{\;\overbrace{\begin{bmatrix} Industry_{1_{Stock_1}} & \cdots & Industry_{26_{Stock_1}} \\ \vdots & \ddots & \vdots \\ Industry_{1_{Stock_{107}}} & \cdots & Industry_{26_{Stock_{107}}} \end{bmatrix}}^{\text{26 Industry Factors}}\right.$$

Picture 3.4.4

3.4.5 Firm characteristic Factor Loading

And the last part is firm characteristic loadings. Since the problem of missing values and the unmatched frequency of firm characteristic data is hard to solved properly, this paper uses the value of those firm characteristic data directly as factor loadings. But as the window is shifting, this kind of loadings changes across time based on the dynamic model estimation. The detail of this part will be included in the dynamic estimation of the factor return part later. In brief, every time we roll the model up we can also get a 107×8 matrix of financial leverage, cash flow, dividend yield, operating expense, profitability, integrated leverage, operating leverage and EPS serving as the corporate factor loadings (in picture 3.4.5).

$$\text{107 Individual Stocks}\left\{\overbrace{\begin{bmatrix} Cash\ Flow_{Stock_1} & \cdots & EPS_{Stock_1} \\ \vdots & \ddots & \vdots \\ Cash\ Flow_{Stock_{107}} & \cdots & EPS_{Stock_{107}} \end{bmatrix}}^{\text{8 Corporate Finance Factors}}\right.$$

Picture 3.4.5

3.4.6 Merge Factor Loadings

It will be illustrated in Section 4.2 to introduce the WLS method. Since the weighing matrix will be related to the flowing market value actually this paper gets rid of it as the technical factor. In this way, all of the selected factors can be merged as a 107×48 matrix containing all of the factors in our MFM. To be clear we also provide a general picture of them.

$$\text{107 Individual Stocks}\left\{\overbrace{\begin{bmatrix} Factor_{1_{Stock_1}} & \cdots & Factor_{48_{Stock_1}} \\ \vdots & \ddots & \vdots \\ Factor_{1_{Stock_{107}}} & \cdots & Factor_{48_{Stock_{107}}} \end{bmatrix}}^{\text{48 Factor Loadings}}\right.$$

Picture 3.4.6

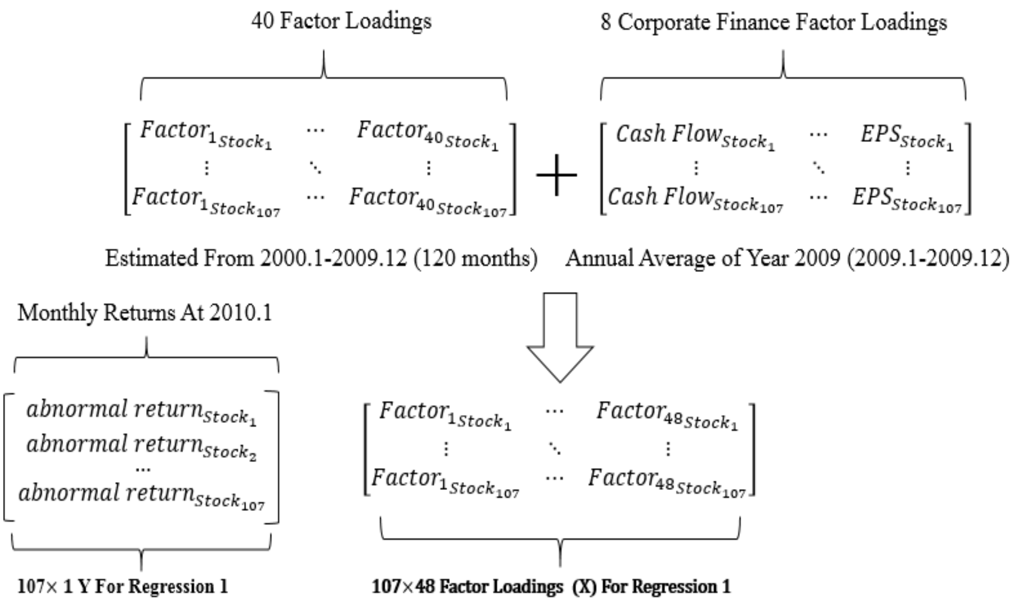
So far the MFM model has been constructed by 48 factors with their loadings. Next we will estimate the factor returns of each rolling period and this will be conducted using the dynamic model with a window.

4. Dynamic Multiple-Factor Model Estimation

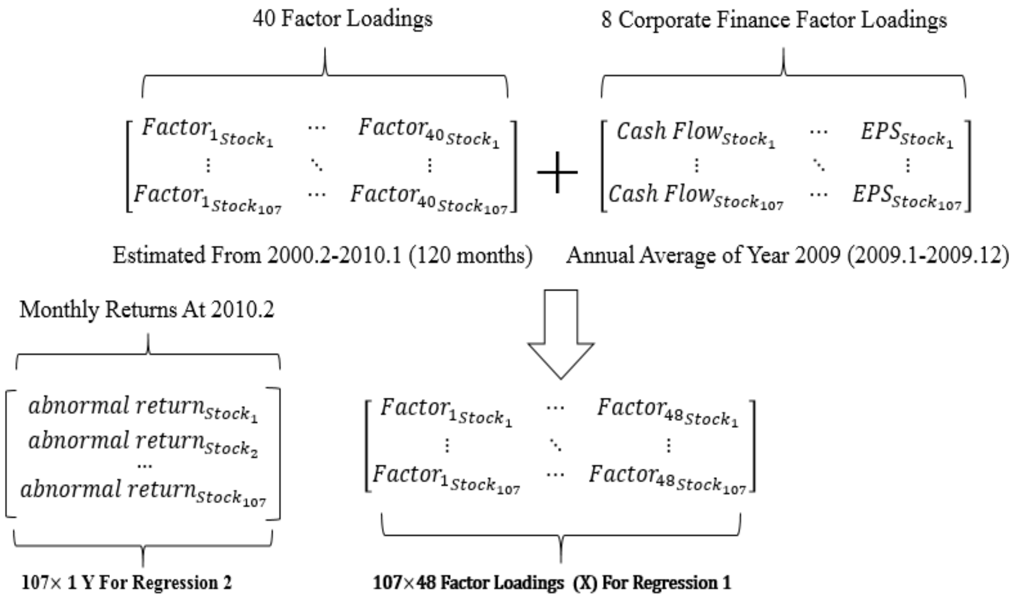
4.1 Dynamic Window for Model Construction

Since in the following part of risk analysis we will estimate the correlation matrix of the specific returns of all the 107 stocks, the individual specific return time series namely the residual time series has to be longer than 107 periods in order to avoid the problem of over identification. Therefore, 120 months (10 years) will serve as the estimation window for the factor loadings of each rolling period. In this way, 73 months are left as the regression periods. The regression process has been illustrated in Picture 4.1.

Regression 1—at 2010.1

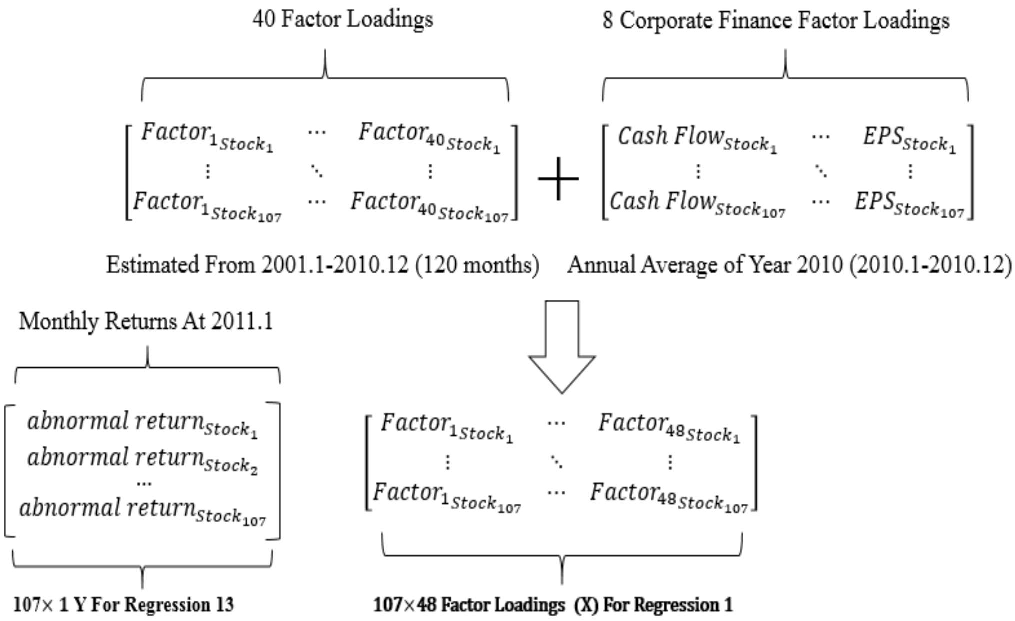


Regression 2—at 2010.2



And as time goes by…

Regression 13—at 2011.1



Picture 4.1

In the picture Y stands for dependent variable while X stands for independent variables. In addition there are two basic rules of rolling the dynamic model:

(1) Firm characteristic loadings will be the annual average of the prior year without changing for 12 months in the same year.

(2) Factor loadings are rolling with a estimation window of 120 months, all the descriptors used to estimate the factor loadings have to be prior to the regression

period of the dependent variable (Y)—individual abnormal returns.

By constructing the dynamic models as above procedure, we have from 2010.1 to 2016.1 in total 73 cross sectional regressions of the individual returns on factor loadings.

In each period of cross sectional regression of the total 73 ones, the 48 coefficients are regarded as the common factor returns while the 107 residual terms are regarded as the specific returns of the stocks in this window period.

4.2 Estimation Method

According to the assumptions in 3.1.3, during the procedure of estimating the factor loadings we have assumed that all of the descriptors are stationary. But in the next step of the cross sectional regression this paper first tests the presence of heteroskedasticity using the dynamic model to estimate the loadings.

Here by using the Lagrange Multiplier Test for heteroscedasticity, we collect the *p*-value of each cross sectional regression. If *p* -value is smaller than 10% we reject the **H0 (null hypothesis)** that the regression model has heteroscedasticity. Therefore, in the 73 periods of regressions there are 62 months with heteroscedasticity. This implies that simply using OLS will produce large bias in MFM.

In order to solve this problem, referring to BARRA's handbook this paper uses the $\frac{1}{\sqrt{Flowing\ Market\ Value}}$ as the weighting matrix to conduct WLS (Weighted Least Square) estimation to get the common factor returns.

4.3 Factor Return Estimation

As has been introduced in Section 3.1.1 in every month every factor will have a 120-period factor return. For example, the 120-period factor returns from 2007.1 to 2015.12 are included in Appendix I, which is the regression result on 2016.1. Every month this is a 120×48 matrix containing the 48 common factor returns from 2010.1 to 2016.1.

5. Risk Analysis

5.1 Robustness Testing of MFM

In each period of estimation, we can get the following results:

(1) 107×48 factor loading matrix X in Picture 5.1.1;

$$X = \overbrace{\begin{bmatrix} Factor_{1_{Stock_1}} & \cdots & Factor_{48_{Stock_1}} \\ \vdots & \ddots & \vdots \\ Factor_{1_{Stock_{107}}} & \cdots & Factor_{48_{Stock_{107}}} \end{bmatrix}}^{\text{48 Factor Loadings}} \left.\vphantom{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}\right\} \text{107 Individual Stocks}$$

Picture 5.1.1

(2) 120×48 factor return matrix B in Picture 5.1.2;

$$B = \overbrace{\begin{bmatrix} Factor\ return_{1,t_1} & \cdots & Factor\ return_{48,t_1} \\ \vdots & \ddots & \vdots \\ Factor\ return_{1,t_{120}} & \cdots & Factor\ return_{48,t_{120}} \end{bmatrix}}^{\text{48 Factors}} \left.\vphantom{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}\right\} \text{120 Months}$$

Picture 5.1.2

(3) 120×107 specific return matrix E in Picture 5.1.3.

$$E = \overbrace{\begin{bmatrix} Residual_{Stock1,t_1} & \cdots & Residual_{Stock107,t_1} \\ \vdots & \ddots & \vdots \\ Residual_{Stock1,t_{120}} & \cdots & Residual_{Stock107,t_{120}} \end{bmatrix}}^{\text{107 Stocks}} \left.\vphantom{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}\right\} \text{120 Months}$$

Picture 5.1.3

(4) 48×48 Factor Return Variance-Covariance Matrix F in Picture 5.1.4.

$$\mathrm{F} = \overbrace{\begin{bmatrix} Var_{1,1} & \cdots & Cov_{1,48} \\ \vdots & \ddots & \vdots \\ Cov_{48,1} & \cdots & Var_{48,48} \end{bmatrix}}^{\text{48 Factors}} \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} \text{48 Factors}$$

Picture 5.1.4

(5) 107×107 Specific Return Variance-Covariance Matrix Δ in Picture 5.1.5.

$$\Delta = \overbrace{\begin{bmatrix} Var_{1,1} & \cdots & Cov_{1,107} \\ \vdots & \ddots & \vdots \\ Cov_{107,1} & \cdots & Var_{107,107} \end{bmatrix}}^{\text{107 Residuals}} \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} \text{107 Residuals}$$

Picture 5.1.5

(6) 107×107 Specific Return Correlation-Coefficient Matrix Γ in Picture 5.1.6.

$$\Gamma = \overbrace{\begin{bmatrix} Corr_{1,1} & \cdots & Corr_{1,107} \\ \vdots & \ddots & \vdots \\ Corr_{107,1} & \cdots & Corr_{107,107} \end{bmatrix}}^{\text{107 Residuals}} \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} \text{107 Residuals}$$

Picture 5.1.6

The robustness testing method in this paper is checking whether Γ (specific return correlation-coefficient matrix) is almost an identity matrix with 1s on the diagonal and other correlation-coefficient $Corr_{i,j}$ of stock i and stock j almost 0. Using the whole 192 months as sample, the 107×107 matrix Γ as in Appendix II.

It can be observed that only 9.8087169% of the values in Γ have an absolute value greater than 0.2 which is very close to an identity matrix.

Furthermore, this paper also conducts the Spearman Correlation-Coefficient test via SPSS for matrix Γ. The **H0 (Null Hypothesis)** is: the two specific returns' correlation-coefficient is significantly different from 0. And the *p*-values of this test is in Appendix III. This shows that only 14.272% of all the values in Γ is significantly different from 0, which to a larger extend proves that the specific return correlation-coefficient matrix can be regarded as an identical matrix.

In other word, these 48 factors chosen indeed divide the individual returns cleanly and thus demonstrating that MFM is robust.

5.2 Total Portfolio Variance

5.2.1 Estimate Total Variance Matrix

The above matrices are the basic results for us to conduct the risk analysis. The general form of calculating the integrated 107×107 variance-covariance matrix is in Equation 5.2.1 as follows.

$$\mathbf{V} = \boldsymbol{X}^T \cdot \boldsymbol{F} \cdot \boldsymbol{X} + \Delta \tag{5.2.1}$$

5.2.2 Calculate Portfolio Total Risk

The next goal is to calculate the total variance of a certain portfolio given the weights within the portfolio. Every portfolio P can be described by a weighing vector. Here we conduct the similar way of calculating the Hushen 300 Index by using the total market value as the weights in these 107 stocks. So the formula of investment weighing vector H_p of the portfolio is in Picture 5.2.1. Every element in the vector H_p is the weight of positions of the stocks held in the portfolio.

$$H_P = \overset{\text{Weighing Vector}}{\left.\begin{bmatrix} \dfrac{1}{Market\ Value_{Stock_1}} \\ \dfrac{1}{Market\ Value_{Stock_2}} \\ \cdots \\ \dfrac{1}{Market\ Value_{Stock_{107}}} \end{bmatrix}\right\}} \text{107 Individual Stocks}$$

Picture 5.2.1

As a result, Portfolio P has its factor loadings given by Equation 5.2.2.

$$\boldsymbol{X_P} = \mathbf{X} \cdot \boldsymbol{H_P} \tag{5.2.2}$$

The following step is to add the variance matrix of all the stocks then we can get the total variance of the portfolio in each period of the 73 months.

The total variance is composed of factor return-explained variance as well as the specific return variance $\sigma_P{}^2$ within the portfolio. The formula is shown in Equation 5.2.3 in each period.

$$\sigma_P{}^2 = \boldsymbol{X_P}^T \cdot \boldsymbol{F} \cdot \boldsymbol{X_P} + \boldsymbol{H_P}^T \cdot \Delta \cdot \boldsymbol{H_P} = \boldsymbol{H_P}^T \cdot \boldsymbol{V} \cdot \boldsymbol{H_P} \tag{5.2.3}$$

In the world of realistic analysis, people tend to use annual volatility (standard deviation) as the direct indicator of risk. From 2010.1 to 2016.1 the results of the total modelled portfolio risk can form a time series as is shown in the blue curve of Figure 5.2.1. In order for easier comparison with the realized risk of Hushen 300, red curve in Figure 5.2.1 presents the true 120-month volatility of Hushen 300 Index.

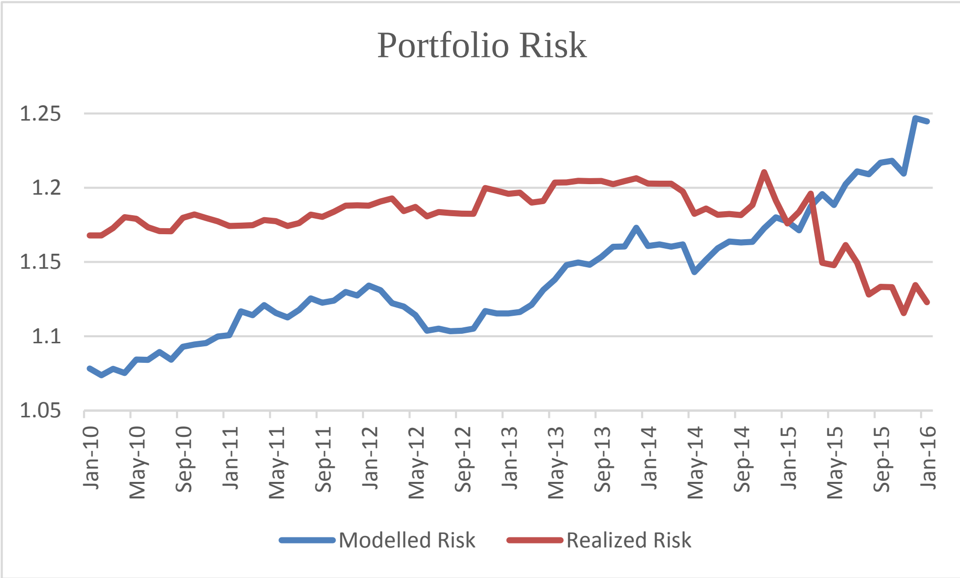


Figure 5.2.1

By observing the movement of the portfolio risk the investor holding it can have a direct felling of how his investment risk is changing, increasing or declining. In this way, the manager of the portfolio can feel the time when their investment position is at great risk and when they are just bearing some accepted risk.

In the figure above, it seems that before the end of 2014 the total risk explained by the MFM indeed goes in line with the realized volatility of the Hushen 300 Index. However, after approximately the spring of 2015 they diverged from each other.

This can be explained by the presence of the market blossom in the beginning of 2015 followed by the stock crash in the mid of Year 2015. Many unexpected noises from the irrational trading operations of both the individual investors and institutional investors brought about unexplained fluctuations towards the Hushen 300 Index. This makes the explanation power of the common factors diminished a lot as well as the specific noisy returns' fluctuation diverge from the index itself to a large degree.

Another explanation is that in this paper the testing portfolio is simply composed of these 107 stocks with their total market values as the investment weights. Compared with the real Hushen 300 Index with 300 changing constituent stocks, some divergence of their risk level can be accepted.

But looking at the previous parts of times, it is true that MFM can provide us with a convincing calculation of risk of the portfolio of an investor.

In this way, given a certain weight of positions held in the portfolio, its total risk can be calculated with MFM relatively accurately. This total risk will be the combination of the explained risk by the common factors and the risk of the specific returns of the portfolio stocks.

5.3 Portfolio Risk Attribution

Since we have acquired the total risk level of a portfolio, we also want to know which stocks contribute greater portfolio risk.

Although generally attributing the risks for the whole portfolio is hard to address, we can still consider the effect of adjusting the position of each stock on the risk level in a certain portfolio. This kind of sensitivity analysis allows us to observe which assets acquire the most significant power affecting the total risk. It can be measured by the partial derivative of total risk with respect to the asset weights.

In this way of thinking we can do the following deductions.

The marginal contribution on total risk (MCTR) of all the assets can form a vector in each period of regression with the formula in Equation 5.3.1.

$$\mathbf{MCTR} = \frac{\partial \sigma_P}{\partial \boldsymbol{H}_{\boldsymbol{P}}^{\boldsymbol{T}}} = \frac{\boldsymbol{V} \cdot \boldsymbol{H}_{\boldsymbol{P}}}{\sigma_P} \tag{5.3.1}$$

Among them, $\mathbf{MCTR}$ is a 107-dimensional vector which is the partial derivative of σ_P with respect to the weighing matrix of a portfolio. Any of its element (for example the nth element) can be understood as: if asset n's weight in the portfolio has changed by 1%, by how much will the total risk change accordingly.

In this way, every period of cross sectional regression will produce one total risk as well as one 107-element vector demonstrating the marginal contribution of the all the assets in the portfolio. In this paper, the result of the 73 months of regressions is listed in Appendix IV.

This result is the portfolio constructing with the weights as the flowing market value of the 107 stocks in each month. It can be seen that in some periods of estimation, Stock 600741 (Huayu Qiche) has the most significant power of impacting the total risk. Therefore, investor holding the portfolio with these assets can observe the position exposure to the total risk of each stock every month. Thus, by adjusting the positions in the most sensitive stocks can secure the whole portfolio value in a stable condition of movement. For example, by observing that Stock 600741 (Huayu

Qiche) attributes to much marginal risk the investor can decrease the values of this stock in his investment.

This is the main contribution of this paper—to provide the investor with a way of calculating the risk level of their investment strategy considering many aspects of potential risks and deciding the solution towards some of the easily affecting stocks.

6. Conclusions

This paper intends to explain the abnormal return of the individual stocks by finding and selecting 48 common factors and their specific return. There are mainly three contributions through the research.

First of all, in the process of constructing the Multiple-Factor Model in total 5 kinds of factors have been taken into consideration, containing technical factors, market factor, macro factors, industry factors and firm characteristic factors. After selecting 48 of them we can completely model the individual returns with an identity matrix of specific returns' correlations. In the specific return correlation-coefficient matrix the majority of the elements' absolute values are close to zero. This robustness check result demonstrates the model construction part is satisfactory.

In addition, by rolling the regressions into a dynamic model with a window of 120 months, this provides a way of estimating the common factor returns in each given window and thus estimating the total variance and the total risk of a certain portfolio. With a weighting vector of positions in the stocks, any portfolio will have a total risk level calculated by the MFM. This serves as an indicator telling the portfolio manager when their position is at large risk.

Last but not least, the part of risk attribution can help the investor who manage the portfolio monitor the specific marginal risk contribution from each stock. As a result, the asset which carries great risks can be found after each period of estimation and thus reaching the goal of portfolio management based on risk control.

The development of BARRA's MFM is still immature in China. It can also be put into further application in other fields such as futures, options and so on.

From what has been discussed above, the concept of risk control can be conducted in many ways. MFM's contribution in risk analysis offers its advantage of being applied by many practitioners.

References

[1] BARRA, 1998, The United States Equity Version3 (E3), *Risk Model Handbook*,

[2] Hicks John Richard, 1935. *Value and Capital*.

[3] Hicks, 1936. Outlook of Economics.

[4] Jacob Marschak, 1959, *Information Economics Comments*.

[5] Harry M. Markowitz, 1952, Portfolio Selection[J], *Journal of Finance*

[6] William Sharpe, 1974, Risk Aversion in the Stock Market, *Journal of Finance*

[7] Rosenberg Barr, 1974, Extra Market Components of Covariance in Security Returns, *Journal of Financial and Quantitative Analysis.*

[8] Stephen Rose, 1976, Arbitrage Theory of Capital Asset Pricing Model, *Journal of Economic Theory.*

[9] Fama, E., and K. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics..*

[10] Grinnold and Ronald N.Kahnm 2000, Active Portfolio Management: Quantitative Theory and Applications.

[11] Thierry Post and Pim van Vliet, 2004, Do multiple factors help or hurt? *SSRN Electronical Journal.*

[12] Lubos Pastor, Lucian A. Taylor and Robert F. Stambaugh, 2014, Scale and Skill in Active Management, *National Bureau of Economic Research Working Paper Series*.

[13]Charoula Daskalaski, Alexandros Kostakis and George Skiadopoulos, 2014, Are there common factors in individual commodity futures returns? J*ournal of Banking and Finance.*

[14] Joseph E.Granville, 1960, The Strategy of Maximum Benefit in the Daily Stock Market.

[15] Data Source: Sina Finance: 西部证券, China: 2016.4.16

http://finance.sina.com.cn/realstock/company/sz002673/nc.shtml

[16] Data Source: Sina Finance: 特力 A, China: 2016.4.16

http://finance.sina.com.cn/realstock/company/sz000025/nc.shtml

[17] Zhang Lu, Yang Yongguo and Yao Songjun, 2010, 基于 ArcGIS 的 GDP 空间插值方法比较分析, *《科技致富向导》*.

[18] Data Source: Federal Reserve Economic Data: Effective Exchange Rate of China, USA: 2014.4.17

https://research.stlouisfed.org/fred2/series/RBCNBIS

[19] 申万一级行业分类标准

http://wenku.baidu.com/link?url=b66OiiAsqElyXHY9L6M6_i_DeHfbBs33eoiH8d2itRmkoso7LHfN_pGhqW8DshmsTKV-i1RNl7OpnELNc7i53fKWdY6Nd9RHWSUIRpzmcDe

[20] Dong Xiangyang, 2012, Efficiency Hypothesis Testing of MACD Index in China's A-share Market, *Orient Securities*

Appendix

Appendix I—Factor Returns from 2007.1 to 2015.12 Every Month

Months	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	…	Factor45	Factor46	Factor47	Factor48
M1	-0.02065	-0.00896	0.014772	-0.01476	0.009072	0.017932	0.005453	…	-0.18672	-0.15283	-0.07426	-0.08592
M2	-0.03854	-0.01153	-0.0306	-0.01633	0.022256	-0.00549	0.033851	…	-0.01515	-0.05382	0.123487	-0.15448
M3	0.022591	0.006077	0.004958	0.010078	-0.01002	0.009078	-0.01039	…	-0.05077	-0.04149	-0.13108	-0.03595
M4	-0.00631	0.021292	-0.02365	-0.01376	0.013215	0.001542	0.022287	…	0.069873	0.040857	-0.00206	0.006305
M5	0.003923	-0.00855	-0.01515	-0.0379	0.032023	-0.01511	0.007172	…	-0.10324	-0.05731	-0.02521	-0.05722
M6	0.045758	-0.01158	-0.00887	0.016513	0.006903	-0.03961	0.002889	…	0.074882	0.110554	0.071263	0.100279
M7	-0.00441	-0.03229	0.025405	-0.00791	0.005902	0.012099	0.020386	…	-0.05162	-0.04686	0.1184	-0.01181
M8	-0.01021	0.002576	0.022258	-1.4E-05	0.017994	0.021703	0.005457	…	0.009077	-0.0163	0.097899	-0.09562
M9	0.014091	-0.01319	-0.01534	0.011175	-0.00869	-0.04811	-0.02704	…	-0.13149	0.005878	-0.19445	0.040829
M10	-0.00581	-0.01886	0.037612	-0.0019	0.003132	0.025464	-0.00417	…	-0.02885	0.027925	0.064972	0.019932
M11	-0.00675	-0.03606	0.007821	-0.00361	-0.02169	-0.03567	0.017043	…	-0.15822	0.074428	0.3054	0.096223
M12	0.005803	0.063662	-0.03595	0.020154	-0.04415	0.033188	-0.00716	…	0.099991	0.106224	0.030001	0.039763
M13	0.002722	-0.02079	0.042143	-0.01461	0.024373	0.010692	0.022052	…	0.006135	0.006237	0.140819	-0.04035
M14	-0.02115	0.005682	0.102038	-0.05009	-0.03554	0.091925	0.062936	…	-0.22382	0.016081	0.164517	0.135282
M15	-0.03252	0.083592	-0.08881	0.087719	-0.01262	-0.01987	0.038949	…	0.287203	0.332441	0.047665	0.165117
M16	-0.01799	0.045863	-0.01657	0.014056	0.016096	0.031973	0.066268	…	0.133251	0.072892	-0.19092	0.105437
…	…	…	…	…	…	…	…	…	…	…	…	…
M117	-0.02491	-0.01601	0.089985	0.008417	0.001707	0.033259	-0.02382	…	0.057019	0.05723	0.017058	0.042729
M118	0.018667	-0.02899	-0.03081	-0.01256	0.008842	-0.06099	-0.00149	…	0.481303	-0.00722	0.259762	0.09225
M119	-0.02112	0.014646	0.010889	0.031182	0.005937	0.029718	0.007647	…	-0.38426	-0.30688	-0.36949	-0.24843
M120	0.008719	-0.00951	-0.0355	-0.00966	-0.00085	-0.04229	-0.00354	…	0.094134	0.093663	0.181825	0.002387

Appendix II—Correlation-Coefficient Matrix of Specific Returns Every Month

Corr	Stock1	Stock2	Stock3	Stock4	Stock5	Stock6	Stock7	…	Stock104	Stock105	Stock106	Stock107
Stock1	1							…				
Stock2	-0.08562	1						…				
Stock3	0.005568	0.02779	1					…				
Stock4	0.091406	0.127001	-0.06117	1				…				
Stock5	-0.00392	0.027954	0.13963	-0.11767	1			…				
Stock6	-0.34431	-0.07817	-0.03455	0.042297	-0.02948	1		…				
Stock7	0.020608	-0.12704	-0.15618	0.107638	-0.13113	-0.08071	1	…				
Stock8	-0.00598	0.004768	-0.04527	-0.23158	0.15535	-0.17989	-0.01824	…				
Stock9	0.001089	0.22546	0.083537	0.016363	0.288839	-0.10864	0.010412	…				
Stock10	0.124863	0.029315	0.038744	-0.09056	0.086917	-0.00363	-0.0169	…				
Stock11	-0.0367	0.244554	-0.0716	-0.00254	0.167674	-0.12443	-0.02192	…				
Stock12	0.139849	0.136323	-0.09862	-0.01259	0.046988	-0.1608	-0.14821	…				
Stock13	0.085618	-0.082475	-0.02779	-0.127	-0.02795	0.078171	0.127045	…				
Stock14	-0.02137	0.034113	-0.02563	0.049524	0.162119	0.001438	0.20785	…				
Stock15	0.041595	-0.12053	0.068819	0.14441	0.325174	-0.01442	-0.0231	…				
Stock16	0.084602	-0.01289	0.186888	-0.00399	-0.01387	-0.14697	-0.1529	…				
…	…	…	…	…	…	…	…	…	…	…	…	…
Stock104	-0.06942	-0.05584	-0.13427	-0.00025	-0.08048	0.069694	-0.02513	…	1	-0.11946	0.087345	0.005704
Stock105	-0.13772	-0.03623	-0.09502	-0.24092	-0.07798	0.023377	0.015225	…	-0.11946	1	-0.05018	-0.13912
Stock106	-0.07884	0.022096	-0.08619	0.015441	0.229058	-0.01589	0.020261	…	0.087345	-0.05018	1	0.013488
Stock107	-0.03514	-0.01693	-0.05282	-0.02989	-0.11311	0.003155	0.133951	…	0.005704	-0.13912	0.013488	1

Notices: There are in total only 9.8087169% correlation coefficients larger than 0.2 in absolute value.

Appendix III—p-value of the Spearman Correlation-Coefficient Matrix Test

p-value	Stock1	Stock2	Stock3	Stock4	Stock5	Stock6	Stock7	…	Stock104	Stock105	Stock106	Stock107
Stock1								…				
Stock2	.205							…				
Stock3	.972	.323						…				
Stock4	.558	.002***	.634					…				
Stock5	.727	.950	.293	.501				…				
Stock6	.917	.102	.383	.802	.505			…				
Stock7	.780	.693	.678	.074*	.262	.438		…				
Stock8	.585	.007***	.049	.680	.108	.859	.085*	…				
Stock9	.285	.353	.583	.187	.742	.394	.434	…				
Stock10	.249	.002***	.408	.314	.029*	.777	.554	…				
Stock11	.126	.043**	.371	.823	.465	.198	.479	…				
Stock12	.091*	.000***	.415	.020**	.540	.072*	.918	…				
Stock13	.285	.188	.787	.011**	.117	.018**	.189	…				
Stock14	.408	.169	.305	.826	.001***	.312	.528	…				
Stock15	.261	.675	.026**	.230	.647	.036**	.822	…				
Stock16	.707	.936	.992	.765	.326	.636	.729	…				
…	…	…	…	…	…	…	…	…	…	…	…	…
Stock104	.272	.937	.812	.384	.019**	.568	.410	…		.166	.309	.902
Stock105	.572	.642	.400	.692	.909	.461	.075**	…	.166		.650	.746
Stock106	.982	.987	.702	.049**	.122	.958	.278	…	.309	.650		.693
Stock107	.839	.001***	.000***	.451	.352	.196	.876	…	.902	.746	.693	

Notices:

1%, 5% and 10% significance level are denoted by ***, ** and *. Significant result indicates significant correlation coefficient.

Appendix IV—Marginal Contribution on Total Risk of Individual Stocks Every Month

Months	Stock1	Stock2	Stock3	Stock4	Stock5	Stock6	Stock7	…	Stock104	Stock105	Stock106	Stock107
M1	0.017912	0.111787	0.080273	0.084837	0.117055	0.078132	0.069213	…	0.117048	0.081353	0.072001	0.062437
M2	0.018301	0.112376	0.077939	0.084171	0.117722	0.074075	0.069052	…	0.114907	0.076401	0.06973	0.067919
M3	0.018877	0.110844	0.079053	0.085227	0.118692	0.076142	0.070379	…	0.117321	0.076255	0.070939	0.066845
M4	0.018533	0.110772	0.079521	0.086516	0.119296	0.076704	0.070044	…	0.118447	0.076393	0.070441	0.06792
M5	0.020762	0.115119	0.077318	0.084115	0.119201	0.07722	0.074524	…	0.118845	0.078376	0.071247	0.068939
M6	0.021155	0.115135	0.079108	0.083271	0.121149	0.077703	0.078605	…	0.115245	0.077925	0.071399	0.068345
M7	0.020239	0.114936	0.079647	0.08403	0.12038	0.078119	0.0777	…	0.11721	0.078893	0.070503	0.068675
M8	0.020462	0.116735	0.07698	0.084772	0.118243	0.077567	0.077487	…	0.119617	0.077457	0.07052	0.07065
M9	0.019716	0.117442	0.077766	0.085361	0.12053	0.079222	0.078191	…	0.118949	0.072388	0.068685	0.071754
M10	0.017709	0.114915	0.077963	0.085722	0.118995	0.078179	0.077058	…	0.114739	0.074057	0.070203	0.069536
M11	0.016592	0.114926	0.077901	0.083659	0.11906	0.077064	0.076372	…	0.114928	0.075803	0.071061	0.070155
M12	0.016611	0.113728	0.079106	0.084733	0.118726	0.078418	0.075687	…	0.112714	0.07545	0.071461	0.069999
M13	0.017121	0.113715	0.079052	0.085088	0.118439	0.077335	0.075982	…	0.113687	0.075522	0.072023	0.069678
M14	0.014953	0.118871	0.079222	0.087542	0.124758	0.083201	0.065025	…	0.099838	0.081346	0.078198	0.052673
M15	0.014987	0.119735	0.077777	0.086966	0.124691	0.084474	0.066117	…	0.101763	0.082421	0.077571	0.053027
M16	0.014853	0.122774	0.075878	0.09142	0.127269	0.084643	0.066242	…	0.108326	0.081062	0.076872	0.052438
…	…	…	…	…	…	…	…	…	…	…	…	…
M70	0.004192	0.119117	0.07386	0.095268	0.125481	0.099625	0.085678	…	0.111831	0.082954	0.075739	0.074501
M71	0.004048	0.119392	0.073369	0.095277	0.125383	0.100496	0.086163	…	0.109309	0.083075	0.076447	0.075295
M72	0.003963	0.122129	0.075583	0.098512	0.128999	0.104551	0.087383	…	0.109797	0.082249	0.079243	0.074803
M73	0.004243	0.122645	0.081631	0.106957	0.126219	0.102591	0.082782	…	0.110283	0.086259	0.079431	0.074286

Notices: This is the results from dynamic model with 73 estimation periods. Every row indicates their MCTR of each month

致谢

历时三个多月的时间，我终于在老师的指导之下，完成了这篇研究多因子模型在 A 股中的应用的论文，也是我大学的最后一篇文章了。还记得去年刚开学的时候，我参加了西南财经大学与倍发合作的联合实验室的项目，结识了一帮跟我一样以量化金融为今后奋斗目标的小伙伴们。首先要感谢黄辰老师，给我们引入了 BARRA 的多因子模型，他作为一个实际的交易员告诉了我们风险控制在实地投资中的重要性，并使我对我毕业论文的选题有了不二的选择——有关风控，我想试一试在这个领域中我的应用能力有多强。在做这篇文章的过程中的一些步骤，我也会跟黄辰老师聊聊，他对于我今后就业的方向也有很多指导。

最重要的，我要感谢我毕业论文的导师——黄霖老师。她在我选题过程中给出了一些奠基性的宝贵意见，包括建议我去阅读当下一些做多因子的论文去提炼一些观点还有调整我自己的建模思路；在写作过程中更是经常给我相当大建设性的建议和意见，让我对于我模型的修正有了更自信可行的方法。在我写做完了之后，黄霖老师非常耐心和细心地给我提出了修改意见和一些写作指导，我被黄霖老师这种对学生认真的精神感动着，也激励着我去完成一篇高质量的更完美的论文。感谢黄霖老师！

另外，在我编程去进行数据处理的期间，姚天昊同学和段威同学对我也有很大的帮助，他们的指导让我对 MATLAB 和 Python 的功能有了更清楚的认识，很大程度上提高了我自己的编程能力。

最后感谢经管院四年对我们悉心的教育和培养，让我们拓宽了视野，掌握了专业知识。写作的过程并非一蹴而就，曲折与心绪都蕴含其中，仅以此篇作为纪念，此时，此地！

经世济民

孜孜以求