Empirical Essays on Asset Pricing in the Chinese Stock Market

Yu Wang

Doctor of Philosophy

University of York
Economics

September 2016

I would like to dedicate this thesis to my loving daughter, wife, and parents.

Abstract

This thesis empirically studies asset pricing in the Chinese stock market. It consists of four separate but closely related essays in Chapters 2-5, respectively.

The first essay tests the validity of the standard Fama-French model in the Chinese Main-board stock market within a relatively new sample period since quarterly financial reports became widely available. The empirical results confirm the good performance of the Fama-French model. By splitting the sample into two sub-periods representing two different market tendencies, the Fama-French model performs better but is less stable in the downward trend than the fluctuating market. The potential explanation is also provided.

The complexity of traditional methods to construct the Fama-French factors limits the model's popularity. The second essay examines a simplified method to construct the factors using the Chinese style indices. The results show that the style index-formed factors lead the Fama-French model to better explaining equity returns and producing less pricing error. The finding is meaningful in the sense that it dramatically lowers the barrier to examining and applying the Fama-French model in both academic research and financial practice.

However, the theoretical foundation behind the Fama-French model, i.e., which macroeconomic variables fundamentally determine the model's success, is not clear. The rest two essays make efforts to find the answer. The third essay examines the extent to which the market, size, value and momentum factors can be linked to fundamental macroeconomic variables. We find that the momentum factor consistently captures the current economic risk as well as predict future GDP growth in China. In contrast, there is no consistent relationship between the market, size and value factors and macroeconomic growth. This finding contradicts Liew and Vassalou's (2000) conclusion that the Fama-French factors can generally predict future real GDP growth in ten developed countries.

The fourth essay examines whether the ability of the above four factors to explain the cross-sectional variation in equity returns, if any, stems from the fact that they are proxies for fundamental macroeconomic risk associated with future GDP growth. The results show that the market, size and value factors together contain some overlapping innovative information about future real GDP growth, and thus, reconcile previous findings in the US (Vassalou, 2003) and Australian (Nguyen et al., 2009) markets. Some possible explanations are also discussed.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my PhD co-supervisors, Prof. Peter Smith and Prof. Zaifu Yang, and PhD Thesis Advisory Panel member, Prof. Jacco Thijssen, for all the invaluable guidance and support from them. Their comments and suggestions improved this thesis a lot and help me develop as a scholar.

I also gratefully acknowledge the financial support by the Departmental PhD Studentship from the Department of Economics and Related Studies, University of York, the State Scholarship Fund from the China Scholarship Council, as well as the Youth Project of the National Social Science Foundation of China: "Mechanism Studies and Policy Suggestions on the Financial Ecological Environment and Spatial agglomeration of Technological Innovation (No. 15CJL052)".

I am indebted to my parents for the best they can do for me, and to all my relatives in China and friends in both China and the UK for their support and faith in me.

I would like to thank my wife, Honglian, for giving a birth to our daughter, Yizhen, before my PhD study and for raising our lovely daughter devotedly in China. I would also like to thank my little Yizhen for her understanding that I cannot company with her every day so I can focus on my PhD study at the University of York, UK. PhD degree is important to me, but since her birth, I have realized that nothing is more important than her.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this thesis are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration.

Chapter 1

Introduction

In finance, the asset pricing theory tries to understand the prices or values of claims to uncertain payments. In academic settings, we commonly use asset pricing theory to give an economic explanation for why prices are what they are, or to predict how prices might change if policy or economic structure changed. The simple concept is that price equals expected discounted payoff. The payoff happens in the future and is uncertain currently. Here, the time effects are not too difficult but the uncertainty or risk makes asset pricing interesting and challenging. In financial literature, you may find that much of finance focuses on expected returns. In fact, a low initial price for a given payoff corresponds to a high expected return, so price and return are different languages for the same phenomenon.

As a branch of the asset pricing, the (linear) factor pricing models have become one of the most popular research areas in finance. The factor pricing models can be tracked back to the modern portfolio theory, which is built on Markowitz's (1952) mean-variance analysis. Based on the earlier work of Markowitz on diversification and modern portfolio theory, Sharpe (1964) and Lintner (1965a,b) independently introduce the capital asset pricing model (CAPM). The CAPM is a model that derives the theoretical required expected return, or discount rate, for an asset in a market, given the risk-free rate available to investors and the risk of the market as a whole. The CAPM is usually expressed as this form:

$$E(R_i) = R_f + \beta_i(E(R_m) - R_f)$$

where β is the ratio of the covariance of returns on a risky asset and a market portfolio to the variance of market portfolio return. It measures the asset's sensitivity to a movement in the overall market. Beta is usually found via regression on historical data. Betas exceeding one signify more than average "riskiness" in the sense of the asset's contribution to overall portfolio risk; betas below one indicate a lower than average risk contribution. $E(R_m) - R_f$ is called the market premium. It's the expected excess return of the market portfolio over the risk-free rate.

The empirical tests on the CAPM show that it is far from the right model to explain the asset returns in the real world. The empirical evidence suggests that a large portion of equity return cannot be explained by the CAPM.[1] For example, using the same sample of ten Chinese sector portfolios as we will use in Chapter 3, CAPM can only explain, on average, about 74% of the returns on sector portfolios. The lowest is about 55% for the Health Care Sector, and the highest is 90% for the Industrials Sector.[2]

To improve the explanatory power of the CAPM, Fama and French (1993) propose a three-factor model including the additional size and value factors. The three-factor model is of this form:

$$E(R_i) = R_f + \beta_{im}(E(R_m) - R_f) + \beta_{is}E(SMB) + \beta_{ih}E(HML)$$

where *SMB*, meaning small minus big, is called the size factor, and it's the difference between the returns on diversified portfolios of small- and big-size stocks; *HML*, meaning high minus low, is the book-to-market equity factor or value factor for short, and it's the difference between the returns on diversified portfolios of high and low book-to-market equity ratio stocks.

The empirical tests on the Fama-French model show a significant improvement over the CAPM in explaining equity returns. Again, using the same Chinese sample as mentioned above, the Fama-French model can explain, on average, about 84% of the returns on sector portfolios, that is, about 10% average improvement over the CAPM. The lowest is still for the Health Care Sector, but it's about 76% now, that is, 21.1% increased. The highest is about 94% for the Industrials Sector.

Since its initiation, the Fama-French model has become a benchmark and been well recognized as a major breakthrough of traditional pricing models, such as the CAPM of Sharpe (1964) and Lintner (1965a,b), the Arbitrage Pricing Theory (APT) of Ross (1976, 1977), and ICAPM of Merton (1973). It is believed that the firm-specific size and book-to-market equity (or value) factors contain some important information that is not captured by the market factor of the CAPM but help the Fama-French model explain the cross-sectional variation in equity

[1] See Fama and French (1992, 1993, 1995, 1996, 1998), Grinold (1993), Davis (1994) and He and Ng (1994), to name a few.

[2] The results of empirical tests on the CAPM, and the Fama-French model, using the sample of the Chinese sector portfolios are not fully reported in this thesis. We report the selected results here only for presenting the improvement of the Fama-French model over the CAPM in explaining the returns on equity portfolios.

returns.

Further efforts have been made to improve the Fama-French model. For example, Carhart (1997) includes the momentum factor into the Fama-French model; Rahim and Nor (2006) suggest the liquidity as a potential factor that can improve the Fama-French model; and most recently, Fama and French (2015a, 2016) propose a five-factor model where the two additional factors of profitability and investment are included. However, current empirical evidence shows that the further improvement of those extensions over the Fama-French model is not as impressive as that over the CAPM.

Since 2010 when its reported GDP in 2009 surpassed Japan's for the first time, China has become the world's second largest economy, following the implementation of economic reform initiated in 1978. China has experienced rapid economic development with GDP growth rate averaging 10% per year over the past three decades. Although China remains a developing country and its market reforms are incomplete, it has been playing an increasingly important and influential role in the global economy.

In tandem with the economic reform and development, China's stock market has also been evolving significantly over the past 30 years. China's modern stock market was established in the early 1990s. There are two stock exchanges nowadays in Mainland China: the Shanghai Stock Exchange (SSE) and the Shenzhen Stock Exchange (SZSE). The SSE was founded on 26th November 1990, while the SZSE was launched on 1st December 1990. Both exchanges comprise the Main Board, but unlike the SSE, the SZSE also introduced the SME Board in May 2004 and the ChiNext market in October 2009. Thus, the SZSE has constructed a framework of China's multi-tiered capital market system including the Main Board, SME Board, and the ChiNext market.

Since the establishment of the two stock exchanges, the Chinese stock market has experienced dramatic development, especially over the past decade. The number of domestic listed companies increased from 53 in 1992 up to 2,613 in 2014 (See Figure 1.1).

At the very beginning of the establishment of the Chinese stock market, the total market capitalization was only RMB 104.81 billion in 1992. The GDP for the same year was RMB

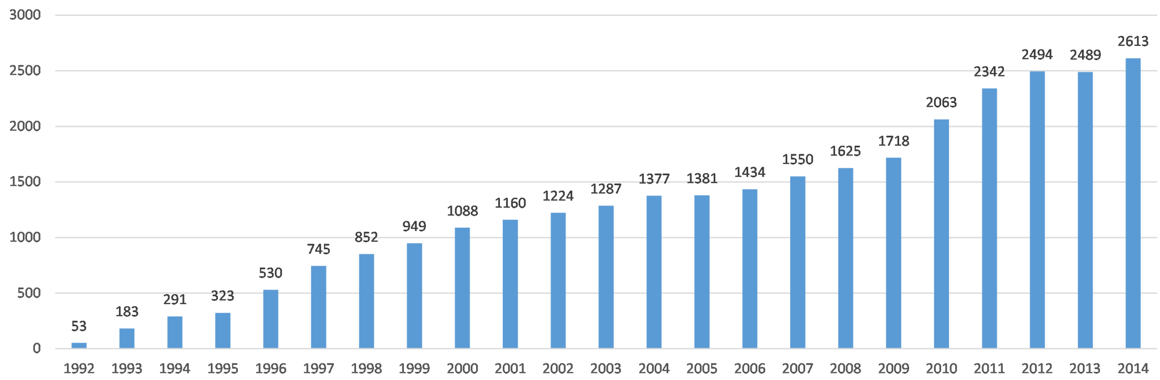

Figure 1.1: Number of domestic listed companies on both SSE and SZSE

Data Source: China Securities Regulatory Commission (CSRC)

2,706.83 billion. This means that China's stock market capitalization was only 4% of its GDP (See Figure 1.2). After 2006, the size of the Chinese stock market expanded dramatically. In 2012, the total market capitalization in the A-share market amounted to RMB 23.04 trillion and was equivalent to 43% of China's GDP in 2012 (See Figure 1.2). This made China become the world's second largest stock market, second to the US. By the end of 2014, the total market capitalization on the two exchanges reached RMB 37.25 trillion, that is, 59% of China's GDP in that year. Unlike the rapid and stable growth of GDP (with a decreasing growth rate), during the past two decades, the Chinese stock market experienced several periods within which financial bubbles boomed and burst. The most notable bubble was in 2007 when the market value was 22% more than the year's GDP, and the market crashed right after it in 2008. The Chinese stock market is generally considered that it is largely independent of the Chinese economy. This is one of the specific features of the Chinese stock market.

As a very young emerging market that is still developing, the Chinese stock market has many features which distinguish itself from many other developed markets and thus make itself interesting.

Firstly, the Chinese stock market is a relatively closed market. This feature can be partially reflected by the types of shares. Both SSE and SZSE have A- and B-share markets. A-shares make up the largest stock class in the Chinese market and are denominated in the Chinese currency (RMB). On the other hand, B-shares are traded on both stock exchanges as well but

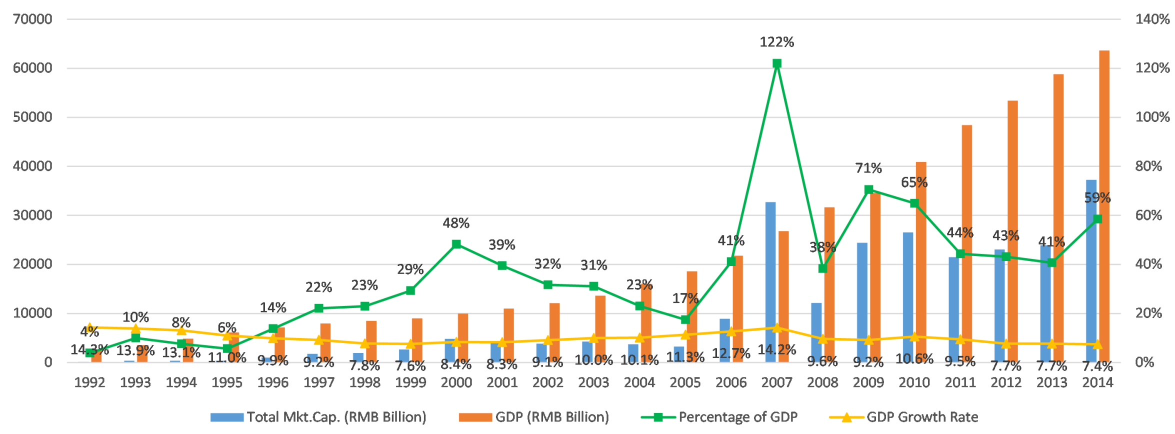

Figure 1.2: Total market cap. of SSE and SZSE and GDP

Data Source: China Securities Regulatory Commission (CSRC) and National Bureau of Statistics of China (NBS)

in US dollars and Hong Kong dollars, respectively. For a long time, only Chinese nationals were allowed to access the A-share market until 2003 when the A-share market was also open to the Qualified Foreign Institutional Investor (QFII). Originally, B-shares were only open to foreign investors, but since 2001 the domestic investors have been allowed to buy the B-shares. However, the relatively limited number of issuers and investors leads the B-share market to be inactive and lagged far behind the A-share market.

Secondly, the Chinese stock market is highly regulated by the authority. The government plays an important role in the market through not only strict regulation but also direct intervention. For example, due to the highly strict regulation on the margin trading/lending, the damage caused by the financial crash in 2007-2008 was limited. Another example is the financial bubble very recently. As presented in Figure 1.3, as of June 2015, the Chinese market hit the second highest peak in history. In early July of 2015, after a nearly 30% fall in the market from the highs on 12th June, the central government made utmost efforts to stabilize the market. The state-owned blue-chip firms and major Chinese securities firms were required by the China Securities Regulatory Commission (CSRC) to buy and hold a substantial amount of securities affected by the downturn. The China Securities Finance Corporation (CSF), a firm created by China's commodities and stock exchanges to finance trades, used funds supplied by the central bank and commercial banks to purchase enough stocks in order to stop the crash.

Thirdly, unlike in the western developed markets where institutional investors predominate, most trades on the Chinese market[3] are triggered by unsophisticated individual investors who trade stock solely based on the rapid increase in stock price. According to the China Securities Regulatory Commission (CSRC) 2014 Annual Report[4], individual investors[5] held 85.78% of A shares by market capitalization. It is exactly different in other developed markets. For example, according to Celik and Isaksson (2013), it is institutional investors who held around 60% of all publicly listed stocks in the US and 82% in Japan in 2011, and the portion was around 89% in the UK in 2012. Although natural persons only held 25.03% of China's A shares in 2014[6], 85.37% of annual turnover was contributed by them. As a result, Chinese individual investors focus more on short-term investment and trade speculatively (Wang and Xu, 2004) and emotionally causing the Chinese market to be much more volatile than other developed markets. As shown in Figures 1.3 and 1.4, China's stock indexes present more volatility compared to the three US indexes, particularly the Standard & Poor's 500 index, in the period from 1992 to 2015.

The fourth feature is the imbalanced industry structure. The Chinese stock market leans heavily towards manufacturing firms, due to China's economy structure. According to the "Results of Industry Classification of Listed Companies" for the fourth quarter of 2014 by CRSC[7], there are more than 60% of listed firms are classified into the manufacturing sector, while other sectors share the rest proportion. In contrast, the dominant sectors are service business-based industries in many developed markets. As a consequence, it is not surprising that China is well known as the World's Factory.

Fifthly, the existence of restricted shares in the Chinese stock market is another feature that should be noticed. The restricted shares were transformed from the non-tradable shares which

[3] 80% of Chinese stocks are owned by individual investors, according to the financial report of New York Times, "China's Market Rout is a Double Threat", by Keith Bradsher and Chris Buckley on 5th July, 2015.

[4] Referred to the CSRC website at http://www.csrc.gov.cn/pub/newsite/zjhjs/zjhnb/.

[5] The CSRC 2014 Annual Report classifies investors into professional institutions, natural persons and general institutions. The former are the commonly called institutional investors who invest on behalf of other people or organizations, while the latter two are retail/individual investors who trade for their personal account.

[6] General institutions held the rest 60.75% of A shares.

[7] http://www.csrc.gov.cn/pub/newsite/scb/ssgshyfljg/, in Chinese.

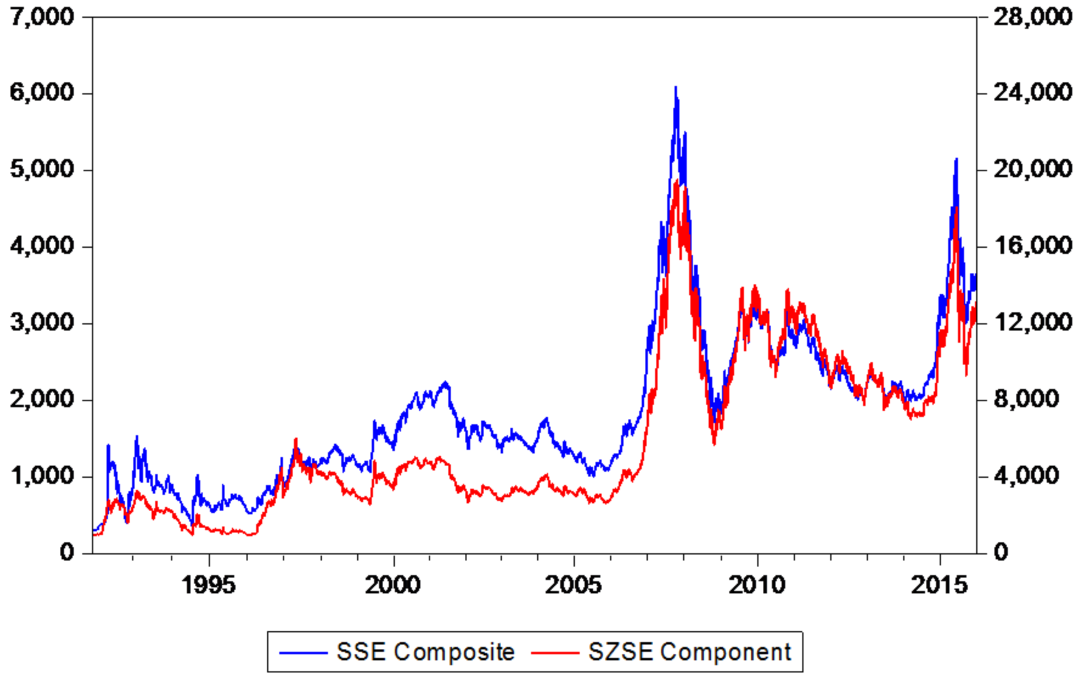

Figure 1.3: SSE Composite Index and SZSE Component Index

Note: The SSE Composite Index is plotted against the left y-axis, while the SZSE Component Index is against the right y-axis.
Data Source: Shanghai DZH Limited

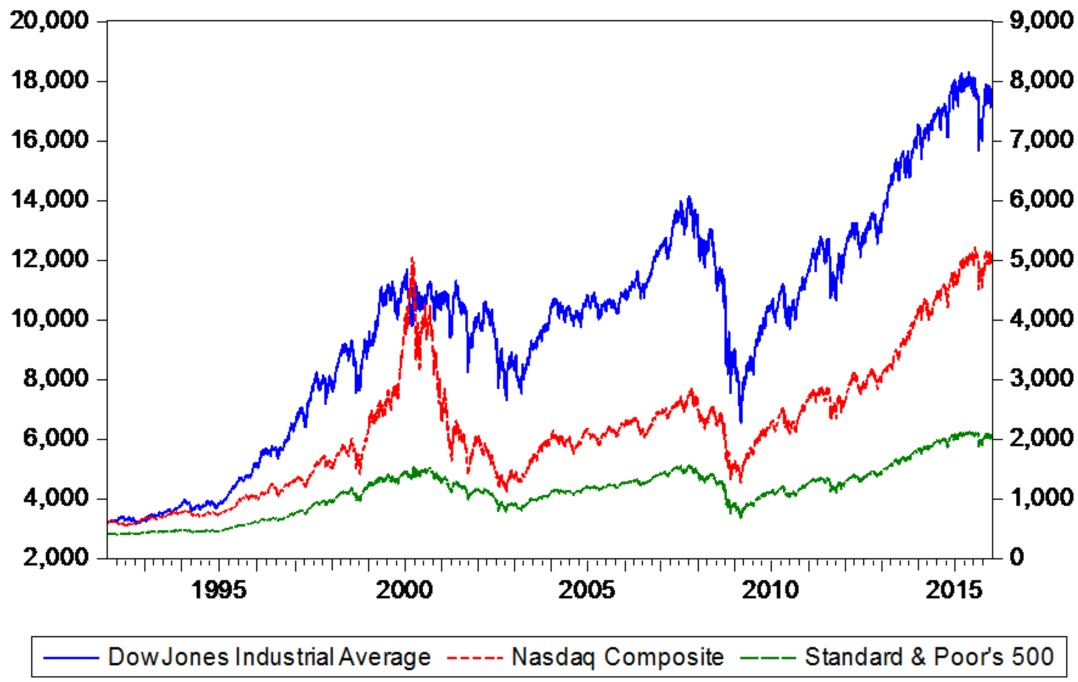

Figure 1.4: Dow Jones Industrial Average, Nasdaq Composite and Standard &Poor's 500

Note: The Dow Jones Industrial Average Index is plotted against the left y-axis, while the Nasdaq Composite and Standard &Poor's 500 Indexes are against the right y-axis.
Data Source: Shanghai DZH Limited

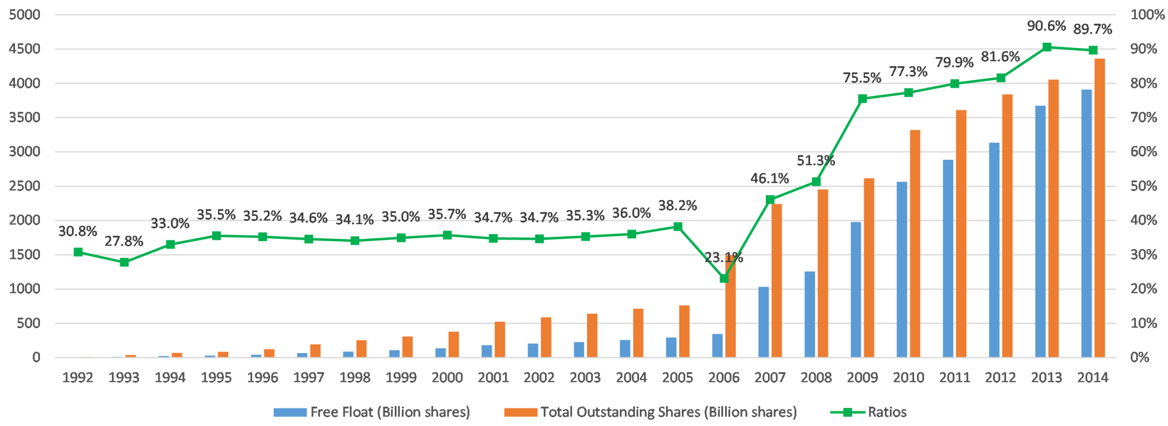

Figure 1.5: Total outstanding shares and free float

Data Source: China Securities Regulatory Commission (CSRC)

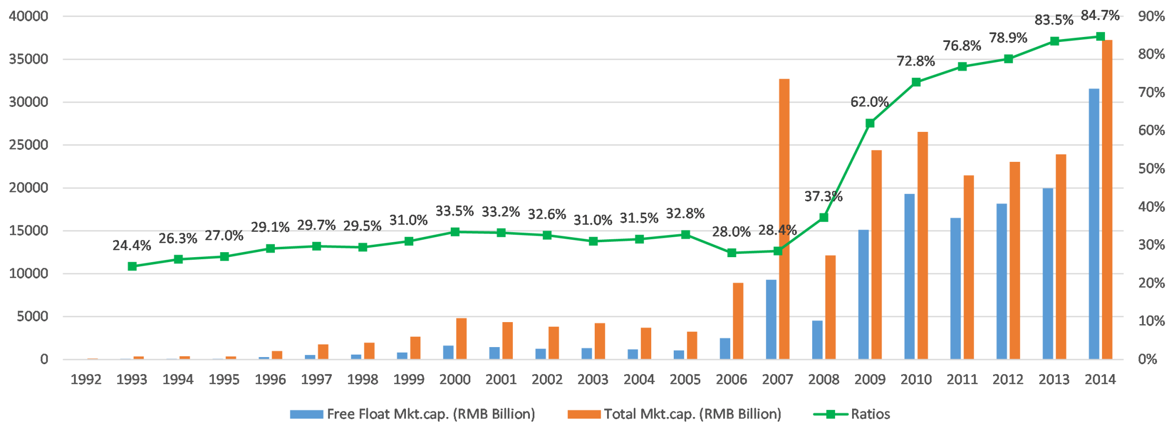

Figure 1.6: Total market cap. and free float market cap.

Data Source: China Securities Regulatory Commission (CSRC)

were owned by the state and were prohibited from selling before the reform of shareholder structure in 2005. The restricted shares were firstly non-floating after the transformation, but would be gradually unleashed after a lockup period. From Figures 1.5 and 1.6, after the shareholder structure reform, both free float shares and their market capitalization experienced a fast increasing period until 2013. Both ratios of free float shares to total outstanding shares and of free float market capitalization to total market capitalization increased to over 80% by the end of 2014. Although the restricted shares have decreased a lot after the reform, the amount of these shares still accounts for a sizable portion of the Chinese stock market.

The last but not the least feature of the Chinese stock market is that the state-holding listed companies have been dominating the market. By the end of 2012, there were 953 state-holding companies listed on the A-share markets of both Shanghai and Shenzhen stock exchanges, accounting for 38.5% of companies listed on the Chinese A-share market. Their market capitalization was worth RMB 13.71 trillion in 2012, occupying 51.4% of market capitalization of all listed companies in the A-share markets. [8]

No matter how special the Chinese stock market is, as a large emerging market that is still quickly developing, there exist lots of investment opportunities. Both domestic and global investors cannot ignore this market. Therefore, the most famous global indexes providers have included or are considering including China's A-shares into their global benchmark index. For example, on 26 May 2015, FTSE started the transition to include China A Shares in its global benchmarks. MSCI is also considering bringing the Chinese stocks into its global emerging market benchmark. Once China's A-shares are included into the global benchmark indexes, the global funds that track those benchmarks will invest a large amount into the Chinese stock market. They will, of course, increasingly care about the performance of the Chinese market.

As a consequence, its special features and the opportunities it provides for global investors make studies on the Chinese stock market interesting and important for both academic research and financial practice. This thesis aims to contribute to the related literature in the context of Chinese stock market. In particular, Chapter 2 tests the validity of the Fama-French model and its extension including the additional momentum and liquidity factors at a level of daily returns. We focus on the daily returns because the daily performance of funds and portfolios is increasingly concerned by the investors, which is especially the case after the boom of Chinese e-finance in recent years. Many mutual funds report their daily returns to customers. Besides, we construct the Fama-French factors with a quarterly reforming frequency after the quarterly financial reports of listed firms were required by the CSRC. The quarterly reforming frequency

[8] Referred to a speech delivered by Wang Yong, the chairman of State-owned Assets Supervision and Administration Commission of the State Council (SASAC), at the National State-owned Assets Supervision and Administration Working Conference at URL: http://www.sasac.gov.cn/n1180/n1566/n259760/n264785/15106589.html, in Chinese.

is one of the main differences from most existing studies. Moreover, we split the sample into two sub-periods which represent a downward and a fluctuating market trends. We confirm the good performance of the Fama-French model in explaining the cross-sectional variation in equity returns for the whole sample and the two sub-samples. The instability tests on the model and its coefficients suggest the model less constant in the downward market trend.

Chapter 3 examines an alternative method to construct the Fama-French factors using the style indexes in the Chinese stock market. By employing the free and publicly accessible style indexes, it dramatically simplifies the traditional ways to construct the two factors. Our results confirm that the Fama-French factors constructed using the simplified method have at least similar properties as those formed using the Fama and French's (1993) original method. The finding is exciting because it significantly eliminates the barrier to applying the Fama-French model in both academic research and financial practice. For academic research, the simplified method makes it much easier to examine and apply the Fama-French model, especially for the countries outside the US where there is no such data available for the Fama-French factors. The simplified method is more meaningful for the practice of risk hedging. We can hedge the risks based on the Fama-French model where the factors are constructed from the style indexes. This means that we only need to long or short the funds, if any, tracking the corresponding style indexes, which would make the hedging strategies transaction cost-efficient and affordable.

Fama and French (1993, 1995, 1996) attribute the success of their three-factor model to the fact that the size and value factors proxy for the state variable in the context of Intertemporal CAPM of Merton (1973). But the two factors lack clear economic links to systematic macroeconomic risk. The next two chapters aim to check the link between the pricing factors and macroeconomic variable, i.e., the GDP growth rate, again in the Chinese context. Chapter 4 examines whether the market, size, value and momentum factors can capture or predict China's GDP growth. We find that the market, size and value factors present a very weak relationship with current or future GDP growth, while the momentum factor can largely capture and predict GDP growth in China. This finding is inconsistent with Liew and Vassalou (2000), where they find that the size and value factors can generally predict future GDP growth in ten developed countries.

Chapter 5 examines whether the success of factor pricing models is attributed to the fact that the priced factors proxy for the risk related to future GDP growth in the Chinese context by answering two questions in the following order: 1) whether the priced factors proxy for the risk related to future GDP growth and 2) whether the future GDP growth risk is priced in the cross-section of equity returns. The future GDP growth risk is captured by a mimicking portfolio such that the risk is quantified in terms of asset returns. The empirical results show that the real, instead of nominal, GDP growth risk is priced in the equity portfolio returns, and that the market, size and value factors together with each other contain the overlapping information about the future real GDP growth risk in China. The Chinese evidence reconciles previous findings in the US and Australia and suggests that the information about macroeconomic risks contained in the priced factors may be country-specific.

Chapter 6 concludes this thesis and provides a brief discussion about possible extensions of the studies conducted in this thesis for future research.

Chapter 2

Testing the Fama-French-Type Model for Different Market Trends - New Evidence from Chinese Stock Market at Daily Level

2.1 Introduction

In a pioneering study, Fama and French (1992) find that the domestic Capital Asset Pricing Model (CAPM) cannot explain the cross-section of asset returns in the US. They proposed an alternative three-factor model which includes, apart from the market factor, a factor related to firms' size or market value (MV) called Small-Minus-Big (SMB), and a factor related to firms' book-to-market equity ratio (BE/ME) which they call High-Minus-Low (HML). Fama and French (1993) further find strong evidence for that the expected excess returns on stocks can be largely explained by their three-factor model in a time-series framework.

Since then, the Fama-French three-factor model (the FF model henceforth) has become a benchmark model. Among a large number of follow-up papers, for example, Fama and French (1995, 1996) document that their model can explain the average equity returns very well not only in the US but also across the world. The FF model has been well recognized as a major breakaway from the traditional asset pricing models such as the CAPM of Treynor (1961, 1962), Sharpe (1964), Lintner (1965a,b), and Mossin (1966), the Arbitrage Pricing Theory (APT) of Ross (1976, 1977), and the ICAPM of Merton (1973).

In spite of its lack of theoretical foundation, the FF model has become very popular for empirical studies. It is believed that the firm-specific size and book-to-market equity factors contain important information that is not captured by the market factor of the CAPM but helps the FF model explain the cross-sectional variation in equity returns on portfolios.

Several extensions of the FF model have also been proposed. For example, Carhart (1997) extends the FF model by adding the momentum factor into the model (a four-factor model) to help explain the mutual fund performance. Liquidity is another important effect in asset pricing models suggested, for example, by Hodrick and Zhang (2001). Unlike the momentum, however, liquidity is an elusive concept that is difficult to define and measure. It is generally defined as that an asset can be traded immediately at low cost with little price impact, which involves four dimensions: trading quantity, trading speed, trading costs and price impact (Liu, 2004). Currently, there is no measurement of liquidity that can capture all the four dimensions together. For example, the bid-ask spread of Amihud and Mendelson (1986) focuses on trading costs; the dollar volume of Brennan et al. (1998) and the share turnover of Datar et al. (1998) capture the trading quantity dimension; while the illiquidity of Amihud (2002) and Pastor and Stambaugh (2003) focuses on the price impact.[1] Among those measurements of liquidity, we choose the turnover as the proxy for liquidity not only because of its simplicity but also because the turnover ratio is one of the common technique analysis indicators used by Chinese investors. Unlike many asset pricing studies, for instance, Amihud (2002) and Pastor and Stambaugh (2003)[2], where the liquidity measures are used directly to price assets, we construct the liquidity factor in line with the idea of constructing the zero-investment mimicking portfolios for the size, book-to-market, and momentum factors. Thus, our liquidity factor is constructed similarly as the one in Rahim and Nor (2006). Moreover, among the literature that studies the importance of liquidity in asset pricing models, empirical evidence also presents that liquidity has an effect on stock returns which is highly correlated with market capitalization or size, and the illiquidity effect is stronger for small firms' stocks.[3] This is because the small firms tend to suffer from the changes in market illiquidity. In times of liquidity drains, the "flight to liquidity" makes large firms' stocks more attractive over small firms' stocks.

Since the establishment of the Shanghai and Shenzhen stock exchanges in the early 1990s, the Chinese stock market has experienced dramatic development, especially over the past decade. Nowadays, the Chinese stock market has become one of the largest markets in the world in terms of the market capitalization of its A-share market. However, it is still relatively young and is considered as an emerging and developing market. There are several special characteristics which distinguish the Chinese stock market from other developed stock markets and make it interesting to examine the application of the FF model in the Chinese stock market.

Firstly, the Chinese stock market is a relatively closed market with strict government regulation. This feature can be partially reflected by the types of shares. For the domestic stock

[1] A more comprehensive summary of empirical studies on the role of liquidity in the US and other countries can be found in Rahim and Nor (2006)

[2] Narayan and Zheng (2010) employ the liquidity measure of Pastor and Stambaugh (2003) and show that a market liquidity helps explain cross-sectional stock returns in the Chinese market.

[3] For example, Amihud (2002), Pastor and Stambaugh (2003), and Rahim and Nor (2006), among others.

market, many Chinese listed companies issue both A-shares and B-shares. A-shares constitute the largest stock class in the Chinese market and are traded on both Shanghai and Shenzhen stock exchanges in the Chinese currency (RMB). Only Chinese nationals and the Qualified Foreign Institutional Investor (QFII) are authorized to trade A-shares. On the other hand, B-shares are traded on both Shanghai and Shenzhen stock exchanges as well but in US dollars and Hong Kong dollars, respectively. Originally, B-shares were only open to foreign investors but now are also open to domestic investors since 2001. However, the relatively limited number of issuers and investors leads the B-share market to be inactive and lagged far behind the A-share market.

Secondly, it is individual investors rather than institutional investors who are heavily involved in daily stock trading. According to the China Securities Regulatory Commission (CSRC) 2014 Annual Report, 85.37% of annual turnover was contributed by natural persons who only held 25.03% of A shares in 2014. The majority of Chinese individual investors lack the experience of equity investment and are relatively less sophisticated compared to the participants in other developed markets. The individual investors in the Chinese stock market generally focus more on short-term investment and trade speculatively and emotionally. (Wang and Xu, 2004) This feature would, in turn, result in a more volatile market.

The third feature is the imbalanced industry structure existing in the Chinese stock market. More than half of the Chinese listed firms are manufacturing-related firms, while other sectors share the rest proportion. In contrast, the service-related sectors dominate many developed markets. This phenomenon is linked to China's economic structure; China is well known as the World's Factory. This feature enforces Chinese investors to put their assets into less diversified portfolios, which makes the idiosyncratic risk hard to eliminate by diversification. Thus, the investment risk is potentially higher in the Chinese market than in many other mature markets.

Fourthly, the existence of restricted shares in the Chinese stock market is another feature that should be noticed. The restricted shares were transformed from the non-tradable shares which were owned by the state and were prohibited from selling before the reform of shareholder structure in 2005. The restricted shares are currently non-floating but will be gradually unleashed after a lockup period. Although the restricted shares have decreased a lot after the reform, the amount of these shares still accounts for a sizable portion of the Chinese stock market. An existing study that examine this feature of the Chinese market is conducted by Wang and Xu (2004). Instead of the book-to-market factor, a factor based on the free float ratio was added to serve as a proxy for corporate governance quality. They find that the model including the market, size, and free float factors significantly improve the explanatory power of the market model. Considering this feature is reasonable, especially before the reform of shareholder structure in 2005 and in the first several years after the reform. We do not consider free float as a pricing factor in this study, because the free float in the Chinese market has increased from 46% in 2007 to over 90% in 2013 by outstanding shares and 28% to 83% by market capitalization as shown in Figures 1.5 and 1.6. That means we cannot distinguish most of firms when ranking by their free float in our sample of this study, for most firms have zero non-free floating shares. But, the reform has not been completed and we should still keep in mind that this feature might affect our empirical results to some extent. Thus, to partially capture this feature in our study, we calculate the weights based on the free-floating market capitalization for the value-weighted returns of portfolios.

Last but not the least, although the restricted shares, most of which are the state-owned shares, have decreased a lot, the state-holding listed companies are still dominating the Chinese A-share markets. By the end of 2012, there were 953 state-holding companies listed on the A-share markets of both Shanghai and Shenzhen exchanges, accounting for 38.5% of companies listed in the Chinese A-share market. Their market capitalization was worth RMB 13.71 trillion in 2012, occupying 51.4% of market capitalization of all listed companies in the A-share markets.[4] The state-holding companies are generally large firms and have market advantages (or even some monopoly power) over other private firms. Their performance is generally better, especially when the market conditions are poor. From this perspective, the

[4]Referred to a speech delivered by Wang Yong, the chairman of State-owned Assets Supervision and Administration Commission of the State Council (SASAC), at the National State-owned Assets Supervision and Administration Working Conference (http://www.sasac.gov.cn/n1180/n1566/n259760/n264785/15106589.html).

state-holding companies could stabilize the market. But, due to the fact that the shares held by the state are generally restricted for sale in the secondary market, the effect of state-holding companies on the market could be passed through the supply-demand pattern. The restriction on sale causes the supply of blue-chip shares in the secondary market is relatively limited. In the good conditions, most investors have to buy the stocks of relatively small firms and push up the prices of those stocks even higher. When the conditions turn round to become worse, the investors pursue the unrestricted shares of state-holding companies' stocks. But due to the limitation of supply of blue chips, the holders of blue chips would like to continue holding those shares. Other holders of small firms' stocks would like to sold out their shares, which will dramatically discount values of those shares. Unfortunately, there is no counterparty in the market want to buy those shares. The fact that no market maker exists in the Chinese market would make this situation worse. As a result, the trading would be frozen in poor market conditions, and the liquidity in the market drains out. This would further intensify the volatility of the Chinese stock market, and also suggests that the liquidity is important to examine the asset pricing in the Chinese market. But, as stated above that liquidity effect is high correlated with size effect, it would not be too surprising if the liquidity factor is insignificant when we also include size factor as in this study. Moreover, due to the multi-dimension of liquidity, we also admit that some measurement of liquidity might be able to explain the return anomalies that cannot be explained by other factors. We leave this open question for future research.

This study aims at testing the FF model in the Chinese stock market within a relatively new sample period from 2007 to 2013 since quarterly financial reports became widely available. Specifically, unlike the existing studies, we construct the Fama-French factors based on quarterly financial reports[5], instead of the annual reports, and examine the model by employing daily portfolio returns. Our empirical test results suggest that the FF model performs well in the Chinese market and that adding the two additional momentum and liquidity factors does not seem to improve the explanatory power of the FF model. Comparing to the empirical results in Fama and French (1993) where the tests are on the monthly basis, our results show

[5]The complete procedure for construction of factors is provided in Section 2.3.1

that both the CAPM and the FF model perform much better on the basis of daily returns. At the first glance, this outstanding improvement seems to mainly come from the stronger explanatory power of the daily market factor. To examine the sources that causes the different findings between ours and Fama and French's (1993). We further examine the models using both monthly and daily data within different periods in the US.[6] We find that during the same period as in Fama and French (1993), i.e., from July 1963 to December 1991, using daily data leads to a slightly poorer performance of the models. But we reach to exactly opposite findings during the period as in our study, i.e., from July 2007 to December 2013. Using daily data results in 1-2% increase in the average adjusted R^2 for both models than using monthly data. More importantly, within the recent period as in our study, the US market witnesses a dramatically improvement of the model performance compared to the previous period within which Fama and French (1993) conduct their study. For the CAPM, the average adjusted R^2 increases from 74% to 89% based on daily data, and from 78% to 87% based on monthly data. While for the FF model, the relevant increases are from 88% to 96% and from 93% to 94% based on daily and monthly data, respectively. This finding suggests that the ability of market factor to explain the equity returns (not the cross-sectional variation in returns) are evolving over time, and the size and book-to-market factors can still explain the portion left by the market factor. While the significant improvement in model performance over time could be attributed to the fact that the diffusion of information over the market is increasingly free and speedy. The results also show that in recent period daily data contain more information which help explain the equity returns than monthly data do, although daily data could also be contaminated more with noise. In the trade-off between more information and more noise, we choose the former and this is one of the returns for employing daily data in our study.

Our study is different from most existing studies in several other aspects. Firstly, most of current studies overwhelmingly examine developed markets, such as the most examined US market (Fama and French, 1992, 1993 and Carhart, 1997), European markets (Gregory

[6]The US data are obtained from Kenneth R. French Data Library. We do not report the detailed empirical results in this study in order to conserve the space.

et al., 2009 and Gregory et al., 2013), Australian market (Faff, 2004[7]), and Japanese market (Chan et al., 1991, and Daniel et al., 2001). There are also studies on emerging markets including, for example, Korea, Malaysian, India, Philippines, Dhaka, Thailand, and Amman. In recent years, researchers have started to pay attention to the Chinese Stock Market. For example, Chen (2004), Liao and Shen (2008), Mao et al. (2008), Xu and Zhang (2014) and, most recently, Chen et al. (2015) examine the application of the FF model in the Chinese market. Our study contributes to the related literature by testing the performance of the FF model in the Chinese A-share (Main Board) markets, financial institutions[8] excluded, of both Shanghai and Shenzhen stock exchanges from 2007 to 2013. During the sample period, there were 1,392 listed firms in total showing up in the A-share markets, but of course the number of listed firms kept changing due to new listing into and delisting from the exchanges.

Secondly, the majority of existing literature on the FF model employs monthly returns, and this is especially the case for earlier studies before the start of this century.[9] Even though some argues that the monthly sampling interval is too short to accurately estimate beta,[10] we believe that testing the FF model based on the daily interval could be more meaningful. One reason is mentioned above that although it would suffer more noise, daily data provide more information for pricing assets over monthly data in recent days. Another reason is related to the features of the Chinese market. As we figure out that the Chinese investors "trade speculatively with very short holding periods" and "ignored long term investment objectives based on future profitability of a firm" (Wang and Xu, 2004), they care more about the daily performance of funds and portfolios. This is especially the case after the boom of Chinese e-fiance in recent years. As a result, almost all mutual funds report their daily returns to customers.

[7]Faff (2004) tests the FF model in Australian market but he does not directly use the Fama-French factors but use the proxies for the factors forming from the style indecies.

[8]Based on 'Guidelines for the Industry Classification of Listed Companies' by the China Securities Regulatory Commission, which can be referred to CSRC's website. In case any discrepancy exists between the Chinese and English context, the Chinese version shall prevail.

[9]Some daily data related to certain country markets can now be accessed online from the Kenneth R. French Data Library.

[10]For example, Kothari et al. (1995) argue that "There are at least three reasons for reexamining the risk-return relation using longer measurement-interval (than monthly interval) returns."

Thirdly, after the initial FF model was proposed by Fama and French (1992, 1993), subsequent researchers followed their portfolio reforming frequency to construct the Fama-French factors, that is, the annual frequency based on the annual financial reports of listed firms. This is feasible if only annual financial reports for most listed firms are available. Now that after January 2007 the quarterly financial reports of listed firms are introduced by the regulator of the Chinese market, it will be of interest to construct the Fama-French factors at the quarterly reforming frequency. In this study, we test the FF model in which the factors are constructed using firms' quarterly financial data. The quarterly financial reports update the company's operation performance more frequently, the investors can better adjust their investment decisions according to the updated information revealed by the financial data. The quarterly disclosure of financial status also forces the firms' operators to work harder to maintain the performance over the competitors throughout the whole year, because a fall in operation performance will be quickly shown in the next financial report, which will make the firm's stock prices decrease immediately within the whole year. Thus, the change in the disclosure frequency of financial reports could potentially influence the stock prices. Although the factor reforming frequency is higher, the conventional 6-month gap is kept to match the financial data to stock trading data when calculating the book-to-market ratio, because 6 months are necessary to wait for all the firms' financial data being widely available.

Fourthly, as far as we are aware, this study is the first to test the FF model in different market trends. The change of trend in the Chinese markets during the sample period in this study is impressive and provides the opportunity to examine the validity of the FF model in the falling and fluctuating trends. After testing the FF model in the whole sample period, we split the sample period into two sub-periods representing the downward and fluctuating trends, respectively. Although this splitting is arbitrary, the results suggest that the FF model performs well in both trends and slightly better in the downward trend. We attribute this slightly better performance in the downward market to the fact that in a tendency market, asset prices tend to move together. This is consistent with the findings of Morck et al. (2000), where they document stocks in emerging markets tend to produce high R^2 in a market model because emerging markets have synchronous stock price movements.

Lastly, comparing the performance of the FF model in different market trends raises another interesting question whether the regression coefficients of the FF model and the model itself are stable over the time, especially at the turning of the trend. Although our stability tests of the coefficients and the FF model indicate that both are not stable over the sample period, the instability seems to augment in the downward market trend even though it is not the case for every portfolio. So it could be considered that the stability of the FF model is somewhat dependent on the trend of the market and we should be more cautious of the validity of the FF model when the market crashes. This finding seems to conflict with the higher explanatory power of the FF model in the downward market. We provide a possible explanation in the concluding section.

The remaining parts of this study are organized as follows: Section 2.2 describes the data sources and sample period. Section 2.3 presents the construction of explanatory and dependent variables, and the regression models. Section 2.4 reports and analyses the results of regressions. Section 2.5 concludes with a summary of the results.

2.2 Data

In the US, data for the Fama-French factors have become increasingly standardized. For example, those data are publicly available at the Kenneth R. French Data Library[11]. However, to our knowledge, unlike in some developed countries, there is no similar database for the Fama-French factors publicly accessible in the Chinese market. Thus, to conduct this study, we must first construct the Fama-French-type factors and portfolios using publicly available data in the Chinese market. The method we choose to construct those factors is in line with Fama and French (1993).

[11] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html, accessed on 30 August 2014.

2.2.1 Data sources

The data to be employed in this study involve four types of data sets: the stock daily trading data, the firms' balance sheet data and capital stock structure data, and the risk-free rate. These data sets are all free and can be obtained from many sources. Specifically, in this study, the stock trading data is exported from a piece of stock software developed by a Chinese company named Great Wise. The balance sheet and capital stock structure data are downloaded from Sina Finance, the financial channel of a Chinese website named Sina with the URL: www.sina.com.cn .[12]

For each firm and stock index, the raw daily trading data consists of the trading date, opening price, highest price, lowest price, closing price, trading volume and value. The return on each stock or index is calculated by taking the natural logarithm of current closing price divided by previous trading day's closing price. The Total Owners/Shareholders' Equity from the balance sheet data set is used to determine the book value of each firm's equity. The Outstanding Shares – the total number of shares issued by a firm, multiplied by the closing price, is used to determine the market value of each firm's equity; while the Circulating Shares (free float) – the number of shares that can be publicly traded in the exchange, multiplied by the closing price, is used to determine the weight of corresponding stock in a value-weighted portfolio. Both Outstanding Shares and Circulating Shares are obtained from the capital stock structure data set.

The Shanghai Interbank Offered Rate (Shibor) plays the role of risk-free rate in this study and can be downloaded from the China Foreign Exchange Trade System & Nation Interbank Funding Center[13].

[12]There are many other sources, for example, similar stock software developed by other companies and difference websites, to acquire these data. Currently, these data for the firms listed on the Shanghai and Shenzhen stock exchanges are publicly available.

[13]URL: www.shibor.org. Both Chinese and English versions of this website are available. Shibor's Code of Conduct can be referred to http://www.shibor.org/shibor/web/html/sszz.html (Chinese version) or http://www.shibor.org/shibor/web/html/sszz_e.html (English version)

2.2.2 Sample period

Balance sheet data

Before 2006 inclusively, there was no quarterly disclosure requirement imposed on listed firms on both Shanghai and Shenzhen stock exchanges. After 30th January 2007, the China Securities Regulatory Commission (CSRC) requires that all listed firms should disclose their previous year's annual report, current year's first-quarter report, interim report, and third-quarter report by the end of April, April, August, and October, respectively, and that the disclosure of the first-quarter report should not be earlier than the disclosure of the last year's annual report.[14] As a result of this requirement, the quarterly accounting data have become widely available after 2007.

The quarterly book values of total owners/shareholders' equity can be firstly obtained from the firm's 2006 annual report up to the 2013 first-quarter report. Although it is required by the CSRC, some firms might not disclose their financial statements in time. To ensure that the accounting variables are known before the returns they are used to explain, as in Fama and French (1992), we match the quarterly equity data for the 4th Quarter of 2006 up to the 1st Quarter of 2013 with the stock trading data for the end of June of 2007 up to September of 2013. The 6-month (minimum) gap is as conservative as in Fama and French (1993), e.g., the first BE/ME using book equity in the 4th Quarter of 2006 matched with the first size (market value) in 2nd Quarter of 2007.

Stock daily trading data

Due to the 6-month gap mentioned above, the whole sample period of interest in the time-series regression starts from 2nd July 2007 (instead of the first trading day in 2007) to 31st December 2013, with 1582 trading days in total. Figure 2.1 presents the trends of the Chinese market during the whole sample period. This period witnessed the historical peak of the Chinese stock market with the highest point of the Shanghai Stock Exchange Composite Index

[14] http://www.csrc.gov.cn/pub/zjhpublic/zjh/200804/t20080418_14481.htm

at around 6,124 on 16th Oct 2007 and with the highest point of the Shenzhen Stock Exchange Component Index at around 19,600 on 10th Oct 2007. After that, the Chinese stock market crashed and the two indexes dropped dramatically and stayed at around 2000 and 8000 points, respectively. This significant turning of trends makes it interesting to examine the performance of the Fama-French-type models in different tendencies of the market.

In addition to the test of the whole period, we split the sample into two sub-periods. The first sub-period starts from 8th October 2007 to 31st December 2008, with 306 trading days. Note that, we exclude the observations in the third quarter of 2007 just to make the first sub-period a typical downward market. From Figure 2.1 we can observe a significant increase in both indexes in the third quarter of 2007. The second sub-period contains the rest of observations which starts from 5th January 2009 to 31st December 2013, and there are 1211 trading days. This period can be considered as a fluctuating market. For comparison, we also present the trends of the US market during the same period in Figure 2.2, where the US market witnessed a V-shape change in market trends.

2.2.3 Alleviating the survivor bias

As time passes, there would be new stocks listed on and old stocks delisted out of the A-Share market. Thus, the survivor bias becomes a concern. To alleviate survivor bias and reflect the overall market performance, at the beginning of sample period (i.e., 2nd July 2007) all the listed non-financial firms' A-share stocks are included in the sample. A newly listed stock will be added in the explained portfolio and the mimicking portfolio at the time of re-forming right after the day when it is first listed. A delisted stock will be included in the portfolios right before the next re-forming day after the stock is delisted, i.e., the last day of the quarter in which the stock is delisted, and then it will be removed for the next portfolio reforming.

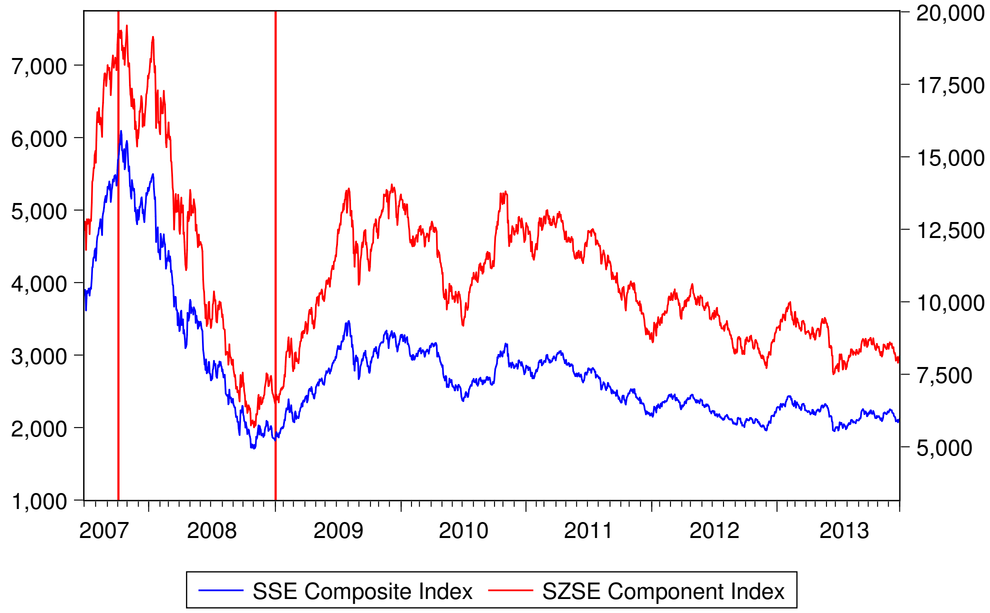

Figure 2.1: Changes in trend of China's stock market from 2nd Jul 2007 to 31st Dec 2013

Note: The SSE Composite Index is plotted against the left y-axis, while the SZSE Component Index is against the right y-axis. The period between the two red vertical lines represents a downward market, while it represents a fluctuating market after the second red line.
Data Source: Shanghai DZH Limited

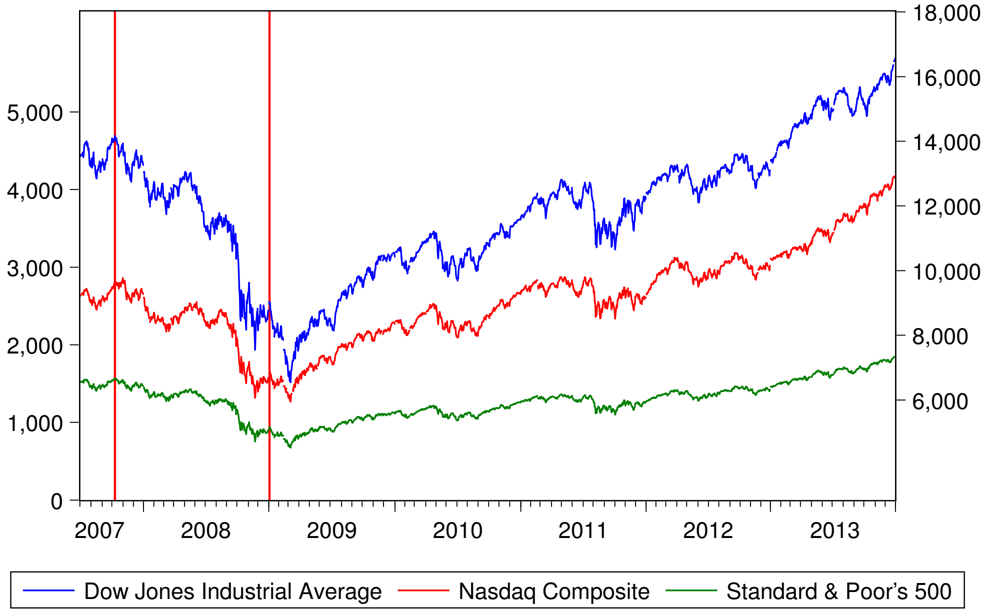

Figure 2.2: Changes in trend of US stock market from 2nd Jul 2007 to 31st Dec 2013

Note: The Dow Jones Industrial Average Index is plotted against the left y-axis, while the Nasdaq Composite and Standard &Poor's 500 Indexes are against the right y-axis. The period between the two red vertical lines is when the Chinese market experienced a downward trend, while a fluctuating market existed in China after the second red line.

Data Source: Shanghai DZH Limited

2.3 Regression Models

In this section we first describe how we construct the explanatory variables including the market portfolio excess returns and four Fama-French-type factors – the Size, BE/ME, Momentum and Liquidity factors – in Section 2.3.1 and the dependent variables including 25 size-and-BE/ME sorted portfolios in Section 2.3.2. The method chosen to construct the factors and the 25 portfolios is in line with Fama and French (1993) except that several specific features of the Chinese market are considered. Three regression models will be explored in Section 2.3.3 and the regression results will be reported in Section 2.4. The first two models are the CAPM and the Fama-French three-factor model. Constructed as the Fama-French factors, two additional factors representing momentum and liquidity are added to make the third model, i.e., a five-factor model. These three models will be examined using the Chinese market data in the whole sample period. The results suggest that the FF model performs much better than the CAPM, and that adding the two additional factors seems not to improve the explanatory power of the five-factor model, compared to the FF model.

2.3.1 Explanatory variables

Before running the regression analysis, we need to construct mimicking portfolios for the explanatory variables including the market excess return and other four risk factors, as well as the portfolios of dependent variables. Except that the Shibor can be accessed immediately, the processes of construction of the relevant portfolios are described in the rest of this section.

Single stock return

Because the stock return data are not immediately available from our data sets, we need to calculate the stock returns first, and then the portfolio returns. The return on each stock on each trading day is calculated as the natural logarithm of the closing price on that day divided by the closing price on the previous trading day, multiplied by 100. The return on each portfolio is the weighted average of all stock returns in that portfolio, and all the returns in the regressions are not annualized but in percentage.

Market portfolio return

The market return is the value-weighted return on the market portfolio of all sample firms during the sample period. The value of each stock is the market value calculated using one-day lagged closing price times the circulating shares (free float) of that stock on each trading day. For the non-trading days, the closing prices are recorded as zero leading to a zero value on those days. Thus, the market portfolio can be viewed as a portfolio with a daily reforming frequency.

Risk-free rate

In many other studies, the one-month T-bill rate is usually used as the proxy for the risk-free rate at both monthly and daily level. For example, the Kenneth R. French Data Library provides the daily T-bill return which "is the simple daily rate that, over the number of trading days in the month, compounds to 1-month T-bill rate from Ibbotson and Associates, Inc." Unlike those studies that use the short-term government bill rate as the proxy for the risk-free rate, in this study, the over-night Shanghai Interbank Offered Rate (Shibor) is chosen as the proxy for the risk-free rate.[15] The over-night Shibor is available by the noon of every working day.

Building blocks for remaining factors

Because the Chinese market is quite different from most of the developed countries, some special treatments should be considered when constructing the portfolios and factors. Specifically, most of the listed firms on both SSE and SZSE issue both circulating and non-circulating (restricted) shares. The circulating shares are those that can be traded by investors in the secondary market, while the non-circulating shares cannot be sold by shareholders (unless approved by the authorizer), thus cannot be traded in the secondary market. The non-circulating

[15] The Shibor is quoted on actual/360 day basis (%, T+0).

shares are mostly state-owned and account for a considerable portion that should not be neglected. When calculating the market value of each firm, the outstanding shares, including the non-circulating shares, are applied; when considering the value weight of each stock in the corresponding portfolio and liquidity factor (the turnover rate in this study), only the circulating shares are considered.[16]

The size, or market value, of a firm is defined as the market value of a firm, which is calculated as the closing price times the firm's outstanding shares (i.e., total number of shares issued, including non-circulating shares). Both closing price and outstanding shares have an influence on the value of a firm's size. Because the closing price fluctuates every trading day, it makes the size of a firm fluctuate as well. To mitigate this fluctuation caused by price, we sort the stocks based on the five-trading-day lagged average size of a firm at the time of reforming the mimicking portfolios. On last day of each quarter from the 2nd Quarter of 2007 to the 3rd Quarter of 2013, all A-share non-financial stocks are ranked by the five-trading-day lagged average size. The median average size is then used as the breakpoint to split all A-share stocks into two groups, i.e., small and big (S and B) groups.

Book-to-market equity is defined as the book value of common equity (BE) divided by the market value of equity. The book equity is the book value of total owners/stockholders' equity.[17] Book-to-market equity (BE/ME) is then the book common equity of a firm divided by the five-trading-day lagged average market equity at the end of the corresponding quarter.[18] All A-share stocks are also broken into three book-to-market equity groups based on the breakpoints for the bottom 30% (Low), middle 40% (Medium), and top 30% (High) of the ranked values of BE/ME.

The construction of SMB and HML is in line with Fama and French (1993). Six portfolios are formed from the sorts of stocks on market equity (ME) and book-to-market equity

[16]Some special features of Chinese stock market can be found in Xu and Zhang (2014).

[17]So far, no firm in China issues preferred stock, so there is no need to adjust the BE for preferred stock as Fama and French (1993) do. Additionally, no adjustment for the deferred taxes and investment tax credit is conducted just for simplicity.

[18]Unlike Fama and French (1993), the negative and zero BE firms are included here to represent the overall market performance, but only positive BE firms are involved in forming the 25 portfolios later as the dependent variables.

(BE/ME) to mimic the underlying risk factors in returns related to size and book-to-market equity. Recall that the conservative six-month gap is used to ensure that the accounting variables are known before the returns they are used to explain, here the first BE/ME in the 4th Quarter of 2006 is matched with the first Size factor in the 2nd Quarter of 2007, while the last BE/ME in the 1st Quarter of 2013 is matched with the last Size factor in the 3rd Quarter of 2013. Thus, six portfolios (S/L, S/M, S/H, B/L, B/M, and B/H) are quarterly constructed from the six-month lagged intersections of the two ME and the three BE/ME groups. Then the daily value-weighted returns on the six portfolios are calculated from first to last trading day of each quarter for the whole sample period, and the portfolios are reformed for the next quarter on the last day of each quarter. Notice that, the value used to calculate the weight is each firm's market value of circulating shares, rather than the outstanding shares, calculated as the closing price times the circulating shares.

Size factor – the portfolio SMB (small minus big), meant to mimic the risk factor in returns related to size, is the difference, on each day, between the simple average of the returns on the three small-stock portfolios (S/L, S/M, and S/H) and the simple average of the returns on the three big-stock portfolios (B/L, B/M, and B/H). Thus, SMB is the difference between the returns on small- and big-stock portfolios with about the same weighted-average book-to-market equity. This difference should be largely free of the influence of BE/ME, focusing instead on the different return behaviors of small and big stocks.

BE/ME factor – the portfolio HML (high minus low), meant to mimic the risk factor in returns related to book-to-market equity, is defined similarly. HML is the difference, on each day, between the simple average of the returns on the two high-BE/ME portfolios (S/H and B/H) and the simple average of the returns on the two low-BE/ME portfolios (S/L and B/L). The two components of HML are returns on high- and low-BE/ME portfolios with about the same weighted-average size. Thus, the difference between the two returns should be largely free of the size factor in returns, focusing instead on the different return behaviors of high- and low-BE/ME firms.

The Momentum and Liquidity (Turnover) factors are constructed at the quarterly level.[19] On every last day of each quarter, the stocks are ranked by the return and the average daily turnover rate for the past quarter, respectively. The past quarter return is just the sum of daily returns in the quarter because of the natural logarithm return. The daily turnover rate is the trading volume divided by the circulating shares. Then both top 30% and bottom 30% stock portfolios are constructed for momentum and turnover factors, respectively.

Momentum – the momentum factor, MOM, is the difference between the equal-weight average of returns on the top 30% stocks and the equal-weight average of returns on bottom 30% stocks sorted by the past quarter return.

Turnover – the turnover factor, TO, is the difference between the equal-weight average of returns on the top 30% stocks and the equal-weight average of returns on bottom 30% stocks sorted by the average turnover rate in the past quarter.

The top panel of Table 2.1 presents the descriptive statistics for the explanatory variables, while the bottom panel shows the correlation between the variables. Several points might be interesting to notice. As a proxy for the risk-free rate, the Shibor is considerably small compared to other variables. It is significantly different from zero, because China's short-term interest rate has never dropped to the zero lower bound. Both return and excess return of market portfolio is negative, and only market excess return is statistically significant at the 10% level. In many other studies, the market excess return is generally positive. For example, in the recent study on the Chinese market by Chen et al. (2015), the market excess return is positive but insignificant. This is not surprising because starting from July 2007, our sample witnesses a historical crash in the Chinese market and then the market fluctuating at relatively low level, while their sample period starts from the very beginning of the Chinese stock market, i.e., July 1997, when the market just started to grow. The size premium (SMB) is positive and highly significant, while the value premium (HML) is positive but insignificant

[19] Unlike in Carhart (1997) where he rolls over the construction of momentum portfolios, that is, he puts focus on the monthly return and the portfolios are re-formed monthly based on the eleven-month returns lagged one month, we focus on the daily return but construct the portfolio with a quarterly frequency as the frequency of construction of SMB and HML portfolios.

Table 2.1: Descriptive statistics and correlation of explanatory variables: 1582 observations

	MR	Shibor	Ex. MR	SMB	HML	MOM	TO
Panel A: Descriptive statistics							
Mean	-0.0753	0.0066***	-0.0819*	0.0618***	0.0246	-0.0580***	-0.0206
t-statistic	-1.5212	80.8095	-1.6548	3.5788	1.6451	-3.5148	-1.0470
Median	0.0364	0.0062	0.0292	0.1167	0.0115	-0.0280	0.0709
Maximum	9.2365	0.0373	9.2282	2.5142	2.9329	2.5626	3.5755
Minimum	-8.7073	0.0022	-8.7168	-3.1094	-2.5783	-4.2016	-3.5717
Std. Dev.	1.9680	0.0033	1.9682	0.6873	0.5944	0.6562	0.7844
Panel B: Correlation							
MR	1						
Shibor	-0.0619	1					
Ex. MR	1.0000	-0.0635	1				
SMB	0.1064	-0.0862	0.1065	1			
HML	0.3402	0.0090	0.3402	0.0056	1		
MOM	-0.2804	0.0324	-0.2805	0.0413	-0.1129	1	
TO	0.7572	-0.0633	0.7572	0.4978	0.2746	-0.1523	1

*** Significance at the 1% level
** Significance at the 5% level
* Significance at the 10% level
Note: MR represents the market return. Shibor is the Shanghai Interbank Offered Rate. Ex.MR is the excess market return over the Shibor. SMB and HML are the Fama-French factors, while MOM and TO are proxies for the momentum and liquidity factors.

with the t-statistic around the 10% critical value. Thus, we find a strong size effect but a weak value premium. This finding is consistent with some other studies, e.g., Wang and Xu (2004) and Chen et al. (2015), but contradicts with most of the existing studies on the Chinese market, e.g., Chen et al. (2010), Cakici et al. (2015) and Carpenter et al. (2015), where both strong size and value effects are found. The different findings about the significance of value premium in the Chinese market might stem from the different choices of sample periods. Different sample periods can lead to different magnitudes and significance of factor premium. Chen et al. (2015) also document that the market, size and value factors are insignificant in the US during the period from July 1997 to December 2013. Even the signs of factor premiums can change in different periods and this is called effect reversal. For example, the reversal of the size effect has been found for some developed markets, such as the US, UK, Australia and Japan in a wave of studies, e.g., Gompers and Metrick (1998), Dimson and Marsh (1998), Gustafson and Miller (1999), Faff (2004) and Pham (2007). This suggests that even if we find a weak

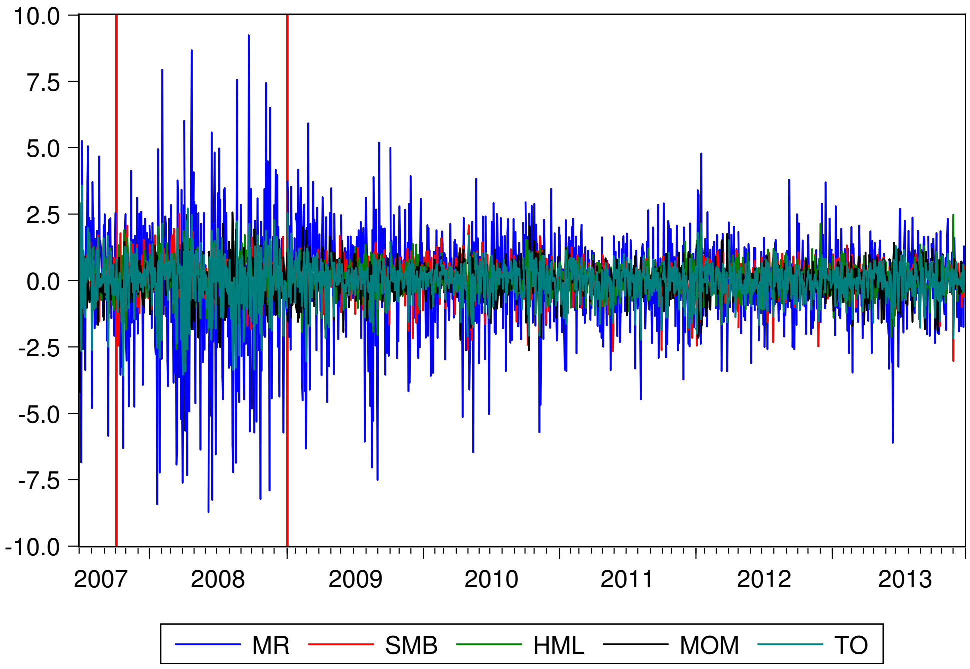

Figure 2.3: Time-series changes in the premiums on pricing factors

Note: This figure plots the time-series changes in the premiums on the five pricing factors. The period between the two red vertical lines represents a downward market, while it represents a fluctuating market after the second red line.

factor premium in our sample, we cannot just drop it from the pricing model, especially we still find generally significant factor betas in time-series regression. The value factor is still an important factor in pricing assets, unless we can find another factor that can completely replace the role of value factor or can better explain equity returns.[20] The correlation between SMB and HML is really low with a coefficient of 0.56%. Thus, the two factors, by construction, are largely orthogonal to each other. The negative premiums on the momentum and turnover factors suggests that the trading based on the two strategies in the Chinese market would fail during our sample period.

Figure 2.3 plots the time-series changes in the premiums on the five pricing factors within the whole sample period, within which the sub-period between the two red vertical lines represents the downward market and the next sub-period after the second red vertical line represents

[20]Fama and French (2015a, 2016) recently find that the two new profitability and investment factors could challenge the role of the value factor, but the similar research is just starting and the evidence is still limit.

a fluctuating market. From this figure we can see that the factors generally present significant volatility during the downward period. Besides, when the market crashes, it is not surprising that the market factor tend to present more downside volatility. Due to this significant difference in the factors' behavior in different trends of market, we also provide the sub-sample descriptive statistics for the five factors in Table 2.2, where the top panel reports the descriptive statistics and correlations of the factors for the downward market and the bottom panel for the fluctuating market. This table confirms the time-series patterns of the factors. The standard deviations of all factors are much higher in the downward market. The market premium in the crash is significantly negative with a value of -0.4416 which is much lower than in the transition period. The magnitudes of other four factors do not present such significant difference in different trends of market. However, the size premium is significantly positive in the transition period, but insignificant in the crash. The value premium presents a different situation that it is significantly positive in the crash, but insignificant in the transition period. The momentum premium is significantly negative in the crash but insignificant in the other period. The turnover premium is negative and insignificant in both sub-periods. The different significance and magnitudes of the factor premiums suggest that the risk premiums are not necessarily preservative over time, and they should be related to different market conditions. The findings here provide fund managers who based their strategies on the risk factors a hint that sticking on a risk factor-based strategy cannot guarantee them a successful investment any time. Thus, they should adjust the investment position according to the changes in the market conditions. The last interesting finding is that the correlations between each pair of factors are also affect by different market conditions. For example, the correlation of the size and value factors is 0.1610 in the downward market, much higher than in the full sample, i.e., 0.0056. It changes to -0.0837 in the fluctuating market.

2.3.2 Dependent variables

The dependent variables in time-series regressions are excess returns on the 25 portfolios over the Shibor, while the 25 portfolios are formed by sorting on size and book-to-market equity. The 25 size-BE/ME portfolios are constructed much like the six size-BE/ME portfolios

Table 2.2: Sub-sample descriptive statistics for the five factors

	Ex.MR	SMB	HML	MOM	TO
Panel A: Summary statistics in the downward market					
Mean	-0.4416**	0.0690	0.0726*	-0.0668	-0.0621
t-statistic	-2.5308	1.3769	1.6696	-1.4691	-0.9345
Median	-0.3204	0.1785	0.0831	-0.0818	0.1161
Maximum	9.2282	2.5142	2.1503	2.5626	2.7170
Minimum	-8.7168	-3.1094	-2.5783	-2.5022	-3.5717
Std. Dev.	3.0521	0.8767	0.7611	0.7958	1.1617
Observations	306	306	306	306	306
Ex.MR	1				
SMB	0.0177	1			
HML	0.4808	0.1610	1		
M	-0.5740	-0.1257	-0.1590	1	
TO	0.8102	0.4218	0.5428	-0.4812	1
Panel B: Summary statistics in the fluctuating market					
Mean	-0.0226	0.0633***	0.0107	-0.0445**	-0.0192
t-statistic	-0.5085	3.4862	0.7099	-2.5657	-1.0649
Median	0.0354	0.1189	-0.0006	-0.0163	0.0668
Maximum	5.9183	2.0770	2.4691	2.0387	2.5418
Minimum	-7.5111	-3.0284	-2.1190	-2.6269	-3.0443
Std. Dev.	1.5436	0.6320	0.5241	0.6032	0.6283
Observations	1211	1211	1211	1211	1211
Ex.MR	1				
SMB	0.1874	1			
HML	0.2260	-0.0837	1		
M	-0.0952	0.1570	-0.0224	1	
TO	0.7160	0.5673	0.0152	0.1272	1

*** Significance at the 1% level
** Significance at the 5% level
* Significance at the 10% level
Note: Ex.MR is the excess market return. SMB and HML are the Fama-French factors, while MOM and TO are proxies for the momentum and liquidity factors. Panel A summarizes the descriptive statistics and correlation in the downward market, and Panel B summarizes those statistics in the fluctuating market.

discussed earlier. Here, we exclude the non-positive book equity stocks. On the last day of each quarter, we sort the stocks by size and book-to-market equity, independently. For the size sort, ME is still measured as the average market equity in the last five trading days of the previous quarter. For the book-to-market sort, where six-month gap employed, ME is the market equity and BE is the book common equity, both measured two-quarter earlier than the previous quarter. We use the breakpoints for ME and BE/ME to allocate all A-share stocks to five size quintiles and five book-to-market quintiles. Thus, 25 portfolios are constructed from the intersections of the size and BE/ME quintiles. Then, we calculate the value-weighted daily returns on the portfolios from 2nd July 2007 to 31st December 2013 with a quarterly reforming frequency during the sample period. The returns on these 25 portfolios for 2nd July 2007 to 31st December 2013 are the dependent variables in time-series regressions. Table 2.3 presents the descriptive statistics for the dependent variables (the 25 portfolio returns). From the table, we can find that except for one portfolio average return being positive, all other portfolios achieve negative average returns in our sample. Considering that during the whole sample period the Chinese market crashed from historical peak and fluctuates at lower level, the negative portfolios returns are not surprising consequence.

2.3.3 The models

The basic models to be examined in the following section are described here. The first model is the CAPM, which is the regression of portfolio excess return on the market excess return.

$$R_{it} - R_{ft} = \alpha_i + \beta_{iM}(R_{Mt} - R_f) + \varepsilon_{it},\ t = 1, \ldots, T \tag{2.1}$$

where R_{it} is the portfolio return to be explained in the regression, R_{Mt} is the market portfolio return, and R_f is the risk-free rate. The CAPM is used as the benchmark to examine the performance of other models.

Adding the size and book-to-market equity factors (SMB and HML) into the CAPM yields the second model, i.e., the FF model, which is the regression of portfolio return on the market return and the size and book-to-market equity factors:

Table 2.3: Descriptive statistics of dependent variables (25 portfolio returns)

Size quintile	Book-to-market equity (BE/ME) quintiles									
	Mean					Standard Deviation				
	Low	2	3	4	High	Low	2	3	4	High
Small	-0.0027	-0.0002	-0.0022	0.0114	-0.0003	1.9827	2.2025	2.2255	2.2689	2.2675
2	-0.0595	-0.0215	-0.0158	-0.0070	-0.0165	2.0632	2.2204	2.2855	2.2835	2.2906
3	-0.0514	-0.0476	-0.0433	-0.0332	-0.0434	2.1002	2.2247	2.2812	2.3208	2.2751
4	-0.0906	-0.0702	-0.0687	-0.0578	-0.0608	2.0474	2.1915	2.2531	2.3107	2.2256
Big	-0.0967	-0.1294	-0.0882	-0.0919	-0.0513	1.8997	2.1725	2.2119	2.0696	1.9865
	Maximum					Minimum				
Small	8.0684	8.9219	8.939	9.1154	8.9044	-8.0830	-11.1445	-9.1501	-9.6729	-9.5572
2	7.9284	8.9816	8.7535	9.4327	9.2218	-8.3131	-9.0638	-9.7495	-9.5655	-10.0210
3	8.7625	9.5379	9.0249	9.5265	9.3925	-9.5973	-9.4138	-9.4449	-12.4466	-10.9211
4	9.3742	8.8806	9.2674	9.1924	9.3871	-9.1217	-8.6635	-9.7255	-11.3764	-10.2319
Big	9.3080	9.4223	9.2664	9.1430	9.5437	-9.1402	-9.3983	-9.7512	-9.0790	-11.1820

This table presents some descriptive statistics for returns on the 25 size-BE/ME sorted portfolios within the whole sample period from 2nd July 2007 to 31st December 2013.

$$R_{it} - R_{ft} = \alpha_i + \beta_{iM}(R_{Mt} - R_f) + \beta_{iSMB}SMB_t + \beta_{iHML}HML_t + \varepsilon_{it},\ t = 1, \ldots, T \tag{2.2}$$

where R_{it}, R_{Mt} and R_f have the same meanings as in regression (2.1), while SMB_t represents the size factor and HML_t the BE/ME factor.

The other models include additional factor(s) – the momentum and liquidity factors. It is the regression of portfolio return on the market return, size, book-to-market equity, momentum and liquidity (turnover) factors:

$$\begin{aligned} R_{it} - R_{ft} = \quad & \alpha_i + \beta_{iM}(R_{Mt} - R_f) + \beta_{iSMB}SMB_t + \beta_{iHML}HML_t \\ & + \beta_{iMOM}MOM_t + \beta_{iTO}TO_t + \varepsilon_{it} \end{aligned} ,\ t = 1, \ldots, T \tag{2.3}$$

where MOM_t is the momentum factor and TO_t is the liquidity (turnover) factor, both of which are constructed from our data sets.

After the time-series regression tests, we also perform the Gibbons et al. (1989) (GRS) F-tests to formally test the null hypothesis that the intercepts for the 25 portfolios are jointly zero in above models. The GRS test statistic is:

$$\left(\frac{T - N - K}{N}\right) \left[\frac{\hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha}}{1 + E(f)'Var(f)^{-1}E(f)}\right] \sim F\left(N, T - N - K\right) \tag{2.4}$$

where T is the number of observations in the time-series; N is the number of assets; K is the number of factors; $E_T(f)$ is the sample mean of the factors; $Var_T(f)$ is their estimated variance-covariance matrix; $\hat{\alpha}$ is a $N \times 1$ vector of estimated intercepts from the individual time-series regressions; and $\hat{\Sigma}$ is the estimated variance-covariance matrix of the intercepts. To derive the GRS statistic, normally distributed errors are assumed. If the null hypothesis holds, then the GRS statistic should be equal to zero.

In addition to the above time-series regressions, we also conduct the tests on the models based on the two-stage approach of Fama and MacBeth (1973). Because our 25 portfolios provide large cross-sectional variation in average returns, for simplicity, we do not re-rank the stocks by their pre-ranking betas to construct new portfolios as in Fama and MacBeth (1973). Instead, we directly use the slope coefficients obtained from the above time-series regression as the first-stage estimates of factor betas for the individual portfolios. Then, taking the FF model as an example, in the second stage, we run the following cross-sectional regression at each trading day t:

$$R_{it} - R_{ft} = \gamma_{0,t} + \gamma_{M,t}\hat{\beta}_{iM} + \gamma_{SMB,t}\hat{\beta}_{iSMB} + \gamma_{HML,t}\hat{\beta}_{iHML} + \varepsilon_{it},\ i = 1, \ldots, N \tag{2.5}$$

Then the factor premiums are estimated as the time-series average of $\hat{\gamma}_{M,t}$, $\hat{\gamma}_{SMB,t}$ and $\hat{\gamma}_{HML,t}$. That is,

$$\begin{aligned} \hat{\gamma}_M &= \frac{1}{T}\sum_{t=1}^{T}\hat{\gamma}_{M,t} \\ \hat{\gamma}_{SMB} &= \frac{1}{T}\sum_{t=1}^{T}\hat{\gamma}_{SMB,t} \\ \hat{\gamma}_{HML} &= \frac{1}{T}\sum_{t=1}^{T}\hat{\gamma}_{HML,t} \end{aligned} \tag{2.6}$$

2.4 Empirical Tests

This section performs the regressions and reports the results. In Section 2.4.1, the regressions of the models (2.1), (2.2) and (2.3) are performed respectively within the whole sample period from 2nd July 2007 to 31st December 2013, and there are 1582 trading days in total. Sections 2.4.2 and 2.4.3 perform the subsample regressions and report the results in downward and fluctuating trends, respectively, followed by the comparison of these results. To make the comparison of the parameters for big and small size portfolios of each BE/ME quintile portfolio and the comparison of the parameters for high and low BE/ME portfolios of each size quintile much clearer, we also report the S-B spread and H-L spread for each parameter along with the corresponding estimated coefficients. The S-B spread is the difference of estimators between the small size portfolio and big size portfolio for each quintile of book-to-market equity. Similarly, the H-L spread reports the difference of estimators between the high book-to-market equity portfolio and low book-to-market equity portfolio for each quintile of size. In Section 2.4.4, we conduct the GRS F-tests on whether the intercepts of the models are jointly zero and report the results in Table 2.11. We also run the Fama-MacBeth regressions for the models in Section 2.4.5. We then make a rough investigation in the stability of estimated coefficients of the models in Section 2.4.6.

2.4.1 Time-series regression in the whole sample period

Tables 2.4-2.6 show the results of the regressions of the models (2.1), (2.2) and (2.3) for the whole sample period from 2nd July 2007 to 31st December 2013 with 1582 trading days.

The first regression performed is for the CAPM, which is treated as the benchmark for the other two models. The regression results, including the estimated coefficients and the corresponding t-statistics, the adjusted R^2 and the standard error of the regression, are summarized in Table 2.4.

There are 15 estimated intercept terms statistically significant at least at the 10% level, 10 of which are significant at the 1% level. From Table 2.4 we can see that the small size and high BE/ME portfolios appear to have the significant intercept terms, as more asterisks show up at the top right panel of α. Most of the intercept terms are positive, while five of them are less than zero. The big size and low BE/ME portfolios tend to have negative intercepts because the negative intercepts show up at the bottom left panel. The positive S-B spreads and H-L spreads suggest that the small size and high BE/ME portfolios tend to have larger intercept terms. This result is consistent with that the small size and high BE/ME firms tend to have higher return than the big size and low BE/ME firms. The average of the 25 intercepts is 0.0336 with a range of -0.0492 to 0.1271.

The slope coefficients of the market excess return for the 25 portfolios are all highly significant at the 1% level, with the smallest t-statistic of 68.3912. The values of the estimated coefficients are around one, and the average is 1.0481 with a range of 0.9001 to 1.1348. Two of the S-B spreads are negative while the other three are positive. Although the S-B spread for the lowest BE/ME quintile is positive, it is relatively small compared to the other two positive S-B spreads. The slopes increase first and then decrease along with the small to big size for each BE/ME quintile. All the H-L spreads are positive and the small size quintile has the biggest value. The values decrease along with each size quintile. The positive H-L spreads suggest that the market excess return would have a greater impact on the excess returns on the high BE/ME portfolios.

The adjusted R^2s of the regressions range from 0.8118 to 0.9341. If we compare the results in Table 2.4 with those results in other literature where monthly returns are exploited, for instance, Table 4 of Fama and French (1993), we can find that the slope coefficients of the market excess return are closer to one and the adjusted R^2s are relatively higher in our study here.[21] It suggests that the performance of the CAPM is better when we examine it at the daily return level. Although the adjusted R^2s are high, there is still a portion that cannot be explained by the CAPM. It leaves the job for the FF model to explain this unexplained portion.

[21] In Table 4 of Fama and French (1993), they report that the range of slope coefficients of the CAPM with monthly returns is from 0.84 to 1.42, while the range of adjusted R^2s is from 0.61 to 0.92. Both ranges are bigger than those reported here.

Table 2.4: Regressions of 25 portfolio daily excess returns on market daily excess returns: From 2nd July 2007 to 31st December 2013, 1582 trading days.

$$R_t - R_f = \alpha + \beta_M (R_{Mt} - R_f) + \varepsilon_t$$

Dependent variable: Returns on 25 stock portfolios formed on Size and Book-to-Market Equity

Size	Book-to-Market Equity (BE/ME) quintiles											
quintiles	Low	2	3	4	High	H-L Spread	Low	2	3	4	High	H-L Spread
	α						t-statistics of α					
Small	0.0650**	0.0780***	0.0774***	0.0925***	0.0805***	0.0154	2.4752	3.2133	3.4492	4.0171	3.4440	0.4386
2	0.0130	0.0583***	0.0670***	0.0763***	0.0675***	0.0545*	0.5446	2.6706	3.0678	3.7624	3.4163	1.7552
3	0.0226	0.0330	0.0406**	0.0527***	0.0408**	0.0183	0.9447	1.6151	2.1430	2.9741	2.4361	0.6259
4	-0.0173	0.0098	0.0152	0.0285*	0.0220	0.0393*	-0.8623	0.5052	1.0174	1.7756	1.5213	1.5897
Big	-0.0296*	-0.0492***	-0.0061	-0.0167	0.0195	0.0491*	-1.6656	-3.2581	-0.4175	-1.0177	1.0459	1.9062
S-B Spread	0.0946***	0.1271***	0.0835***	0.1092***	0.0610**		2.9834	4.4495	3.1167	3.863	2.0403	
	β_M						t-statistics of β_M					
Small	0.9079***	1.0355***	1.0530***	1.0715***	1.0670***	0.1591***	69.8903	75.2589	82.8509	80.1218	81.5768	8.6331
2	0.9668***	1.0544***	1.0916***	1.0974***	1.1064***	0.1396***	75.9377	87.6488	88.4645	91.4059	95.3421	8.1021
3	0.9839***	1.0658***	1.1050***	1.1294***	1.1091***	0.1252***	70.6141	88.8584	100.0187	98.9766	107.5672	7.2244
4	0.9764***	1.0581***	1.1057***	1.1348***	1.0920***	0.1156***	81.8683	100.9322	129.8757	121.6015	95.5952	6.9981
Big	0.9001***	1.0610***	1.0838***	1.0000***	0.9456***	0.0455***	88.0442	96.6040	104.0833	92.6463	68.3912	2.6447
S-B Spread	0.0078	-0.0254	-0.0308*	0.0715***	0.1214***		0.4706	-1.4452	-1.8741	4.1632	6.3811	

Table 2.4: (Continued)

$$R_t - R_f = \alpha + \beta_M (R_{Mt} - R_f) + \varepsilon_t$$

Dependent variable: Returns on 25 stock portfolios formed on Size and Book-to-Market Equity

Size	Book-to-Market Equity (BE/ME) quintiles											
quintiles	Low	2	3	4	High	H-L Spread	Low	2	3	4	High	H-L Spread
	Adjusted R^2						Standard Error of regression					
Small	0.8118	0.8560	0.8670	0.8637	0.8575		0.8602	0.8359	0.8117	0.8378	0.8559	
2	0.8503	0.8733	0.8835	0.8945	0.9035		0.7984	0.7904	0.7802	0.7417	0.7116	
3	0.8500	0.8888	0.9087	0.9172	0.9205		0.8136	0.7419	0.6895	0.6680	0.6416	
4	0.8808	0.9029	0.9328	0.9341	0.9323		0.707	0.6831	0.5843	0.5930	0.5790	
Big	0.8694	0.9238	0.9299	0.9041	0.8774		0.6865	0.5997	0.5858	0.6410	0.6956	

*** Significance at the 1% level
** Significance at the 5% level
* Significance at the 10% level

Note: The t-statistics reported are calculated from standard errors corrected for serial correlation and White (1980) heteroskedasticity using the Newey-West (1987) estimator. The rule for choosing the lag parameter q is that $q = 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression (see Stock and Watson, 2007, page 607).

The S-B Spread reports the difference of estimations between the small size portfolio and big size portfolio for each quintile of BE/ME. Similarly, the H-L Spread reports the difference of estimations between the high BE/ME portfolio and low BE/ME portfolio for each quintile of size.

Adding the factors related to size and BE/ME into the CAPM, we examine the FF model in Table 2.5. We can see that the intercept terms as a whole are not as statistically significant as those of the CAPM in Table 2.4. There are only 7 estimated intercept terms statistically significant at least at the 10% level, among which only one of them is significant at the 1% level. The number of positive intercept terms decreases to 13, while 12 of intercept terms are negative. The middle three S-B spreads are positive while the left and right S-B spreads are negative. Three H-L spreads are greater than zero while the other two are not. The average of the 25 intercepts is -0.0014 with a range of -0.0304 to 0.0249. Three findings here are that the significance of estimated intercepts decreases, that the number of positive intercepts is equal to the number of negative ones, and that the estimated values of intercepts are closer to zero. These findings provide the first evidence for that the FF model performs better than the CAPM in explaining the cross section of excess returns.

The slope coefficients of the market excess return for the 25 portfolios are still all highly significant at the 1% level, with the smallest t-statistic of 92.6099 increased from 68.3912. The values of these slopes are around one, and the average is 1.0303 with a range of 0.8910 to 1.1031. Compared to the estimated slopes of the market excess return of the CAPM in Table 2.4, the FF model seems to have the market excess return slopes closer to one although some of them are a bit far away from one. The middle quintiles of both size and BE/ME portfolios tend to have greater slopes although it is not the case for some portfolios. Three S-B spreads for low BE/ME quintiles are less than zero but the other two are positive. The H-L spreads decrease along with the size quintiles from 0.0959 to -0.1077. The top four H-L spreads are positive and less than those in Table 2.4.

The slope coefficients of the size factor are all statistically significant at the 1% level, with the smallest absolute value of t-statistic of -5.6001. The average of these slopes is 0.5532 with a range of -0.4335 to 1.0064. The slopes of the biggest size portfolios are all negative for each BE/ME quintile while the rest slopes are positive. Thus, a negative relationship between the returns on the biggest size portfolios for each BE/ME quintile and the returns on the mimicking portfolios sorted by size is suggested. The value of slopes for each BE/ME quintile decreases along with the size quintiles from small to big. The smaller size portfolios have greater slopes related to the size factor while the bigger size portfolios have smaller or even negative slopes. Thus, all the S-B spreads are greater than one. Up to here, the findings are consistent with those reported in Table 6 of Fama and French (1993). The two H-L spreads for small size quintiles are positive while the other three for middle and big size quintiles are negative. The middle BE/ME quintile portfolios tend to have greater slopes related to the size factor while the low and high BE/ME quintile portfolios have smaller slopes. However, these are just opposite to Fama and French's (1993) results. It suggests that the size factor has a greater and positive impact on the returns on small size portfolios while a less or even negative impact on the returns on big size portfolios, and that this conclusion is robust at both monthly and daily level.

The slope coefficients of the BE/ME factor are almost all statistically significant at the 1% level, except that three are significant at the 5% level and two are not significant. The average of these slopes is -0.0268 with a range of -0.8232 to 0.6876. The estimated slopes for the two high BE/ME quintile portfolios are all positive, while those for the middle and low BE/ME quintile portfolios are almost all negative with only one exception. Thus, a positive relationship between the returns on the high BE/ME portfolios for each size quintile and the returns on the mimicking portfolios sorted by BE/ME is suggested. It also suggests that the returns on the middle and low BE/ME portfolios for each size quintile are negatively related to the returns on the mimicking portfolios sorted by BE/ME. The value of slopes for each size quintile increases along with the BE/ME quintiles from low to high. Thus, all H-L spreads are all positive. The two S-B spreads for high BE/ME quintiles are negative while the other three for the median and low BE/ME quintiles are positive. Up to here, the conclusion is similar to, but more significant than, what they find in Table 6 of Fama and French (1993). It suggests that the BE/ME factor has a greater and positive impact on the returns on high BE/ME portfolios but a less or even negative impact on the returns on low BE/ME portfolios, and that this conclusion is robust at both monthly and daily level.

The adjusted R^2s of the regressions range from 0.9159 to 0.9746, which are higher than the corresponding results of the CAPM reported in Table 2.4. Thus, adding the size and BE/ME factors into the model improves the power of the FF model to explain the cross section of returns on the 25 portfolios at the daily level. This is consistent with the findings in related studies at the monthly level, such as Fama and French (1993).

Table 2.5: Regressions of natural logarithm stock daily returns on the natural logarithm stock-market daily returns: From 2nd July 2007 to 31st December 2013, 1582 trading days.

$$R_t - R_f = \alpha + \beta_M(R_{Mt} - R_f) + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \varepsilon_t$$

Dependent variable: Returns on 25 stock portfolios formed on Size and Book-to-Market Equity

Size	Book-to-Market Equity (BE/ME) quintiles											
quintiles	Low	2	3	4	High	H-L Spread	Low	2	3	4	High	H-L Spread
	α						t-statistics of α					
Small	0.0175	0.0163	0.0144	0.0214*	0.0107	-0.0067	1.2961	1.3555	1.2978	1.7353	0.716	-0.3339
2	-0.0281**	0.0039	0.0038	0.0132	0.0003	0.0284*	-2.1697	0.3823	0.3224	1.3303	0.0343	1.7779
3	-0.0094	-0.0126	-0.0053	0.0006	-0.0143	-0.0049	-0.5788	-0.9931	-0.4006	0.0524	-1.5101	-0.2607
4	-0.0304**	-0.0206	-0.0167	-0.0073	-0.0117	0.0187	-2.0455	-1.3816	-1.4202	-0.5651	-0.9703	0.9804
Big	0.0222**	-0.0272*	0.0082	-0.0083	0.0249*	0.0028	1.9721	-1.9149	0.5688	-0.5783	1.8599	0.1586
S-B Spread	-0.0047	0.0435**	0.0062	0.0297	-0.0142		-0.2678	2.3375	0.3416	1.5717	-0.7062	
	β_M						t-statistics of β_M					
Small	0.9146***	1.0076***	1.0173***	1.0161***	1.0105***	0.0959***	92.6099	114.0145	135.7894	128.4087	133.2884	7.7018
2	0.9773***	1.0368***	1.0501***	1.0505***	1.0383***	0.0610***	109.7867	137.1971	146.7942	143.2231	185.4248	5.7987
3	1.0002***	1.0544***	1.0840***	1.0862***	1.0501***	0.0500***	102.4803	111.9844	127.1986	142.5802	165.8168	4.2935
4	1.0092***	1.0586***	1.0930***	1.1031***	1.0406***	0.0314**	95.1721	117.9121	158.8987	138.455	106.2953	2.1755
Big	0.9987***	1.1017***	1.0956***	0.9710***	0.8910***	-0.1077***	144.3434	101.7295	101.9455	93.1662	100.5529	-9.5789
S-B Spread	-0.0841***	-0.0941***	-0.0783***	0.0451***	0.1194***		-6.9777	-6.7286	-5.9737	3.4447	10.2421	
	β_{SMB}						t-statistics of β_{SMB}					
Small	0.9398***	0.9959***	0.9721***	1.0064***	0.9745***	0.0347	34.6525	50.9559	42.2129	47.139	47.0017	1.0152
2	0.8396***	0.9197***	0.9414***	0.9059***	0.8554***	0.0158	37.4601	44.2315	45.4639	51.2005	58.0573	0.5878
3	0.7028***	0.7926***	0.7381***	0.7215***	0.6837***	-0.0191	29.2639	29.3811	35.1696	36.5587	37.3908	-0.6318
4	0.4465***	0.5763***	0.5255***	0.4821***	0.3246***	-0.1219***	16.2453	20.7977	22.662	23.2368	12.7895	-3.259
Big	-0.3788***	-0.1681***	-0.1992***	-0.3341***	-0.4335***	-0.0548	-18.8344	-5.6001	-7.0236	-11.0973	-15.9859	-1.6224
S-B Spread	1.3186***	1.1639***	1.1713***	1.3405***	1.4080***		39.0535	32.5006	32.064	36.3194	41.2463	

Table 2.5: (Continued)

$$R_t - R_f = \alpha + \beta_M(R_{Mt} - R_f) + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \varepsilon_t$$

Dependent variable: Returns on 25 stock portfolios formed on Size and Book-to-Market Equity

Size	Book-to-Market Equity (BE/ME) quintiles											
quintiles	Low	2	3	4	High	H-L Spread	Low	2	3	4	High	H-L Spread
	β_{HML}						t-statistics of β_{HML}					
Small	-0.4060***	-0.0891***	-0.0048	0.1746***	0.1972***	0.6032***	-13.7229	-2.9342	-0.1756	5.45	6.6385	14.3872
2	-0.4060***	-0.1620***	0.0632**	0.1284***	0.3533***	0.7593***	-13.3956	-5.3531	2.0193	5.0585	14.6917	19.6251
3	-0.4128***	-0.1757***	-0.0626**	0.1592***	0.3269***	0.7396***	-14.2281	-4.8725	-1.9881	5.8051	13.6276	19.6491
4	-0.4807***	-0.2137***	-0.0666**	0.1346***	0.3828***	0.8635***	-12.8400	-6.0147	-2.2022	4.8617	14.2014	18.7177
Big	-0.8232***	-0.3355***	-0.0426	0.4024***	0.6876***	1.5108***	-25.7074	-9.7370	-1.1642	11.6965	17.8058	30.1159
S-B Spread	0.4172***	0.2463***	0.0379	-0.2278***	-0.4904***		9.5707	5.3634	0.8311	-4.8454	-10.0654	
	Adjusted R^2						Standard Error of regression					
Small	0.9322	0.9524	0.9561	0.9565	0.9452		0.5162	0.4806	0.4664	0.473	0.5309	
2	0.9417	0.9558	0.9627	0.9684	0.9746		0.4983	0.4667	0.4417	0.4058	0.365	
3	0.9159	0.9507	0.9580	0.9632	0.9680		0.6092	0.4940	0.4677	0.4452	0.407	
4	0.9214	0.9387	0.9586	0.9552	0.9508		0.5740	0.5426	0.4587	0.4892	0.4936	
Big	0.9445	0.9337	0.9336	0.9288	0.9389		0.4477	0.5596	0.5699	0.5523	0.4909	

*** Significance at the 1% level
** Significance at the 5% level
* Significance at the 10% level
Note: The t-statistics reported are calculated from standard errors corrected for serial correlation and White (1980) heteroskedasticity using the Newey-West (1987) estimator. The rule for choosing the lag parameter q is that $q = 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression (see Stock and Watson, 2007, page 607).
The S-B Spread reports the difference of estimations between the small size portfolio and big size portfolio for each quintile of BE/ME. Similarly, the H-L Spread reports the difference of estimations between the high BE/ME portfolio and low BE/ME portfolio for each quintile of size.

By comparing the regression results in Table 2.4 and 2.5, we find that the intercept term of the FF model is not as significant as that of the CAPM. This suggests that the Fama-French factors can explain a proportion of the CAPM's alpha. We can also find that the adjusted R^2 increases after adding the Fama-French factors into the CAPM. These findings suggest the FF model outperforms the CAPM. But a small portion of the cross-sectional portfolio returns still cannot be explained by the FF model, which remains to be explored.

Table 2.6 presents the results of the third regression model (2.3) where two more factors, i.e., the momentum and liquidity (turnover) factors, are added. From the table, we can see that the estimated intercepts and slopes of market, size and BE/ME factors and their corresponding t-statistics are rather similar to those of the FF model in Table 2.5. The intercept terms have an average value of 0.0010, with a range of -0.0293 to 0.0267, and the significance as a whole does not change. The slope coefficients related to the market excess return and the size factor for the 25 portfolios are all highly significant at the 1% level. The average of the market slopes is 1.0072, with a range of 0.9332 to 1.0810, while the average of the size slopes is 0.5133, with a range of -0.4231 to 1.0109. There are 24 slope coefficients related to the BE/ME factor now significant at the 1% level, and the average of the slopes is -0.0305, with a range of -0.8283 to 0.6943. To save the space, attention will be put on the estimated slopes related to the new factors.

For the two new added factors, only 12 slopes related to the momentum factor are statistically significant at the 10% level, four of which are significant at the 1% level and six at the 5% level. The average of these slopes is 0.0003, with a range of -0.0647 to 0.0841. There are 20 estimated slopes related to the turnover factor significant at the 10% level, 15 of which are significant at the 1% level and two at the 5% level. The average of these slopes is 0.0826, with a range of -0.1520 to 0.2199. The adjusted R^2s of the regressions range from 0.9162 to 0.9749, just a slight increase of 0.0003 over the FF model.

Comparing the estimated parameters and the adjusted R^2s of this five-factor model with those of the FF model, we can find that although the estimated slopes related to the momentum and turnover factors are somewhat statistically significant, adding these two additional factors seems not very helpful in improving the explanatory power of the model.

Table 2.6: Regressions of natural logarithm stock daily returns on the natural logarithm stock-market daily returns: From 2nd July 2007 to 31st December 2013, 1582 trading days.

$$R_t - R_f = \alpha + \beta_M(R_{Mt} - R_f) + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \beta_{MOM}Momentum_t + \beta_{TO}Turnover_t + \varepsilon_t$$

Dependent variable: Returns on 25 stock portfolios formed on Size and Book-to-Market Equity

Size	Book-to-Market Equity (BE/ME) quintiles											
quintiles	Low	2	3	4	High	H-L Spread	Low	2	3	4	High	H-L Spread
	α						t-statistics of α					
Small	0.0155	0.0195	0.0154	0.0267**	0.0153	-0.0001	1.1301	1.6500	1.3827	2.2596	1.0519	-0.0069
2	-0.0288**	0.0108	0.0062	0.0180*	-0.0033	0.0255	-2.2361	1.0475	0.5303	1.8012	-0.3480	1.5969
3	-0.0045	-0.0050	0.0029	0.0044	-0.0192**	-0.0146	-0.2876	-0.4053	0.2198	0.3772	-2.0389	-0.7994
4	-0.0293*	-0.0097	-0.0115	0.0001	-0.0187	0.0106*	-1.9142	-0.6616	-0.9595	0.0076	-1.5515	0.5454
Big	0.0222**	-0.0264*	0.0136	-0.0087	0.0204	-0.0018	1.9897	-1.8959	0.9426	-0.5988	1.5639	-0.1042
S-B Spread	-0.0067	0.0459**	0.0017	0.0355*	-0.0051		-0.3817	2.5132	0.0942	1.8879	-0.2601	
	β_M						t-statistics of β_M					
Small	0.9603***	0.9420***	0.9776***	0.9765***	0.9753***	0.0151	67.4537	73.1277	85.7748	74.2731	67.4667	0.7434
2	0.9596***	1.0045***	0.9939***	1.0207***	1.0236***	0.0640***	74.7219	88.0718	76.4640	88.3665	96.9525	3.8520
3	0.9761***	1.0024***	1.0355***	1.0640***	1.0574***	0.0813***	65.1525	72.4095	84.7917	102.5154	80.8392	4.0869
4	0.9839***	1.0177***	1.0506***	1.0536***	1.0652***	0.0813***	69.5156	77.4072	82.4208	74.9049	71.7276	3.9643
Big	0.9678***	1.0810***	1.0748***	0.9835***	0.9332***	-0.0346*	84.1223	77.1716	62.0554	66.0843	61.9424	-1.8246
S-B Spread	-0.0075	-0.1389***	-0.0973***	-0.0070	0.0421**		-0.4113	-7.3013	-4.6918	-0.3544	2.0183	
	β_{SMB}						t-statistics of β_{SMB}					
Small	1.0109***	0.8928***	0.9126***	0.9343***	0.9109***	-0.1000**	29.0240	37.0892	34.4165	33.1144	31.3549	-2.2052
2	0.8166***	0.8535***	0.8542***	0.8493***	0.8450***	0.0285	30.1696	34.8824	31.6556	35.2965	42.6973	0.8487
3	0.6542***	0.6963***	0.6452***	0.6788***	0.7083***	0.0541	21.5269	20.1655	25.5252	28.2186	28.8144	1.3847
4	0.4070***	0.4862***	0.4496***	0.3901***	0.3802***	-0.0269	10.6108	13.8653	15.2516	13.6283	12.4466	-0.5482
Big	-0.4231***	-0.2000***	-0.2446***	-0.3150***	-0.3602***	0.0630	-14.6542	-6.0088	-6.8824	-8.4310	-10.4528	1.4010
S-B Spread	1.4340***	1.0928***	1.1573***	1.2493***	1.2710***		31.6967	26.6061	26.0961	26.6849	28.2026	

Table 2.6: (Continued)

$$R_t - R_f = \alpha + \beta_M(R_{Mt} - R_f) + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \beta_{MOM}Momentum_t + \beta_{TO}Turnover_t + \varepsilon_t$$

Dependent variable: Returns on 25 stock portfolios formed on Size and Book-to-Market Equity

Size	Book-to-Market Equity (BE/ME) quintiles											
quintiles	Low	2	3	4	High	H-L Spread	Low	2	3	4	High	H-L Spread
	β_{HML}						t-statistics of β_{HML}					
Small	-0.3985***	-0.0998***	-0.0113	0.1683***	0.1916***	0.5901***	-13.8887	-3.5880	-0.4327	5.5945	6.6388	14.4998
2	-0.4090***	-0.1670***	0.0540*	0.1237***	0.3506***	0.7596***	-13.5715	-5.7398	1.9318	5.0714	15.1451	19.9889
3	-0.4165***	-0.1840***	-0.0702**	0.1557***	0.3278***	0.7443***	-14.6077	-5.4636	-2.3717	5.8879	14.1481	20.2590
4	-0.4849***	-0.2199***	-0.0734***	0.1268***	0.3865***	0.8714***	-13.1084	-6.5472	-2.5979	4.6225	14.3412	19.0396
Big	-0.8283***	-0.3389***	-0.0458	0.4045***	0.6943***	1.5227***	-26.9943	-9.8752	-1.2753	11.6189	18.0781	30.9733
S-B Spread	0.4298***	0.2390***	0.0345	-0.2361***	-0.5027***		10.2315	5.4106	0.7767	-5.1322	-10.4638	
	β_{MOM}						t-statistics of β_{MOM}					
Small	0.0342	-0.0450**	-0.0404*	0.0151	0.0124	-0.0218	1.2935	-2.0029	-1.8258	0.6252	0.4726	-0.5860
2	-0.0326	0.0448**	-0.0428*	0.0217	-0.0647***	-0.0321	-1.5502	2.0689	-1.8694	1.0866	-3.8145	-1.1865
3	0.0304	0.0278	0.0395*	0.0189	-0.0520**	-0.0824**	1.0527	1.0875	1.8272	0.9083	-2.5441	-2.3283
4	-0.0189	0.0841***	0.0107	0.0286	-0.0566**	-0.0377	-0.6467	3.2536	0.5128	1.2652	-1.9772	-0.9217
Big	-0.0396	-0.0171	0.0414	0.0104	-0.0022	0.0374	-1.6390	-0.5816	1.2041	0.3689	-0.0869	1.0643
S-B Spread	0.0738**	-0.0279	-0.0818**	0.0047	0.0146		2.0607	-0.7546	-1.9998	0.1261	0.3992	
	β_{TO}						t-statistics of β_{TO}					
Small	-0.1520***	0.2199***	0.1289***	0.1471***	0.1301***	0.2820***	-3.6192	5.9160	3.7373	3.8963	3.0518	4.7139
2	0.0523	0.1308***	0.1867***	0.1141***	0.0306	-0.0218	1.5013	3.5967	4.8961	3.4885	0.9724	-0.4635
3	0.0964**	0.1955***	0.1867***	0.0858***	-0.0435	-0.1400**	2.1352	4.9470	5.3869	2.6690	-1.1073	-2.3377
4	0.0843*	0.1749***	0.1556***	0.1866***	-0.1072***	-0.1915***	1.9261	4.0857	3.7606	4.6481	-2.8736	-3.3295
Big	0.0975**	0.0685*	0.0883*	-0.0411	-0.1517***	-0.2492***	2.4193	1.8088	1.7726	-0.9101	-3.8240	-4.4065
S-B Spread	-0.2495***	0.1514***	0.0406	0.1882***	0.2818***		-4.2862	2.8524	0.6693	3.1976	4.8392	

Table 2.6: (Continued)

$$R_t - R_f = \alpha + \beta_M(R_{Mt} - R_f) + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \beta_{MOM}Momentum_t + \beta_{TO}Turnover_t + \varepsilon_t$$

Dependent variable: Returns on 25 stock portfolios formed on Size and Book-to-Market Equity

Size	Book-to-Market Equity (BE/ME) quintiles											
quintiles	Low	2	3	4	High	H-L Spread	Low	2	3	4	High	H-L Spread
	Adjusted R^2						Standard Error of regression					
Small	0.9331	0.9540	0.9567	0.9572	0.9456		0.5128	0.4725	0.4634	0.4696	0.5287	
2	0.9418	0.9565	0.9637	0.9688	0.9749		0.4978	0.4631	0.4353	0.4033	0.3629	
3	0.9162	0.9519	0.9591	0.9634	0.9682		0.6080	0.4878	0.4613	0.4440	0.4055	
4	0.9216	0.9403	0.9593	0.9562	0.9514		0.5733	0.5355	0.4548	0.4835	0.4906	
Big	0.9450	0.9338	0.9340	0.9288	0.9398		0.4457	0.5592	0.5685	0.5524	0.4876	

*** Significance at the 1% level
** Significance at the 5% level
* Significance at the 10% level
Note: The t-statistics reported are calculated from standard errors corrected for serial correlation and White (1980) heteroskedasticity using the Newey-West (1987) estimator. The rule for choosing the lag parameter q is that $q = 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression (see Stock and Watson, 2007, page 607).
The S-B Spread reports the difference of estimations between the small size portfolio and big size portfolio for each quintile of BE/ME. Similarly, the H-L Spread reports the difference of estimations between the high BE/ME portfolio and low BE/ME portfolio for each quintile of size.

2.4.2 Subsample time-series regression in a downward market

Tables 2.7 and 2.8 show the results of regressions on two models – the CAPM and the FF model – in the downward period from 3rd September 2007 to 31st December 2008, with 326 trading days.

In this subsection, the first regression performed is the CAPM, which is treated as the benchmark for the FF models but will be compared with the results in Section 2.4.1.

Shown in Table 2.7, eight estimated intercept terms are statistically significant at least at the 10% level, two of which are significant at the 1% level and the other two at the 5% level. The number of significant intercepts is less than that shown in Table 2.4. It suggests that the small size and high BE/ME portfolios have the significant constant terms. Among the 25 portfolios, four portfolios have negative intercepts, while 21 portfolios have intercepts greater than zero. The big size and low BE/ME portfolios tend to have negative intercepts. The average of the 25 intercepts is 0.0492, with a range of -0.0724 to 0.1544. The positive S-B spread and H-L spread report that the small size and high BE/ME portfolios tend to have larger constant terms. Compared to Table 2.4, the average and the range are greater and the increments could not be ignored. Other features of estimations here are very similar to that in Table 2.4.

The slope coefficients of the market excess return are all highly significant at the 1% level, with the smallest t-statistic of 39.8969. The values of these slopes are around one, and the average is 1.0247, with a range of 0.8538 to 1.1171. Compared to Table 2.4, the average decreases but the range widens. Two S-B spreads are negative while the other three are positive. The slopes increase first and then decrease from small to big size for each BE/ME quintile. All the H-L spreads are positive and the small size quintile has the biggest value. The values decrease monotonically along with the size quintiles. The positive H-L spreads suggest that the market excess return would have a greater impact on the excess returns on the high BE/ME portfolios.

The adjusted R^2s of the regressions range from 0.8304 to 0.9634. Compared to Table 2.4, the adjusted R^2s as a whole are greater. This suggests that the CAPM performs better in the downward trend. The adjusted R^2s are also greater than those at the monthly return level when compared to Table 4 of Fama and French (1993). Although the highest value of the adjusted R^2s reaches 0.9634, the lowest value of the adjusted R^2s is only 0.8304. There is still a portion that cannot be explained by the CAPM.

Table 2.7: Regressions of natural logarithm stock daily returns on the natural logarithm stock-market daily returns: From 8rd October 2007 to 31st December 2008, 306 trading days.

$$R_t - R_f = \alpha + \beta_M (R_{Mt} - R_f) + \varepsilon_t$$

Dependent variable: Returns on 25 stock portfolios formed on Size and Book-to-Market Equity

Size	Book-to-Market Equity (BE/ME) quintiles											
quintiles	Low	2	3	4	High	H-L Spread	Low	2	3	4	High	H-L Spread
	α						t-statistics of α					
Small	0.0198	0.1214*	0.0919	0.1809***	0.1179*	0.0980	0.2430	1.6876	1.3718	2.8500	1.8590	0.9483
2	-0.0523	0.0517	0.1348**	0.1599***	0.1456**	0.1979**	-0.7368	0.7669	2.3225	2.6622	2.4763	2.1472
3	0.0354	0.0295	0.0672	0.0984*	0.0657	0.0303	0.5151	0.5366	1.3075	1.9521	1.2911	0.3544
4	-0.0444	0.0687	0.0278	0.0462	0.0605	0.1049	-0.8836	1.4851	0.7619	0.9390	1.1964	1.4714
Big	-0.0356	-0.078*	-0.0217	0.0382	0.0652	0.1008	-0.8348	-1.7409	-0.6100	0.9230	1.2021	1.4610
S-B Spread	0.0554	0.1994**	0.1136	0.1426*	0.0527		0.6018	2.3528	1.4976	1.8820	0.6315	
	β_M						t-statistics of β_M					
Small	0.8600***	1.0184***	1.0293***	1.0420***	1.0353***	0.1754***	47.7234	46.9503	55.0595	50.7721	53.5464	6.6354
2	0.9231***	1.0166***	1.0669***	1.0653***	1.0913***	0.1681***	54.5009	57.3550	60.0457	55.5868	59.4425	6.7305
3	0.9482***	1.0394***	1.0644***	1.0914***	1.0909***	0.1427***	47.9530	60.1031	67.6360	70.4345	66.4733	5.5546
4	0.9771***	1.0370***	1.0760***	1.1187***	1.0780***	0.1010***	53.6487	71.1384	92.7850	75.3025	53.7846	3.7276
Big	0.9425***	1.0117***	1.0390***	0.9817***	0.9777***	0.0352	68.1740	76.3474	87.9391	83.5425	50.2223	1.4724
S-B Spread	-0.0826***	0.0067	-0.0097	0.0603**	0.0576**		-3.6361	0.2635	-0.4368	2.5510	2.1005	

Table 2.7: (Continued)

$$R_t - R_f = \alpha + \beta_M (R_{Mt} - R_f) + \varepsilon_t$$

Dependent variable: Returns on 25 stock portfolios formed on Size and Book-to-Market Equity

Size	Book-to-Market Equity (BE/ME) quintiles											
quintiles	Low	2	3	4	High	H-L Spread	Low	2	3	4	High	H-L Spread
	Adjusted R^2						Standard Error of regression					
Small	0.8347	0.8909	0.9055	0.8987	0.9073		1.1678	1.0874	1.0147	1.0675	1.0096	
2	0.8824	0.9038	0.9238	0.9188	0.9254		1.0285	1.0120	0.9354	0.9665	0.9459	
3	0.9003	0.9307	0.9416	0.9428	0.9391		0.9629	0.8656	0.8089	0.8206	0.8478	
4	0.9420	0.9485	0.9641	0.9511	0.9460		0.7398	0.7372	0.6336	0.7740	0.7857	
Big	0.9488	0.9444	0.958	0.9452	0.9133		0.6680	0.7489	0.6637	0.7214	0.9195	

*** Significance at the 1% level
** Significance at the 5% level
* Significance at the 10% level
Note: The t-statistics reported are calculated from standard errors corrected for serial correlation and White (1980) heteroskedasticity using the Newey-West (1987) estimator. The rule for choosing the lag parameter q is that $q = 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression (see Stock and Watson, 2007, page 607).
The S-B Spread reports the difference of estimations between the small size portfolio and big size portfolio for each quintile of BE/ME. Similarly, the H-L Spread reports the difference of estimations between the high BE/ME portfolio and low BE/ME portfolio for each quintile of size.

The test results of the FF model are shown in Table 2.8. We can see that the intercept terms as a whole are not as statistically significant as those of the CAPM in Table 2.7. There are only three estimated intercept terms statistically significant at least at the 10% level, two of which is significant at the 5% level. The number of positive intercept terms decreases to 12, while 13 of intercept terms are negative. The average of the 25 intercepts is 0.0049, with a range of -0.0574 to 0.0539. Two S-B spreads and three H-L spreads are negative, and the rest ones are positive. The decreasing significance, the equal numbers of positive and negative intercepts, and being closer to zero provide the first evidence that the FF model performs better than the CAPM in explaining the cross-sectional excess returns.

The slope coefficients of the market excess return are all highly significant at the 1% level, with the smallest t-statistic of 54.4327. The values of these slopes are around one, and the average is 1.0083, with a range of 0.8875 to 1.0808. Compared to the estimated slopes of the market excess return of the CAPM in Table 2.7, the FF model seems to have the market excess return slopes closer to one although some of them are a bit far away from one. The middle quintiles of both size and BE/ME portfolios tend to have greater slopes although it is not the case for some portfolios. Three S-B spreads for low BE/ME quintiles are less than zero but the other two are positive. The H-L spreads decrease along with the size quintiles from 0.0724 to -0.1199. The top four H-L spreads are positive and less than those in Table 2.7.

The slope coefficients of the size factor are all statistically significant at the 1% level, with the smallest absolute value of t-statistic of -5.0095. The average of these slopes is 0.5121, with a range of -0.5421 to 0.9786. The slopes of the biggest size portfolios are all negative for each BE/ME quintile while the rest slopes are positive. The value of slopes for each BE/ME quintile decreases along with the size quintiles from small to big. The smaller size portfolios have greater slopes related to the size factor while the bigger size portfolios have smaller or even negative slopes. Thus, all S-B spreads are greater than one. Up to here, the conclusion is consistent with that reported in Table 6 of Fama and French (1993). All the H-L spreads are negative. It suggests that the size factor has a greater and positive impact on the returns on small size portfolios but a less or even negative impact on the returns on big size portfolios, and that this conclusion is robust at both monthly and daily level.

The slope coefficients of the BE/ME factor are almost all statistically significant at the 1% level, except that two are significant at the 5% level, one at 10% level, and the other two are not at the 10% level. The average of these slopes is 0.1202 with a range of -0.5448 to 0.7920. The estimated slopes for the middle and high BE/ME quintile portfolios are all positive, while those for the lowest BE/ME quintile portfolios are all negative. The value of slopes for each size quintile increases along with the BE/ME quintiles from low to high. Thus, all H-L spreads are all positive. Two S-B spreads for high BE/ME quintiles are negative while the other three for median and low BE/ME quintile are positive. Up to here, the conclusion is similar to those found in Table 6 of Fama and French (1993). It suggests that the BE/ME factor has a greater and positive impact on the returns on high BE/ME portfolios while a less or even negative impact on the returns on low BE/ME portfolios, and that this conclusion is robust at both monthly and daily level.

The adjusted R^2 of the regressions range from 0.9216 to 0.9863, which are higher than the corresponding results of the CAPM reported in Table 2.7. Thus, adding the size and BE/ME factors into the model improves the power of the FF model to explain the cross-sectional returns on the 25 portfolios at the daily level. This is consistent with the findings in related studies at the monthly level, such as Fama and French (1993).

Table 2.8: Regressions of natural logarithm stock daily returns on the natural logarithm stock-market daily returns: From 8rd October 2007 to 31st December 2008, 306 trading days.

$$R_t - R_f = \alpha + \beta_M(R_{Mt} - R_f) + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \varepsilon_t$$

Dependent variable: Returns on 25 stock portfolios formed on Size and Book-to-Market Equity

Size	Book-to-Market Equity (BE/ME) quintiles											
quintiles	Low	2	3	4	High	H-L Spread	Low	2	3	4	High	H-L Spread
	α						t-statistics of α					
Small	0.0092	0.0391	0.0117	0.0678**	0.0030	-0.0061	0.1983	0.9753	0.3755	2.1353	0.1131	-0.1151
2	-0.0568	-0.0090	0.0419	0.0507*	0.0212	0.0780	-1.3041	-0.2478	1.4917	1.9079	0.8729	1.5641
3	0.0329	-0.0341	0.0030	0.0020	-0.0369	-0.0698	0.7314	-0.9506	0.0879	0.0731	-1.5079	-1.3622
4	-0.0355	0.0191	-0.0156	-0.0262	-0.0217	0.0138	-0.9011	0.5273	-0.5337	-0.7619	-0.5350	0.2449
Big	0.0552**	-0.0203	-0.0069	0.0102	0.0034	-0.0518	2.0659	-0.5147	-0.2004	0.2553	0.0840	-1.0701
S-B Spread	-0.0461	0.0595	0.0187	0.0576	-0.0004		-0.8636	1.0560	0.4006	1.1247	-0.0078	
	β_M						t-statistics of β_M					
Small	0.9111***	1.0003***	1.0125***	0.9956***	0.9841***	0.0730***	49.5004	70.3817	82.2578	71.9739	88.5640	3.3965
2	0.9738***	1.0144***	1.0330***	1.0156***	1.0249***	0.0512***	55.2378	78.3189	101.8692	93.3196	130.1163	2.6507
3	0.9942***	1.0226***	1.0426***	1.0404***	1.0343***	0.0402**	57.9504	64.5754	79.9846	106.7652	112.0131	2.0612
4	1.0088***	1.0182***	1.0582***	1.0810***	1.0208***	0.0120	63.3003	76.3390	118.2896	85.1026	58.6743	0.5095
Big	1.0078***	1.0473***	1.0407***	0.9361***	0.8849***	-0.1230***	102.1197	73.6413	76.7022	71.1173	69.5450	-7.6356
S-B Spread	-0.0968***	-0.0470**	-0.0282	0.0595***	0.0992***		-4.6331	-2.3362	-1.5418	3.1142	5.8733	
	β_{SMB}						t-statistics of β_{SMB}					
Small	0.9740***	0.9602***	0.9478***	0.9777***	0.9279***	-0.0461	15.8858	22.5927	20.4401	27.9540	33.7942	-0.6863
2	0.8726***	0.8852***	0.8706***	0.8669***	0.8336***	-0.0390	16.7692	21.414	24.9525	26.8299	31.6066	-0.6688
3	0.7686***	0.6972***	0.6279***	0.6525***	0.6574***	-0.1112*	14.9498	18.9435	14.2672	17.4008	20.9244	-1.8455
4	0.3691***	0.4551***	0.3758***	0.4987***	0.3373***	-0.0318	9.8912	9.3179	9.2440	14.6526	7.6891	-0.5526
Big	-0.3400***	-0.3100***	-0.1974***	-0.2990***	-0.5375***	-0.1975***	-8.9118	-7.7622	-3.6601	-6.1210	-8.7499	-2.7314
S-B Spread	1.3140***	1.2701***	1.1452***	1.2767***	1.4654***		18.1961	21.7796	16.1006	21.2519	21.7785	

Table 2.8: (Continued)

$$R_t - R_f = \alpha + \beta_M(R_{Mt} - R_f) + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \varepsilon_t$$

Dependent variable: Returns on 25 stock portfolios formed on Size and Book-to-Market Equity

Size	Book-to-Market Equity (BE/ME) quintiles											
quintiles	Low	2	3	4	High	H-L Spread	Low	2	3	4	High	H-L Spread
	β_{HML}						t-statistics of β_{HML}					
Small	-0.4675***	0.1105	0.1005*	0.3457***	0.3881***	0.8556***	-6.1684	1.6468	1.7468	5.2787	8.6976	9.7283
2	-0.4591***	-0.0192	0.2457***	0.3777***	0.5179***	0.9770***	-5.7359	-0.3400	6.2474	8.4840	14.4283	11.1374
3	-0.4164***	0.1109	0.1548**	0.3976***	0.4438***	0.8602***	-6.9949	1.5331	2.5629	8.2416	10.9120	11.9308
4	-0.2804***	0.1370***	0.1321***	0.2933***	0.4626***	0.7431***	-4.7181	2.9462	2.9983	5.4771	9.1160	9.5075
Big	-0.5301***	-0.2833***	-0.0060	0.3925***	0.7970***	1.3271***	-11.7062	-4.7260	-0.0967	6.9308	9.0319	13.3801
S-B Spread	0.0626	0.3938***	0.1065	-0.0469	-0.4089***		0.7087	4.3768	1.2636	-0.5413	-4.1349	
	Adjusted R^2						Standard Error of regression					
Small	0.9234	0.9587	0.9711	0.9749	0.9803		0.7949	0.6689	0.5615	0.5311	0.4650	
2	0.9487	0.9600	0.9807	0.9804	0.9873		0.6792	0.6529	0.4706	0.4750	0.3898	
3	0.9503	0.9671	0.9713	0.9811	0.9797		0.6801	0.5963	0.5674	0.4722	0.4895	
4	0.9543	0.9654	0.9753	0.9722	0.9646		0.6566	0.6043	0.5258	0.5842	0.6362	
Big	0.9775	0.9568	0.9607	0.9569	0.9560		0.4427	0.6602	0.6425	0.6398	0.6552	

*** Significance at the 1% level
** Significance at the 5% level
* Significance at the 10% level

Note: The t-statistics reported are calculated from standard errors corrected for serial correlation and White (1980) heteroskedasticity using the Newey-West (1987) estimator. The rule for choosing the lag parameter q is that $q = 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression (see Stock and Watson, 2007, page 607).

The S-B Spread reports the difference of estimations between the small size portfolio and big size portfolio for each quintile of BE/ME. Similarly, the H-L Spread reports the difference of estimations between the high BE/ME portfolio and low BE/ME portfolio for each quintile of size.

2.4.3 Subsample time-series regression in a fluctuating market

Tables 2.9 and 2.10 report the regression results in the fluctuating sub-sample period from 5th January 2009 to 31st December 2013, with 1211 trading days. In this subsection, the first regression performed is the CAPM, which is treated as the benchmark for the FF model but will be compared with the results in Section 2.4.1. Table 2.9 outputs the regression results of the CAPM in the fluctuating trend.

There are 15 estimated intercept terms statistically significant at least at the 10% level, ten of which are significant at the 1% level and the other three at the 5% level. It suggests that the small size and high BE/ME portfolios have the significant intercept terms. But the number is close to that shown in Table 2.4. Among the 25 portfolios, five portfolios have negative intercepts while 20 portfolios have intercepts greater than zero. The big size and low BE/ME portfolios tend to have negative intercepts. The average of the 25 intercepts is 0.0293, with a range of -0.0438 to 0.0791. Both the average and the range are less than those reported in Table 2.4. The positive S-B spreads and H-L spreads report that the small size and high BE/ME portfolios tend to have larger intercept terms. This is consistent with that in Table 2.4. Compared to Table 2.4, the average and the range decrease and the increments could not be ignored. Other features of estimations here are very similar to those reported in Table 2.4.

The slope coefficients of the market excess return are all highly statistically significant at the 1% level, with the smallest t-statistic of 66.7060. The values of these slopes are around one, and the average is 1.0757, with a range of 0.8564 to 1.1638. Compared to Table 2.4, the average decreases but the range widens. Two S-B spreads are negative while the other three are positive. The positive S-B spread for the lowest BE/ME quintile is not small now, compared to the other two positive S-B spreads. All the H-L spreads are positive and the small size quintile has the biggest value. The value does not decrease monotonically along with the size quintiles here. The positive H-L spreads suggest that the market excess return would have a greater impact on the excess returns on the high BE/ME portfolios.

The adjusted R^2s of the regressions range from 0.7862 to 0.9209, which are lower compared to the corresponding results in Table 2.4 but still higher than those in Fama and French

(1993). It suggests that the performance of the CAPM deteriorates in the fluctuating market, but is still better at the daily level than at the monthly level.

Table 2.9: Regressions of natural logarithm stock daily returns on the natural logarithm stock-market daily returns: From 5th January 2009 to 31st December 2013, 1211 trading days.

$$R_t - R_f = \alpha + \beta_M(R_{Mt} - R_f) + \varepsilon_t$$

Dependent variable: Returns on 25 stock portfolios formed on Size and Book-to-Market Equity

Size	Book-to-Market Equity (BE/ME) quintiles											
quintiles	Low	2	3	4	High	H-L Spread	Low	2	3	4	High	H-L Spread
	α						t-statistics of α					
Small	0.0703***	0.0683***	0.0791***	0.0769***	0.0742***	0.0038	2.7255	2.7648	3.4921	3.1888	2.9594	0.1067
2	0.0317	0.0595***	0.0521**	0.0567***	0.0540***	0.0223	1.3396	2.6803	2.2570	2.7667	2.7988	0.7294
3	0.0277	0.0366	0.0297	0.0494***	0.0373**	0.0096	1.2026	1.6073	1.4887	2.7599	2.2060	0.3374
4	-0.0089	0.0017	0.0174	0.0262*	0.0133	0.0221	-0.3920	0.0761	1.0487	1.6550	0.9974	0.8441
Big	-0.0238	-0.0438***	-0.0167	-0.0409**	0.0047	0.0285	-1.1972	-2.7991	-1.1070	-2.2959	0.2434	1.0269
S-B Spread	0.0942***	0.1121***	0.0958***	0.1178***	0.0695**		2.8894	3.8335	3.5201	3.9291	2.1908	
	β_M						t-statistics of β_M					
Small	0.9703***	1.0652***	1.0787***	1.1048***	1.1073***	0.1370***	59.7063	61.3353	62.3503	60.2332	60.1313	5.5776
2	1.0164***	1.0972***	1.1175***	1.1384***	1.1241***	0.1077***	57.2546	69.1790	63.5022	73.6609	74.3419	4.6172
3	1.0353***	1.0994***	1.1527***	1.1638***	1.1290***	0.0937***	59.1767	63.8553	75.4593	82.8472	87.9819	4.3176
4	0.9806***	1.0907***	1.1389***	1.1623***	1.1032***	0.1226***	56.3944	72.1303	87.7256	96.0469	111.0283	6.1229
Big	0.8564***	1.1229***	1.1269***	1.0141***	0.8974***	0.0409*	68.6813	74.9612	67.9553	58.9991	48.4531	1.8330
S-B Spread	0.1139***	-0.0577**	-0.0482**	0.0908***	0.2100***		5.5598	-2.5173	-2.0127	3.6106	8.0389	

Table 2.9: (Continued)

$$R_t - R_f = \alpha + \beta_M (R_{Mt} - R_f) + \varepsilon_t$$

Dependent variable: Returns on 25 stock portfolios formed on Size and Book-to-Market Equity

Size	Book-to-Market Equity (BE/ME) quintiles											
quintiles	Low	2	3	4	High	H-L Spread	Low	2	3	4	High	H-L Spread
	Adjusted						Standard Error of regression					
Small	0.8039	0.8263	0.8306	0.8337	0.8159		0.7396	0.7539	0.7520	0.7616	0.8119	
2	0.8331	0.8474	0.8518	0.8731	0.8852		0.7022	0.7186	0.7195	0.6698	0.6247	
3	0.8409	0.8505	0.8842	0.9014	0.9037		0.6951	0.7113	0.6440	0.5943	0.5688	
4	0.8209	0.8641	0.9053	0.9209	0.9202		0.7070	0.6676	0.5686	0.5257	0.5014	
Big	0.7862	0.9099	0.9152	0.8683	0.8383		0.6894	0.5454	0.5295	0.6097	0.6082	

*** Significance at the 1% level
** Significance at the 5% level
* Significance at the 10% level

Note: The t-statistics reported are calculated from standard errors corrected for serial correlation and White (1980) heteroskedasticity using the Newey-West (1987) estimator. The rule for choosing the lag parameter q is that $q = 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression (see Stock and Watson, 2007, page 607).

The S-B Spread reports the difference of estimations between the small size portfolio and big size portfolio for each quintile of BE/ME. Similarly, the H-L Spread reports the difference of estimations between the high BE/ME portfolio and low BE/ME portfolio for each quintile of size.

The test results of the FF model are reported in Table 2.10. We can see that the intercept terms as a whole are not as statistically significant as those of the CAPM in Table 2.9. Only three intercept terms are significant at the 5% level. The number of positive intercept terms decreases to 9, while 16 of intercept terms are negative. The average of the 25 intercepts is -0.0056, with a range of -0.0333 to 0.0219. The middle three S-B spreads are positive while the left and right S-B spreads are negative. Four H-L spreads are greater than zero while the other for the smallest size quintile is not. The decreasing significance, the equal numbers of positive and negative intercepts, and being closer to zero provide the first evidence that the FF model performs better than the CAPM in explaining the cross-sectional excess returns.

The slope coefficients of the market excess return are all highly significant at the 1% level, with the smallest t-statistic of 96.8157. The values of these slopes are around one, and the average is 1.0410, with a range of 0.8771 to 1.1574. Compared to the estimated slopes of the market excess return of the CAPM in Table 2.9, the FF model seems to have the market excess return slopes closer to one although some of them are a bit far away from one. The middle quintiles of both size and BE/ME portfolios tend to have greater slopes although it is not the case for some portfolios. Three S-B spreads for low BE/ME quintiles are less than zero but the other two are positive. The H-L spreads decrease along with the size quintiles from 0.0944 to -0.0901. The top four H-L spreads are positive and less than those in Table 2.9.

The slope coefficients of the size factor are all statistically significant at the 1% level, with the smallest absolute value of t-statistic of -8.9010. The average of these slopes is 0.5557, with a range of -0.4258 to 1.0011. The slopes of the biggest size portfolios are all negative for each BE/ME quintile while the rest of slopes are positive. The value of slopes for each BE/ME quintile decreases along with the size quintiles from small to big. The smaller size portfolios have greater slopes related to the size factor while the bigger size portfolios have smaller or even negative slopes. Thus, all S-B spreads are greater than one. Up to here, the conclusion is consistent with that reported in Table 6 of Fama and French (1993). Four H-L spreads are positive and the negative one has bigger absolute value. The middle BE/ME quintile portfolios tend to have greater slopes related to the size factor while the low and high BE/ME quintile portfolios have smaller slopes. However, these are just opposite to Fama and French's (1993) results. It suggests that the size factor has a greater and positive impact on the returns on small size portfolios but a less or even negative impact on the returns on big size portfolios, and that this conclusion is robust at both monthly and daily level.

The slope coefficients of the BE/ME factor are almost all statistically significant at the 1% level, except that two are significant at the 5% level and one is not. The average of these slopes is -0.1030, with a range of -1.0173 to 0.6524. The slopes for the three low and median BE/ME quintile portfolios are all negative, while the slopes for the two high BE/ME quintile portfolios are all positive. The value of slopes for each size quintile increases along with the BE/ME quintiles from low to high. Thus, all H-L spreads are all positive. Two S-B spreads for low BE/ME quintiles are positive while the other three for median and high BE/ME quintile are positive. Up to here, the conclusion is similar to those in Table 6 of Fama and French (1993). It suggests that the BE/ME factor has a greater and positive impact on the returns on high BE/ME portfolios but a less or even negative impact on the returns on low BE/ME portfolios, and that this conclusion is robust at both monthly and daily level.

The adjusted R^2s of the regressions range from 0.8939 to 0.9651, which are higher than the corresponding results of the CAPM reported in Table 2.9, but lower than the corresponding results of the FF model reported in Table 2.5. Thus, adding the size and BE/ME factors into the model improves the explanatory power of the FF model in cross-sectional returns on the 25 portfolios at the daily level. This is consistent with the findings in related studies at the monthly level, such as Fama and French (1993).

Table 2.10: Regressions of natural logarithm stock daily returns on the natural logarithm stock-market daily returns: From 5th January 2009 to 31st December 2013, 1211 trading days.

$$R_t - R_f = \alpha + \beta_M(R_{Mt} - R_f) + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \varepsilon_t$$

Dependent variable: Returns on 25 stock portfolios formed on Size and Book-to-Market Equity

Size	Book-to-Market Equity (BE/ME) quintiles											
quintiles	Low	2	3	4	High	H-L Spread	Low	2	3	4	High	H-L Spread
	α						t-statistics of α					
Small	0.0161	0.0059	0.0163	0.0111	0.0089	-0.0072	1.2862	0.5155	1.3562	0.8275	0.503	-0.3297
2	-0.0168	0.0024	-0.0082	-0.0020	-0.0038	0.0130	-1.4256	0.2336	-0.6553	-0.2000	-0.3736	0.8384
3	-0.0104	-0.0127	-0.0189	0.0015	-0.0104	0.0000	-0.7208	-0.9731	-1.5360	0.1250	-0.9721	0.0006
4	-0.0333**	-0.0328**	-0.0183	-0.0040	-0.0104	0.0229	-2.0671	-2.0970	-1.4714	-0.2958	-0.8600	1.1356
Big	0.0165	-0.0321**	-0.0023	-0.0230	0.0219	0.0054	1.4462	-2.1675	-0.1565	-1.5475	1.6414	0.3067
S-B Spread	-0.0004	0.0381**	0.0186	0.0341*	-0.0130		-0.0235	2.0284	0.9735	1.7035	-0.5829	
	β_M						t-statistics of β_M					
Small	0.9297***	1.0042***	1.0091***	1.0235***	1.0242***	0.0944***	88.7124	106.4613	121.1241	101.3316	94.2101	6.2534
2	0.9834***	1.0439***	1.0504***	1.0679***	1.0405***	0.0570***	125.6503	110.0110	106.4805	123.8483	135.6460	5.2026
3	1.0178***	1.0609***	1.1059***	1.1046***	1.0565***	0.0387***	106.6660	101.6979	107.5665	132.1691	133.7509	3.1280
4	0.9880***	1.0734***	1.1065***	1.1230***	1.0547***	0.0667***	75.7682	93.4288	133.2045	111.4816	120.9372	4.2510
Big	0.9672***	1.1574***	1.1463***	1.0089***	0.8771***	-0.0901***	104.1141	80.8200	71.7289	69.0355	69.6852	-5.7568
S-B Spread	-0.0374***	-0.1533***	-0.1372***	0.0147	0.1471***		-2.6717	-8.9379	-7.6127	0.8248	8.8428	
	β_{SMB}						t-statistics of β_{SMB}					
Small	0.9056***	0.9975***	0.9792***	1.0011***	0.9839***	0.0782**	33.1518	45.7558	37.1581	40.7305	35.2644	2.0039
2	0.8211***	0.9195***	0.9393***	0.9001***	0.8415***	0.0204	39.4181	36.8919	36.5043	42.8361	46.8865	0.7425
3	0.6704***	0.8188***	0.7789***	0.7289***	0.6843***	0.0139	25.7662	22.3695	39.3553	32.6658	31.3575	0.4097
4	0.4871***	0.6022***	0.5783***	0.4534***	0.3008***	-0.1863***	12.5705	16.9884	22.0947	17.1503	9.8398	-3.7745
Big	-0.4258***	-0.1156***	-0.2135***	-0.3554***	-0.3885***	0.0372	-18.3573	-2.8631	-6.3193	-9.9463	-13.6836	1.0159
S-B Spread	1.3314***	1.1131***	1.1927***	1.3566***	1.3724***		37.1532	24.2530	27.8381	31.2782	34.4768	

Table 2.10: (Continued)

$$R_t - R_f = \alpha + \beta_M(R_{Mt} - R_f) + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \varepsilon_t$$

Dependent variable: Returns on 25 stock portfolios formed on Size and Book-to-Market Equity

Size	Book-to-Market Equity (BE/ME) quintiles											
quintiles	Low	2	3	4	High	H-L Spread	Low	2	3	4	High	H-L Spread
	β_{HML}						t-statistics of β_{HML}					
Small	-0.3768***	-0.2021***	-0.0721***	0.0589**	0.0996***	0.4764***	-12.8094	-7.2215	-2.7122	1.8374	2.7795	10.2762
2	-0.3912***	-0.2249***	-0.0647**	0.0179	0.2485***	0.6397***	-15.6780	-6.5026	-2.0893	0.6944	9.3867	17.5843
3	-0.4421***	-0.3175***	-0.1686***	0.0420*	0.2598***	0.7019***	-12.6340	-9.0346	-5.9799	1.7921	9.9740	16.0900
4	-0.5842***	-0.3770***	-0.1561***	0.0595*	0.3310***	0.9152***	-12.4394	-9.5288	-4.7147	1.8519	9.5763	15.6949
Big	-1.0173***	-0.3341***	-0.0389	0.4236***	0.6524***	1.6697***	-31.4302	-8.0997	-0.9602	9.8021	17.0941	33.3656
S-B Spread	0.6405***	0.1320***	-0.0333	-0.3646***	-0.5528***		14.6435	2.6479	-0.6875	-6.7751	-10.5600	
	Adjusted R^2						Standard Error of regression					
Small	0.9407	0.9519	0.9435	0.9433	0.9185		0.4067	0.3966	0.4343	0.4447	0.5402	
2	0.9436	0.9527	0.9509	0.9611	0.9651		0.4080	0.4000	0.4141	0.3709	0.3445	
3	0.9228	0.9410	0.9546	0.9579	0.9582		0.4842	0.4469	0.4032	0.3880	0.3747	
4	0.8939	0.9237	0.9471	0.9431	0.9376		0.5441	0.5002	0.4251	0.4458	0.4433	
Big	0.9227	0.9192	0.9203	0.9064	0.9214		0.4146	0.5166	0.5134	0.5139	0.4240	

*** Significance at the 1% level
** Significance at the 5% level
* Significance at the 10% level

Note: The t-statistics reported are calculated from standard errors corrected for serial correlation and White (1980) heteroskedasticity using the Newey-West (1987) estimator. The rule for choosing the lag parameter q is that $q = 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression (see Stock and Watson, 2007, page 607).

The S-B Spread reports the difference of estimations between the small size portfolio and big size portfolio for each quintile of BE/ME. Similarly, the H-L Spread reports the difference of estimations between the high BE/ME portfolio and low BE/ME portfolio for each quintile of size.

2.4.4 GRS tests for three models in different periods

In the three previous subsections, we conduct the time-series regressions of within the full sample period and two sub-periods. To better compare the performance of each model in different periods, we conduct GRS F-tests to examine whether the intercepts for the 25 portfolios are jointly zero in the three models within each period. The results are reported in the three panels of Table 2.11. Panel A of Table 2.11 reports the average values of intercepts, the GRS statistics with associated p-values, and the average adjusted R^2s in the full sample. The CAPM produces the largest average alpha of 0.0336 and the smallest average adjusted R^2 of 88.93%. The associated GRS statistic and the p-value suggests that the null hypothesis that the intercepts are jointly zero for the CAPM has to be rejected at very low significant level. On the other hand, both the FF model and the five-factor model produce small absolute values of mean alphas and high average adjusted R^2 over 94.8%. Although the GRS tests reject the hypothesis of joint-zero intercepts for both models at the 10% level, the five-factor model shows a relatively more significant GRS statistic making it rejected at the 5% level. Remember that the pricing model's alpha represents the pricing error of the model. Thus, the GRS tests suggest that, among the three models, the FF model produces the least significant pricing errors, while the CAPM produces the most.

The GRS F-tests results within the first sub-sample representing a downward market are reported in Panel B. Again, the CAPM produces the largest positive mean alpha and the lowest average adjusted R^2 with a increase to 92.43%. The average alphas for the other two model are small and the average adjusted R^2s are over 96.6%. According to the GRS tests, the null of joint-zero intercepts cannot be rejected at the 10% level for all three models in the downward market. Even so, the FF model presents the least GRS statistic associated with the highest p-value. Thus, the results suggest that the FF model produces the least pricing errors.

Panel C reports the GRS test results in the fluctuating market. The CAPM still produces the largest alphas and lowest adjusted R^2. However, for all three models, the adjusted R^2s are less than those in both Panels A and B. The GRS tests reject the intercepts of the CAPM being jointly zero at the 5% level, but cannot reject the null for other two models. Again, the FF

Table 2.11: GRS tests for all three models

Panel A: GRS tests within the full sample			
	CAPM	FF	Five-factor
Mean alpha	0.0336	-0.0014	0.0010
GRS statistic	2.1847	1.3835	1.5801
p-value	0.0006	0.0984	0.0344
Mean adj. R^2	0.8893	0.9482	0.9488
Panel B: GRS tests within the downward market			
	CAPM	FF	Five-factor
Mean alpha	0.0558	0.0043	0.0029
GRS statistic	1.3840	0.9485	1.0541
p-value	0.1092	0.5378	0.3968
Mean adj. R^2	0.9243	0.9662	0.9664
Panel C: GRS test within the fluctuating market			
	CAPM	FF	Five-factor
Mean alpha	0.0293	-0.0056	-0.0030
GRS statistic	1.6528	1.0536	1.1693
p-value	0.0229	0.3916	0.2574
Mean adj. R^2	0.8612	0.9377	0.9386

This table reports the GRS F-test results on the null hypothesis that the intercepts are jointly zero across the 25 size-BE/ME sorted portfolios using the daily return data in the Chinese stock market. Panel A summarizes the results within full sample period from July 2007 to December 2013; Panel B presents the results within the downward sub-period from October 2007 to December 2008 when the market experiences a historical crash; and Panel C within the fluctuating sub-period from January 2009 to December 2013 when the market slowly recovers from the previous crash and fluctuates at low level. The mean alpha in terms of returns are in percent but not annualized.

model produces the least pricing errors.

Comparing the GRS test results in all three panels, we find that the GRS statistics for all three models are least significant and the associated p-values are largest in Panel B, which suggests that all three models perform better in the downward market than in the fluctuating market and in the whole period. Moreover, the FF model produces the least pricing errors among the three models, while the CAPM produces the most.

2.4.5 Fama-MacBeth regressions

In this subsection, we conduct the Fama-MacBeth regressions for all three models in the whole sample period, as well as the two sub-periods. The results for the different samples are summarized in Panels A, B and C of Table 2.12, respectively. As for the whole sample period, Panel A shows that the premiums on the market, size and value factors are all significant at least at the 5% level in the FF model and the five-factor model. The size premium tends to be the most significant one. The magnitudes of all the three factors are close to the mean returns on the risk factors reported in Table 2.1, and the signs are the same. This panel also shows that the intercepts of the three models estimated from the Fama-MacBeth regressions are all significant at the 10% level. But the intercept of the FF model is less significant than the other two. Thus, the Fama-MacBeth regression results suggest that, among the three models considered in this study, the FF model seems the best one in explaining the cross-sectional variation in equity returns, while the CAPM seems to have the least explanatory power.

The results of the two sub-sample Fama-MacBeth regressions are reported in Panels B and C, which are more interesting. Comparing with the mean returns on the risk factors in Table 2.2, the estimated premiums on the market, size and value factors in the FF and five-factor models are generally similar to their counterparts in both sub-samples. But this is not the case for the momentum and turnover factors. Most interesting, we confirm the findings in Table 2.2 that we find insignificant size premium but significant market and value premiums in the downward period and significant size premium but insignificant value premium in the fluctuating period. The estimated market premium is significant in the FF model but insignificant in the five-factor model. The FF model tend to be the best model again in the downward period, but it is not easy to tell which one, the FF or five-factor model is better in the fluctuating period. The results also suggest that the FF model performs best in the crash.

2.4.6 Stabilities of estimated slope coefficients

The evidence from the above time-series regressions, GRS F-test, and Fama-MacBeth regressions largely suggests that the FF model is the best model for pricing equity returns in our sample of the Chinese stock market, and it performs better in the downward market than in the fluctuating market. Due to this different performance in different market conditions, it might be of interest to investigate the stability of the parameters and the model itself over time, and

Table 2.12: Fama-MacBeth regressions for all three models

	Intercept	$R_M - R_f$	SMB	HML	MOM	TO
Panel A: Full sample with 1582 observations						
CAPM	-0.1626**	0.1053				
	(-2.4698)	(1.2522)				
FF	0.1228*	-0.2039**	0.0651***	0.0348**		
	(1.8418)	(-2.4814)	(3.6960)	(2.1985)		
Five-factor	0.1937**	-0.2752***	0.0604***	0.0407**	0.0399	0.0183
	(2.4989)	(-3.0098)	(3.4117)	(2.5265)	(0.5656)	(0.4788)
Panel B: Downward market with 306 obs						
CAPM	-0.5060	0.1065				
	(-2.3827)	(0.3815)				
FF	0.1258	-0.5682**	0.0792	0.0807*		
	(0.5427)	(-1.9691)	(1.5539)	(1.7571)		
Five-factor	0.1543	-0.5910*	0.0684	0.0767*	0.2393**	-0.0200*
	(0.5068)	(-1.6916)	(1.3372)	(1.6760)	(2.0796)	(-0.1853)
Panel C: Fluctuating market with 1211 obs						
CAPM	-0.1200**	0.1163				
	(-1.9736)	(1.5036)				
FF	0.1207**	-0.1453**	0.0676***	0.0188		
	(2.1213)	(-2.0419)	(3.6604)	(1.1709)		
Five-factor	0.0933	-0.1172	0.0684***	0.0183	0.0631	-0.0275
	(1.4561)	(-1.5253)	(3.6957)	(1.1191)	(0.8360)	(-0.9373)

*** Significance at the 1% level
** Significance at the 5% level
* Significance at the 10% level
Note: This table reports the estimated risk premiums using the Fama-MacBeth regressions. The betas obtained from the first-stage time-series regressions are estimated as the betas of 25 size and BE/ME sorted portfolios. We then run cross-sectional day-by-day regressions of portfolio returns on their exposures to $R_M - R_f$, SMB, and HML. Panel A summarizes the results for the full sample, and Panels B and C summarize for the two sub-sample periods representing a downward market and a fluctuating market, respectively.

time-series patterns of changes in them.

Because the five-factor model seems not able to beat the FF model, in this subsection we focus on the FF model, but also consider the CAPM as a benchmark. We then perform recursive regressions of the CAPM and the FF model over the whole sample period with a starting window of 90 days, i.e., approximately a quarter. Based on the recursive estimation, we conduct the instability tests on the CAPM and the FF model based on Hansen's (1992) instability statistics for linear models. Hansen's stability test extends Nyblom's (1989) Lagrange multiplier test for constancy of all parameters to individual coefficients. For the individual coefficient tests, the null hypothesis is

$$H_0\text{: } \beta_i \text{ is constant for } i = 1, \ldots, k \text{ or } H_0\text{: } \sigma^2 \text{ is constant,}$$

and the corresponding test statistic is

$$L_i = \frac{1}{nV_i} \sum_{t=1}^{n} S_{it}^2, i = 1, \ldots, k \tag{2.7}$$

where $V_i = \sum_{t=1}^{n} f_{it}^2$,

$f_{it} = \begin{cases} x_{it}\hat{\varepsilon}_t & i = 1, \ldots, k \\ \hat{\varepsilon}_t^2 - \hat{\sigma}^2 & i = k+1 \end{cases}$ and, by assumption, $\sum_{t=1}^{n} f_{it} = 0$, $i = 1, \ldots, k+1$, and

$S_{it} = \sum_{j=1}^{t} f_{ij}$, $i = 1, \ldots, k+1$.

The limiting distribution of L_i under the null is a Cramer-von Mises distribution:

$$L_i \Rightarrow \int_0^1 B_1^u(\lambda) B_1^u(\lambda)\, d\lambda \tag{2.8}$$

We reject the null hypothesis if L_i is greater than the critical value.

Similarly, for testing the joint hypothesis

$$H_0\text{: } \beta \text{ and } \sigma^2 \text{ are constant}$$

Hansen's Lagrange multiplier statistic is

$$L_c = \frac{1}{n}\sum_{t=1}^{n} \mathbf{S}_t' \mathbf{V}^{-1} \mathbf{S}_t = \frac{1}{n} tr\left(\mathbf{V}^{-1}\sum_{t=1}^{n} \mathbf{S}_t' \mathbf{S}_t\right) \tag{2.9}$$

where $\mathbf{V} = \sum_{t=1}^{n} \mathbf{f}_t \mathbf{f}_t'$,

$\mathbf{f}_t = \left(f_{it}, \ldots, f_{k+1,t}\right)'$, and

$\mathbf{S}_t = \left(S_{it}, \ldots, S_{k+1,t}\right)'$.

Under the null of no-structural change

$$L_c \Rightarrow \int_0^1 \mathbf{B}_{k+1}^u(\lambda)\,\mathbf{B}_{k+1}^u(\lambda)\,d\lambda \tag{2.10}$$

where $\mathbf{B}_{k+1}^u(\lambda) = \mathbf{W}_{k+1}(\lambda) - \lambda \mathbf{W}_{k+1}(1)$, and $\mathbf{W}_{k+1}(\lambda)$ is $k+1$ dimensional Brownian motion. Again, we reject the null if L_c is greater than the critical value.

Hansen's (1992) instability tests are robust to heteroskedasticity. The null distribution is non-standard and depends upon number of parameters tested for stability. Nyblom (1989) and Hansen (1992) provide the critical values computed by simulation.

The values of Hansen's (1992) statistics for joint and individual instability tests for the CAPM and the FF model are reported in Tables 2.13 and 2.14, respectively. For each instability statistic across the 25 portfolios, the number of significant ones are summarized under "No. of Sig.". According to joint instability tests, both the CAPM and the FF model are non-constant over time for all the 25 portfolio. For the tests on each individual coefficient, we find that the constant terms in both models are rather insignificant, only two are merely significant at the 10% level. In the CAPM, both the model's variance and market beta are largely non-constant across the 25 portfolios over time. Although adding the additional size and value does not decrease the level of instability of the FF model as a whole, it does reduce the number of significant market beta from 24 in the CAPM to 11 in the FF model. Besides, only 12 size factor loadings present significant instability at the 10% level. However, the coefficients related to the value factor are generally non-constant.

Hansen's (1992) instability tests are informative about the type of structural change in

Table 2.13: Instability tests for the CAPM

	Small					2					3				
	Low	2	3	4	High	Low	2	3	4	High	Low	2	3	4	High
Joint instability tests:															
	8.07***	6.65***	3.92***	5.49***	2.19***	6.82***	5.61***	4.91***	6.75***	9.36***	3.80***	3.37***	6.13***	6.89***	6.33***
Individual instability tests:															
Variance	4.43***	3.65***	1.99***	3.51***	0.78**	3.62***	3.28***	2.45***	3.85***	8.09***	1.53***	1.08***	1.66***	2.46***	5.58***
Constant	0.12	0.23	0.14	0.33	0.14	0.08	0.08	0.31	0.42*	0.29	0.06	0.08	0.23	0.19	0.10
MR	2.80***	1.43***	1.22***	0.98***	1.15***	3.10***	1.85***	1.21***	1.83***	0.65**	2.16***	2.21***	3.52***	1.63***	0.37*

	4					Big					No. of Sig.			
	Low	2	3	4	High	Low	2	3	4	High	Total	1%	1~5%	5~10%
Joint instability tests:														
	1.83***	3.58***	3.30***	5.53***	7.80***	2.32***	6.54***	8.61***	3.71***	5.62***	25	25	0	0
Individual instability tests:														
Variance	0.06	0.23	0.24	3.21***	7.15***	1.17***	0.68**	2.72***	2.35***	4.41***	22	20	2	0
Constant	0.17	0.12	0.05	0.16	0.11	0.14	0.17	0.32	0.41*	0.12	2	0	0	2
MR	1.47***	3.08***	2.89***	1.77***	0.12	1.00***	5.11***	5.51***	0.66**	2.20***	24	21	2	1

*** Significance at the 1% level
** Significance at the 5% level
* Significance at the 10% level
This table reports the Hansen's (1992) statistics for individual and joint parameters instability tests with the null hypotheses that individual coefficients are constant and that all parameters are jointly constant, respectively. The tests are robust to heteroskedasticity. The null distribution is non-standard and depends upon number of parameters tested for instability. The critical values are computed by simulation in Nyblom (1989) and Hansen (1992).

Table 2.14: Instability tests for the FF model

	Small					2					3				
	Low	2	3	4	High	Low	2	3	4	High	Low	2	3	4	High
Joint instability tests:															
	16.20***	12.70***	4.94***	5.78***	5.77***	8.01***	9.92***	7.51***	9.86***	5.85***	3.56***	8.55***	11.50***	9.23***	3.66***
Individual instability tests:															
Variance	14.80***	8.45***	1.22***	1.36***	0.20	7.30***	4.19***	0.74**	1.73***	0.74**	1.33***	2.09***	4.82***	1.02***	1.74***
Constant	0.13	0.17	0.03	0.19	0.02	0.11	0.04	0.22	0.43*	0.12	0.05	0.08	0.22	0.10	0.07
MR	0.24	0.19	0.31	0.23	0.21	0.23	0.05	0.31	0.55**	0.27	0.15	0.32	1.17***	0.67**	0.05
SMB	0.34	0.23	0.20	0.26	0.35	0.27	0.06	0.14	0.21	0.15	0.13	0.49**	1.90***	0.50**	0.10
HML	0.16	4.43***	3.32***	4.73***	4.07***	0.19	3.27***	5.14***	5.23***	3.87***	1.41***	5.73***	4.07***	3.35***	1.36***

	4					Big					No. of Sig.			
	Low	2	3	4	High	Low	2	3	4	High				
Joint instability tests:											Total	1%	1~5%	5~10%
	8.94***	10.70***	10.70***	5.69***	5.93***	14.70***	8.12***	10.90***	5.25***	8.33***	25	25	0	0
Individual instability tests:														
Variance	1.34***	0.52**	0.66**	0.94***	4.29***	0.49**	0.54**	2.29***	1.80***	6.09***	24	18	6	0
Constant	0.19	0.11	0.07	0.10	0.08	0.24	0.21	0.42	0.39*	0.05	2	0	0	2
MR	0.31	0.62**	0.87***	1.00***	0.22	5.21***	4.78***	7.05***	1.92***	0.38*	11	7	3	1
SMB	0.59**	0.39*	1.94***	0.25	0.40*	2.43***	2.51***	1.02***	0.86***	1.75***	12	7	3	2
HML	4.96***	7.12***	4.79***	1.54***	0.68**	11.40***	0.12	0.34	0.59**	0.63**	21	18	3	0

*** Significance at the 1% level
** Significance at the 5% level
* Significance at the 10% level
This table reports the Hansen's (1992) statistics for individual and joint parameters instability tests with the null hypotheses that individual coefficients are constant and that all parameters are jointly constant, respectively. The tests are robust to heteroskedasticity. The null distribution is non-standard and depends upon number of parameters tested for instability. The critical values are computed by simulation in Nyblom (1989) and Hansen (1992).

models, but the tests are not informative about the date of structural change. Thus, we also plot the graphs of recursive estimations to have a sketchy look at the structural change within different periods. Because the FF model can better explain the equity returns and is less unstable than the CAPM over our sample period, we next only focus on the graphs for the FF model. Also, in order to save the space, we only present the typical graphs of recursive regressions of several portfolios for the purpose of interpretation. According to Table 2.14, for example, the FF model presents the most significant instability as a whole for the smallest size and lowest BE/ME portfolio. Although we cannot reject the null hypothesis that each individual coefficient is constant, the graphs of the coefficients of three factors over the sample in Figure 2.4 show that the coefficients fluctuate more during the shaded sub-period when the Chinese market experiences the historical crash. Further, the graph of 1-step residuals with $\pm 2\hat{\sigma}_t$ plotted on either side of zero shows the major outliers within the crashed period and that the two $\pm 2\hat{\sigma}_t$ lines also enlarge within this period and narrow down after the crash. Note that the residuals outside of the error bars are either outliers or are associated with changes in $\hat{\sigma}_t$. The 1-step Chow tests partially reflects this. The graph of "1up CHOWs" shows that the dates on which the values of the Chow F-test are over its 1% critical value are mostly gathered within the shaded period. Figure 2.5 shows the graphs for the other end of the 25 portfolios, i.e., the biggest size and highest BE/ME portfolio. The graphs present clearer non-constant patterns of the coefficients than those for the smallest size and lowest BE/ME portfolio. The evidence in the graphs of 1-step residuals and Chow tests is not as significant as that in Figure 2.4, but the graphs show major outliers of residuals gathering in the crash and less than one year after the market bottom. More importantly, the two $\pm 2\hat{\sigma}_t$ lines do enlarge within the shaded period.

Generally speaking, the graphs, including those not presented in this study, suggest that the FF model is largely non-constant over time and it tend to present slightly more instability in the crash than in the transition period, although the evidence is not extremely clear for all the 25 portfolios.[22]

[22]For some portfolios, the instability shows up in both downward and fluctuating periods, and hard to tell in

Because recursive regression increases the length of window by one as it moves forward along the time with a new observation added into the regression, the coefficient estimates will be smoothed as the window increases. For robust tests on the above findings, we also run inverse recursive regressions from the last day backward to the first day within our sample. We do find that the coefficients remain higher level of non-constant and the errors of residuals tend to enlarge in the downward market. Thus, the FF model tends to present more instability in the period of market crash.

The instability tests on the FF model based on Hansen (1992) as well as the graphs of recursive regressions provide a sketch for the time-series changes in the model's coefficients and the model itself. A further investigation on the regime changes in the FF model under different market conditions (e.g., the bull market, the bear market and the transition market) should be interesting for future research. A related extension is to examine the behavior of return anomalies under different market conditions, which could provide some explanation for the reversal of factor effects and, for example, the weak value premium found in our sample.

2.5 Conclusion

This study aims to test the Fama-French three-factor pricing model and its extension in the Chinese stock market within a relatively recent sample period after quarterly financial reports of listed companies became widely available in 2007. To distinguish from the construction of factors using the annual financial data in the majority of related literature, we construct the factors based on the firms' quarterly financial data. We believe that reforming the factors at a quarterly frequency can lead the factors to capture the most updated information on firms' operation performance, which could influence the stock prices of listed firms. Instead of examining the models at a level of monthly returns as in most existing studies, we then examine the models using daily portfolio returns. This is because the Chinese investors care more about the short-term investment performance, which is especially the case since the e-finance boom.

---

which period the FF model presents higher level of instability.

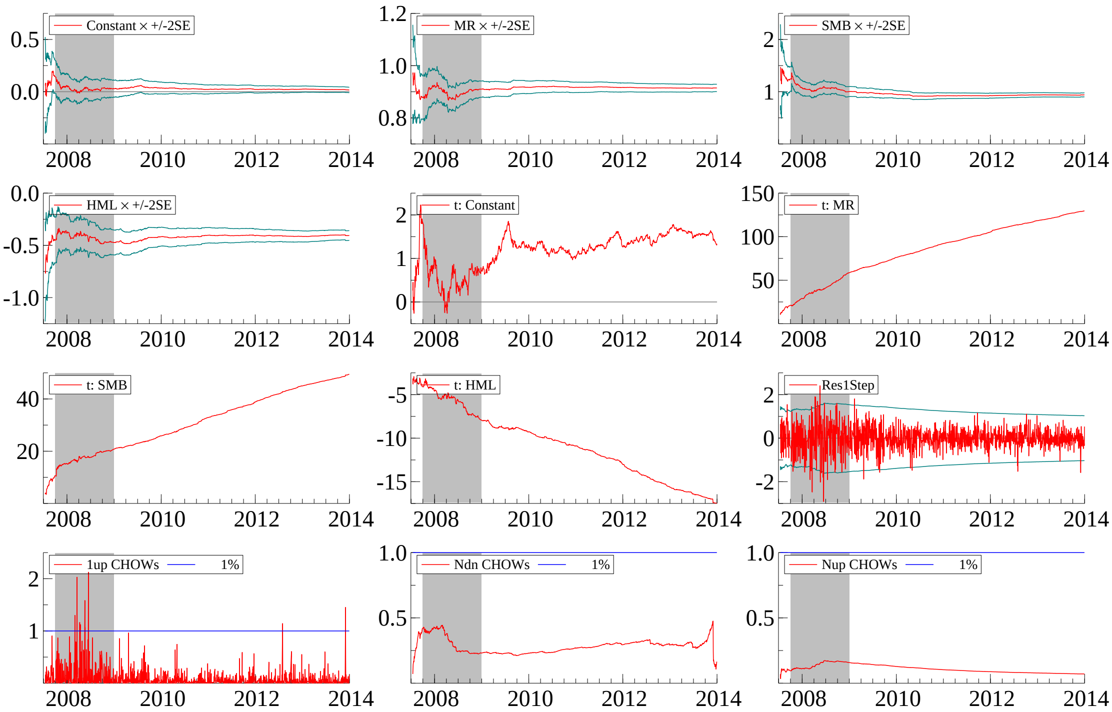

Figure 2.4: Recursive estimations of the smallest size and lowest BE/ME portfolio

Note: Sample period is July 2007 to December 2012. The shaded area represent the first sub-period when the Chinese market experiences a historical crash, after that the market fluctuates at low level. The starting window of the recursive regression is 90 days.

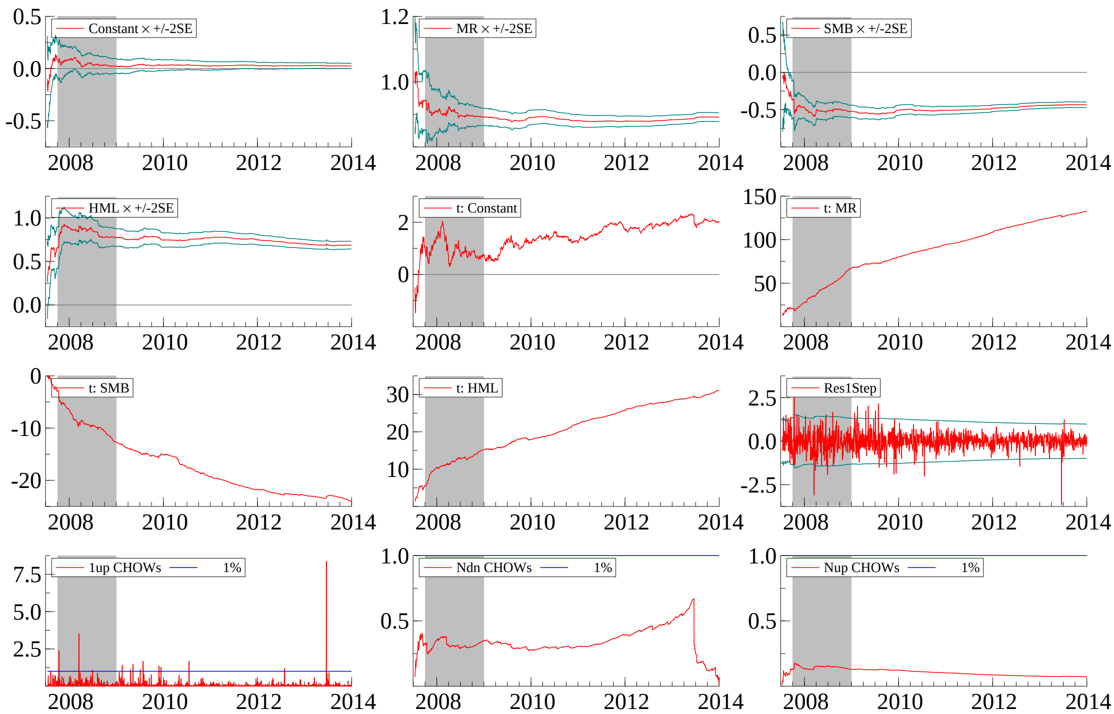

Figure 2.5: Recursive estimations of the FF model

Note: Sample period is July 2007 to December 2012. The shaded area represent the first sub-period when the Chinese market experiences a historical crash, after that the market fluctuates at low level. The starting window of the recursive regression is 90 days.

Nowadays, almost all the mutual funds reveal their daily returns to customers, even though most funds have a much longer investment periods. Although using the daily returns, instead of monthly returns, could introduce more noise, we believe it can include more information related to pricing assets. We choose more information over more noise between the trade-off. According to our unreported time-series regression results, using daily returns produces slightly higher adjusted R^2 than using monthly returns by 1-2% during the same period as our sample in this study for both China and the US. But the conclusion is opposite in the much earlier period as in Fama and French (1993).

The results of full-sample time-series regressions suggest that the Fama-French model performs better than the CAPM in explaining equity returns, with an average adjusted R^2 over 94.8%. Adding the additional momentum and turnover factors seems not able to further improve the explanatory power. By comparing the unreported regression results for different periods in both China and the US, we can see that the adjusted R^2 increase dramatically in the recent period, and this increase mainly stems from the market factor. This might suggest that the information diffusion over the market in nowadays is much more speedy with little obstruction. We highlight this here for future research.

Our sample witnesses the historical crash of the Chinese market from October 2007 to November 2008. After the crash, the market fluctuates at low level. Thus, it provides the opportunity to examine the pricing models under different market conditions. As far as we are aware, this study is the first to examine the FF model in different market trends. By splitting our sample into two sub-periods representing two different market trends, the results suggest that the pricing models seems to perform better in the downward period, and among the three models, the FF model continue to perform well in both sub-periods. The findings inspire us to further examine whether the FF model performs better in the crash than in the transition period.

The GRS F-tests confirm that the FF model is superior to the CAPM and the five-factor model because it produces the least pricing errors in both full sample and sub-samples. Moreover, the FF model also performs better in the crash, according to the least significant GRS statistic for the FF model in the first sub-period.

The Fama-MacBeth regressions also suggest the FF model performs better in the crash. But from the results of the Fama-MacBeth regression, the evidence for whether the FF model is better than the five-factor model is less clear. Because the Fama-MacBeth approach estimates the factor premiums, we also compare our estimated factor premiums in different sample periods with those found in other studies. We found significantly negative estimated market premium, which is consistent with the fact that the Chinese market crashes in our sample and remains at low level. As for the size and value premiums, most studies find both are significantly positive in China, e.g., Chen et al. (2010), Cakici et al. (2015) and Carpenter et al. (2015); while some studies, including Wang and Xu (2004) and Chen et al. (2015), find strong size effect but no value effect in the Chinese market. By splitting our sample into two sub-samples representing the downward and fluctuating market trends, our interesting findings are that there is no size effect but significant value effect in the crash and it is opposite situation in the transition period. The inference of our findings could be that in the crash both big and small firms suffer the poor economic and financial conditions to the same extent, but value firms are more easy to suffer the poor conditions than growth firms. Thus, investors require higher returns to hold the stocks of value firms over growth firms. Therefore, our findings suggest that the premiums on risk factors could vary with the market or economic conditions over time. Our findings also partially support the so-called reversal of factor effects in different periods, which could be due to the changes in market conditions.

Although the FF model presents very good performance in explaining equity returns, the instability tests based on Hansen (1992) show that the model is hardly to be constant over time. The graphs of recursive regressions provides a sketch of the time-series changes in the model's coefficients and the model itself. The graphs show that the model tends to present more instability in the crash. This finding seems to conflict with its better performance in the downward market. The possible explanation is as follows. The better performance of the FF model in explaining equity returns (i.e., the higher adjusted R^2 and the less pricing errors) in the crash is mainly attributed to the co-movements of asset prices. In a tendency market, the prices of most stocks tend to move together in the same direction. This is consistent with Morck et al. (2000), where they document that high R^2 in a market model tending to show up in emerging markets is due to the synchronous stock price movements. Thus, when stock prices move together, the information related to pricing assets diffuses more easily and is widely known to investors over the market. On the other hand, in the crash, the extent to which the price each stock would move is largely different from one another. Although moving in the same direction, the different magnitudes of price movements causes the enlarged volatility of returns on the risk factors as shown in Table 2.3. The more volatile the risk factors are, the less constant the associated factor betas should be. As a result, the models present more instability in the crash than in the transition period.

The interesting findings of this study can be summarized as follows. First, the pricing models present different performance in explaining equity returns in different trends of market. Second, the significance, magnitudes and even signs of factor premiums could change under different market conditions and the volatility of factor premiums tends to enlarge in the downward market. Third, the instability of pricing models, both individual parameters and the models themselves, varies over time and tends to present higher level of instability in the crash. Based on the above findings, this study leaves open questions for future research, i.e., how the factor effects would behave in different market conditions, how the behavior of the risk factors could affect the performance of pricing models in different trends of market, and most importantly, what are the fundamental sources potentially causing the different behavior of risk factors. Answering the last question would be helpful for figuring out the sources of fundamental risks.

Chapter 3

A Simplified Method to Construct the Fama-French Factors using the "Off-the-shelf" Chinese Style Indexes

3.1 Introduction

The capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965a,b), and Black (1972) assumes that a positive and linear relationship exists between the expected excess return on stock or portfolio and its market beta or factor loading, i.e., the slope coefficient in the regression of a stock's excess return on the market portfolio's excess return. If the CAPM correctly prices the expected stock or portfolio excess return, then the intercept term in the regression should also be zero.

Fama and French (1992) employ a two-stage methodology in which the main test involves the Fama and MacBeth (1973) cross-sectional regression technique. They find that the market beta has little or no power to explain cross-sectional variation in equity returns, while the other two factors related to fundamental variables of the size (or market value) and the ratio of book value to market value of equity (BE/ME) play important roles in pricing the cross-sectional returns. Based on these findings, Fama and French (1993) move to a time-series regression framework and propose a three-factor model comprising a market factor, a size factor, and a BE/ME factor to explain the variation in stock returns. They find that the expected stock excess return is linearly related to the exposures to the factors. Meanwhile, the significant incremental improvement in the explanatory power of their three-factor model over the CAPM has made their model popular. Numerous follow-up studies, such as Grinold (1993), Davis (1994), He and Ng (1994) and Fama and French (1996, 1998), provide further supportive evidence for Fama and French's original findings. In their recent work, Fama and French (2015a,b) find that a five-factor model which includes the original Fama-French three factors and two additional factors of profitability and investment performs even better than the original three-factor model and that the book-to-market equity factor becomes redundant in the sense that its average return is fully captured by its exposures to the other factors of the five-factor model. But it does not imply that there is no value premium. The two additional factors in their five-factor model are relatively new discoveries and the research of these factors in different markets and time periods is still limited. Their finding may be specific to the period and country examined.

There are several explanations that try to understand the impressive performance of the Fama-French model in explaining the cross-sectional variation in excess returns on stocks and portfolios. Those different explanations lead to intense controversies[1] on the Fama-French model and most are concentrated on the value premium. Some common explanations are summarized as follows. The first explanation attributes to sample selection problem by arguing that the Fama-French three-factor model is a result found by chance, and it is unlikely to be observed out of sample (e.g., Black, 1993 and MacKinlay, 1995). Kothari et al. (1995) also show that the value effect presents a considerable change in its strength over time. However, major studies provide evidence against the sample-specific explanation. For example, Chan et al. (1991), Capaul et al. (1993), Davis (1994), Fama and French (1998), and Davis et al. (2000) document strong relations between average return and size and value premiums in markets across the world and over different time periods.

The second one is the behavior-bias explanation saying that the size and value premiums stem from overreactions of investors to firms' performance. For example, the big (large market capitalization) or growth (low book-to-market equity ratio or BE/ME) firms are generally good on financial fundamentals, while the small-cap or value (high BE/ME) firms are associated with poor fundamentals. Investors tend to overreact to firms' financial status. As a result, irrationally low values are assigned to poor firms while high values to good firms. When the overreaction is corrected, high returns are generated to stocks of poor firms and low returns to good firms. Some selected papers for this behavior-bias explanation are DeBondt and Thaler (1987), and Lakonishok et al. (1994).

The third explanation comes from Fama and French. In sequential papers, Fama and French (1993, 1995, 1996, 1998) claim that the additional size and value premiums are not anomalies at all. These premiums are just compensation for risk in a multifactor version of Merton's (1973) intertemporal capital asset pricing model (ICAPM) or Ross's (1976) arbitrage pricing theory (APT). Thus, they attribute the success of the Fama-French model to the fact

[1] Faff (2004) provides a literature review summarizing the controversy and debate on the Fama-French model.

that returns on the size and value effects are proxies for systematic risk factors.

The last explanation is suggested by Daniel and Titman (1997). They argue that the size and value premiums are actually not risk factors. In fact, the expected returns of assets are directly related to their characteristics for reasons, such as behavioral biases or liquidity. An example is that investors like growth stocks but dislike value stocks. This would result in a value premium, but it may have nothing to do with the systematic risk. The above example of overreaction can also be viewed as a variant of this explanation. However, Davis et al. (2000) argue that the evidence in favor of the characteristics model of Daniel and Titman (1997) is special to the rather short sample period. In contrast, they show that in a much longer period, the risk-based explanation for the size and value premiums is more appropriate for the success of the Fama-French model.

The Chinese market has just become a focus of empirical studies in recent years after the rapid development of China's economy and stock market. Thus, the related literature for the Chinese market is limited compared to other developed markets. Among current studies, the existing empirical evidence in the Chinese market generally conclude that the market factor itself cannot explain the cross-sectional variation in equity returns, while adding the Fama-French factors largely improve the model's explanatory power in terms of higher adjusted R^2 and less significant pricing error (alpha). As for the factor premiums, different significance, magnitudes or even signs have been found over different sample periods. For example, in the recent study on the Chinese market by Chen et al. (2015), the market excess return is positive but insignificant. But in the sample period of Chapter 2, the market premium is negative and slightly significant. We also find a strong size effect but a weak value premium in Chapter 2, which is consistent with, for example, Wang and Xu (2004) and Chen et al. (2015), but contradicts with most studies on the Chinese market, e.g., Chen et al. (2010), Cakici et al. (2015) and Carpenter et al. (2015), where both strong size and value effects are found. We also find that the factor premiums can change in different periods in Chapter 2. The insignificant or negative factor premiums in some specific periods do not necessarily mean no factor premiums in the market, considering that the risk loadings on the three factors are almost all significant over different sample periods. We would more like to attribute the insignificant or negative factor premiums to the so-called “reversal of factor effects” (see, e.g., Gompers and Metrick, 1998, Dimson and Marsh, 1998, and Gustafson and Miller, 1999) that factors can have different rewards or no rewards under different market conditions as concluded in the previous chapter.

Although the existing evidence tends to support the Fama-French model as superior factor pricing model over the CAPM, one feature that makes it less appealing than the CAPM is the difficulty and complexity to construct the size and value factors, as well as the availability of required data. It is the fact that there is a wide range of indices readily available as a reasonable proxy for market portfolio, which enhances the charm of the CAPM on the grounds of simplicity. Unlike the market factor, however, there are no analogous proxies for the size and BE/ME factors either directly observable or publicly accessible. This would lead to a non-trivial challenge for applying the Fama-French model to other countries outside of the US, which is especially the case for unsophisticated individual investors. This would also limit the application and examination of Fama-French model if the researchers cannot have the readily prepared data for the size and BE/ME factors to hand. Due to this shortcoming, Faff (2001) proposes a method of constructing simple proxies for the size and BE/ME factors by using four “off-the-shelf” commercial style indices for the Australian market. These indices are produced by the Frank Russell Company using the Australian Stock Exchange (ASX) data. Faff (2003, 2004) further examines the Fama-French model where the factors are constructed from style indices for the US and Australian markets and Pham (2007) presents the Japanese evidence. They all claim that the proxies for both size and BE/ME factors constructed from style indices have similar properties of the sophisticatedly constructed Fama-French factors and that the evidence obtained from the formal time-series tests is quite favorable to the model. When taking into account the estimated risk premiums based on a system application of the Generalized Method of Moments (GMM) technique, however, the support for the Fama-French model turns out to be less persuasive.

Faff’s method provides a way to simplify the construction of the Fama-French factors.

This method makes the Fama-French model more approachable if the style index-formed Fama-French factors have at least a similar performance to those constructed using Fama and French's (1993) original method. But, the simplicity of this method is not the only reason for us to conduct this study. Instead, we aim to examine the performance of the Fama-French model in which the factors are constructed from the commercially available indices. Cremers et al. (2012) show that in the US the alpha of the Fama-French model primarily arises from the disproportionate weight that the Fama-French factors place on small value stocks, which have higher average returns. The commercial indices for the Chinese market generally exclude the smallest capitalization firms and those with poorest financial fundamentals (thus, the distressed value firms) as the constituents. If the finding of Cremers et al. (2012) is also applied in the Chinese market, then we should expect smaller pricing errors when we choose index-formed factors over the sophisticate counterparts. Thus, our main purpose is not just to extend Faff's work of examining the performance of the style index-formed Fama-French model in the Chinese market, but more importantly, to provide evidence for whether the Fama-French model using the style index-formed factors has at least similar properties to the model using sophisticatedly constructed factors. If this is the case, then the style index-formed factors will be better alternatives to the sophisticate ones. Besides, there are other benefits of using the index-formed factors. Unlike proxies for the Fama-French factors provided by, for example, the RESSET database, the index-formed factors are constructed from the publicly accessible commercial style indices, thus available for every one. Also, the commercial style indices can produce daily updated proxies for the two factors, thus one can apply the Fama-French factors up to the most recent day. Therefore, the dramatically decreased complexity and the above benefits of this simplified method to construct the Fama-French factors will enhance the charm of the Fama-French model not only in the academy but also in practice.

It is very meaningful to make the application of the Fama-French model much easier in financial practice. We can consider this in an example of hedging. As we present in Chapter 2, the Fama-French model can explain on average about 90% of expected returns, which is 10%-20% over the CAPM. Thus, the hedging strategies based on the Fama-French model can be more efficient than those based on the CAPM in the sense that less variance in expected returns (alphas) cannot be hedged by the Fama-French model. However, Fama and French's original method is complex and time-consuming. It is necessary to deal with all the listed stocks, which makes hedging based on this method cost-inefficient. Suppose that an investor holds a portfolio of small-cap and value stocks. If she expects the market condition will change against the performance of small firms but in favor of value firms, this would suggest a risk exposure on size effect. The investor wants to avoid this risk. One way is to hedge the risk only related to the size effect by short-selling a portfolio which best captures the size effect. For Fama and French's original method, the investor needs to short-sell all stocks falling in three small size groups and takes a long position in those falling in three big size groups. But, by applying Faff's method, our investor only needs to take short and long positions in the corresponding style indices; provided the existence of products based on those style indices. Even if there is no such style index product, taking short-long positions in constituents of style indices is still much easier.[2] We can extend the strategy in the above-mentioned example to a broader category of strategies in portfolio management, i.e., the "smart beta"[3], a latest popular jargon from the fund-management industry. Basically, smart beta strategies aim to beat a cap-weighted index and improve a portfolio's risk-adjusted return over the long term by tracking multi-factor indices based on multiple empirically rewarded factors and utilizing multi-weighting strategies which potentially maximize diversification.

To our best knowledge, there is no published study examining whether the simple method of constructing Fama-French factors is suitable for the Chinese stock market. This might be partially due to the fact that the Chinese stock market is so young that the system of Chinese stock indices is limited. The existing indices mainly focus on the listed stocks with good performance, and by selection, the creation of the majority of indices filters the smallest and poorest performed stocks from being the constituents of indices. Thus survivorship bias and

[2]Here, if Faff's method is generally valid, we may expect a boom of financial products based on such style indices, such as funds tracking style indices (e.g. ETF's) and futures with underlying assets based on style indices, to name a few.

[3]Smart beta is also known as “factor-based investing”, “advanced beta” or “enhanced beta”.

success bias widely exist in almost all indices, which might lead to a lack of representativeness of the overall market index. In Faff (2001, 2003, 2004) and Pham (2007), they claim that the proxies for both size and BE/ME factors constructed from style indices have the similar properties of the sophisticated construction of Fama-French factors, but they do not provide a direct comparison between the two methods in their work. We believe that a direct comparison would give the investors certain amount of confidence in applying the Fama-French model based on the simple method of constructing the size and BE/ME factors. As a by-product, our tests provide a hint for whether the surviorship or success bias potentially explain the success of the Fama-French model as an early controversy concerned by Kothari et al. (1995).

The aims of this study are therefore to 1) construct the simple proxies for the Fama-French size and BE/ME factors from a set of style indices with the widest coverage on the Chinese A-share market; 2) show that all three index-formed factors adequately capture the features of the corresponding sophisticatedly constructed factors by running regressions of returns on index-formed factors on returns on sophisticate factors; 3) provide a direct comparison of the performances of the Fama-French model using both methods of constructing factors based on formal individual regressions of returns on industry portfolios on both sets of factors; 4) provide a GMM system regression analysis as a superior alternative to the standard Fama-MacBeth regression, which allows us to estimate the risk loadings and premiums at the same time; 5) conduct the robustness test by examine whether the model's alpha for common benchmark indices is closer to zero when the index-formed factors are used. Based on our findings, we confirm that the index-formed Fama-French factors are superior to those constructed from Fama and French's original method.

The remainder of this study is structured as follows. The next section provides a brief description of the common indices, especially the style indices, available to investors in the Chinese stock market. Section 3.3 shows the method to construct the proxies for size and BE/ME factors and run the regressions of index-formed factors on sophisticate factors. Section 3.4 presents the empirical results of a formal time-series regression and a GMM based test on a system of the restricted Fama-French model. Section 3.5 concludes the study.

3.2 Description of related indices and data issues

There are (at least) four sets of indices accessible for Chinese stock investors, namely, CSI, CNI, SSE and SZSE indices. The first two cover the A-share market of both the Shanghai Stock Exchange (SSE) and the Shenzhen Stock Exchange (SZSE). SSE indices focus only on the A-share market of SSE and SZSE indices only on the A-share market of SZSE (including the Main Board, SME board and ChiNext market). In this study, we choose six CNI style indices to construct the proxies for Fama and French's size and BE/ME factors because these indices are free and publicly accessible, i.e., "off-the-shelf", and cover both SSE and SZSE.[4] The CNI1000 index will be used as the market index because the six CNI style indices are constructed from the constituents of CNI1000 index. Thus, all the indices allow the existence of survivorship bias and success bias to construct the index-formed factors. Next, we provide a brief discussion of the related indices.[5]

3.2.1 Scale (size) and style indices

The "CNI Indices", developed by the Shenzhen Securities Information Co., Ltd, is a cross market index system, which covers both Shenzhen and Shanghai security markets and contains about 50 indices categorized as "CNI Core Indices", "CNI Scale Indices", "CNI Style Indices", "CNI Thematic Indices", "CNI Region Indices", "CNI Sector Indices", "CNI Strategy Indices" and "CNI Composite Indices". Although our focus is on the CNI Style Indices in this study, we first describe some of the CNI Scale Indices, because the style indices are constructed after the construction of the corresponding scale indices. The CNI 1000 Index is a free-float capitalization-weighted index [6] and comprises 1000 largest and most liquid A-share stocks listed and traded on the Shanghai Stock Exchange and the Shenzhen Stock Exchange

[4]Although the CSI indices also cover both SSE and SZSE, but the CSI small size style indices are not free and accessible unless from a database with certain amount of fees. Usually, this amount of fees is too high for an individual investor.

[5]For more detailed information of the four sets of indices, please refer to the index providers' websites: http://www.csindex.com.cn and http://www.cnindex.com.cn .

[6]The free float shares are the amount of tradable shares outstanding in the open stock market. It is defined as all negotiable shares of the company minus the negotiable shares held by the 3 types of shareholders (government holdings, strategic holdings, or long term holdings by founders, families or senior executives) with holding proportion exceed 5% of the company's total shares.

(including Shenzhen Main Board, SME Board and ChiNext Market). All A-share stocks[7], excluding those identified as ST (Special Treatment) and *ST[8], listed for more than 6 months (unless its total market capitalization and free-float market capitalization comprehensively ranks within the top 10 among all stocks listed on Shanghai Stock Exchange and Shenzhen Stock Exchange) on the Main Board of Shanghai Stock Exchange (SSE) and Main Board, SME board and ChiNext market of Shenzhen Stock Exchanges (SZSE) and with no penalty by the China Securities Regulatory Commission (CSRC) in the most recent year are eligible for the candidate constituents in the CNI 1000 Index. The top 1000 of these eligible stocks, ranked by the arithmetic average of three ratios – the daily average total market capitalization ratio, free-float market capitalization ratio and turnover ratio of each eligible stock in the most recent 6 months, are selected as the constituents of the CNI 1000 Index. The CNI 1000 Index represents approximately 78% of the total market capitalization of the A-share market and is considered as a comprehensive barometer of the large and middle-cap segment of the A-share universe.

The constituents of the CNI 1000 Index, ranked using the same rule as in the selection method for the CNI 1000 Index, are further segmented into three size indices: the CNI Large Cap Index, the CNI Middle Cap Index, and the CNI Small Cap Index. Specifically, the top 200 stocks occupying about 51.96% of the A-share market capitalization constitute the CNI Large Cap Index, whilst the middle 300 stocks, about 13.64% of the A-share market capitalization, make up the CNI Middle Cap Index. The bottom 500 stocks, around 13.57% of the A-share market capitalization, compose the CNI Small Cap Index. The constituents of the CNI Large Cap, Middle Cap and Small Cap indices are the index universe of the corresponding style indices. The growth investment style is captured by three variables which are internal growth rate, income growth rate and net profit growth rate. A growth score for each stock is calculated based on these three variables. The value investment style characteristic is defined by four variables: D/P, B/P, CF/P and E/P, and a value score for each stock is computed based on

[7] Convertible preference shares and loan stocks are excluded until converted.

[8] ST and *ST represent those stocks issued with a delisting-risk alert or the alert for 'special treatment for other reasons' by the SSE and SZSE.

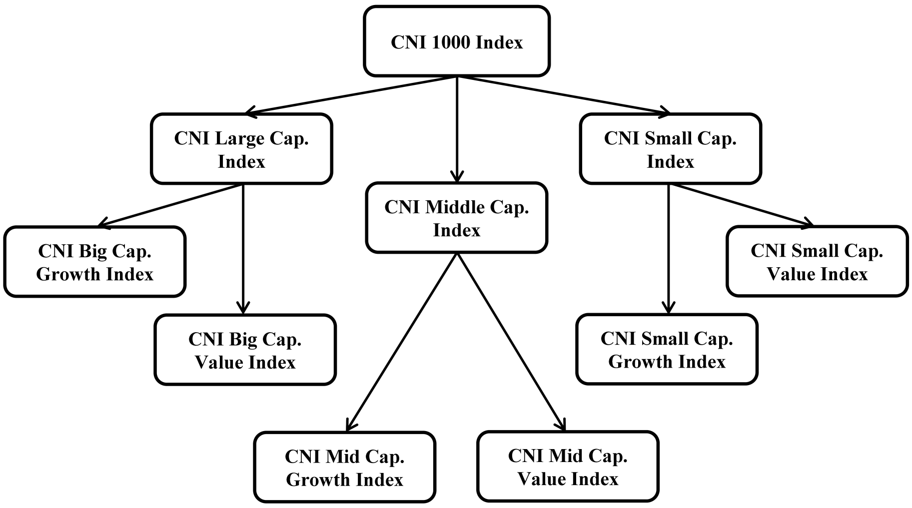

Figure 3.1: Relationship between CNI size indices and style indices

these four variables. The stocks are then ranked by growth scores and value scores, respectively. Then a certain number of stocks in one of the above three CNI scale indices with top growth scores will be selected as the constituents of growth style index; and similarly, the same number of stocks in one of the three CNI scale indices with top value scores will be selected as the constituents of value style index. As a result, there are 66 stocks for the CNI Large Cap growth and value indices, respectively; 100 for the CNI Middle Cap growth and value indices, respectively; and 166 for the CNI Small Cap growth and value indices, respectively. Figure 3.1 draws a relationship between the scale/size indices and the style indices.

Considering that the number of Chinese A-share listed stocks is currently over 2,500, it is obvious that, by construction, the constituents of the CNI style indices exclude delisted stocks and more than half of the listed stocks whose sizes are very small and/or who have the poorest performance in the A-share market. Thus, we should keep in mind that both survivorship bias and success bias must exist in those indices.

3.2.2 Sector indices

We choose the CSI All Share sector indices as the proxies for the sector portfolios. The CSI indices are provided by the China Securities Index Company Limited. The CSI All Share Index consists of all the A shares, except for the ST and *ST stocks and the temporarily suspended stocks from trading, satisfying the following condition that the listing time of a stock is more than three months unless the daily average total market value of a stock since its initial listing is ranked top 30 in all the A shares. According to the latest Shanghai Stock Exchange Statistics Annuals 2015 and 2014, there are 24 out of 986 and 31 out of 944 A-Share stocks listed on the Main Board of SSE being marked as ST or *ST stocks by the ends of 2014 and 2013, respectively; whilst the latest Shenzhen Stock Exchange Fact Books 2014 and 2013 show that there are 22 out of 1606 and 29 out of 1524 A-Share stocks listed on the Main Board, SME Board and ChiNext Markets of SZSE being marked as ST or *ST stocks by the ends of 2014 and 2013, respectively.[9] The numbers of ST and *ST stocks are generally negligible in each year, compared to the sizes of both SSE and SZSE A-Share markets.

The constituent stocks of the CSI All Share Index are classified into ten sectors[10] which are energy, materials, industrials, consumer discretionary, consume staples, health care, financials, information technology, telecommunication services and utilities.[11] For sectors that consist of less than or equal to 50 stocks, all the stocks enter into the sector index; while for sectors that consist more than 50 stocks, the stocks in each sector will firstly be ranked by the daily average trading value and the daily average total market capitalization, respectively, in descending order, then the stocks at the bottom 10% by the trading value and stocks at the bottom 2% by the cumulative market cap will be deleted (the deletion process will be terminated if the number of stocks decreases to 50), and the remaining stocks enter into the sector index.[12]

Choosing the CNI All Share sector indices is because it has the widest coverage over the

---

[9] The Shanghai Stock Exchange Statistics Annuals are downloadable at http://www.sse.com.cn/researchpublications/publication/yearly/; whilst the Shenzhen Stock Exchange Fact Books are available at http://www.szse.cn/main/marketdata/wbw/marketstat/, but the volumes earlier than 2012 are partially publicly accessible.

[10] This ten-sector classification is sometimes called the level-one classification and is popular for constructing the sector indices by many index-providing companies.

[11] The names (codes) of these 10 CSI All Share Sector Indices are CSI All Share Energy (000986), Materials (000987), Industrials (000988), Consumer Discretionary (000989), Consumer Staples (000990), Health Care (000991), Financials (000992), Information Technology (000993), Telecommunication Services (000994), Utilities (000995).

[12] An examination of the FF model in line with the original idea of Fama and French (1993) might use the CNI 1000 Index as the proxy for market factor, and testing the FF model on CNI 1000 Sector Indices. But this is not the end of this study. Recall that this study tries to find if the available style indices which cover a small part of the market in terms of number of stocks can be used to evaluate the overall market performance. This may be more meaningful for financial practice.

A-share stocks of both Shanghai and Shenzhen Stock Exchanges. Thus, the CSI All Share sector indices should be a good representative for the Chinese A-share market. Of course, by construction, survivorship bias and success bias also exist in these indices.

3.2.3 Common benchmark indices

The benchmark indices to be examined in the robustness tests are 23 commonly mentioned indices in the Chinese market. To save space, we only provide some brief descriptions of those indices in Table 3.1 without detailed discussion. The CNI1000 and its subsequent scale and style indices are discussed in more detail in the first subsection before.

3.2.4 Risk-free rate

The overnight Shanghai Interbank Offered Rate (Shibor) plays the role of risk-free rate in this study. The data can be downloaded from the China Foreign Exchange Trade System & Nation Interbank Funding Center.[13] The reasons we choose the over-night Shibor as the risk-free rate are as follows. First, this study examines the Fama-French model at the daily level, thus we need a proxy for the daily risk-free rate of return. Unlike other developed countries where the one-month Treasury-bill rate is available, the shortest term of Chinese government bond/bill is three months. The three-month rate might not be a good candidate to be transferred to the daily risk-free rate. Second, the data for Chinese government bonds are not publicly available, whilst the Shibor is publicly accessible. Third, the Shibor is the market based risk-free rate and the benchmark for China's money market. For each business day, the Shibor is released before the noon. This means that although it fluctuates every business day, it is common knowledge for every market participant no later than the noon of each day. Last, but also importantly, Shibor is also the risk-free rate in the RESSET database since its initiation on 8th October 2006. The candidate for the risk-free rate in the RESSET database was the three-month deposit rates from 1st February 1989 to 6th August 2002, and then changed to the

[13]URL: http://www.shibor.org/. Both Chinese and English versions of this website are available.
Shibor's Code of Conduct can be referred to http://www.shibor.org/shibor/web/html/sszz.html (Chinese version) or http://www.shibor.org/shibor/web/html/sszz_e.html (English version).

Table 3.1: Descriptions of 23 common benchmark indices in the Chinese market

Short Name	Full Name	Brief Description
SSE	SSE Composite	Its constituents are all listed stocks (A shares and B shares) at Shanghai Stock Exchange. One of the most well-known indices in the Chinese A-share stock market.
SSE50	SSE 50	It consists of 50 largest and most influential stocks of good liquidity and representativeness from Shanghai security market.
SZSE	SZSE Component	It comprises the 500 largest and most liquid A-share stocks listed on the Shenzhen Stock Exchange and represents the performance of the multi-tier (i.e., main board, SME board and ChiNext market) Shenzhen stock market. One of the most well-known indices in the Chinese A-share stock market.
CSI300	CSI 300	It consists of 300 stocks with the largest market capitalization and liquidity from the entire universe of listed A-share companies in China.
CSI500	CSI 500	Excluding the constituents of CSI300, it consist of the top 500 stocks in the rest universe of stocks listed on both Shanghai and Shenzhen exchanges.
CSI800	CSI 800	It consists of CSI300 and CSI500, and reflects the performance of the large, mid, and small-cap companies in Shanghai and Shenzhen securities market.
CSI300G	CSI 300 Growth	It consists of 100 stocks with top growth scores from the constituents of CSI300.
CSI300V	CSI 300 Value	It consists of 100 stocks with top value scores from the constituents of CSI300.
CSI500G	CSI 500 Growth	It consists of 150 stocks with top growth scores from the constituents of CSI500.
CSI500V	CSI 500 Value	It consists of 150 stocks with top value scores from the constituents of CSI500.
CSI800G	CSI 800 Growth	It consists of 250 stocks with top growth scores from the constituents of CSI800.
CSI800V	CSI 800 Value	It consists of 250 stocks with top value scores from the constituents of CSI800.
CNI1000	CNI 1000	It consists of top 1000 stocks listed on both Shanghai and Shenzhen exchanges.*
CNILarge	CNI Large Cap.	It consists of the stocks ranked top 200 from the constituents of CNI1000.
CNIMid	CNI Middle Cap.	It consists of the stocks ranked 201-500 from the constituents of CNI1000.
CNISmall	CNI Small Cap.	It consists of the stocks ranked 501-1000 from the constituents of CNI1000.
CNIGrowth	CNI Growth	It consists of 332 stocks with top growth scores from the constituents of CNI1000.
CNIValue	CNI Value	It consists of 332 stocks with top value scores from the constituents of CNI1000.
CNILargeG	CNI Large Growth	It consists of 66 stocks with top growth scores from the constituents of CNILarge.
CNILargeV	CNI Large Value	It consists of 66 stocks with top value scores from the constituents of CNILarge.
CNIMidG	CNI Mid Growth	It consists of 100 stocks with top growth scores from the constituents of CNIMid.
CNIMidV	CNI Mid Value	It consists of 100 stocks with top value scores from the constituents of CNIMid.
CNISmallG	CNI Small Growth	It consists of 166 stocks with top growth scores from the constituents of CNISmall.
CNISmallV	CNI Small Value	It consists of 166 stocks with top value scores from the constituents of CNISmall.

Note: Because CNI1000 is used as the index-formed proxy for market factor in this study, we will not use it as a benchmark to examine the performance of two factor sets.

coupon rates of three-month central bank bills until being replaced by Shibor. Although the RESSET database chooses the three-month Shibor to calculate the excess return on market portfolio, we think the overnight Shibor might reflect more daily information.[14]

3.2.5 RESSET database

In this chapter, we resort to the RESSET database to obtain the data for sophisticate factors instead of the one constructed by ourselves in the previous chapter. The reasons are as follows. Firstly, the RESSET database is one of the commonly used databases to obtain the Chinese financial data. Because we aim to compare the index-formed factors with the sophisticate ones, thus choosing the one widely used by other studies is more reasonable. Secondly, the factors constructed in Chapter 2 only employ the stocks listed on the main boards of Shanghai and Shenzhen stock exchanges, excluding the SME board and ChiNext market of Shenzhen exchange. Focusing on only the main-board market is due to the fact that our sample spans from July 2007 to December 2013. The SME board was built in May 2004, slightly earlier than our sample, and the ChiNext market built in October 2009, right in the middle of the sample. The financial data for the stocks on both the SME board and the ChiNext market used to construct the quarterly-reformed factors are not as complete as those for the stocks on the main-board market. Thus, excluding the SME and ChiNext would lead to more reliable results in the previous chapter. On the contrary, the indices chosen here to construct the index-formed factors are only available after 2010. Thus, the sample period in this chapter starts from January 2010 to December 2014, that is, after the establishment of both the SME and ChiNext markets. Besides, these indices cover full A-share market, including the two main boards of SSE and SZSE, as well as the SME board and ChiNext market of SZSE. As a parallel counterpart, the sophisticate factors should also cover full A-share market.

The RESSET[15] database provides daily, weekly and monthly data for the market, size and

[14]In fact, the choice of over-night or three-month Shibor leads to rather similar results in our study and will not change our main conclusions.

[15]URL: http://www.resset.cn/

BE/ME factors in the A-share, B-share and A- and B-share markets of the SSE, the SZSE and the combination of both SSE and SZSE, respectively. As the relevant indices in this study focus on the A-share stocks of both SSE and SZSE, we choose the data for three factors provided by RESSET with the same coverage on the Chinese A shares as the above-mentioned indices in order to make a reasonable comparison. Besides, RESSET provides the data for both the tradable (free-float) market value-weighted and market capitalization-weighted market, size and BE/ME factors. Because the CSI and CNI indices are all calculated using the tradable market value weights, in this study we will choose the data for tradable market value-weighted factors. Of course, the sample period is the same for both data sets.

3.3 Data issues and construction of proxies for SMB and HML

This section provides a discussion of data issues as well as the methodology of constructing the style index-formed Fama-French factors.

3.3.1 Returns on ten sector indices

Returns on the ten sector indices are the proxies for equity portfolios to be priced on the left-hand side of individual pricing model. The returns are calculated as the difference between the natural logarithm of current index points and the natural logarithm of one-period lag index points as follows

$$R_{p,t} = ln\left(I_{p,t}\right) - ln\left(I_{p,t-1}\right) = ln\left(\frac{I_{p,t}}{I_{p,t-1}}\right) \tag{3.1}$$

where $R_{p,t}$ is the return on portfolio $p = 1,2,\ldots,10$, and $I_{p,t}$ represents the closing price of index or portfolio p at time t. The closing price of each index in this study is not adjusted for dividends or other similar distributions, because we cannot find the data for the adjusted ones. Because our analysis is on the daily base, we believe the dividend reinvestment is not an important issue and would not change our conclusions. The dependent variables in formal regressions of pricing model are the excess returns on the ten sector portfolios over

Table 3.2: Descriptive statistics of excess returns on the CSI All Share sector portfolios

CSI Sectors	Mean	Std. Dev.	t-statistic	Min	Max
Energy	-0.0567	1.7070	-1.1562	-7.0257	7.0829
Materials	-0.0285	1.5910	-0.6230	-7.4150	7.3185
Industrials	0.0013	1.4740	0.0308	-7.3031	5.2730
Cons. Disc	-0.0017	1.4530	-0.0417	-6.5602	4.8150
Cons. Staples	0.0004	1.3537	0.0099	-5.5297	3.9250
Health Care	0.0226	1.5007	0.5251	-6.7829	5.2490
Financials	0.0120	1.5765	0.2656	-7.8727	6.4788
IT	0.0301	1.7796	0.5892	-6.5415	5.3633
Telecom	0.0034	1.7556	0.0676	-7.4111	5.5376
Utilities	0.0068	1.2710	0.1873	-5.9809	4.8753

Note: The sample spans from 05 January 2010 to 31 December 2014, with 1211 observations. Daily % returns are given.

the risk-free rate. The span of these returns is from 05 January 2010 to 31 December 2014. Table 3.2 presents the basic descriptive statistics of excess returns on the ten CSI All Share sector portfolios.[16] Notice that all the means of returns on sector portfolios are statistically insignificant from zero during the sample period. This is not surprising provided that the market during the same period gains an average daily return around zero, and the standard deviation of daily returns is relatively high.

3.3.2 Two sets of returns on three factors

Now we turn to the two sets of factors. We first compute the returns on the related indices. Denote the CNI Big Cap. Growth Index as I_{BG}, the CNI Big Cap. Value Index as I_{BV}, the CNI Mid Cap. Growth Index as I_{MG}, the CNI Mid Cap. Value Index as I_{MV}, the CNI Small Cap. Growth Index as I_{SG}, the CNI Small Cap. Value Index as I_{SV}, and the market CNI1000 Index as I_m. Let $I_{i,t}$ be the closing price/points for index i on trading day t, then the daily return on the index is calculated as

$$R_{i,t} = ln\left(\frac{I_{i,t}}{I_{i,t-1}}\right), \; i = SV, SG, MV, MG, BV, BG \text{ and } m \tag{3.2}$$

[16]Notice that all returns reported in this study are in daily percentages but not annualized.

Table 3.3: Descriptive statistics of two sets of factors and correlations

	$R_{mt}^{CNI}-R_f$	$R_{mt}-R_f$	SMB_t^{CNI}	HML_t^{CNI}	SMB_t	HML_t
Panel A: Basic descriptive statistics						
Mean	0.0071	0.0098	0.0110	0.0149	0.0521	0.0003
Std. Dev.	1.3596	1.2583	1.0424	0.3486	0.6515	0.4882
t-statistic	0.1809	0.2711	0.3668	1.4888	2.7817	0.0200
Max	5.0549	4.4400	4.7150	2.0400	2.5100	2.8000
Min	-6.2857	-5.7700	-5.8150	-1.5833	-4.0600	-1.9900
Obs.	1211	1211	1211	1211	1211	1211
Panel B: Correlations						
$R_{mt}^{CNI}-R_f$	1					
$R_{mt}-R_f$	0.9939	1				
SMB_t^{CNI}	0.1334	0.1212	1			
HML_t^{CNI}	-0.0664	-0.0441	-0.5519	1		
SMB_t	0.2217	0.2047	0.8150	-0.4245	1	
HML_t	-0.0586	-0.0175	-0.6002	0.6804	-0.5527	1

Note: $R_{mt}^{CNI}-R_f$ is the excess market return on the CNI1000 Index over Shibor. SMB_t^{CNI} and HML_t^{CNI} are the proxies for size and BE/ME factors constructed from the CNI style indices. $R_{mt}-R_f$, SMB_t and HML_t are the market factor, size factor and BE/ME factor obtained from RESSET database. Returns are daily percentage returns. The sample period starts from 05 Jan 2010 to 31 Dec 2014 with 1211 trading days.

Once we have the returns on each style index, we can use them to construct the proxies for the Fama-French factors as

$$SMB_t^{CNI} = \left(\frac{R_{SV,t}+R_{SG,t}}{2}\right) - \left(\frac{R_{BV,t}+R_{BG,t}}{2}\right) \tag{3.3}$$

$$HML_t^{CNI} = \left(\frac{R_{SV,t}+R_{MV,t}+R_{BV,t}}{3}\right) - \left(\frac{R_{SG,t}+R_{MG,t}+R_{BG,t}}{3}\right) \tag{3.4}$$

where SMB represents the factor related to size (Small-Minus-Low), and HML represents the factor related to book-to-market equity (High-Minus-Low or Value-Minus-Growth). Here, we use the superscript CNI to denote that the size and BE/ME factors are constructed from the CNI style indices in order to distinguish them from those provided by RESSET.

Table 3.3 presents some basic descriptive statistics of the proxies for the Fama-French factors and the data from the RESSET database in Panel A, followed by the correlations between each pair of these variables reported in Panel B. All the means of six variables are positive during the sample period, but only the sophisticated size factor is statistically significant. If we pair each factor from different data sets, the findings would be of more interest. The values of these basic descriptive statistics for each factor from the different data sets are quite different from each other. The means of both market and size factors from RESSET are rather larger than those constructed from style indices, while the opposite situation happens for the two BE/ME factors.

The correlation between the market factors from the two data sets is really high with a value of 0.9939. The correlation between the two size factors is not as high as that between the market factors, but it has a relatively large value of 0.8150. The pair of two BE/ME returns has a relatively lower relationship between each other; the correlation is 0.6804. From the original idea of constructing the mimicking portfolios related to the size and BE/ME factors in Fama and French (1993), we would expect a rather low correlation between the size and BE/ME factors. That is, the size factor should be largely free of the influence of the BE/ME factor, focusing instead on the different return behaviors of small and big stocks, while the BE/ME factor should be largely free of the size factor, focusing instead on the different return behaviors of high- and low-BE/ME firms. The value of the correlation is -0.08 in Fama and French (1993), and 0.0056 in the previous chapter. In many other studies where Fama and French's (1993) original method is applied, the size and BE/ME factors generally present a low correlation between each other. We also test the data provided by Kenneth French Data Library and find that the low correlation exists both in the US and in other countries. This is especially true for relatively higher frequency data, such as monthly and daily factor returns. However, this expectation of a low correlation does not show up here. The correlations between the size and BE/ME factors from the two different data sets are both negative with relatively large magnitudes of about -0.55. It might not be very surprising that the size and BE/ME factors constructed from style indices do not present a low correlation because they are constructed distinguishable from the original method of Fama and French (1993). But due to the lack of detailed explanation on the construction of the Fama-French factors from the RESSET database, we cannot judge what leads to a relatively high negative relationship between the size and BE/ME factors constructed by RESSET, and this question goes beyond the scope of this study.

The above preliminary analysis is only descriptive. It might be more interesting to find out whether and how much each of those index-formed factors can represent its corresponding factor constructed by the original method, by running regressions of the index-formed factors on those constructed by RESSET, respectively. The estimated coefficients together with the adjusted R^2 and the root mean square error are reported in Table 3.4. If the index-formed factors can proxy for the corresponding factors constructed by the method of Fama and French (1993), then we would expect that each index-formed factor can be largely explained by all three sophisticatedly formed factors and, meanwhile, the explanatory power for each index-formed factor should mainly comes from its corresponding sophisticatedly formed counterpart, rather than the other two factors. This expectation is confirmed by the results presented in Table 3.4.

Panel A of Table 3.4 reports the results from the regression of the CNI market factor on the three RESSET factors. The multivariate regression shows that the estimated slopes related to the market and BE/ME factors are statistically significant at the 1% level, and the adjusted R^2 is about 99%. When we run the univariate regression of the index-formed market factor on each of the RESSET factors, we can find that although the slope estimators related to the market and size factors are statistically significant at the 1% level, the adjusted R^2s from the univariate regression on the RESSET size and BE/ME factors are really small, while the adjusted R^2 from the univariate regression on the RESSET market factor is slightly less than that from the multivariate regression. This suggests that the excess return on the CNI1000 Index can well proxy for the market factor constructed by the Fama and French's original method.

Panel B presents similar findings for the CNI size factor. The results from the multivariate regression of the CNI size factor on the three RESSET factors show that the estimated slope coefficients of the RESSET size and BE/ME factors in addition to the intercept term are statistically significant at the 1% level, with a relatively large adjusted R^2 of almost 70%. The results from the univariate regression show that the highest adjusted R^2 of over 66%, slightly less than that of the multivariate regression, comes from the univariate regression on the RESSET size factor. Although the adjusted R^2 from the univariate regression on the RESSET BE/ME factor is around 36%, it is dramatically less than that from the univariate regression on the RESSET size factor. This finding suggests that the index-formed size factor can capture a large portion of the characteristics of the sophisticated size factor, with a relatively small portion of the characteristics of the sophisticated BE/ME factor. Thus, we can conclude that the CNI size factor is a good proxy for the RESSET size factor, although it is not as good as the CNI market factor as a proxy for the RESSET market factor.

Panel C tells a similar story for the CNI BE/ME factor. The slope estimators related to the RESSET size and BE/ME factors and the constant term of the multivariate regression are statistically significant at the 10% level. The corresponding adjusted R^2 is about 47%, which is less impressive compared to those mentioned above. The estimated coefficients of the univariate regression of the CNI BE/ME factor on the RESSET market factor are statistically insignificant, and the adjusted R^2 is only 0.11%. This suggests that the CNI BE/ME factor excludes the main features of the RESSET market factor, which is the first supportive evidence for the CNI BE/ME factor. The estimated parameters of both univariate regressions on the RESSET size and BE/ME factors are statistically significant at the 1% level, however, the adjusted R^2 of 46% from the univariate regression on the RESSET BE/ME factor is much larger and closer to that from the multivariate regression, compared to the adjusted R^2 from the univariate regression on the RESSET size factor. Thus, the CNI BE/ME factor could be a good proxy for the sophisticated BE/ME factor because it mainly captures the properties of the RESSET BE/ME factor, rather than those of the RESSET market and size factors.

3.4 Empirical Tests

This section examines the Fama-French model using the proxies for both size and BE/ME factors that constructed from the CNI style indices and the data for the three factors provided by RESSET, and compares the performance of both sets of factors in two steps. The first step is

Table 3.4: Estimated coefficients of regressions of index-formed factors on RESSET factors

$R_{mt}-R_f$	SMB_t	HML_t	Constant	adj. R^2	RMSE
Pannel A: Regression of $R_{mt}^{CNI}-R_f$ on three factors from RESSET database					
1.0746***	-0.0138	-0.1250***	-0.0027	0.9896	0.1384
(196.88)	(-1.39)	(-10.48)	(-0.64)		
1.0740***			-0.0035	0.9879	0.1494
(170.26)			(-0.78)		
	0.4627***		-0.0170	0.0484	1.3263
	(5.47)		(-0.42)		
		-0.1633	0.0071	0.0026	1.3578
		(-1.31)	(0.18)		
Pannel B: Regression of SMB_t^{CNI} on three factors from RESSET database					
-0.0220	1.1253***	-0.4525***	-0.0473***	0.6965	0.5743
(-1.13)	(30.41)	(-9.58)	(-2.82)		
0.1004***			0.0100	0.0139	1.0352
(2.72)			(0.29)		
	1.3040***		-0.0569***	0.6640	0.6042
	(37.29)		(-3.11)		
		-1.2815***	0.0113	0.3597	0.8342
		(-20.36)	(0.46)		
Pannel C: Regression of HML_t^{CNI} on three factors from RESSET database					
-0.0055	-0.0343*	0.4604***	0.0166**	0.4654	0.2549
(-0.73)	(-1.75)	(18.88)	(2.00)		
-0.0122			0.0150	0.0011	0.3485
(-1.02)			(1.27)		
	-0.2271***		0.0267**	0.1795	0.3158
	(-9.92)		(2.47)		
		0.4859***	0.0148*	0.4625	0.2556
		(18.87)	(1.81)		

Note: ***, ** and * represent the statistical significance at the 1%, 5% and 10% level, respectively.
The sample period starts from 05 Jan 2010 to 31 Dec 2014 with 1211 trading days.
The associated t-statistic is contained in the parentheses below the coefficient estimate, and it is corrected by the Newey-West standard errors. The rule used to choose the maximum order of any significant autocorrelation in the disturbance process is following Stock and Watson (2007, p.607) as that # of lags $= 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression. RMSE is the root mean square error of the regression.

to run the formal individual regressions of an unrestricted model wherein there is no restricted condition on the intercept term. In the second step, following Faff (2003, 2004), we employ the Generalized Method of Moments (GMM) approach of MacKinlay and Richardson (1991) to a restricted model with the condition of zero intercept term. We will discuss both steps in more details below. In each step, both factor sets will be tested, and based on the empirical results, a direct comparison will be conducted for checking whether the index-formed Fama-French factors have the similar properties as the sophisticatedly constructed ones in line with Fama and French (1993). We find that the simple method of constructing the Fama-French factors from the "off-the-shelf" style indices have a better performance than the sophisticated ones in the Chinese stock market.

3.4.1 Unrestricted model

The commonly examined Fama-French model has the following form:

$$R_{it} - R_{ft} = \alpha_i + b_i(R_{mt} - R_{ft}) + s_i SMB_t + h_i HML_t + \varepsilon_t \tag{3.5}$$

where

R_{it} is the realized return on asset i at time t;

R_{ft} is the realized return on the risk-free asset at time t;

R_{mt} is the realized return on the market portfolio at time t;

SMB_t is the realized return on the mimicking portfolio for the size factor at time t;

HML_t is the realized return on the mimicking portfolio for the BE/ME factor at time t, and

ε_t is the error term, which is assumed to be independent of other factors.

Table 3.5 reports the estimated outcome of regressions of the unrestricted Fama-French model across the industry portfolios using the factors constructed from the CNI style indices; while Table 3.6 presents those results using the factors provided by the RESSET database. The main findings from both tables are summarized as follows. First, for both sets of factors, all the market betas are statistically significant at the 1% level. The market betas are quite close to unity, with a mean of 0.9658 for the index-formed market factor which is less than the mean of 1.0282 for the RESSET market factor. This is consistent with Fama and French (1993) and many other studies and suggests that the market factor has little ability to explain the cross-sectional variation in equity returns. Second, all SMB betas are statistically significant at the 1% level. There are eight positive SMB betas for the index-formed data set and seven for the RESSET data set. The mean SMB beta is positive for both data sets. Third, there are eight index-formed HML betas which are statistically significant at the 5% level and half of them are negative, compared to nine statistically significant and seven negative for the RESSET HML betas. The means of the HML betas are negative for both data sets. Fourth, the mean intercepts are negative for both data sets, but only three intercepts are statistically significant for the index-formed data set while five for RESSET. Because the intercept term of the Fama-French model captures the model's pricing error, the insignificance of the intercept suggests a good performance of the model. Lastly, the adjusted R^2 is high with a mean value of about 85% for both data sets. The mean adjusted R^2 is higher, while the root mean squared error is less for the CNI indexed-formed factors. These findings suggest that based on the formal individual regression analysis, the Fama-French model performs well in explaining the returns on the ten sector portfolios. The three index-formed factors present a slightly better performance supported by the statistically insignificant intercept, the slightly higher adjusted R^2 and the lower average root mean square error.

3.4.2 Restricted model

Fama and French (1996) propose an empirical three-factor asset pricing model as:

$$E(R_i - R_f) = b_i E(R_m - R_f) + s_i E(SMB) + h_i E(HML) \tag{3.6}$$

where

$E(R_i - R_f)$ is the expected excess return on asset i;

Table 3.5: Estimated coefficients of regression of unrestricted FF model using proxies from CNI style indices

CSI Sectors	b_i	s_i	h_i	α_i	adj.R^2	RMSE
Energy	1.1193***	-0.1638***	-0.0599	-0.0622***	0.7801	0.8005
	(49.6305)	(-3.8644)	(-0.5601)	(-2.8202)		
Materials	1.0708***	0.2515***	0.0047	-0.0384***	0.9053	0.4896
	(71.2232)	(9.2294)	(0.0718)	(-2.7084)		
Industrials	1.0052***	0.3407***	0.2397***	-0.0125	0.9601	0.2944
	(138.6000)	(17.4177)	(5.1374)	(-1.3931)		
Cons. Disc	0.9574***	0.3410***	-0.1676***	-0.0091	0.9398	0.3566
	(100.7287)	(21.0677)	(-3.0755)	(-0.8678)		
Cons. Staples	0.7822***	0.2093***	-0.8079***	0.0051	0.7792	0.6360
	(58.8868)	(5.8286)	(-9.1275)	(0.2720)		
Health Care	0.7741***	0.4291***	-0.9253***	0.0272	0.7745	0.7127
	(56.2429)	(12.0729)	(-10.6713)	(1.3992)		
Financials	1.0697***	-0.5968***	0.3976***	0.0039	0.9423	0.3787
	(100.0315)	(-38.7098)	(7.0110)	(0.3899)		
IT	0.9960***	0.6271***	-0.8513***	0.0302*	0.9029	0.5546
	(70.2332)	(25.6487)	(-11.1110)	(1.8302)		
Telecom	0.9808***	0.5977***	-0.4292***	-0.0025	0.8259	0.7326
	(51.8245)	(18.0654)	(-4.4271)	(-0.1184)		
Utilities	0.8083***	0.1994***	0.6583***	-0.0106	0.7918	0.5800
	(53.2429)	(6.5419)	(9.9914)	(-0.6401)		
Mean	0.9564	0.2235	-0.1941	-0.0069	0.8602	0.5536

Note: ***, ** and * represent the statistical significance at the 1%, 5% and 10% level, respectively.
The sample period starts from 05 Jan 2010 to 31 Dec 2014 with 1211 trading days.
The associated t-statistic is contained in the parentheses below the coefficient estimate, and it is corrected by the Newey-West standard errors. The rule used to choose the maximum order of any significant autocorrelation in the disturbance process is following Stock and Watson (2007, p.607) as that # of lags $= 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression. RMSE is the root mean square error of the regression.

Table 3.6: Estimated coefficients of regression of unrestricted FF model using RESSET three-factor data

CSI Sectors	b_i	s_i	h_i	α_i	adj.R^2	RMSE
Energy	1.2502***	-0.4402***	-0.0433	-0.0462**	0.8119	0.7403
	(51.64)	(-7.49)	(-0.62)	(-2.13)		
Materials	1.1583***	0.1595***	-0.3006***	-0.0484***	0.8860	0.5373
	(54.49)	(4.48)	(-6.22)	(-3.20)		
Industrials	1.0719***	0.3941***	-0.0697**	-0.0300***	0.9384	0.3660
	(105.15)	(15.84)	(-2.39)	(-2.59)		
Cons. Disc	1.0178***	0.4164***	-0.3293***	-0.0336***	0.9175	0.4173
	(76.13)	(15.38)	(-7.33)	(-2.72)		
Cons. Staples	0.8702***	-0.0854**	-1.1537***	-0.0036	0.8089	0.5918
	(55.87)	(-1.98)	(-19.56)	(-0.20)		
Health Care	0.8405***	0.3415***	-1.1135***	-0.0034	0.7611	0.7335
	(51.24)	(7.47)	(-14.76)	(-0.16)		
Financials	1.1296***	-0.4350***	0.5440***	0.0233	0.8343	0.6417
	(50.68)	(-9.92)	(9.16)	(1.36)		
IT	1.0506***	0.9188***	-0.6749***	-0.0283	0.8751	0.6289
	(63.48)	(20.18)	(-14.33)	(-1.47)		
Telecom	1.0278***	0.9281***	-0.3655***	-0.0553**	0.8163	0.7523
	(49.17)	(16.07)	(-6.76)	(-2.40)		
Utilities	0.8656***	0.2073***	0.2452***	-0.0127	0.7769	0.6003
	(56.52)	(4.78)	(3.51)	(-0.69)		
Mean	1.0282	0.2405	-0.3261	-0.0238	0.8426	0.6009

Note: ***, ** and * represent the statistical significance at the 1%, 5% and 10% level, respectively.
The sample period starts from 05 Jan 2010 to 31 Dec 2014 with 1211 trading days.
The associated t-statistic is contained in the parentheses below the coefficient estimate, and it is corrected by the Newey-West standard errors. The rule used to choose the maximum order of any significant autocorrelation in the disturbance process is following Stock and Watson (2007, p.607) as that # of lags $= 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression. RMSE is the root mean square error of the regression.

$E(R_m - R_f)$ is the expected excess return on the market portfolio;

R_f is the return on the risk-free asset;

$E(SMB)$ is the expected return on the mimicking portfolio related to the size factor;

$E(HML)$ is the expected return on the mimicking portfolio related to the BE/ME factor.

This form of asset pricing model suggests that if the model is correct or has no pricing error, then there should be no intercept term. By taking the expectation of unrestricted model (3.5) and comparing it to model (3.6), we can observe that the zero intercept restriction in model (3.6) constitutes the null hypothesis of zero alpha, i.e., $\mathrm{H}_0 : \alpha_i = 0$, for $i = 1, 2, \ldots, N$, in model (3.5). Thus, in the next empirical tests, we impose the zero intercept restriction on the regression model. Then, the regression is in the following restricted form without intercept:

$$R_{it} - R_f = b_i(R_{mt} - R_f) + s_i SMB_t + h_i HML_t + \varepsilon_{it},\ i = 1, 2, \ldots, N \tag{3.7}$$

In addition, we introduce the mean adjustment process to each of the explanatory variables of equation (3.7) as follows:

$$R_{mt} - R_f = \lambda_m + \xi_{mt} \tag{3.8}$$

$$SMB_t = \lambda_s + \xi_{st} \tag{3.9}$$

$$HML_t = \lambda_h + \xi_{ht} \tag{3.10}$$

We combine the above equations (3.7), (3.8), (3.9) and (3.10) into a system of equations, in order to estimate the coefficients and to examine the validity of the pricing model (3.6) as a whole based on the GMM approach.[17] If we substitute the factors in equation (3.7) by their corresponding mean adjustments of equations (3.8), (3.9) and (3.10), then this specification

[17]This method is employed in Faff (2003, 2004), and can also be referred to Ferson and Harvey (1994). A modification of this method can be found in Pham (2007), where the system is separately set up for each asset *i*.

effectively yields that

$$R_{it} - R_f = b_i\lambda_m + s_i\lambda_s + h_i\lambda_h + b_i\xi_{mt} + s_i\xi_{st} + h_i\xi_{ht} + \varepsilon_{it},\ i = 1,2,\ldots,N \tag{3.11}$$

which is analogy to the second stage of Fama-MacBeth regression.

As an alternative to the standard Fama-MacBeth two-stage regression, this setup allows us to estimate the risk loadings, i.e., b_i, s_i and h_i, for $i = 1,\ldots,N$, through equation (3.7) as well as the risk premiums, i.e., λ_m, λ_s and λ_h, through equations (3.8), (3.9) and (3.10) on three factors at the same time. The individual null hypothesis for each coefficient is that each coefficient is equal to zero, using a z-statistic to test under the asymptotically normal distribution. The advantage of the GMM system regression over the Fama-MacBeth approach is that the GMM approach allows for conditional heteroscedasticity, serial correlation, and non-normal distribution, while the Fama-MacBeth regressions provide standard errors corrected only for cross-sectional correlation but not for time-series autocorrelation.

We can rewrite the above system in the form of matrices as follows:

$$\underset{(N+3)\times 1}{\begin{bmatrix} R_{1t}-R_f \\ R_{2t}-R_f \\ \vdots \\ R_{Nt}-R_f \\ R_{mt}-R_f \\ SMB_t \\ HML_t \end{bmatrix}} = \underset{(N+3)\times 4}{\begin{bmatrix} b_1 & s_1 & h_1 & 0 \\ b_2 & s_2 & h_2 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ b_N & s_N & h_N & 0 \\ 0 & 0 & 0 & \lambda_m \\ 0 & 0 & 0 & \lambda_s \\ 0 & 0 & 0 & \lambda_h \end{bmatrix}} \cdot \underset{4\times 1}{\begin{bmatrix} R_{mt}-R_f \\ SMB_t \\ HML_t \\ 1 \end{bmatrix}} + \underset{(N+3)\times 1}{\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{Nt} \\ \xi_{mt} \\ \xi_{st} \\ \xi_{ht} \end{bmatrix}} \tag{3.12}$$

or in a compact matrix form:

$$\begin{bmatrix} \mathbf{R}_t \\ \mathbf{F}_t \end{bmatrix} = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \Lambda \end{bmatrix} \begin{bmatrix} \mathbf{F}_t \\ 1 \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ \xi_t \end{bmatrix} \tag{3.13}$$

$$\begin{aligned}
\text{where } \mathbf{R}_t &= \left[R_{1t}-R_f, R_{2t}-R_f, \ldots, R_{Nt}-R_f\right]' \\
\mathbf{F}_t &= \left[R_{mt}-R_f, SMB_t, HML_t\right]' \\
\mathbf{B} &= \begin{bmatrix} b_1 & s_1 & h_1 \\ b_2 & s_2 & h_2 \\ \vdots & \vdots & \vdots \\ b_N & s_N & h_N \end{bmatrix} \\
\Lambda &= [\lambda_m, \lambda_s, \lambda_h]' \\
\varepsilon_t &= [\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{Nt}]' \\
\xi_t &= [\xi_{mt}, \xi_{st}, \xi_{ht}]'
\end{aligned}$$

Next, we examine the above system based on the GMM approach. The above system of equations contains $4N+3$ sample moments ($1/T \sum_{t=1}^{T} \varepsilon_{it}$, $1/T \sum_{t=1}^{T} \varepsilon_{it}(R_{mt}-R_f)$, $1/T \sum_{t=1}^{T} \varepsilon_{it} SMB_t$, $1/T \sum_{t=1}^{T} \varepsilon_{it} HML_t$ for $i = 1, 2, \ldots, N$, and $1/T \sum_{t=1}^{T} \xi_{jt}$ for $j = m, s, h$), and $3N+3$ unknown parameters to be estimated. So this system is obviously over-identified, and Hansen's (1982) J-statistic is:

$$J = T \cdot g_T\left(\hat{\phi}\right)' \cdot S_T^{-1} \cdot g_T\left(\hat{\phi}\right) \sim \chi^2(N) \tag{3.14}$$

where $g_T\left(\hat{\phi}\right) = \frac{1}{T}\Sigma_{t=1}^{T} f_t\left(\hat{\phi}\right)$ is the empirical moment condition vector. The J-statistic involves $N \; (= (4N+3)-(3N+3))$ over-identifying restrictions, and is asymptotically distributed as a chi-square statistic with N degrees of freedom. Furthermore, the estimation technique employs the heteroskedasticity and autocorrelation (HAC) weighting matrix and asymptotically consistent covariance matrices, and uses an iterated procedure (Ferson and Foerster, 1994) with an initial identity weighting matrix. The HAC standard errors are based on Bartlett kernel with lags chosen by Newey-West method.[18]

Table 3.7 reports the estimated risk premiums, i.e., $\Lambda = [\lambda_m, \lambda_s, \lambda_h]'$, and Hansen's (1982) J-statistics for the above system based on the GMM test for both data sets in Panel A and B,

[18] We attach a sample version of the Stata code for the GMM system analysis in Appendix for reference.

Table 3.7: Restricted Fama-French three-factor model based on GMM test

Test year	λ_m	λ_s	λ_h	Hansen's J	obs.
Panel A: Factors from CNI Indices					
2010	0.2707***	0.1810***	-0.0618***	14.3907	241
	(3.55)	(2.67)	(-5.24)	[0.1559]	
2011	-0.1641**	-0.1666***	0.0529***	11.5485	244
	(-2.15)	(-3.12)	(2.75)	[0.3164]	
2012	0.0273	-0.0301	0.0268	9.4840	243
	(0.41)	(-0.70)	(1.60)	[0.4869]	
2013	0.0599	0.2265***	-0.0539***	10.8815	238
	(0.84)	(6.36)	(-2.89)	[0.3668]	
2014	0.2116***	0.0149	0.0296	10.2194	245
	(3.57)	(0.24)	(1.15)	[0.4215]	
2010-2014	0.0047	0.0191	0.0102	12.1555	1211
	(0.12)	(0.59)	(0.80)	[0.2748]	
Panel B: Factors from RESSET database					
2010	-0.1133	0.0971**	-0.1145***	13.1863	241
	(-1.57)	(2.00)	(-3.09)	[0.2134]	
2011	-0.0487	-0.0125	-0.0164	11.9636	244
	(-0.73)	(-0.31)	(-0.48)	[0.2875]	
2012	-0.0007	0.0526	0.0107	9.8570	243
	(-0.01)	(1.57)	(0.35)	[0.4531]	
2013	0.0790	0.1699***	-0.0053	6.8466	238
	(1.33)	(6.11)	(-0.26)	[0.7398]	
2014	0.2544***	0.0867**	0.0613**	11.2956	245
	(4.72)	(2.14)	(2.19)	[0.3350]	
2010-2014	0.0295	0.0391*	0.0058	21.8930	1211
	(0.84)	(1.77)	(0.34)	[0.0157]	

Note: This table reports the results of testing the Fama-French model in the system of regression equations (3.7), (3.8), (3.9) and (3.10) based on the Generalized Method of Moments. The associated z-statistic is in parentheses () below the estimated coefficients. GMM reports the Hansen's J-statistic, which is used to determine the validity of the overidentifying restrictions in a GMM model. Hansen's J-statistic is distributed as a chi square with N $(=(4N+3)-(3N+3))$ degrees of freedom. The associated p-value is in square brackets [ ]. If the p-value is high, then we cannot reject the null hypothesis of the model being correctly specified.
The estimation technique employs the heteroskedasticity and autocorrelation (HAC) weighting matrix and consistent covariance matrices, and uses an iterated procedure (Ferson and Foerster, 1994) with an initial identity matrix. The HAC standard errors are based on Bartlett kernel with lags chosen by Newey-West method.

Table 3.8: Estimated coefficients of Fama-French model in the overidentifying system

CSI Sectors	b_i	s_i	h_i
Panel A: Factors from CNI Indices			
Energy	1.1470***	-0.2258***	-0.0920
	(46.39)	(-5.36)	(-0.94)
Materials	1.1025***	0.2093***	0.0589
	(54.70)	(7.68)	(0.84)
Industrials	1.0012***	0.3073***	0.1799***
	(174.10)	(15.52)	(3.85)
Cons. Disc	0.9649***	0.3457***	-0.0557
	(93.27)	(19.99)	(-0.89)
Cons. Staples	0.7703***	0.2009***	-0.9695***
	(59.91)	(5.20)	(-9.06)
Health Care	0.7625***	0.4030***	-1.0134***
	(51.61)	(9.89)	(-12.13)
Financials	1.0624***	-0.5630***	0.5030***
	(94.16)	(-35.62)	(8.83)
IT	0.9747***	0.6418***	-0.7679***
	(67.56)	(25.54)	(-9.20)
Telecom	0.9540***	0.6580***	-0.3277***
	(50.81)	(17.75)	(-2.87)
Utilities	0.7807***	0.2237***	0.6417***
	(59.86)	(8.84)	(10.69)
Panel B: Factors from RESSET database			
Energy	1.2591***	-0.4990***	-0.0959
	(52.62)	(-8.55)	(-1.61)
Materials	1.1780***	0.1205***	-0.3169***
	(52.65)	(3.72)	(-6.69)
Industrials	1.0737***	0.3730***	-0.0590**
	(120.95)	(16.78)	(-2.06)
Cons. Disc	1.0179***	0.4263***	-0.3082***
	(69.02)	(17.41)	(-6.05)
Cons. Staples	0.8629***	-0.0985**	-1.2484***
	(50.51)	(-2.09)	(-20.58)
Health Care	0.8460***	0.3426***	-1.0803***
	(48.67)	(6.81)	(-13.39)
Financials	1.1214***	-0.3345***	0.5962***
	(45.14)	(-6.37)	(11.17)
IT	1.0505***	0.9125***	-0.6612***
	(61.55)	(15.99)	(-13.87)
Telecom	1.0296***	0.9475***	-0.3503***
	(53.58)	(14.04)	(-6.15)
Utilities	0.8582***	0.2101***	0.2393***
	(58.25)	(4.77)	(3.83)

Note: This table is complementary to Table 3.7 for the whole sample years 2010 to 2014. Almost all the estimated coefficients of the market, size and BE/ME factors are statistically significant at the 5% level. The associated z-statistic is in parentheses ( ) below the estimated coefficients.

respectively. We examine the Fama-French model for each year from 2010 to 2014, followed by the tests over the whole sample period. We only report the estimated risk loadings on three factors for the whole sample period in Table 3.8 in order to save space. The results of the coefficient estimators for each year are similar to those for the whole sample period.

The major features in Table 3.7 are summarized as follows. As a whole, we can see that the estimated mean returns (i.e., the risk premiums) of all the three factors related to market, size and BE/ME are generally statistically insignificant. This is not a too surprising result, because from Table 3.3 we can see that the means of all factors from both data sets are generally indifferent from zero during the whole sample period from 2010 to 2014. From both Panels A and B, the magnitudes and signs of estimated risk premiums are similar for both data sets. Specifically, for the index-formed data, there are only three years for the market, size and BE/ME premiums are statistically significant at least at the 10% level, while the number of years for the RESSET data set is one, three and two, respectively. As to the GMM test on the whole sample period, there is no statistically significant risk premium for the index-formed data set, while only the size premium is significant at the 10% level for the RESSET data. According to the Hansen's J-statistics for each sample year, the Fama-French model cannot be rejected at any conventional level of significance for both data sets. When we examine the model for all sample years as a whole, the p-value reduces for both sets of factors, but we still cannot reject the model in the case of the index-formed factors. However, this is not the case for the RESSET factors, and the p-value decreases to only 0.0157. We observe that the estimated market, SMB and HML risk premiums are of mixed signs. There are several negative estimated risk premiums presented in both panels. This finding suggests the existence of the effect reversal of the three factors. In particular, the reversal of the size effect has been found in a wave of recent studies for some developed markets, such as the US, UK, Australia and Japan (see Gompers and Metrick, 1998, Dimson and Marsh, 1998, Gustafson and Miller, 1999, Faff, 2004 and Pham, 2007). From the results of this table, the evidence for the Fama-French model is still supportive in the case of the index-formed factors but less persuasive in the case of the sophisticatedly constructed factors. A comparison between Panels

A and B suggests that the Fama-French model based on the index-formed factors has better performance than that based on the sophisticatedly constructed factors.

The magnitudes, statistical significance and signs of the estimated coefficients from the GMM system regression presented in Table 3.8 are very similar to those of coefficient estimators from the formal individual regressions reported in Tables 3.5 and 3.6 for both sets of factors, respectively. The difference is mainly caused by the zero-intercept restriction imposed on the GMM system test. This suggests the GMM test on the Fama-French model yields similar coefficient estimators to those based on the formal individual regressions. Moreover, a comparison between Panels A and B of Table 3.8 also supports the conclusion that the Fama-French model based on the index-formed factors at least has the similar performance to that based on the sophisticatedly constructed factors.

3.4.3 Robustness tests on passive benchmark indices

As robustness tests to more formally compare the ability of each set of factors to explain stock returns, we follow Cremers et al. (2012) to check whether the index-formed factors or their sophisticate counterparts produce zero (or close to zero) alphas for passive benchmark indices in the Chinese market. As mentioned in Cremers et al. (2012), the factor pricing models, such as the Fama-French three-factor model or the Carhart four-factor model, generally assign economically and statistically large non-zero alphas to passive benchmark indices for extended periods of time. As a result of that, even pure index funds tracking the benchmark indices appear to present large positive or negative "skill", which will lead to mistakenly evaluating the performance of passive index funds. Thus, it is important to find the factor pricing models or the proxies for the risk factors that would produce less alphas in the passive benchmark indices. In this subsection, we aim to provide evidence for which set of factors in this study would produce less alphas in the regressions of passive benchmark indices. In order to conserve the space, we will focus on the alphas, adjusted R^2s and root mean squared errors of the regressions. The magnitudes, significance and signs of the slope coefficients are similar for both sets of factors.

Table 3.9: Summarized results of regressions of passive benchmarks for two factor sets

	CNI factors				RESSET factors			
Panel A: Summarized regression results								
	Alpha	t-stat.	Adj. R^2	RMSE	Alpha	t-stat.	Adj. R^2	RMSE
SSE	-0.0141	-2.93	0.9788	0.1266	-0.0061	-1.47	0.9887	0.1737
SSE50	-0.0041	-0.71	0.9780	0.4429	0.0188	1.56	0.9027	0.2108
CSI300	-0.0076	-3.24	0.9971	0.2069	0.0032	0.54	0.9770	0.0735
CSI500	0.0016	0.35	0.9891	0.2794	-0.0279	-3.35	0.9665	0.1592
CSI800	-0.0060	-4.66	0.9992	0.1529	-0.0056	-1.23	0.9874	0.0379
CSI300G	-0.0034	-0.56	0.9789	0.3498	0.0062	0.60	0.9387	0.2053
CSI300V	-0.0071	-1.37	0.9838	0.3826	0.0137	1.24	0.9260	0.1791
SZSE	-0.0172	-1.73	0.9425	0.3984	-0.0066	-0.59	0.9258	0.3509
CNILarge	-0.0009	-0.40	0.9969	0.2457	0.0136	1.99	0.9671	0.0752
CNIMid	-0.0001	-0.03	0.9886	0.2861	-0.0192	-2.25	0.9629	0.1588
CNISmall	0.0073	1.54	0.9901	0.2728	-0.0282	-3.57	0.9690	0.1543
CNIGrowth	-0.0014	-0.35	0.9893	0.2549	-0.0075	-0.98	0.9667	0.1443
CNIValue	0.0047	1.14	0.9894	0.2971	0.0166	1.96	0.9519	0.1398
CNILargeG	-0.0028	-0.49	0.9805	0.3855	0.0074	0.66	0.9271	0.1995
CNILargeV	0.0101	1.74	0.9803	0.4592	0.0331	2.59	0.8932	0.1971
CNIMidG	-0.0043	-0.57	0.9732	0.3799	-0.0291	-2.43	0.9361	0.2458
CNIMidV	-0.0022	-0.26	0.9640	0.4016	-0.0047	-0.38	0.9295	0.2870
CNISmallG	0.0112	2.20	0.9861	0.3554	-0.0269	-2.58	0.9477	0.1834
CNISmallV	-0.0039	-0.77	0.9867	0.3335	-0.0271	-2.52	0.9523	0.1765
CSI500G	0.0098	1.28	0.9722	0.3707	-0.0207	-1.70	0.9409	0.2541
CSI500V	-0.0011	-0.15	0.9726	0.3699	-0.0191	-1.58	0.9404	0.2506
CSI800G	-0.0053	-1.09	0.9861	0.2655	-0.0014	-0.18	0.9619	0.1603
CSI800V	-0.0068	-1.52	0.9885	0.3017	0.0079	0.91	0.9523	0.1479
Mean	-0.0019		0.9823	0.1811	-0.0048		0.9483	0.3182
Panel B: GRS tests and Root mean squared alpha								
	GRS stat.	p-value	RMS α		GRS stat.	p-value	RMS α	
	3.5820	2.79e-8	0.0072		4.7858	1.19e-12	0.0181	

This table summarizes the estimated alphas and associated t-statistics, adjusted R^2s and root mean squared errors for the regressions of 23 common benchmark indices of the Chinese market in Panel A. The GRS tests for the null hypothesis of all alphas being jointly equal to zero and the root mean squared alpha are reported in Panel B.

Table 3.9 presents the related test results. The alphas and associated t-statistics, adjusted R^2s and root mean squared errors for the regressions of excess returns on the 23 common benchmarks indices of the Chinese market are summarized in Panel A for both sets of factors. We can see that most of the individual alphas are statistically insignificant for both factor sets. The average alpha is -0.0019 for the CNI index-formed factor set, and is -0.0048 for the RESSET factors. The average adjusted R^2 is higher and the average root mean squared error is lower for the CNI factors. Although the GRS tests reject the null hypothesis that the intercepts are jointly equal to zero for both sets of factors, the GRS statistic is smaller and the associated p-value is larger for the index-formed factors than those for the sophisticate factors. Moreover, the root mean squared alpha is also smaller in the case of the CNI factors. All the results suggest that the index-formed factors are superior to the sophisticate factors in the sense of producing less pricing errors or alphas in the Fama-French model.

Our results here support the findings of Cremers et al. (2012) that the pricing errors or alphas of the Fama-French model primarily arise from the disproportionate weight to small and value stocks in the construction of the factors. Remember that the index-formed market factor in this study is CNI1000 index. Its constituent stocks exclude more than half of listed stocks in the Chinese market. Those stocks are generally small size (market capitalization) and have relatively poor financial performance, i.e., the distressed or value firms. Our index-formed size and value factors are constructed using the constituent stocks of the CNI1000 index. Thus, all three index-formed factors excluding the small and value stocks. By doing so, our index-formed factors mitigate the disproportion weight to small and value stocks in constructing the Fama and French's original factors, which is a problem that causes the more significant alpha in the Fama-French model according to Cremers et al. (2012). Thus, the index-formed market, size and value factors can reduce more pricing errors than their sophisticate counterparts in the pricing models.

3.5 Conclusion

The efficacy of Faff's simplified method of constructing the usable proxies for the Fama-French factors and the Fama-French model based on this method has been verified in the US (Faff, 2003), Australian (Faff, 2004) and Japanese (Pham, 2007) markets. In those developed markets, the results of a formal time-series asset pricing test are favorable to the Fama-French model even using the style indices to construct the factors. However, when one takes into account the estimated risk premiums based on a system application of the Generalized Method of Moments (GMM), the support for the Fama-French model turns out to be less persuasive. However, whether the set of index-formed factors or the set of sophisticate factors is superior to the other in explaining equity returns is unknown in the existing studies.

This study extends the examination of Faff's method to the Chinese market by constructing the Fama-French factors from the off-the-shelf style indices covering the A-share market of both Shanghai and Shenzhen stock exchanges. More importantly, we provide direct evidence for that the index-formed factors performs better than their sophisticate counterparts. Specifically, we first run the regression of index-formed factors on the sophisticate factors and find that each of the index-formed factors can largely capture the features of its sophisticate counterpart. We then compare the ability of both sets of factors to explain equity returns in the time-series regressions. The comparison of results from the unrestricted time-series regression of the Fama-French model for the two factor sets shows that the index-formed factors do share at least similar properties to those constructed using Fama and French's original method. The findings of the smaller magnitude of insignificant alpha, higher adjusted R^2, and lower root mean square error of the regression on the index-formed factors provide the first evidence for that the Fama-French model using the index-forms factors performs better than using the sophisticate factors in explaining returns on sector portfolios.

We also examine the two factor sets in a system, which includes the Fama-French model with the zero-intercept restriction imposed, as well as three mean adjustment processes to market, size and value factors. We conduct the tests based on the GMM approach. This setup allows us to estimate the risk loadings and risk premiums on three factors at the same time. Besides, it is a superior alternative to the standard Fama-MacBeth two-stage regression because the GMM approach allows for conditional heteroscedasticity, serial correlation, and non-normal distribution, while the standard errors of Fama-MacBeth regressions are only corrected for cross-sectional correlation but not for time-series autocorrelation. The significance and signs of estimated risk loadings and risk premiums show very similar results for both factor sets. Hansen's J-statistic of GMM tests cannot reject the Fama-French model in each sample year for both sets of factors. However, we fail to reject the model for the whole sample period from 2010 to 2014 only for the index-formed factors but not for the factors constructed using Fama and French's original method. Thus, the results generally support the Fama-French model with only an exception in the case of sophisticate factors. Although the estimated risk premiums on both sets of factors from the GMM system regression are generally insignificant, this is consistent with the insignificant means of daily returns on the factors during the sample period. The estimated risk premiums on all three factors are also of mixed signs, which suggests the existence of reversal of factor effects.

As a robustness test, we examine whether the index-formed or sophisticate factors can produce zero or close to zero alphas for passive benchmark indices in the Chinese market. Our results show that for the passive benchmark indices, the index-formed factors produce less alphas, higher adjusted R^2s and lower root mean squared error for the benchmarks. Although the GRS tests reject the null hypothesis of alphas jointly being zero, the GRS statistics is smaller, the associated p-value is higher, and the root mean squared alpha is less for the index-formed factors. Our findings here support the argument of Cremers et al. (2012) that the disproportionate weight to small and value stocks in the construction of the factors is the main source of the pricing errors or alphas of the Fama-French model. Using the indices that include only big capitalization and growth style to construct the factors lead to less pricing errors of the Fama-French model as shown in our study.

To sum up, our empirical results suggest that the index-formed factors are superior alternative to the sophisticate ones constructed using Fama and French's original method in explaining equity returns. This finding is meaningful for both the academia and the financial industry. It suggests that we can construct the Fama-French factors in a much simpler way by using the commercially available style indices. This simplified method will lower the barrier to exploiting the Fama-French model and will benefit the financial industry by, for example, enriching the financial products based on style indices and hedging the risks related to the size and book-to-market effects in a much more cost-efficient way.

Appendix

The following is a sample version of Stata codes that we used to run the GMM system analysis on equation 21 or 22.

```
set more off
gmm (eq1: CSIAllEn - Rmrf*{b1} - Smb*{s1} - Hml*{h1}) ///
(eq2: CSIAllMa - Rmrf*{b2} - Smb*{s2} - Hml*{h2}) ///
(eq3: CSIAllIn - Rmrf*{b3} - Smb*{s3} - Hml*{h3}) ///
(eq4: CSIAllCD - Rmrf*{b4} - Smb*{s4} - Hml*{h4}) ///
(eq5: CSIAllCS - Rmrf*{b5} - Smb*{s5} - Hml*{h5}) ///
(eq6: CSIAllHC - Rmrf*{b6} - Smb*{s6} - Hml*{h6}) ///
(eq7: CSIAllFi - Rmrf*{b7} - Smb*{s7} - Hml*{h7}) ///
(eq8: CSIAllIT - Rmrf*{b8} - Smb*{s8} - Hml*{h8}) ///
(eq9: CSIAllTS - Rmrf*{b9} - Smb*{s9} - Hml*{h9}) ///
(eq10: CSIAllUt - Rmrf*{b10} - Smb*{s10} - Hml*{h10}) ///
(eq11: Rmrf - {a1}) ///
(eq12: Smb - {a2}) ///
(eq13: Hml - {a3}) , ///
igmm instruments(eq1 eq2 eq3 eq4 eq5 eq6 eq7 eq8 eq9 eq10: Rmrf Smb Hml) ///
wmatrix(hac bartlett opt) winitial(identity) vce(hac bartlett opt) ///
igmmeps(1e-4) igmmweps(1e-4)
estat overid
```

Note:

Rmrf is the excess returns on market portfolio over the risk-free rate;

Smb is the size factor;

Hml is the value factor;

CNIAll** represents each of 10 CNI All Share sector indices.

The parameters to be estimated are included in braces ({}). Those {b1}...{b10}, {s1}...{s10}, and {h1}...{h10} represent the risk loadings on three factors for each of sector portfolios, while those {a1}, {a2}, {a3} represent the risk premiums on three factors. The command, "estat overid", performs test of overidentifying restrictions. Note that we number the 13 equations and choose Rmrf, Smb, and Hml as instruments for equations 1-10. This is to avoid the redundant moment conditions.

Chapter 4

Can Market, Size, BE/ME and Momentum Factors Predict Future GDP Growth in China?

4.1 Introduction

Fama and French (1992) propose a three-factor model that includes a size factor (SMB) and a book-to-market equity factor (HML) in addition to the market factor, and show that their model can largely explain the cross-sectional variation in the equity returns which cannot be explained by the CAPM in the US. The further supportive evidence for their model is found not only in the US but also across the international markets in their sequential papers, e.g., Fama and French (1993, 1995, 1996, 1998).

The dramatical improvement of the Fama-French model over the CAPM has attract considerable attention of academic researchers to the economic explanation for the success of the SMB and HML factors in the Fama-French model. For example, some interpret the success of the Fama-French model as data snooping, thus, it might not show up during different sample periods (e.g., Black, 1993, MacKinlay, 1995, and Kothari et al., 1995). Another is the behavior-bias explanation that investors' overreactions to the firms' performance by overvaluing well-performed and undervaluing poorly-performed firms rise the size and value premiums (e.g., DeBondt and Thaler, 1987, and Lakonishok et al., 1994). Moreover, the characteristics explanation of Daniel and Titman (1997) says that the expected returns of assets are directly related to their characteristics, such as the size and book-to-market ratio, which are not related to the fundamental risk. Instead of the above explanations, Fama and French (1993, 1995, 1996) and other studies, e.g., Davis et al. (2000), prefer a risk-based explanation, which attributes the impressive performance of the Fama-French model to the fact that SMB and HML act as the state variables in the context of Intertemporal Capital Asset Priding Model (ICAPM) of Merton (1973). However, the supportive evidence for their argument is relatively limited, and the economic intuition behind the model about what fundamental economic risk variables the size and book-to-market equity factors capture is not clear.

If Fama and French's above argument is the truth, then a significant relationship between the factors and certain fundamental economic variables or state variables should be found. Some existing empirical research provides supportive evidence for the risk-based explanation behind the SMB and HML factors by examining the relationship between the Fama-French factors and macroeconomic variables. The mechanism behind the relationship between the Fama-French factors and macroeconomic variables has been examined in some studies. For example, Vassalou and Xing (2004) show that both the size and book-to-market effects contains some intimate information related to default risk. Petkova (2006) find that SMB proxies for innovative information in default spread and HML for innovative information in term spread. Both default and term spreads are systematic risks, and considered to be able to predict time variation in the investment opportunity set. As an aggregated macroeconomic variable, GDP growth should contain some summarized information about the future investment opportunity set. Thus, the Fama-French factors are potentially related to GDP growth. As a candidate for state variables for which both SMB and HML proxy in the context of Merton's (1973) ICAPM, GDP growth is the first natural macroeconomic variable to be examined in order to find out whether the risk-based explanation for the success of the Fama-French model holds.

Among the empirical studies providing support for the risk-based explanation for the Fama-French model, Liew and Vassalou (2000) relate the pricing factors, including the market, size, book-to-market equity, and momentum factors, to future GDP growth. Specifically, they examine whether the profitability of the strategies based on the size, book-to-market equity and momentum effects can predict future GDP growth in ten developed countries. They find that the Fama-French factors do contain significant information about macroeconomic risk. The information contained in SMB and HML is to a large degree independent of that contained in the domestic market factor. Moreover, the ability of SMB and HML to predict future GDP growth is not weakened even in the presence of popular business cycle variables in some countries examined. As to the momentum factor, there is little evidence supporting the similar fundamental risk-based explanation for it. Australia is one of the ten examined countries, in which they find that based on the regressions of future GDP growth on those factors, the coefficient related to SMB is positive and statistically significant and the coefficient on HML is insignificant. However, Nguyen et al. (2009) present an inconsistent finding in the Australia that the coefficient on SMB is negatively significant while the coefficient on HML is positive but insignificant.

The robustness of findings of Liew and Vassalou (2000) has not been examined in the Chinese market. This study attempts to implement the analysis in the context of the Chinese market by linking current and future GDP growth with the four pricing factors, i.e., the market, size, value and momentum factors. Moreover, we distinguish the nominal from real GDP growth in the analysis. The results suggest an opposite conclusion compared to that in Liew and Vassalou (2000). We find that it is the momentum factor, instead of other factors, that captures more information about current and future economic risk, and that methods of constructing the momentum factor would influence the link between the momentum factor and GDP growth. By comparing the regressions of nominal and real GDP growth, the factors perform better in the regressions of nominal GDP growth, which suggests that the factors as a whole seem to capture more information about nominal GDP growth. This might be because the information about inflation is embedded in the equity returns.

The rest of this study is organized as follows. Section 4.2 describes the data. Section 4.3 implements the regression analysis. The regression analysis in the presence of business cycle variables is conducted in Section 4.4. Section 4.5 concludes this study.

4.2 Data

We obtain different data from different sources. The financial data are obtained from a well-known Chinese financial database, named RESSET database. The macroeconomic data are from the National Bureau of Statistics of China (NBS). The proxies for the term and default yield spreads come from the People's Bank of China (PBC, which is China's central bank) and China's biggest five state-own commercial banks. Our choice of RESSET database here as the source of financial data, instead of those in the previous chapters, is only due to the availability of data. In the previous two chapters, the data frequency for returns is daily. In Chapter 2, we use the quarterly financial data of firms to construct the common pricing factors. The quarterly financial data are only available after 2006. As a result, the sample period starts from July 2007. In Chapter 3, we use commercial style indices to construct pricing factors, and the data for style indices are only available after 2010. Thus, the sample period begins after 2010. In this chapter, we examine the relationship between the pricing factors and GDP growth. GDP growth is a macroeconomic variable with quarterly observations. In order to obtain enough sample size for observations of GDP growth, our sample in this chapter starts as early as 1993. Considering the availability of relevant data, we resort to the RESSET database as the source of data for pricing factors. In the rest of this section, we provide a rough description of the data to be employed in the following analysis.

4.2.1 Factors data

The data for the risk-free rate (RF), the market excess return (MKT), and returns on the size factor (SMB), the book-to-market equity factor (HML) and the momentum factor (WML) are obtained from the RESSET Database, which is a well-known Chinese financial database. We choose the data for the returns on these factors that are constructed from the A-share stocks in both Shanghai and Shenzhen stock exchanges with the tradable market value weight.

The proxy for the risk-free rate has changed in different periods as follows: starting from 1st February 1989 to 6th August 2002, the risk-free rate series is derived from the three-month deposit rates; between 7th August 2002 and 7th October 2006, the coupon rates of three-month central bank bills are used as the risk-free rate; after 7th October 2006 the proxy for the risk-free rate is the three-month Shanghai Interbank Offered Rate (Shibor).

The market excess return is the difference between the market stock portfolio return and the risk-free rate. The return on the size factor is calculated as the difference between the small market equity companies' return and the large market equity companies' return. The return on the book-to-market equity factor is the difference between the high book-to-market equity companies' return and the low book-to-market equity companies' return.

As to the momentum factor, the RESSET Database provides several momentum factor series based on different methods of construction. In this study, we examine two momentum factor series noted as "WML11" and "WML12", respectively. "WML11" represents Carhart's (1997) classical momentum factor, which is constructed as the equal-weighted average returns on the portfolio of firms with the highest 30 percentile eleven-month returns lagged one month minus the equal-weighted average returns on the portfolio of firms with the lowest 30 percentile eleven-month returns lagged one month. "WML12" is another commonly used momentum factor (e.g. Liew and Vassalou, 2000) and it is calculated as the difference in the tradable market value-weighted average returns on the portfolio of firms with the highest 30 percentile twelve-month returns lagged one month and the tradable market value-weighted average returns on the portfolio of firms with the lowest 30 percentile twelve-month returns lagged one month.

Daily, weekly and monthly data for the MKT, SMB and HML factors are available from the RESSET Database since 1st July 1992. As for the momentum factor, only monthly data are available starting from July 1992. We choose the monthly data for the factor returns and continuously compound monthly returns into quarterly, semi-annual and annual returns.

We next examine the performance of MKT, SMB, HML, WML11 and WML12 in the Chinese A-share market during the time period since the annual return data are first available, i.e., the sample period is from Quarter 2 1993 up to Quarter 4 2014. Table 4.1 presents the summary of the performance of the market, size, book-to-market equity, and two momentum factors in Panel A, along with the correlation of these factors with each other in Panel B. Note that all the returns in Table 4.1 are the quarterly observed past-year returns in percent. This means that the holding period for the strategies based on these mimicking risk factors is one year. Thus, it is not surprising that the magnitudes of returns on these factors are much higher than those reported in the previous chapters where daily frequency for factor returns is applied. In the next, we also compare the significance and/or signs of the factors with those in the previous chapters.

We first find a positive value for the market excess return which is statistically significant at the 5% level. This finding suggests that taking the long position in the market portfolio with a one-year holding period will produce an average return of about 13.79% over the risk-free return since the establishment of the Chinese stock market. Compared to the previous chapters, the market premium is positive but insignificant in Chapter 3, while it is slightly significant and negative in Chapter 2 due to the historical crash in the Chinese market included in the sample.

We also confirm the positive size premium of 16.14% statistically significant at the 1% level. The positive and highly significant size premium is also found in the two previous chapters, even for the very short daily holding periods. This suggests that small size stocks tend to consistently produce higher returns over large size stocks no matter whether the holding periods are short or long. Although the premium on the style index-formed size factor in Chapter 3 is positive but insignificant, we attribute it to the fact that the relatively small size stocks within the constituents of indices are still the large or middle size stocks in the market, because the index providers usually choose the large firms with good performance records as the constituents of indices. Thus, the difference in returns on small and large size stocks within the constituents of indices is less significant than those within the universe of stocks over the whole market.

As in previous chapters, we do not verify a statistically significant value premium, but we find a negative average premium within the sample period. One possible explanation could be that high book-to-market stocks face a higher risk of distress and they are more likely to survive when the economic outlook is good rather than bad. Given that China's economy grows impressively over the past three decades, the prices of high book-to-market stocks tend to be higher within most part of our relatively long sample period from 1993 to 2014 in this chapter, compared to other developed markets where economic growth is less rapid. This in turn leads to lower returns on those stocks. In contrast, the low book-to-market stocks might be relatively underpriced within our sample period. As a result, the value effect yield an insignificantly negative average return over 1993 to 2014. The sample periods in the two previous chapters only cover the very recent periods, i.e., 2007 to 2013 in Chapter 2 and 2010 to 2014 in Chapter 3, when the China's economic growth has started to slow down (See Figure

4.1).

Although the average returns on the two momentum strategies are both negative, only the return on WML11 is statistically significant at the 1% level while the other is insignificant. This finding suggests that the profitability of the momentum strategy is sensitive to either the weight of each asset in the portfolio or the number of months that the momentum strategy is based on. However, the question which element dominates the performance of different momentum strategies is out of the scope of this study. We also find significantly negative daily momentum premium in Chapter 2. This suggests that investors following the momentum strategies are very unlikely to make profit in the Chinese market, no matter whether short or long holding period they choose.

The correlation between SMB and HML is negative with a relatively high absolute value of 63.72%. This negative and high correlation between size and value effects is also found for the daily premiums in Chapter 3, where the factors are provided by the RESSET database or constructed from style indices. But this is not the case in Chapter 2, where the correlation is only 0.56% for the two factors. The factors in Chapter 2 are constructed following the original idea of Fama and French (1993) such that the factors are largely free of or orthogonal to one another. It might be interesting to explore the correlation between the factors from the perspective of diversified investment. We suspect this highly close correlation between the size and value factors provided by the RESSET database coming from the sequential sorts, instead of independent sorts, by size and book-to-market equity.[1] However, the RESSET Database does not provide a detailed instruction for the construction procedure of portfolios, thus we cannot confirm whether our guess is correct or not. As for the two momentum factors, it is not surprising that the correlation between WML11 and WML12 is very high with a value of 71.63%. The size, book-to-market equity and momentum factors have a relatively low correlation with the market factor, respectively. The low correlation between market factor

[1] There are two popular portfolio sorting procedures to construct the factors. One is that used in Fama and French (1993) where two independent sorts by size and book-to-market equity create the SMB and HML mimicking zero-investment portfolios. The other is the one used by, for example, Liew and Vassalou (2000), where they use three sequential sorts to create SMB, HML and WML because of the small number of securities in some countries.

Table 4.1: A summary of descriptive statistics of and correlation between pricing factors

(in percent)

	MKT	SMB	HML	WML11	WML12
Panel A: Discriptive statistics of pricing factors					
Mean	13.79	16.14	-0.15	-2.58	-0.34
Median	-3.96	10.02	1.03	-1.74	1.53
Std. Dev.	57.30	32.02	16.04	12.67	20.76
T-value	2.25**	4.70***	-0.09	-3.98***	-0.15
Min	-65.13	-20.41	-58.33	-69.71	-97.89
Max	263.63	174.71	29.50	42.83	40.97
Skewness	2.06	2.40	-1.15	-1.93	-2.41
Kurtosis	7.81	10.71	5.96	14.42	12.24
Panel B: Correlations between factors					
MKT	1				
SMB	-0.2054	1			
HML	0.1481	**-0.6372**	1		
WML11	0.1167	-0.3235	0.2735	1	
WML12	0.3567	-0.1369	0.2277	**0.7163**	1

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level
Note: all returns in this table are the quarterly observed past-year returns in percent. Thus, holding period for mimicking risk factor portfolios is one year. Sample period is 1993Q2 to 2014Q4.

and other factors suggests that other factors might contain some extra information which is not captured by the market excess return. We will discuss this a bit further later.

4.2.2 Macroeconomic data

The macroeconomic data, such as the Gross Domestic Product (GDP) and the Value-added of the Secondary Industry (VASI), are obtained from the National Bureau of Statistics of China (NBS).

GDP data

The NBS provides three series of quarterly observed GDP-related data titled as "Gross Domestic Product, Accumulated (100 million yuan)", "Indices of Gross Domestic Product (preceding year=100), Accumulated" and "Gross Domestic Product, Growth Rate (preceding quarter=100)(%)". Firstly, "Gross Domestic Product, Accumulated (100 million yuan)" is the value (in 100 million RMB Yuan) of GDP over the accumulative quarter(s), such as the first quarter, the first to second quarters, the first to third quarters and the first to fourth quarters in a certain year, calculated at the prices in the corresponding year.[2] This GDP series starts from the first quarter of 1992. Because this series reports the GDP at the prices in each corresponding year, we can use this series to calculate the accumulative GDP growth rate without inflation adjustment. We call this GDP growth rate as the nominal GDP growth rate (GDPGR-N). Secondly, "Indices of Gross Domestic Product (preceding year=100), Accumulated" refers to the relative indices that reflect the trend and degree of changes in GDP for a certain period of time, which treats the last year as the base year and is calculated at constant prices. This series is also available since the first quarter of 1992. Subtracting 100 from the reported indices of GDP obtains the commonly reported GDP growth rate. We call this GDP growth rate as the real GDP growth rate (GDPGR-R). Lastly, "Gross Domestic Product, Growth Rate (preceding quarter=100)(%)" is the seasonally adjusted quarter-on-quarter GDP growth rate of this quarter over the last quarter, calculated at constant prices. But this series is only available from the first quarter of 2011, thus we can only have a very small number of observations for this series. We drop this series in this study and focus on the first two series of GDP growth rate.

Although the value of GDP presents the remarkable feature of seasonality, we believe the seasonality would not influence our analysis. This is because in our study we consider the quarterly observed annual accumulated GDP growth rate, that is, the interval to calculate the GDP growth rate spans a year. Both nominal and real GDP growth series utilized in our study are year-on-year changes during the same period of each year, thus, we can avoid contamination by seasonality. We confirm this by running the regressions of our GDP growth rates on seasonal dummy variables and find that the estimated dummy coefficients are highly

Table 4.2: Seasonality test on the GDP and VASI growth rates

	Q1	Q2	Q3	Cons.	Adj. R^2	No.
Panel 1: Seasonality test on GDP growth rate						
GDPGR-N	0.6733	0.2234	-0.0037	15.6796	-0.0345	88
	(0.27)	(0.09)	(0.00)	(8.99)		
GDPGR-R	0.3217	0.0913	-0.0043	10.1130	-0.0303	92
	(0.49)	(0.14)	(-0.01)	(21.62)		
Panel 2: Seasonality test on VASI growth rate						
VASIGR-N	0.4932	0.2052	-0.0191	15.7640	-0.0353	88
	(0.16)	(0.07)	(-0.01)	(7.11)		
VASIGR-R	-0.2043	-0.1957	-0.2174	11.8044	-0.0334	92
	(-0.20)	(-0.19)	(-0.21)	(16.09)		

This table summarizes the results of seasonality tests on both GDP and VASI growth rates. Q1, Q2 and Q3 are the seasonal dummy variables which represent the first, second and third quarter, respectively. The corresponding t-statistics are reported in the parentheses. "No." is the number of observations.

insignificant.[3] The regression results are presented in Panel 1 of Table 4.2.

VASI data

In their work, Liew and Vassalou (2000) use the Industrial Production as one of the business cycle variables. We do not have this data for China, thus we use the Value-added of the Secondary Industry (VASI), also provided by the NBS, as a substitution. "Value-added of the Secondary Industry refers to the final products at market prices produced by all resident units of a country or a region engaged in the secondary industry during a certain period of time. The secondary industry refers to mining and quarrying (except auxiliary activities of mining and quarrying), manufacturing (except repairs for metal products, machinery and equipment), production and supply of electricity, steam, gas and water, and construction."[4] Quite similar to the GDP data,[5] the NBS also provides three series of quarterly observed Value-added of the Secondary Industry. We also use the first two series to calculate the two series of VASI growth

[2]The historical quarterly data are revised on the basis of the Third National Economic Census of China.

[3]Another way to consider seasonality is to remove it from the series by using X-13ARIMA-SEATS, which is a notable software package for seasonal adjustment produced, distributed, and maintained by the U.S. Census Bureau. X-13ARIMA-SEATS is the successor to X-12-ARIMA and X-11-ARIMA. The source code can be found on the U.S. Census Bureau's website. It can also be used together with many statistical packages, such as EViews and Gretl.

[4]Explanatory Notes of Indicators, NBS.

[5]In fact, the GDP is the sum of the Value-added of the Primary Industry, Value-added of the Secondary Industry and Value-added of Tertiary Industry.

rates, i.e., the nominal VASI growth rate (VASIGR-N) and real VASI growth rate (VASIGR-R), and drop the third series. In the regressions we will examine in Section 4.4, we match nominal GDP growth rate with nominal VASI growth rate, and real GDP growth rate with real VASI growth rate. By doing this, it is not necessary to consider the inflation or level of prices as a business cycle variable in our regressions. As for the issue of seasonality, because of the same reason as that for GDP growth, we do not need to worry about this issue for VASI growth in our study. The seasonality test on VASI growth is conducted to confirm it, and the results are summarized in Panel 2 of Table 4.2.

Term spread and Default spread data

One common problem with which would be encountered when studying the developing countries, such as China, is the unavailability of recorded data, especially for early years. Due to the unavailability of data for government and corporate bond yields during the early period that this study covers, there are not enough observations to calculate the common term and default spreads on the quarterly basis. Thus, we use the term-deposit rates and loan interest rates obtained from the People's Bank of China (PBC) and China's biggest five state-own commercial banks[6] as proxies to compute the term spread (TERM) and default spread (DEF). Specifically, TERM is calculated as 5-year RMB term-deposit rates minus 6-month RMB term-deposit rates (lump-sum deposit for lump-sum withdrawal), and DEF as 5-year RMB Loan Interest Rates of Financial Institutions minus 5-year RMB term-deposit rates.

[6]The People's Bank of China (PBC) is China's central bank and provides guidance and floating range of term-deposit rates and loan interest rates for commercial banks from time to time, which is one of the monetary policy tools of the PBC. China's biggest five state-own commercial banks are the Industrial and Commercial Bank of China (ICBC), the Bank of China (BOC), the China Construction Bank (CCB), the Agricultural Bank of China (ABC) and the Bank OF Communications (BOCOM). They implemented the same deposit rates and loan interest rates as the guidance by the PBC and as each other. Since 8th June 2012, all five banks started to implement a bit higher deposit rates than the guidance, but these rates are still the same among the five banks. We use the deposit rates and loan interest rates implemented by the five banks in order to reflect the market rates.

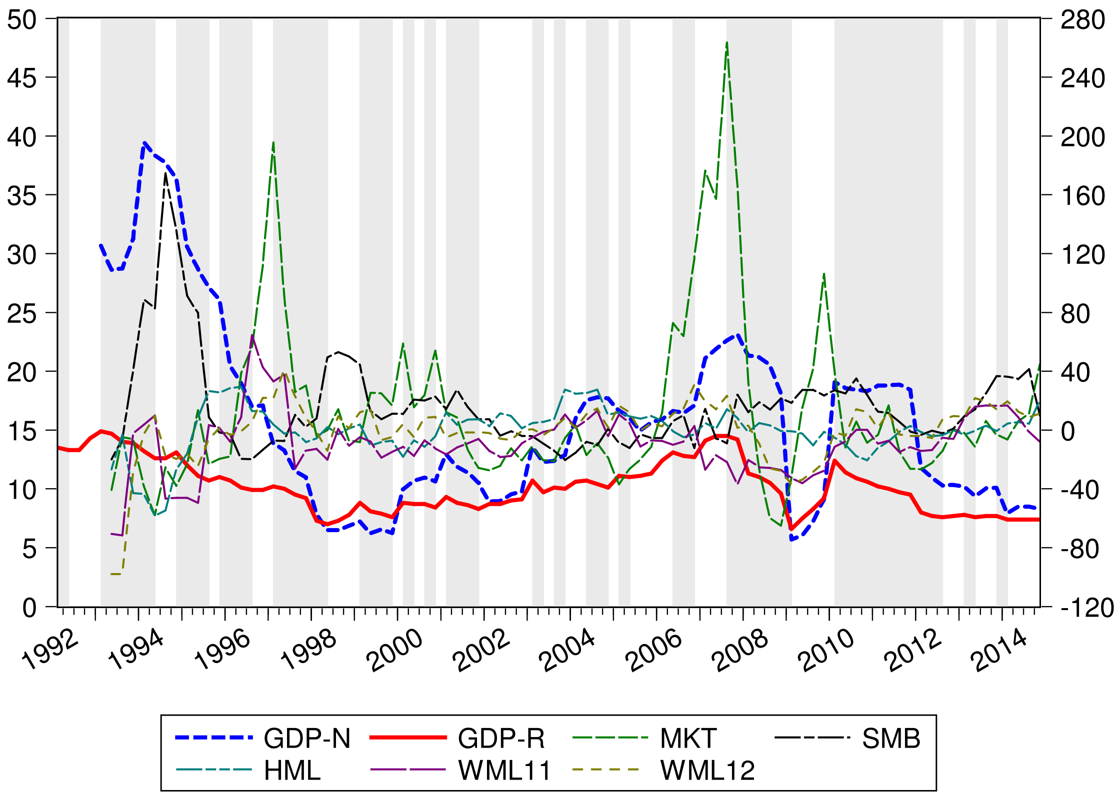

Figure 4.1: Changes in GDP growth rates and factor returns
Note: The nominal and real GDP growth rates are plotted against the left y-axis, while the returns on factors are against the right y-axis. The shaded periods represents economic recessionary periods, within which the real GDP growth rates are less than the previous ones.
Data Source: RESSET Database and National Bureau of Statistics of China (NBS)

4.2.3 Returns on the trading strategies at different states of the macroeconomy

We examine the returns on the MKT, SMB, HML, WML11 and WML12 trading strategies at different states of current and future economic growth. Figure 4.1 depicts the changes in the two GDP growth rates as well as the returns on the market, size, value factors and the two momentum factors. The two thick lines draw the GDP growth rates with the scale on the left vertical axis. The factor returns are represented by five thin lines, and the scale is on the right vertical axis.

Specifically, in this subsection we associate past year's annual returns on each of the factors with current and next year's GDP growth rates, respectively. Then we sort the returns by current and future GDP growth. In line with Liew and Vassalou (2000), we call "good" states of the economy if the states exhibit the highest 25% GDP growth, and "bad" states if the lowest 25% GDP growth. The rest 50% are called "normal" states. Table 4.3 presents the results.

Panel 1 of Table 4.3 reports the mean return on each factor in three economic states measured by current GDP growth in terms of nominal and real rates. The returns on MKT and SMB are all positive in all three economic states. It suggests that return on the market portfolio is consistently over the risk-free rate and that small size stocks consistently produce higher average returns over large size stocks no matter what the economic condition is. This is strong evidence in favour of both market and size effects across economic cycle. The returns on MKT show a positive relationship with current GDP growth in both nominal and real cases. Holding the market portfolio will produce the highest average return over the risk-free rate during the periods of high growth. This suggests the existence of a positive correlation between the financial market performance and the economic status in China. The average return on SMB is highest in good states, but this strategy produces the lowest average return in normal, instead of bad, status of economy.

The small firms tend to perform much better in good state over any other states, and investors know this. The high return on SMB in good state might be because that in good time investors prefer to hold small stocks, which in turn pushes up the prices of small stocks. Thus, the high return on SMB in good state may come from the appreciation in market value of small stocks. In contrast, the high return on SMB in bad time over normal time may be because the investors require higher return to compensate for higher risk that small firms would face in poor economic condition. The high return on SMB in bad time suggests the small firms are very likely to survive when China's economy turns poor.

The returns on HML are mixed signs in three economic states, which suggests weak evidence for the value premium. The strategy of HML produces the highest and positive return in normal states, while negative in both good and bad states with the lowest return linking with high economic growth. These findings suggest that the positive value premium only exists in normal time. In good state, investors might prefer growth to value stocks, which rises up the prices of growth stocks. The negative return on HML in good state might be due to the high returns on growth stocks which is the results of appreciation in market values of growth stocks. On the other hand, the negative premium on HML might be because value firms are much likely to fail in poor economy. Investors know that value firms are very likely to suffer from financial distress. If the value firms are unlikely to survive in bad time, then investors would not like to hold the stocks of those firms.

WML12 shows the similar results to HML, except for a small positive return in bad states in the case of real growth. WML11 produces negative returns in all states with the poorest performance showing up in good states. The findings here suggest that the momentum strategy based on WML12 is the one that can provide positive average return only in normal time in both cases of normal and real GDP growth. There is no guarantee for positive returns in other states. The momentum strategy based on WML11 fails in all the economic states.

Panel 2 links the future GDP growth to the performance of each strategy. The results show that the market excess returns and the returns on SMB are all positive leading each state of future economic growth. The returns on HML present a similar relationship with the state of future GDP growth to that of current GDP growth. As we show in the introduction, several studies find that the size and value effects capture contains some intimate information about the systematic risk associated with default and/or term spreads. Those information may be aggregated in the macroeconomic variable, for example, the GDP growth. Thus, SMB and HML should contain some information helpful for predicting future GDP growth, as shown in Liew and Vassalou (2000). However, Panel 2 of this table does not present supportive evidence for the above argument, because there is no clear positively monotonic relationship between the size and value factors and future GDP growth. On the other hand, the returns on WML11 seem to have a negatively monotonic relation with future real GDP growth, although the returns are all negative followed by all states of future growth. The negatively monotonic relation also turns up between the returns on WML12 and future GDP growth in both nominal and real cases. Based on these findings from Table 4.3, it suggests that the momentum factor might have the strongest power to capture the macroeconomic fundamental risks and to predict future economic growth in China. We will further explore it in the following sections by conducting regression analysis.

4.3 Regressions analysis

We examine the relation between the factors of MKT, SMB, HML, WML11 and WML12 and the nominal and real GDP growth rates by employing univariate and multivariate regression analysis.

4.3.1 Univariate regression

In this subsection, we examine the univariate regressions of current and future GDP growth rates (both nominal and real rates) on returns of the MKT, SMB, HML, WML11 and WML12 strategies over three different past holding periods, and reports the results in Table 4.4 and 4.5, respectively. In each table, Panel 1 reports the results from regressions that use quarterly holding period returns, while Panels 2 and 3 report the results from regressions that employ semi-annual and annual holding period returns, respectively.

All regressions based on the quarterly observed data are of the following forms:

$$GDPGR^{i}_{(t-4,t)} = a + b * Factor_{(t-n,t)} + \varepsilon_{(t-4,t)} \tag{4.1}$$

$$GDPGR^{i}_{(t,t+4)} = a + b * Factor_{(t-n,t)} + \varepsilon_{(t,t+4)} \tag{4.2}$$

where $n = 1, 2, \text{or} 4$ and the subscript $(t-n,t)$ represents the past n quarters holding periods; $i = R \, \text{or} \, N$ which represents Real or Nominal GDP growth; $Factor$ represents the returns on one of the factors, i.e., MKT, SMB, HML, WML11 or WML12; ε is the residual term of the regression and is assumed to have zero expectation and is uncorrelated with the factor returns. Equation (4.1) represents the regression of current annual nominal or real GDP growth rate on each of the factors. It examines the extent to which each factor can capture the macroeconomic

Table 4.3: Past-year returns on factors sorted by current and future GDP growth: quarterly observations

	Good	Normal	Bad	Diff.	Good	Normal	Bad	Diff.	Good	Normal	Bad	Diff.
Panel 1: Current GDP growth												
	MKT				SMB				HML			
GDPGR-N	22.26	11.63	9.55	12.71	34.20	4.06	21.71	12.49	-5.75	2.69	-0.12	-5.63
	[87.63]	[48.08]	[32.28]	(0.64)	[52.37]	[15.14]	[18.02]	(1.06)	[26.34]	[11.01]	[8.25]	(-0.96)
GDPGR-R	37.16	7.11	3.47	33.69*	33.59	4.63	21.19	12.41	-10.65	5.65	-0.99	-9.66*
	[87.08]	[48.43]	[19.93]	(1.77)	[52.77]	[15.49]	[17.82]	(1.04)	[21.98]	[13.09]	[6.50]	(-1.98)
	WML11				WML12							
GDPGR-N	-23.40	-0.21	-11.61	-11.79*	-12.89	7.04	-2.22	-10.67				
	[24.09]	[17.90]	[14.19]	(-1.98)	[30.84]	[11.50]	[16.42]	(-1.43)				
GDPGR-R	-21.86	-3.44	-6.84	-15.02**	-7.26	2.52	0.99	-8.25				
	[24.37]	[18.90]	[15.98]	(-2.42)	[33.25]	[12.51]	[16.68]	(-1.04)				
Panel 2: Future GDP growth												
	MKT				SMB				HML			
GDPGR-N	31.26	7.65	8.29	22.97	37.98	2.73	17.77	20.21	-13.51	5.37	0.78	-14.29
	[83.78]	[42.94]	[53.95]	(1.06)	[52.21]	[14.80]	[17.38]	(1.68)	[21.02]	[13.55]	[4.76]	(-3.04)
GDPGR-R	4.80	19.25	12.11	-7.31	15.13	16.48	13.77	1.36	-3.68	1.56	-1.62	-3.20
	[57.60]	[63.61]	[49.87]	(-0.44)	[40.48]	[34.20]	[18.79]	(0.14)	[23.35]	[15.79]	[4.14]	(-0.14)
	WML11				WML12							
GDPGR-N	-23.61	-4.58	-5.33	-18.28**	-11.50	2.04	3.65	-15.15*				
	[24.21]	[17.35]	[19.41]	(-2.70)	[34.62]	[9.61]	[16.52]	(-1.81)				
GDPGR-R	-13.39	-10.80	-3.39	-10.01	-8.02	-1.80	7.66	-15.68*				
	[24.96]	[20.41]	[18.09]	(-1.49)	[34.67]	[12.51]	[12.92]	(-1.94)				

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level

GDPGR-N represents the nominal GDP growth rate and GDPGR-R represents the real GDP growth rate. Panel 1 and 2 report the mean returns (in percent) on each factor in different economic states measured by current and future GDP growth, respectively. We characterize “good” states of the economy if the states exhibit the highest 25% GDP growth, and “bad” states if the lowest 25% GDP growth. The rest 50% are called “normal” states. “Diff.” represents the difference between the average returns in good states and those in bad states. The t-statistics are reported in the parentheses, while the standard deviations are reported in the square brackets.

fundamental risk. Equation (4.2) conducts the regression of future annual nominal or real GDP growth rate on one of the factors, and examines the ability of each factor to predict future economic growth.

Same as the argument of Liew and Vassalou (2000), GDP growth rates are observed at quarterly frequencies, and therefore, consecutive annual growth rates have three overlapping quarters and this would induces serial correlation in the residuals of our regressions. To correct for this, we use the Newey and West (1987) estimator and set the parameter q equal to three.[7]

Table 4.4 reports the results from equation (4.1) and the results from equation (4.2) are summarized in Table 4.5. Because the intercept terms are not our focus, we do not report the estimated intercepts for all regressions of GDP growth involved in this study. Based on the results in Table 4.4, there is no significant change in the slope coefficients related to individual factors for different holding periods. But we find that for the regressions of current nominal GDP growth rate, the adjusted R^2s of the univariate regressions of current GDP growth rate on SMB, HML, WML11 and WML12 increase as the holding period increases from over past one quarter to over past year. However, it is not exactly the case for current real GDP growth, except that the adjusted R^2 for the univariate regression of current real GDP growth on the past year's return on the market factor increases significantly to 7.37%. The similar findings are also shown in Table 4.5 for the univariate regressions of future GDP growth rate. In general, the quarterly year-on-year returns on the factors seem to have greatest power to capture the macro-economic risk and to predict future economic growth as presented by the adjusted R^2s reported in the third panels of Tables 4.4 and 4.5. Thus, the analysis hereafter will focus on the annual holding period, i.e., Panel 3 of each table.

From Panel 3 of Table 4.4, we find negative relations between current nominal GDP growth and the market excess return, the book-to-market equity factor and the momentum factors, respectively, and only the slope coefficient of WML12 is statistically significant. We also

[7]Setting the parameter q is also based on the rule for choosing the lag parameter q that $q = 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression (see Stock and Watson, 2007, page 607). Our results remain qualitatively the same when the parameter q in the Newey-West estimator takes the alternative values of 4, 5, or 6.

Table 4.4: Univariate regressions of current annual GDP growth rates based on different calculation methods on past one-quarter, two-quarter and four-quarter factor returns

$$GDPGR^{i}_{(t-4,t)} = a + b * Factor_{(t-n,t)} + \varepsilon_{(t-4,t)} \text{ where } n = 1, 2, \text{or} 4 \text{ and } i = N \text{ or } R.$$

	Slope cofficients					Adjusted R^2					No.
	MKT	SMB	HML	WML11	WML12	MKT	SMB	HML	WML11	WML12	
Panel 1: past one-quarter holding period											
GDPGR-N	-0.0481	0.1952	-0.2503	-0.0391	-0.1550	0.0026	0.0559	0.0400	-0.0090	0.0113	88
	(-1.01)	(1.43)	(-1.39)	(-0.48)	(-1.40)						
GDPGR-R	0.0008	0.0095	-0.0612**	-0.0248*	-0.0339**	-0.0113	-0.0092	0.0373	0.0097	0.0290	90
	(0.05)	(0.32)	(-2.45)	(-1.67)	(-2.21)						
Panel 2: past two-quarter holding period											
GDPGR-N	-0.0327	0.1608	-0.2260	-0.0851	-0.1362***	0.0044	0.1005	0.0740	0.0160	0.0507	88
	(-0.81)	(1.45)	(-1.32)	(-1.23)	(-2.67)						
GDPGR-R	0.0076	0.0044	-0.0453*	-0.0337**	-0.0329**	0.0006	-0.0103	0.0386	0.0595	0.0626	89
	(0.57)	(0.18)	(-1.90)	(-2.10)	(-2.14)						
Panel 3: past four-quarter holding period											
GDPGR-N	-0.0079	0.1046**	-0.1937	-0.0919	-0.1062**	-0.0085	**0.1686**	**0.1434**	**0.0481**	**0.0664**	87
	(-0.32)	(2.17)	(-1.60)	(-1.55)	(-2.10)						
GDPGR-R	0.0106	0.0054	-0.0276	-0.0258*	-0.0183	**0.0737**	-0.0049	0.0339	0.0568	0.0219	87
	(1.57)	(0.47)	(-1.41)	(-1.74)	(-0.90)						

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level
GDPGR-N represents the nominal GDP growth rate calculated from the series of "Gross Domestic Product, Accumulated(100 million yuan)" and GDPGR-R represents the real GDP growth rate calculated from the series of "Indices of Gross Domestic Product (preceding year=100)".
The t-statistics reported are calculated from standard errors corrected for serial correlation and White (1980) heteroskedasticity using the Newey-West (1987) estimator. The rule for choosing the lag parameter q is that $q = 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression (see the text by Stock and Watson, 2007, page 607). The t-statistics are reported in the parentheses.

Table 4.5: Univariate regressions of future GDP growth rates based on different calculation methods on past one-quarter, two-quarter and four-quarter factor returns

$GDPGR^i_{(t,t+4)} = a + b * Factor_{(t-n,t)} + \varepsilon_{(t,t+4)}$ where $n = 1, 2, \text{or} 4$, and $i = N \text{ or } R$.

	Slope cofficients					Adjusted R^2					No.
	MKT	SMB	HML	WML11	WML12	MKT	SMB	HML	WML11	WML12	
Panel 1: past one-quarter holding period											
GDPGR-N	0.0228	0.0910	-0.3197**	-0.0946*	-0.1523***	-0.0083	0.0036	0.0873	0.0126	0.0537	86
	(0.48)	(0.69)	(-2.57)	(-1.95)	(-3.10)						
GDPGR-R	0.0133	-0.0124	-0.0475**	-0.0182	-0.0299*	0.0062	-0.0076	0.0208	0.0016	0.0260	86
	(0.95)	(-0.46)	(-2.06)	(-1.22)	(-1.96)						
Panel 2: past two-quarter holding period											
GDPGR-N	0.0163	0.0747	-0.2700***	-0.1361**	-0.1738***	-0.0076	0.0152	0.1267	0.0813	0.1555	85
	(0.45)	(0.80)	(-2.85)	(-2.25)	(-3.62)						
GDPGR-R	0.0089	-0.0059	-0.0300	-0.0212	-0.0297**	0.0081	-0.0094	0.0141	0.0223	0.0622	85
	(0.95)	(-0.32)	(-1.54)	(-1.56)	(-2.48)						
Panel 3: past four-quarter holding period											
GDPGR-N	0.0013	0.0526	-0.1814**	-0.1247**	-0.1675***	-0.0122	**0.0452**	**0.1582**	**0.1250**	**0.2316**	83
	(0.06)	(1.52)	(-2.35)	(-2.49)	(-2.96)						
GDPGR-R	0.0007	-0.0013	-0.0062	-0.0169*	-0.0230*	-0.0119	-0.0119	-0.0096	**0.0225**	0.0514	83
	(0.13)	(-0.15)	(-0.37)	(-1.81)	(-1.86)						

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level
GDPGR-N represents the nominal GDP growth rate calculated from the series of "Gross Domestic Product, Accumulated(100 million yuan)" and GDPGR-R represents the real GDP growth rate calculated from the series of "Indices of Gross Domestic Product (preceding year=100)".
The t-statistics reported are calculated from standard errors corrected for serial correlation and White (1980) heteroskedasticity using the Newey-West (1987) estimator. The rule for choosing the lag parameter q is that $q = 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression (see the text by Stock and Watson, 2007, page 607). The t-statistics are reported in the parentheses.

find a positive and statistically significant relation between current nominal GDP growth and SMB. For the current real GDP growth, the relations between current real GDP growth and the market and size factors are both positive but not statistically significant. The relations for the book-to-market equity and momentum factors are also negative but the statistically significant slope coefficient is the one related to WML11. These findings suggest that SMB and WML12 alone can capture only a small portion of information related to current nominal economic growth, while WML11 can only capture an even smaller portion of information related to current real GDP growth.

From Panel 3 of Table 4.5, we find negative relations between future nominal GDP growth and the book-to-market equity and momentum factors which are statistically significant, and positive relations but not statistically significant for the market excess return and the size factor. In the case of future real GDP growth, except for a positive but not statistically significant slope coefficient related to the market factor, the rest slope coefficients are negative, among which only those related to the two momentum factors are statistically significant. The findings here suggest that the book-to-market equity and momentum factors have greater ability to predict future nominal economic growth compared to other factors. For future real GDP growth, the two momentum factors, WML11 and WML12, still have a certain amount of ability to predict economic growth, but this ability seems not as strong as that in the case of nominal growth.

It is well known that the market factor is a leading indicator for future economic growth. Fama (1981) finds that the relationship between future economic growth and the market factor is significantly positive in the United States. Aylward and Glen (2000) extend the similar findings to international evidence. This relationship is also confirmed by Liew and Vassalou (2000) where the positive relationship exists in nine out of ten examined countries, six countries have a statistically significant relationship, and eight countries have positive adjusted R^2. We also find a positive relationship between future GDP growth and returns on the market factor in both nominal and real cases, but this relationship in both cases is statistically insignificant. The magnitude of the estimated slope coefficient is very small, and the corresponding adjusted R^2 is negative. Based on these findings from Table 4.5, it suggests that the market factor is not a good leading indicator for China's future economic growth. Moreover, our finding of statistically insignificant relationships of current and future GDP growth with respect to the market factor is consistent with the reality in the Chinese economy and stock market in the sense that China has a very high average GDP growth rate in the past three decades, but the stock market does not reflect this economic achievement. This might be attributed to the fact that much of trading in the Chinese market is triggered by unsophisticated individual investors who behave speculatively and focus on short-term investment and that the profits made by investors more rely on the release of information about the government policy, instead of the fundamental performance of listed firms.

The findings from Table 4.5 suggest that among those factors examined in this study the momentum factor has the greatest explanatory power for China's future economic growth and the relationship between future GDP growth and the momentum factor seems negative. This is quite different with many countries examined in Liew and Vassalou (2000). Thus, it would be interesting to compare our findings with theirs, although they only examine the relationship between the factors and real GDP growth.

We find a positive relation between SMB and future nominal GDP growth but a negative relation between SMB and future real GDP growth, while both relations are not statistically significant. In their studies, the slope coefficients with respect to the size factor are positive in all of the ten countries and seven coefficients are statistically significant.

In our study, the slope coefficients related to HML are all negative in both cases of future nominal and real GDP growth, which suggests a higher return on HML is followed by a poorer future GDP growth. One of the negative slopes is statistically significant in the case of future nominal, but not real, GDP growth. In their findings, the coefficients are positive in eight of the ten countries for HML and five coefficients are statistically significant.

The relationships between future GDP growth rate and the momentum factors of WML11 and WML12 are negative and statistically significant. The results suggest that WML12 has the strongest explanatory power for future nominal GDP growth with a statistically significant slope coefficient and the highest adjusted R^2 of 23.16%. This finding contradicts that in Liew and Vassalou (2000), where they find that the coefficients of WML are positive in six of the ten countries, and only two coefficients are statistically significant. In general, they find that the momentum strategy appears to contain little information about future economic growth, while a positive and equally strong relation between the performance of HML and SMB and future economic growth exists in most countries.

4.3.2 Bivariate regression

The information contained in the returns of SMB, HML, WML11 and WML12 about future economic growth is largely independent of the information contained in the market factor. To confirm this, we first run the regressions of each of trading strategies - SMB, HML, WML11 and WML12 - on the market excess return, then run the bivariate regressions of GDP growth on the market factor together with one of the other factors, respectively.

The results of the regressions of each of trading strategies on the market excess return are reported in Table 4.6. The results show that the beta coefficients are generally statistically insignificant and that the adjusted R^2s are small, except that for the regression of WML12 on the market factor, the beta coefficient is statistically significant at the 1% level and the adjusted R^2 is 11.7%, relatively higher than others. Therefore, the relationship between SMB, HML, WML11 or WML12, and future GDP growth seems unlikely to be induced by the market factor.

We next formally compare the information content of SMB, HML, WML11 and WML12 with that of MKT by examining the bivariate regressions in the rest of this subsection. The bivariate regressions include the market factor and one of the other factors, respectively, and have the following forms:

$$GDPGR^i_{(t-4,t)} = a + b * MKT_{(t-4,t)} + c * Factor_{(t-4,t)} + \varepsilon_{(t-4,t)} \quad (4.3)$$

$$GDPGR^i_{(t,t+4)} = a + b * MKT_{(t-4,t)} + c * Factor_{(t-4,t)} + \varepsilon_{(t,t+4)} \quad (4.4)$$

where the subscript $(t-4,t)$ represents the past four quarters holding periods; $i = R \, \text{or} \, N$

Table 4.6: Regressions of returns on trading strategies on excess market return

$$Factor_t = alpha_F + beta_F * MKT_t + \varepsilon_t$$

SMB			HML		
beta	alpha	adj.R^2	beta	alpha	adj.R^2
-0.1148	17.7256**	0.0309	0.0415	-0.7240	0.0104
(-1.54)	(2.65)		(1.13)	(-0.21)	

WML11			WML12		
beta	alpha	adj.R^2	beta	alpha	adj.R^2
0.0428	-9.5447***	0.0020	0.1293***	-2.1225	0.1170
(0.56)	(-2.97)		(3.42)	(-0.62)	

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level
The t-statistics are reported in the parentheses.

which represents Real or Nominal GDP growth; MKT is the quarterly market excess return over the risk-free rate; $Factor$ represents the returns on one of the trading strategies - SMB, HML, WML11 and WML12; ε is the residual term of the regression and is assumed to have zero expectation and is uncorrelated with the factor returns. Equation (4.3) represents the regression of current annual nominal or real GDP growth rate on the market factor and one of other factors, which examines the extent to which the market factor together with one of other factors can capture the macro-economic fundamental risk. Equation (4.4) conducts the regression of future annual nominal or real GDP growth rate on the market factor and one of other factors, and examines the ability of both the market factor and one of the other factors to predict future economic growth.

Table 4.7 summarizes the regression results with the left panel showing the results from the regressions of current GDP growth and the right panel presenting those of future GDP growth. For the current GDP growth reported in Panel 1 of Table 4.7, adding the market factor into the regressions on one of the other factors slightly reduces the adjusted R^2 in the case of nominal GDP growth while significantly increases the adjusted R^2 in the case of real GDP growth. For the future GDP growth in Panel 2, adding the market factor into the bivariate regressions slightly reduces the adjusted R^2 in the case of nominal GDP growth, except for the one on WML12 whose adjusted R^2 increases to 26.4%. The adjusted R^2 does not change too much in the case of real GDP growth.

In the case of current nominal GDP growth, the sign and statistical significance of slope coefficients related to SMB, HML, WML11 and WML12 remain the same; while for current real GDP growth, the corresponding coefficients have the similar situation, but the slopes related to the market factor turn to be statistically significant. There is no change in the signs and statistically significance of slope coefficients related to all factors happening in the case of future GDP growth.

In general, based on the above findings, even in the presence of the market factor, the coefficients of the size, BE/ME and momentum factors remain the same as those in the univariate regressions, while their magnitude is largely unchanged. As expected, the adjusted R^2s are now larger than those from the univariate regressions that include only the market factor in Section 4.3.1. Therefore, the returns of SMB, HML and WML contain some extra information about current and future state of the macro-economy, over the information contained in the market factor. Once again, WML appears to have the greatest ability to explain current and future economic growth.

4.3.3 Multivariate regression

The above finding that, compared to the market, size and BE/ME factors, the momentum factor constructed from the Chinese A-share stock market seems to have the greatest ability to explain current and future economic growth in China can be further confirmed by running the multivariate regressions of current and future GDP growth rates on all of the four factors, respectively. The results are summarized in Table 4.8 where Panel 1 presents the ability of all factors together to capture the macroeconomic fundamental risks and Panel 2 shows the overall ability of these factors to predict future growth of the economy. The regressions take the following forms:

$$GDPGR^{i}_{(t-4,t)} = a + b * MKT_{(t-4,t)} + s * SMB_{(t-4,t)} + h * HML_{(t-4,t)} + w * WML_{(t-4,t)} + \varepsilon_{(t-4,t)} \quad (4.5)$$

Table 4.7: Bivariate regressions of GDP growth rates based on different calculation methods on past four-quarter factor returns

Panel 1: $GDPGR^i_{(t-4,t)} = a + b * MKT_{(t-4,t)} + c * Factor_{(t-4,t)} + \varepsilon_{(t-4,t)}$;

Panel 2: $GDPGR^i_{(t,t+4)} = a + b * MKT_{(t-4,t)} + c * Factor_{(t-4,t)} + \varepsilon_{(t,t+4)}$; where $i = N$, or R.

	Panel 1: Current GDP growth				*Panel 2: Future GDP growth*			
	MKT	SMB	Adj. R^2	No.	MKT	SMB	Adj. R^2	No.
GDPGR-N	0.0043	0.1062**	0.1596	87	0.0075	0.0553	0.0369	83
	(0.22)	(2.24)			(0.39)	(1.59)		
GDPGR-R	0.0117*	0.0096	0.0843	87	0.0006	-0.0011	-0.0242	83
	(1.83)	(0.83)			(0.11)	(-0.13)		
	MKT	HML	Adj. R^2	No.	MKT	HML	Adj. R^2	No.
GDPGR-N	0.0001	-0.1938	0.1332		0.0087	-0.1859**	0.1528	
	(0.01)	(-1.60)			(0.47)	(-2.33)		
GDPGR-R	0.0120*	-0.0339	0.1311		0.0010	-0.0067	-0.0213	
	(1.88)	(-1.51)			(0.18)	(-0.39)		
	MKT	WML11	Adj. R^2	No.	MKT	WML11	Adj. R^2	No.
GDPGR-N	-0.0040	-0.0906	0.0376		0.0070	-0.1271**	0.1173	
	(-0.18)	(-1.52)			(0.45)	(-2.58)		
GDPGR-R	0.0118**	-0.0296**	0.1526		0.0015	-0.0174*	0.0124	
	(2.17)	(-2.19)			(0.31)	(-1.79)		
	MKT	WML12	Adj. R^2	No.	MKT	WML12	Adj. R^2	No.
GDPGR-N	0.0067	-0.1127**	0.0573		0.0266	-0.1944***	**0.2640**	
	(0.27)	(-2.20)			(1.44)	(-4.46)		
GDPGR-R	0.0148**	-0.0329*	0.1588		0.0043	-0.0274**	0.0545	
	(2.20)	(-1.93)			(0.82)	(-2.58)		

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level
GDPGR-N represents the nominal GDP growth rate calculated from the series of "Gross Domestic Product, Accumulated(100 million yuan)" and GDPGR-R represents the real GDP growth rate calculated from the series of "Indices of Gross Domestic Product (preceding year=100)".
The t-statistics reported are calculated from standard errors corrected for serial correlation and White (1980) heteroskedasticity using the Newey-West (1987) estimator. The rule for choosing the lag parameter q is that $q = 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression (see Stock and Watson, 2007, page 607).The t-statistics are reported in the parentheses.

$$GDPGR^i_{(t,t+4)} = a + b * MKT_{(t-4,t)} + s * SMB_{(t-4,t)} + h * HML_{(t-4,t)} + w * WML_{(t-4,t)} + \varepsilon_{(t,t+4)} \quad (4.6)$$

The adjusted R^2 increases in both cases of current nominal and real GDP growth. Compared with the results in Panel 1 of Table 4.7, the slope coefficients related to MKT are still statistically significant in the regression of real GDP growth rate. The slopes related to SMB remain statistically insignificant but the signs change in the regression of real growth. The signs and significance of slopes with respect to the BE/ME and momentum factors remain unchanged. The adjusted R^2 also increases for the future nominal GDP growth, but decreases for the future real growth. The positive relationship between future GDP growth and the market factor still exists and the slope turns to be statistically significant in the regression of future nominal GDP growth rate on the factors with WML12 as the proxy for the momentum factor. The slopes related to SMB and HML are now all negative and still statistically insignificant. The negative and statistically significant slope coefficients related to the two momentum factors confirm that the momentum factor has the greatest ability to predict future economic growth in both nominal and real cases, and that it performs better in predicting nominal growth inferred by the highest adjusted R^2 of 35.77%.

4.4 Regressions that include business cycle variables

In this section, we examine how much the information contained in the factors regarding future economic growth is left after considering the presence of some popular business cycle variables, such as the risk-free rate, TERM, DEF and VASI growth. We do not run similar regressions for current GDP growth in this section, because VASI is an important component of GDP and involving VASI into the regression of current GDP growth will make little sense. The multivariate regressions to be estimated are of the form:

$$\begin{aligned} GDPGR^i_{(t,t+4)} &= a + b * MKT_{(t-4,t)} + s * SMB_{(t-4,t)} + h * HML_{(t-4,t)} \\ &\quad + w * WML_{(t-4,t)} + d * RF_{(t)} + e * TERM_{(t)} \\ &\quad + f * DEF_{(t)} + g * VASI_{(t)} + \varepsilon_{(t,t+4)} \end{aligned} \quad (4.7)$$

Table 4.8: Multivariate regressions of GDP growth rates based on different calculation methods on past four-quarter returns

Panel 1: $GDPGR^i_{(t-4,t)} = a + b*MKT_{(t-4,t)} + s*SMB_{(t-4,t)} + h*HML_{(t-4,t)} + w*WML_{(t-4,t)} + \varepsilon_{(t-4,t)}$;

Panel 2: $GDPGR^i_{(t,t+4)} = a + b*MKT_{(t-4,t)} + s*SMB_{(t-4,t)} + h*HML_{(t-4,t)} + w*WML_{(t-4,t)} + \varepsilon_{(t,t+4)}$;

where $i = N$, or R.

Panel 1: Current GDP growth						
	MKT	SMB	HML	WML11	Adj. R^2	No.
GDPGR-N	0.0056	0.0673	-0.0966	-0.0403	0.1765	87
	(0.30)	(1.15)	(-0.69)	(-0.65)		
GDPGR-R	0.0123**	-0.0059	-0.0324	-0.0259*	0.1737	87
	(2.30)	(-0.44)	(-1.05)	(-1.87)		
	MKT	SMB	HML	WML12	Adj. R^2	
GDPGR-N	0.0161	0.0783	-0.0755	-0.0923	**0.2172**	
	(0.76)	(1.60)	(-0.58)	(-1.53)		
GDPGR-R	0.0154**	-0.0004	-0.0278	-0.0286*	**0.1820**	
	(2.30)	(-0.03)	(-1.05)	(-1.73)		

Panel 2: Future GDP growth						
	MKT	SMB	HML	WML11	Adj. R^2	No.
GDPGR-N	0.0103	-0.0288	-0.1874	-0.1048**	0.2199	83
	(0.72)	(-0.56)	(-1.52)	(-2.00)		
GDPGR-R	0.0010	-0.0094	-0.0121	-0.0198**	0.0010	83
	(0.21)	(-0.91)	(-0.50)	(-2.04)		
	MKT	SMB	HML	WML12	Adj. R^2	
GDPGR-N	0.0295*	-0.0040	-0.1536	**-0.1725*** **	**0.3577**	
	(1.72)	(-0.13)	(-1.62)	(-4.27)		
GDPGR-R	0.0040	-0.0049	-0.0071	-0.0270**	0.0342	
	(0.74)	(-0.52)	(-0.33)	(-2.57)		

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level
GDPGR-N represents the nominal GDP growth rate calculated from the series of "Gross Domestic Product, Accumulated(100 million yuan)" and GDPGR-R represents the real GDP growth rate calculated from the series of "Indices of Gross Domestic Product (preceding year=100)".
The t-statistics reported are calculated from standard errors corrected for serial correlation and White (1980) heteroskedasticity using the Newey-West (1987) estimator. The rule for choosing the lag parameter q is that $q = 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression (see Stock and Watson, 2007, page 607).The t-statistics are reported in the parentheses.

where RF is the risk-free rate; $TERM$ is the term spread calculated as 5-year RMB term-deposit rates minus 6-month RMB term-deposit rates; DEF is the default spread calculated as 5-year RMB Loan Interest Rates of Financial Institutions minus 5-year RMB term-deposit rates. $VASI$ is the Value-added of the Secondary Industry. Our choice of the business cycle variables is again largely constrained by the availability of relevant data. In Liew and Vassalou (2000), the business cycle variables are Treasury bill yield or Call Money Rate, dividend yield on country market capitalization index, ten-year government yield minus T-Bill yield (term yield spread), and past one-year growth in country industrial production. The relevant data for those variables in China are generally only available for recent years. Thus, we choose the most similar variables to substitute those in Liew and Vassalou (2000). We use the risk-free rate as proxy for Treasury bill yield, term-deposit rate spread for term yield spread, and value-added of the secondary industry for industrial production. The data for dividend yield in the Chinese market is also not available over our long sample period, thus we drop it from, but add default spread into, our regression model in this section.

These business cycle variables are believed to be able to predict economic output growth. It is not surprising that VASI can predict future GDP growth as it is the main component of China's GDP. As described in Section 4.2, three-month deposit rate, coupon rate of central bank bill, and Shibor are proxies for the risk-free rate in China during three sequential periods. All these are short end of yield curve that can be influenced by central bank as the monetary instrument. Thus, the ability of the risk-free rate to predict future output is not surprising. The ability of term spread and default spread to predict are verified in many studies, for example, Stock and Watson (1989). The standard economic explanation for the ability of term spread to predict output growth is that term spread is an indicator of an effective monetary policy. That is, the monetary tightening results in high short-term interest rates, relative to long-term interest rates, which in turn causes the economy to slow down (Bernanke and Blinder (1992)). The predictive content of default spread may arise from expectations of default risk based on the expectations of sales and profits (Stock and Watson, 1989), or from that it is a sensitive measure of monetary policy (Bernanke and Blinder, 1992) or it detects the influences of supply and demand for private debt (Friedman and Kuttner, 1992).[8] Thus, all the four variables are leading indicators of economic output growth.

We include the business cycle variables into the multivariate regressions in Section 4.3.3, and report the results in Table 4.9. Compared with those in Panel 2 of Table 4.8, in the presence of business cycle variables in addition to the market, size, value and momentum factors, the adjusted R^2 increases dramatically in both cases of nominal and real GDP growth. This confirms the ability of our business cycle variables to forecast future economic activity. The slope coefficients related to MKT are positive and significant for the regressions of nominal GDP growth. But for the regressions of real GDP growth, the slopes related to MKT are negative and only the one in the regression with WML12 included is significant. The slope coefficients related to the size factor are still insignificant in both nominal and real cases, and they are negative for nominal and positive for real GDP growth. As for the value factor, the slope coefficients are negative and insignificant in the nominal case, but positive and significant in the real case. This is an interesting finding and might be because some mixed information content of HML helpful for forecasting future real GDP growth is filtered by business cycle variables. On the other hand, in the presence of business cycle variables, the significance of slope coefficients related to momentum factor tends to decrease. The only significant one is in the regression of nominal GDP growth when WML12 is included. The slopes keep the negative sign in the nominal case, but change to positive in the real case. Thus, including the business cycle variables, the slope coefficients related to the pricing factors change signs in the regression of real GDP growth, while keep signs in the regression of nominal GDP growth, compared to the regressions without business cycle variables.

The slopes related to the business cycle variables are generally significant and their signs are also in line with our expectations. The risk-free rate is the short-term interest rate and can be seen as a monetary instrument of central bank. A high short-term interest rate implies monetary tightening which would slows down the economic growth, and vice versa. This

[8]The more comprehensive review for the relevant literature is given by, for example, Stock and Watson (2003) and Wheelock and Wohar (2009).

relationship between short-term rate and future economy is confirmed by the negative slope of risk-free rate. The positive slope of term spread and negative slope of default spread are consistent with the economic interpretation for the ability of term and default spreads to predict output growth that we just mentioned above. An increase in term spread implies the ease of monetary policy (a decrease in short-term interest rate) which leads to future output growth. In contrast, an increase in default spread means that the sales and profits are expected to decrease, which suggests the future economic activity would slow down. The positive slope of VASI is consist with the fact that it is the main component of GDP in China.

As we find above, including the business cycle variables could change the significance and signs of factor slopes. This suggests that some overlap exists in the information content of the market, size, BE/ME and momentum factors and business cycle variables. Same as Liew and Vassalou (2000), the aim of this section is to show that size, BE/ME and momentum factors contain some information similar in nature to that contained in popular business cycle variables regarding future economic growth. Fama and French (1992, 1993, 1995, 1996, 1998) imply that it is reasonable to consider SMB and HML as state variables in the context of Merton's (1973) intertemporal capital asset pricing model (ICAPM), because they can predict future changes in the investment opportunity set. Because of the generally insignificant slope coefficients related to the factors, our findings in China only weakly support their hypothesis, but show that it is more possible that the momentum factor can serve as the state variable in the context of ICAPM. The findings here requires us to further investigate Fama-French argument in next chapter.

4.5 Conclusion

Liew and Vassalou (2000) examine ten developed countries and find that the size and value factors do contain significant information about future GDP growth. The information contained in the Fama-French factors is to a large degree independent of that in the domestic market factor. Moreover, the ability of SMB and HML to predict future GDP growth is not weakened in the presence of popular business cycle variables in some countries examined.

Table 4.9: Multivariate regressions: the ability of factors predicting the GDP growth in the presence of business cycle variables

$$\begin{aligned} GDPGR^i_{(t,t+4)} &= a + b*MKT_{(t-4,t)} + s*SMB_{(t-4,t)} + h*HML_{(t-4,t)} + w*WML_{(t-4,t)} \\ &\quad + d*RF_{(t)} + e*TERM_{(t)} + f*DEF_{(t)} + g*VASI_{(t)} + \varepsilon_{(t,t+4)} \end{aligned}$$

Panel 1:	MKT	SMB	HML	WML11	RF	TERM	DEF	VASI	Adj. R^2
GDPGR-N	0.0138*	-0.0224	-0.0420	-0.0224	-14.8625***	6.2295*	-19.8314***	0.4501***	0.7698
	(1.84)	(-1.33)	(-1.06)	(-0.95)	(-4.00)	(1.67)	(-4.24)	(4.73)	
GDPGR-R	-0.0042	0.0041	**0.0311*** **	0.0075	-5.8902***	2.4724**	-5.2408***	0.3938***	0.6349
	(-1.61)	(0.72)	(2.70)	(0.88)	(-6.41)	(2.47)	(-4.44)	(4.03)	
Panel 2:	MKT	SMB	HML	WML12	RF	TERM	DEF	VASI	Adj. R^2
GDPGR-N	0.0204***	-0.0143	-0.0403	**-0.0560** **	-13.3704***	4.4998	-17.6379***	0.4558***	0.7845
	(2.76)	(-1.17)	(-1.09)	(-2.16)	(-3.70)	(1.14)	(-4.17)	(5.13)	
GDPGR-R	-0.0053**	0.0019	**0.0301** **	0.0090	-6.0267***	2.7513**	-5.3495***	0.3903***	0.6360
	(-2.42)	(0.39)	(2.55)	(0.86)	(-6.21)	(2.59)	(-4.53)	(3.96)	

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level
GDPGR-N represents the nominal GDP growth rate calculated from the series of “Gross Domestic Product, Accumulated(100 million yuan)” and GDPGR-R represents the real GDP growth rate calculated from the series of “Indices of Gross Domestic Product (preceding year=100)”.
The t-statistics reported are calculated from standard errors corrected for serial correlation and White (1980) heteroskedasticity using the Newey-West (1987) estimator. The rule for choosing the lag parameter q is that $q = 0.75 \times \sqrt[3]{T}$, rounded to an integer, where T is the number of observations used in the regression (see Stock and Watson, 2007, page 607).The t-statistics are reported in the parentheses.

There is little evidence supporting the fundamental risk-based explanation for the momentum factor. Thus, they argue that their findings support the risk-based explanation for the Fama-French model that the size and value factors act as the state variables in the context of ICAPM.

This study extends the robust examination of Liew and Vassalou's (2000) findings to the context of China. Specifically, we examine whether the four pricing factors, i.e., the market, size, value and momentum factors are able to capture the information about current and future GDP growth in both nominal and real terms. We reach to a very different conclusion. The relationship between the market, size and value factors and current and future GDP growth, no matter in nominal or real term, is rather weak. Instead, we find that the momentum factor captures most information about current and future economic growth in the context of the Chinese stock market. Thus, our findings could not be considered as supportive evidence for Liew and Vassalou's (2000) argument that the Fama-French factors are state variables related to macroeconomy. Even so, we still obtain some interesting findings in the Chinese market.

As we find in the two previous chapters, the momentum factor is less helpful for explaining equity returns, but the other three factors largely explain both time-series and cross-sectional equity returns. The rather weak relationship between the market, size and value factors and GDP growth suggests that there is no significant link of the Chinese stock market to macroeconomic. This might be because the trading in the Chinese stock market is dominated by individual investors[9]. This suggests that the majority of individual investors care less about the firm's fundamentals and long-term prospects, but seek for short-term gains through speculative and irrational trading with strong herding behaviour in buying winners and selling losers. This is one feature of the Chinese stock market that distinguishes itself from other developed market. This feature may be able to explain the strong link of the momentum factor to the economic activity. For example, the firms that perform well in good economic state would first produce high payoff for the investors (the "Winners"). Because of the herd behaviour in the market, more and more investors would pursue those firm's stocks, which pushes up

[9]Referred to the China Securities Regulatory Commission (CSRC) 2014 Annual Report, at http://www.csrc.gov.cn/pub/newsite/zjhjs/zjhnb/.

the prices of stocks. The overpriced stocks lead to less returns on those stocks. The similar situation is applied to those firms that perform poorly in good state (the "Losers"). As such, a negative relation between the momentum factor and economic output exists in the Chinese market.

Also, the weak link between the market factor and future GDP growth suggests that the market factor has such little ability to predict China's future economic output that it is not a good leading indicator of China's future economic activity.

By including the popular business cycle variables in the analysis, we find some changes in the significance and signs of slope coefficients related to the factors. Although those changes are not dramatic, this still suggests some overlapping information content of the factors and business cycle variables related to future GDP growth contained in the factors. However, the evidence is again rather weak.

Moreover, our regressions of future GDP growth including the business cycle variables confirms the ability of the risk-free rate, term spread and default spread to predict future GDP growth in both nominal and real terms. The relationships between the business cycle variables and GDP growth are consistent with the conventional economic explanations for why those business cycle variables can forecast economic output growth or recessions.

Chapter 5

Can Priced Factors be Proxies for News Related to GDP Growth? - The Chinese Evidence

5.1 Introduction

The fundamental task of asset pricing theory is to understand the prices of or returns on assets whose future payoffs are uncertain. To achieve this goal, researchers have been trying to find common factors which can explain the prices of different assets. Building on the earlier work of Markowitz (1952) on portfolio investment, Sharpe (1964), Lintner (1965b) and Mossin (1966) independently introduced the capital asset pricing model (CAPM). The CAPM states that the individual risk premium equals the product of market premium times β. This suggests that the asset returns can be thoroughly explained by the excess return on a market portfolio over a risk-free rate, which is also called market factor, together with its sensitivity to non-diversifiable risk (β). Although the CAPM still remains popular due to its simplicity and utility in a variety of situations, its failure in empirical tests cannot be overlooked.

The size (Banz, 1981) and book-to-market equity (Rosenberg et al., 1985) effects are two common and well-known anomalies in the CAPM. Fama and French (1992, 1993, 1995, 1996) propose an alternative model including a market factor (MKT), a size factor (SMB) and a book-to-market factor (HML) and show that their model can well explain the cross-sectional variation in equity returns.

The Fama-French model has been extended and improved in various ways. For example, Carhart (1997) adds a momentum factor (WML); while Fama and French (2015a,b, 2016) include other two additional factors of profitability and investment. Most of the extensions empirically outperform the CAPM. The factors, or more precisely, the proxies for the factors chosen in empirical tests, are constructed directly from a bundle of equity assets. It is generally acknowledged that the equity asset returns are highly correlated with, or mainly determined by economic activities. But these pricing models per se do not identify which economic variables determine the movement of asset prices.

Fama and French (1993, 1995, 1996) claim that SMB and HML play the role in capturing the risk associated with an unidentified state variable in the context of the Intertemporal Capital Asset Pricing Model (ICAPM) in Merton (1973). If this risk-based explanation is applied, SMB and HML should capture information related to fundamental economic risks which influence the time-varying investment opportunity set causing the movement of equity prices. Unfortunately, there is a lack of evidence to support which fundamental economic variables have the most influence on the movement of asset prices.

Some relevant empirical studies relate the Fama-French factors to macroeconomic variables. Liew and Vassalou (2000) find that the Fama-French factors can generally predict the future GDP growth in ten developed countries. Their findings further inspire Vassalou (2003) to examine whether the ability of the Fama-French factors to explain equity returns stems from their role as state variables in proxying for one of macroeconomic risks, that is, news related to future GDP growth. She shows that the evidence from the US market supports the argument that SMB and HML are the state variables which mainly capture the risks about future GDP growth. Thus, based on the evidence of those studies, the information of time-variation in the investment opportunity set for which the size and value factors proxy are summarized by the risk related to future GDP growth.

Instead of macroeconomic variables, another stream of risk-based explanation for the success of the Fama-French model argues that the changes in financial investment opportunities are not necessarily exclusively related to macroeconomic risks, such as news about future GDP growth, but should be related to innovations that forecast future investment opportunities. The studies in this stream include Brennan et al. (2004) where the state variables are the real interest rate and the Sharpe ratio and Petkova (2006) who chooses the dividend yield, term spread, default spread and short-term T-bill as the state variables. The evidence from these studies to some extent supports that the Fama-French factors act as the state variables in the ICAPM because they capture the risks related to certain specific economic variables.

As a direct extension of the previous chapter, this study follows the first stream by examining the robustness of Vassalou's (2003) conclusion. That is, we examine whether the success of the Fama-French/Carhart model is due to the fact that the pricing factors contain information about the macroeconomic risk in future GDP growth in the context of China. Moreover, because the GDP growth is an aggregated economic variable that summarizes information about the economic activity, it is a natural candidate as the starting point for examining the risk-based explanation for the success of the Fama-French model in a country. Our choice of GDP growth to be examined is also constrained by the data availability which is a common and important difficulty in order to study the young emerging markets. China's GDP data are released by the National Bureau of Statistics of China (NBS) and publicly accessible on the NBS's website. Unlike the GDP data, the data for other specific variables that could be potential candidates for the state variables in the ICAPM are generally incomplete and only available in recent years. Thus, the sample sizes of observations for those variables are too small to make reasonable inference.[1]

To the best of our knowledge, the study of Nguyen et al. (2009) is the first and only one to conduct the robustness test on Vassalou's conclusion outside the US. They present opposing evidence against Vassalou (2003) that the Fama-French factors are unlikely to be the state variables which capture the news related to future GDP growth in Australia. Specifically, in the presence of a GDP factor capturing news related to future GDP growth, SMB and HML do not lose their ability to explain the cross section of equity returns; and news related to future GDP growth is not a risk factor priced in equity returns. They also compare the ability of the three factors to forecast next year's GDP growth in Australia with that in Liew and Vassalou (2000). Liew and Vassalou (2000) show that in the US both market and book-to-equity factors present a significantly positive relationship with future GDP growth while the size factor appears to have a positive but insignificant relationship. Meanwhile, in Australia, it is the size factor that significantly and positively relates to future GDP growth whilst the other two factors exhibit a positive but insignificant relationship. Nguyen et al. (2009) show that the relationship between the size factor and future GDP growth is significantly negative, while for the market factor the relationship is positively significant at the 10% level, and is positive but insignificant for the book-to-equity factor. Accordingly, they suggest that the

[1] Due to the same reason of unavailable data, in somewhere of this study we have to calibrate some unrecorded data in early period of the Chinese stock market using the recorded data or use different data series from the conventional ones to proxy for some variables. How we deal with this difficulty is discussed in detailed in the next section. We admit the quality of data would bias our conclusion, but we have to suffer it in order to conduct this study.

different relationships between future GDP and the factors found in the US and Australia bring about their different findings from Vassalou's (2003). The existing evidence reaches to a contradictory conclusion for the argument whether the Fama-French factors are state variables that capture the risk in future GDP growth. This study aims to provide empirical evidence for the aforementioned contention whether future GDP growth is the fundamental driver of asset prices in the context of an increasingly important emerging Chinese stock market.

Since 2010, China has become the second-largest economy and its A-share stock market has become the world's second-largest stock market by market value. Many famous global indices providers have included or are considering including China A shares into their global benchmark index.[2] Once China A shares are included into the global benchmark indices, the funds tracking the global benchmarks will invest a large amount into the Chinese stock market. These facts suggest that the Chinese economy and stock market are increasingly important to global investors. Moreover, because of its special features which distinguish it from other developed markets, the Chinese market is also interesting to study. The Chinese government plays an important role in the market not only through strict regulation and intervention, but also through the dominating state holding of listed companies. Much of the trading in the Chinese market is not triggered by institutional investors, but by the inexperienced individual investors, who focus more on short term investment and trade speculatively and emotionally leading to a more volatile market. Contrast to its importance and specificity, research on the Chinese market generally falls behind the development of the Chinese market. Many studies have shown that the Fama-French three-factor model performs reasonably well in the Chinese stock market. For example, more than 85% of variation in returns on China A shares can be explained by the Fama-French model according to Chen (2004); while Xu and Zhang (2014) show that, on average, the Fama-French model is able to explain more than 93% of the variation in portfolio returns on China A shares. However, the Chinese market is an uncultivated land for the test of what fundamental or macroeconomic risk the priced factors, such as the

[2]For example, on 26 May 2015 FTSE starts transition to include China A Shares in its global benchmarks. MSCI will also bring Chinese stocks into its global emerging-markets benchmark, once the opening up of Chinese market is satisfied.

market excess return, the size factor, the book-to-market equity factor and/or the momentum factor, are proxies for. One reason may be attributed to the lack of low-frequency, especially quarterly, data for empirical studies. For example, the GDP growth data are quarterly observed, in order to connect the GDP growth with pricing factors, the quarterly returns on, say, the equity and/or fixed income portfolios are necessary. However, the Chinese financial market is relatively young, and the quarterly observations are very limited, because the two major Chinese financial markets - the Shanghai and Shenzhen stock exchanges were established in 1990.[3] Considering the fact that the data are not well recorded and that the variety of data is limited in the early stage of the Chinese market, lack of data is the first difficulty we need to overcome in order to carry out this study. The ways we solve this difficulty are presented in Section 5.2.

Since their findings do not favor the argument that the priced factors capture information about future GDP growth, we follow the approach of Nguyen et al. (2009) in order to have a direct comparison between the Chinese and Australian evidences. We also include a fourth popular risk factor, i.e., the momentum factor (WML), because in the previous chapter we find that the momentum factor can better predict future GDP growth than other factors in the Chinese context.

To conduct this study, we need to construct a mimicking portfolio, called the GDP factor, that captures news related to future GDP growth. The method is to regress future GDP growth on returns on financial assets (called base assets) but also control for some variables believed to be able to predict the returns on financial assets. This method is first presented in Lamont (2001) and also employed by Vassalou (2003) and Nguyen et al. (2009). The logic behind this method is as follows. Investors employ the information diffused over the financial market to forecast future economic conditions. The prices of financial assets reflect investors' expectation of changes in economy. By linking future GDP growth to financial asset returns, we create an explicit relationship between macroeconomic activity and financial market. Because

[3]The Shanghai Stock Exchange was founded on 26 November 1990 and the Shenzhen Stock Exchange on 1 December 1990.

in the context of asset pricing, only the innovations should be rewarded returns or premiums, we can use the control variables to “filter” the expected component of future GDP growth from the GDP mimicking portfolio. As such, the component left in the GDP factor is the innovation that cannot be expected by investors using the financial market information. This method is different from the one employed in Petkova (2006), where he applies a first-order vector autoregression (VAR) process for his state variables and uses the residuals of the VAR to proxy for innovations in state variables.

In this study, we construct two GDP factors capturing news related to future nominal and real GDP growth by using nominal and real accumulative GDP data series, respectively. Then we add the two GDP factors into standard pricing models and run the formal individual regressions as well as a system regression based on the General Method of Moment (GMM) approach. This setup enables us to examine whether the factors are proxies for the risk associated with future GDP growth. If the pricing factors mainly contain the information about future GDP growth, then in the presence of GDP factors, not only the four factors should lose ability, if any, to explain equity returns in the setup of individual regressions, but also the estimated risk premiums on those factors should be statistically insignificant in the setup of the GMM system regression.

The results of individual regressions show that the momentum factor is not a priced factor and that including the two GDP factors can only marginally improve the standard models. Thus, the individual regressions do not provide any evidence to support the argument that the pricing factors are the proxies for news related to future GDP growth. However, on the other hand, the results of GMM system regression present a significantly positive estimator of the risk premium on the real GDP factor with a magnitude close to the mean value of returns on the real GDP factor, which suggests that the news about future real GDP growth is priced in equity returns. But it is not the case for nominal GDP growth. The significantly positive risk premium on the real GDP factor, together with the decreased magnitudes of risk premiums on MKT, SMB and HML, suggests that the market, size and book-to-market equity factors, instead of the momentum factor, contain some overlapping information about future real GDP growth, but a large proportion of information that they contain are unknown.

The rest of the study is organized as follows. Section 5.2 describes the data related issue. Section 5.3 examines the ability of base assets to predict future GDP growth and constructs the mimicking portfolio mainly capturing news related to future GDP growth. Our main questions and methods are presented in Section 5.4. Section 5.5 reports the empirical results. Section 5.6 concludes.

5.2 Data

This section provides a description of the different types of data to be employed in this study. The data have different sources. The method of data calibration for fitting the missing data when necessary is also discussed in detail. In this study, the data of returns obtained from a well-known Chinese financial database, called RESSET, are on a monthly basis. We continuously compound monthly returns into quarterly returns to construct the mimicking portfolio capturing news related to future GDP growth in Section 5.3.2. All returns reported in this study are in percentage but not annualized.

5.2.1 Returns on the four risk factors and risk-free rates

The RESSET database provides the data series of market excess return (MKT), Fama-French factors (SMB and HML), momentum factor (WML), and risk-free rate in daily, weekly and monthly frequencies for the Chinese stock market. In this study, we choose the monthly returns on those factors constructed from the A-share stocks listed on both the Shanghai Stock Exchange (SSE) and the Shenzhen Stock Exchange (SZSE) with the weights based on the tradable (free-float) market value.

The risk-free rate series starts from 1st February 1989 and the choice of the candidate for the risk-free rate has changed twice since then. Starting from 1st February 1989 to 6th August

2002, the risk-free rate series is derived from the three-month deposit rates. After that, the coupon rates of three-month central bank bills are used as the risk-free rate till 7th October 2006. From 8th October 2006 up to now, the proxy for the risk-free rate is the three-month Shanghai Interbank Offered Rate (Shibor)[4].

The data for the four factors are available since July 1992, about one year and a half after the establishment of the Shanghai and Shenzhen stock exchanges. The market excess return is calculated as the difference between the return on the stock market portfolio and the risk-free rate. The return on the size factor is the difference between small market equity companies' return and large market equity companies' return. The return on the book-to-market equity factor is the difference between the return on stocks with a high ratio of book equity to market equity and the return on stocks with a low ratio of book equity to market equity. As to the momentum factor, the RESSET database provides several series of momentum factors based on different methods of construction. In this study, we choose a commonly used momentum factor (e.g. Liew and Vassalou, 2000), instead of the classical Carhart's (1997), because we found that the former has a slightly better ability to predict future GDP growth.[5] The chosen momentum factor is calculated as the difference between the tradable market value-weighted average returns on a portfolio of firms with the highest 30 percentile twelve-month returns lagged one month and the tradable market value-weighted average returns on a portfolio of firms with the lowest 30 percentile twelve-month returns lagged one month.[6] Table 5.1 summarizes the descriptive statistics for the four pricing factors and the correlation between them.

[4]The Shibor rates can also be obtained from the China Foreign Exchange Trade System & Nation Inter-bank Funding Center (http://www.shibor.org).

[5]The results for the comparison of the ability to predict future GDP growth between the two momentum factors have been suppressed to conserve space.

[6]Notice that the chosen momentum factor in this study is different from the classical Carhart's (1997). In his study, the momentum factor is constructed as the equal-weighted average returns on portfolio of firms with the highest 30 percentile eleven-month returns lagged one month minus the equal-weighted average returns on portfolio of firms with the lowest 30 percentile eleven-month returns lagged one month.

Table 5.1: A summary of descriptive statistics of and correlation between monthly returns on four pricing factors

	MKT	SMB	HML	WML
Panel A: Discriptive statistics of pricing factors				
Mean	0.8390	1.0386	-0.0640	-0.2240
Median	0.5050	1.0550	0.1800	0.0350
Std. Dev.	11.8260	5.6623	4.8662	7.6566
T-value	1.1657	3.0138	-0.2162	-0.4806
Min	-38.33	-18.65	-26.80	-97.27
Max	106.88	42.59	28.95	23.94
Skewness	2.6544	1.5715	-0.6975	-7.6672
Kurtosis	26.3204	14.8291	14.8263	97.9753
Panel B: Correlations between factors				
MKT	1			
SMB	0.2762	1		
HML	0.1490	-0.1880	1	
WML	-0.0370	0.0285	-0.3551	1

Sample period is July 1992 to December 2014, with 270 monthly observations.

5.2.2 Returns on the proxies for base assets

Although RESSET provides the returns on the above-mentioned four factors, it does not offer more detailed information about the building blocks, i.e., the six size and book-to-market sorted portfolios and two momentum sorted ("Winner" and "Loser") portfolios, for constructing the size, book-to-market and momentum factors. That is, there is no data available for returns on equity base assets directly from the RESSET database. It is also impossible to construct the building blocks for the Fama-French factors by ourselves because we cannot find the complete financial statement data for all listed firms on both Shanghai and Shenzhen stock exchanges in the 1990s to the first half of 2000s, the early stage of the Chinese stock market.[7] To solve this problem, we resort to the idea of using commercially available style indices as the proxies for base equity assets. This idea is based on the findings of Faff (2001, 2003, 2004), where he proposes a simple method to construct the Fama-French size and book-to-market equity factors using the style index and shows that both factors have the similar properties to those constructed from the original method proposed by Fama and French (1993). In Chapter

[7]Because we will need quarterly data for base assets returns, if there is a lack of these financial statement data in early stage, then we cannot construct enough number of observations of factor returns for our analysis.

3, we also confirm the similar properties in the Chinese stock market. Unfortunately, even if we can use the style (as well as momentum) indices as the proxies for base equity assets, the data for those indices are only available in recent years. If we exclude the early period, then we will not have enough observations for analysis. We will describe the process of fixing this problem after the brief introduction of the style and momentum indices.

There are (at least) four sets of indices accessible for Chinese stock investors, which are CSI Indices, CNI Indices, SSE Indices and SZSE Indices. The first two cover the A-share markets of both SSE and SZSE, while SSE Indices focus only on the A-share market of SSE and SZSE Indices only on the A-share market of SZSE (including the Main Board, SME Board and ChiNext markets). The constituents of each index are chosen from the typical stocks satisfying the criteria set by index providers with a limit on the number of stocks included. In order to get better coverage on both stock exchanges, we choose the style and momentum indices that have the widest sample coverage over both stock exchanges from the CSI and CNI Indices. Specifically, we choose six style indices, i.e., the Small-Cap Growth (SG), Small-Cap Value (SV), Middle-Cap Growth (MG), Middle-Cap Value (MV), Big-Cap Growth (BG) and Big-Cap Value (BV) indices from CNI Indices[8] and the 800-Momentum index from CSI Indices. The data for CNI style indices is available since 4th January 2010 and for 800-Momentum index since 4th January 2005. Obviously, these data are limited and we need to calibrate the missing data to continue this study. Next, we describe the way of fitting the data for returns on the base assets.

Fitting the returns on CNI style indices

The RESSET database records the data for daily returns on the six CNI style indices (Big-cap Growth, Big-cap Value, Mid-cap Growth, Mid-cap Value, Small-cap Growth, and Small-cap

[8]Notice that these six style indices to be used as base assets are not the same as those in Vassalou (2003) and Nguyen et al. (2009). In their papers, the stocks are sorted into two size groups and three book-to-market groups and six portfolios are constructed from the intersections of these groups. Here, there are three size groups and two growth-value groups.

Table 5.2: Estimated coefficients from the regressions of returns on six CNI style indices on MKT, SMB and HML

	MKT	SMB	HML	Cons.	Adj. R^2
BG	1.1606***	-0.6114***	-0.1877*	0.6862***	0.9532
	(39.08)	(-7.18)	(-1.83)	(3.20)	
BV	1.0825***	-0.5298***	0.5681***	1.1560***	0.9430
	(21.43)	(-3.93)	(3.86)	(3.82)	
MG	1.0463***	0.1844**	-0.6044***	-0.1373	0.9248
	(23.69)	(2.10)	(-5.09)	(-0.57)	
MV	1.2359***	0.0319	-0.1998*	0.4852**	0.9450
	(29.34)	(0.41)	(-1.80)	(2.51)	
SG	1.0783***	0.5539***	-0.4799***	-0.1084	0.9441
	(26.25)	(5.45)	(-4.26)	(-0.51)	
SV	1.1170***	0.4091***	-0.2463***	0.0285	0.9601
	(31.24)	(7.30)	(-3.70)	(0.16)	

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level
This table reports the results of regressions of returns on the style indices on the market, size and book-to-market equity factors.

$$R_{it} = a_i + b_i * MKT_t + s_i * SMB_t + h_i * HML_t + \varepsilon_{it},\ i = BG, BV, MG, MV, SG, SV$$

where MKT_t is the market excess return over the risk-free rate; SMB_t is the return on size factor; HML_t is the return on book-to-market factor; and ε_{it} is the error term.
Standard errors are corrected for serial correlation up to three lags and White's (1980) heteroskedasticity using the Newey-West (1987) estimator to calculate the t-statistics.
Sample period is January 2010 to December 2014, with 60 monthly observations.

Value) since 5 January 2010. Thus, the monthly returns provided by RESSET start from January 2010. RESSET also provides daily data for the Fama-French factors and the market excess return since 1 July 1992 and monthly data since July 1992. The method we used to calibrate the unrecorded data of returns on the style indices is as follows.

We first run the regressions of returns on the six CNI style indices on the market excess return and the Fama-French factors during the period when the data for returns on the six style indices is available (i.e., January 2010 to December 2014). Then we use the estimated coefficients from the regressions, together with the data for MKT, SMB, and HML provided by RESSET, to calculate the fitted values for the unrecorded style index returns before those data are available (i.e., July 1992 to December 2009). Table 5.2 presents the estimated coefficients and the corresponding adjusted R^2 of the regressions.

We can see that almost all factor loadings are statistically significant at least at the 10%

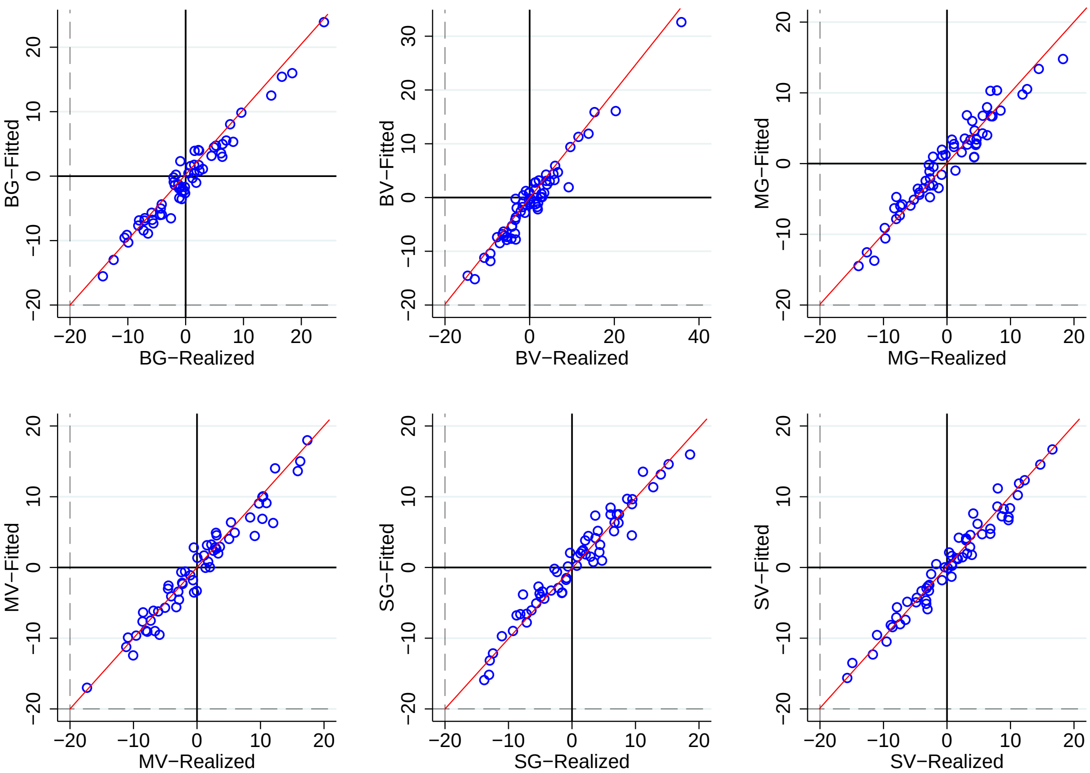

Figure 5.1: Dispersion of fitted from actual returns on style indices during the period from Jan 2010 to Dec 2014

level, 14 of which are statistically significant at the 1% level. Three of the constant terms are not statistically significant, which positively supports the ability of the three factors to explain the cross-sectional variation in returns on equity assets. The values of the adjusted R^2s are greater than 92%, which is high and thus suggests the dispersion of the fitted values from the actual values must be small. This dispersion is presented in Figure 5.1 in a visual way, where each point on the red straight line suggests the actual value is equal to the fitted value, and each blue circle pairs an actual value with its corresponding fitted value. The blue circles are generally around the red straight line, which indicates that the dispersion of the fitted values from the actual values is small.

In order to use the results from the above regressions to obtain the fitted values for unrecorded returns on the CNI style indices, we assume that the estimated coefficients would not change too much before and after the date when the returns on the style indices were initially recorded, or say, the relationship between each style index and the three factors provided by the RESSET database would not change during the period that we are interested in.[9] The summary statistics for returns on the CNI style indices are reported in Table 5.3. Panel 1 reports the summary statistics for actual returns from January 2010 to December 2014 with 60 observations; Panel 2 presents those for fitted returns from July 1992 to December 2009 with 210 observations; and Panel 3 summarizes the combination of fitted and actual returns. We can find that the four moments, i.e., the mean, standard deviation, skewness and kurtosis, for the actual and fitted returns on the six style indices (base assets) are quite different in Panels 1 and 2. However, this difference comes from the different patterns of the market, size and value factors within different periods. To show this, we separate the return series of three factors into two sub-periods same as the periods for the actual and fitted returns on base assets, respectively. The summary statistics for the three factors in the two sub-periods are reported in Panels 1 and 2 of Table 5.4, respectively.

Fitting the returns on CSI 800-Momentum index

We fit the missing data for returns on the 800-Momentum (800-MOM) index in a similar way. That is, we first run the regression of returns on the 800-MOM index on MKT, SMB, HML, and WML, and then use the estimated coefficients to calculate the fitted returns on the index. Table 5.5 reports the results of the regression and Figure 5.2 visually presents the dispersion of fitted values from actual values during the period from January 2005 to December 2014. Table 5.6 reports the summary statistics for actual and fitted returns on the 800-MOM index, as well as their combination. The different patterns in four moments for actual and fitted returns are due to the same reason for the style indices.

[9]It should be noticed that this is a strong assumption, because many studies have shown that the risk factor exposures are time-variant. We hope the estimated coefficients at least can largely capture the features of each style index from the three pricing factors during the whole period that we are interested in.

Table 5.3: Summary statistics for returns on CNI style indices

	BG	BV	MG	MV	SG	SV
Panel 1: Summary statistics for returns on the CNI style indices (obs. 60)						
Mean	0.1087	0.6628	0.1620	0.6317	0.6058	0.5848
Median	-0.6800	-0.0150	-0.4700	0.0100	1.0800	0.7700
Std. Dev.	7.2604	7.9929	6.7382	7.6079	7.4518	7.2149
T-value	0.1159	0.6424	0.1862	0.6431	0.6297	0.6279
Min	-14.31	-14.64	-13.92	-17.32	-13.81	-15.70
Max	23.92	35.84	18.28	17.38	18.60	16.61
Skewness	0.8390	1.5852	0.2033	0.2114	0.0653	-0.0547
Kurtosis	4.3828	8.1702	2.8865	2.5665	2.5407	2.5926
Panel 2: Summary statistics for fitted returns on the CNI style indices (obs. 210)						
Mean	0.6228	0.5660	1.3377	1.3552	1.7372	1.6128
Median	0.0516	0.9075	0.3435	0.5255	0.9269	0.8036
Std. Dev.	14.2929	14.3299	13.9339	16.0285	15.3124	15.3617
T-value	0.6315	0.5724	1.3912	1.2252	1.6440	1.5214
Min	-42.1753	-46.5685	-34.3938	-45.4570	-37.1211	-40.7505
Max	97.0781	95.9628	116.6844	132.4672	136.4658	135.5889
Skewness	1.5102	1.1994	2.7857	2.6073	3.2686	3.1545
Kurtosis	12.5095	12.0289	24.4089	23.4140	30.3436	29.4123
Panel 3: Summary statistics for combined returns on the CNI style indices (obs. 270)						
Mean	0.3576	0.3332	1.1066	1.0877	1.5096	1.3781
Median	-0.1838	0.1243	0.5616	0.4027	0.9701	0.8036
Std. Dev.	13.0390	13.1525	12.6603	14.5565	13.9243	13.9472
T-value	0.4507	0.4163	1.4363	1.2278	1.7814	1.6236
Min	-42.1753	-46.5685	-34.3938	-45.4570	-37.1211	-40.7505
Max	97.0781	95.9628	116.6844	132.4672	136.4658	135.5889
Skewness	1.6172	1.3004	2.9218	2.7487	3.4077	3.3099
Kurtosis	14.2260	13.4228	28.0526	26.9543	34.7089	33.8557

This table reports the summary statistics for returns on the CNI style indices. Panel 1 reports the summary statistics for actual returns from January 2010 to December 2014 with 60 observations; Panel 2 presents those for fitted returns from July 1992 to December 2009 with 210 observations; and Panel 3 summarizes the combination of fitted and actual returns with 270 observations.

Table 5.4: Summary statistics for returns on the market, size and value factors with two sub-periods

	Panel 1: Jan 2010 to Dec 2014			*Panel 2: Jul 1992 to Dec 2009*		
	Rmrf	Smb	Hml	Rmrf	Smb	Hml
Mean	0.0898	1.1150	0.0005	1.0530	1.0167	-0.0825
Median	-0.1850	1.1600	-0.1150	0.5700	1.0050	0.2500
Std.Dev.	6.0650	4.0388	3.2180	13.0159	6.0547	5.2491
T-value	0.1147	2.1384	0.0012	1.1724	2.4334	-0.2277
Min	-13.75	-17.30	-6.43	-38.33	-18.65	-26.80
Max	14.75	8.52	14.84	106.88	42.59	28.95
Skewness	0.2824	-1.5525	1.6071	2.5170	1.7906	-0.8129
Kurtosis	2.8198	8.7238	9.1178	22.9069	14.1653	13.7338

This table reports the summary statistics for returns on the market, size and value factors. Panel 1 reports the summary statistics within the period from January 2010 to December 2014 with 60 observations; Panel 2 presents those for the period from July 1992 to December 2009 with 210 observations. The summary statistics within the complete period are in Table 5.1.

Table 5.5: Estimated coefficients from the regressions of returns on CSI momentum index on MKT, SMB, HML and MOM

	MKT	SMB	HML	WML	Cons.	Adj. R^2
800-MOM	0.9787***	-0.2185***	-0.4355***	0.3253***	0.2142	0.9675
	(45.23)	(-5.62)	(-5.66)	(6.87)	(1.34)	

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level
This table reports the results of regressions of returns on the momentum index on the market, size, book-to-market equity, and momentum factors.

$$R_{it} = a_i + b_i * MKT_t + s_i * SMB_t + h_i * HML_t + w_i * WML_t + \varepsilon_{it},\ i = 800 - \text{MOM}$$

where MKT_t is the market excess return over the risk-free rate; SMB_t is the return on size factor; HML_t is the return on book-to-market factor; WML_t is the return on momentum factor; and ε_{it} is the error term.
Standard errors are corrected for serial correlation up to three lags and White's (1980) heteroskedasticity using the Newey-West (1987) estimator to calculate the t-statistics.
Sample period is January 2005 to December 2014, with 120 monthly observations.

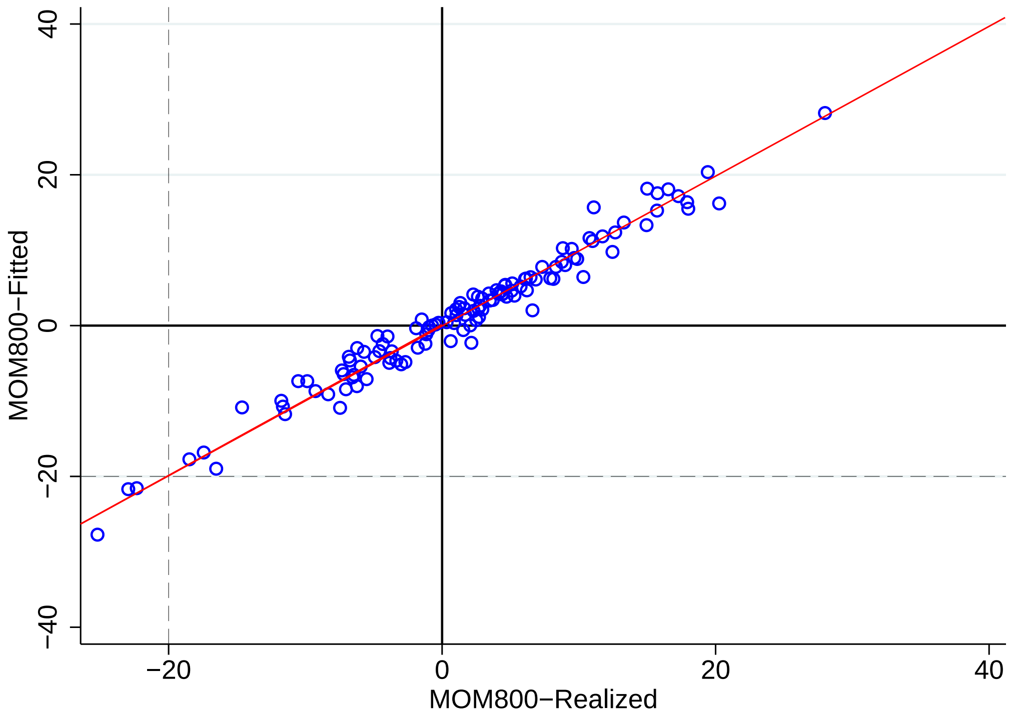

Figure 5.2: Dispersion of fitted from actual returns on 800-MOM index during the period from Jan 2005 to Dec 2014

Table 5.6: Summary statistics of returns on 800-Momentum

	800-MOM	800-MOM fitted	Combined
Obs.	120	150	270
Mean	1.4604	0.2057	0.7634
Median	2.10	-0.4717	0.4089
Std. Dev.	9.2993	13.0279	11.4636
T-value	1.7203	0.1934	1.0942
Min	-25.20	-38.0966	-38.0968
Max	28.00	84.8104	84.8104
Skewness	-0.2125	1.9791	1.5029
Kurtosis	3.5300	15.0324	14.2256

This table reports the summary statistics for returns on the CSI 800-Momentum index. Column 1 reports the summary statistics for actual returns from January 2005 to December 2014 with 120 observations; Column 2 presents those for fitted returns from July 1992 to December 2004 with 150 observations; and Column 3 summarizes the combination of fitted and actual returns with 270 observations.

Returns on the two fixed income portfolios

As to the data for returns on the two base assets of fixed income portfolios in Vassalou (2003) and Nguyen et al. (2009), i.e., TERM and DEF, the former is the difference between returns on long term and short term government fixed income securities, while the latter is the difference in returns on long-term corporate bonds and long-term government bonds. Again, we do not have enough data recorded for the Chinese bond market because we need quarterly observations of these data.[10] Instead, we choose the term-deposit rates and the loan interest rates obtained from the People's Bank of China (PBC) and China's five biggest state-owned commercial banks[11] to compute TERM and DEF. That is, TERM is calculated as 5-year RMB term-deposit rates minus 6-month RMB term-deposit rates (lump-sum deposit for lump-sum withdrawal), and DEF as 5-year RMB loan interest rates of financial institutions minus 5-year RMB term-deposit rates.

5.2.3 GDP growth rate

We will examine two quarterly observed Gross Domestic Product (GDP) growth series in this study. The GDP data are originally obtained from the National Bureau of Statistics of China (NBS). The first series is accumulative value of GDP calculated at current prices in the corresponding year, which will be used to compute the nominal accumulative GDP growth rate (GDPGR-N), i.e., the one with no inflation adjustment. The second series is indices of GDP calculated at constant prices with the preceding year equal to 100, thus subtracting 100

[10] In fact, we only have the quarterly data for returns on and yields of fixed income portfolios in Chinese fixed income market in recent years, thus, the numbers of available quarterly observations cannot produce a reliable regression analysis.

[11] The People's Bank of China (PBC) is the China's central bank and provides guidance and floating range of term-deposit rates and loan interest rates for commercial banks from time to time, which is one of the monetary policy tools of PBC. China's biggest five state-own commercial banks are Industrial and Commercial Bank of China (ICBC), Bank of China (BOC), China Construction Bank (CCB), Agricultural Bank of China (ABC) and Bank OF Communications (BOCOM). They implemented the same deposit rates and loan interest rates as the guidance set by PBC. Since 8th June 2012, all five banks started to implement a bit higher deposit rates than the guidance, but these rates are still the same among five banks. We use the deposit rates and loan interest rates implemented by the five banks in order to reflect the market rates.

from this series yields the commonly reported GDP growth rate. We name this series as the real GDP growth rate (GDPGR-R). The data for both GDP series are accessible since the first quarter of 1992.

Unlike in Vassalou (2003) and Nguyen et al. (2009) where the seasonally adjusted quarter GDP data are continuously compounded to get the quarterly observed annual GDP growth rate, we do not have enough observations for the seasonally adjusted data since NBS only provides this series since Quarter1 of 2011. We believe the seasonality would not influence our analysis, because in this study, as well as in their papers, we consider the quarterly observed annual GDP growth rate, that is, the GDP growth rate spans a year covering four quarters. Both nominal and real GDP growth series utilized in our study are year-on-year changes during the same period of each year, thus, can avoid the contamination with seasonality. We confirm this by running the regressions of our GDP growth rates on seasonal dummy variables and find that the estimated dummy coefficients are highly insignificant.[12]

We report the summary statistics of nominal and real GDP growth rates in Table 5.7. Some points should be noticed in this table. The reason for the difference between the two numbers of observations is because the accumulative value of GDP is available from the first quarter of 1992, and thus the yearly accumulative GDP growth rate can only be calculated one year later from the first quarter of 1993. The larger mean of GDPGR-N over GDPGR-R is not a surprise for the latter removes inflation in the series. The standard deviation of GDPGR-N is also larger than that of GDPGR-R, and the maximum value of GDPGR-N is extremely high. In fact, the computed nominal GDP growth rates (GDPGR-N) are over 25% up to almost 40% in years 1993 to 1995. During the same period, the real GDP growth rates are also high with the values over 10% up to 14.9%. The difference between nominal and real GDP growth is considered to be mainly caused by high inflation rates with the values of Consumer Price Indices (CPI, same period last year = 100) over 110 up to 128 in those years.

[12]The regression results are suppressed here to save space, but are available from the author, on request.

Table 5.7: Summary statistics for nominal and real GDP growth rates

	GDPGR-N	GDPGR-R
Obs.	88	92
Mean	15.9029	10.2152
Median	14.3870	10
Std. Dev.	8.0465	2.2102
T-value	18.5401	44.3320
Min	5.6928	6.6
Max	39.4779	14.9
Skewness	1.0872	0.4470
Kurtosis	3.7515	2.1832

This table reports the summary statistics of nominal and real GDP growth rates. Sample period of GDPGR-N is 1993 Q1 to 2014 Q4; while sample period of GDPGR-R is 1992 Q1 to 2014 Q4. The reason for this difference is because that the accumulative value of GDP is available since 1992 Q1, thus the yearly accumulative GDP growth rate can only be calculated one year later since 1993 Q1.

5.2.4 Control variables

Vassalou (2003) chooses four control variables when constructing the mimicking portfolio to capture the news related to future GDP growth. These control variables include the term yield spread (TERMY), the default yield spread (DEFY), the risk-free rate and the variable CAY[13]. There is no CAY data for the Australian market, thus Nguyen et al. (2009) drop the CAY variable and use the other three as the control variable. In our study, we also drop the CAY for its unavailability in China. Due to the same reason mentioned in Section 5.2.2, we do not have enough quarterly observations for yields of the Chinese government and corporate bonds, thus, we again use the term-deposit rates and the loan interest rates to calculate the proxies for term spread and default spread. This means that the control variables of term and default yield spreads share the same data series with the returns on fixed income base assets of TERM and DEF. As in their papers, the values of control variables are known at the point of time to construct the mimicking portfolio, thus in our study, the control variables of TERM and DEF, marked as TERM-lag and DEF-lag, are one period (quarter) lagged compared to the fixed income base assets.

[13]The CAY is a detrended wealth variable computed by Lettau and Ludvigson (2001a), which represents deviations from a common trend found in consumption, asset wealth, and labor income.

5.2.5 Dependent variables

As in the studies of Vassalou (2003) and Nguyen et al. (2009), our dependent variables in the formal individual regressions and the GMM system regression are monthly returns on 25 equity portfolios with the weights based on tradable market value constructed by using Fama and French's (1993) method. These data are also provided by the RESSET database since July 1992. But in the first three years, i.e., before June 1995, many data are missing and RESSET does not explain this. The numbers of observations for returns on the 25 portfolios are summarized in Column "Obs." of Table 5.8 together with a summary for the number of stocks in each of the 25 portfolios reported in the last three columns. In order to ensure the quality of data, we exclude the data in the first three years and use the data from July 1995 to December 2014 for the individual regression and GMM system regression analysis, leading to a total of 234 observations for each series. Another benefit from excluding the data in the early years from the sample could be that it alleviates the distortion of portfolio returns due to the lack of sample stocks in each of the 25 portfolios. We can see from Table 5.8 that there are 23 portfolios consisting of only one stock in the beginning of the establishment of the Chinese stock market. The summary statistics for returns on the 25 equity portfolios excluding the data in the first three years are reported in Table 5.9.

5.3 Construction of GDP mimicking portfolio

We follow the method employed by Vassalou (2003) and Nguyen et al. (2009) to construct a GDP mimicking portfolio which mainly captures news related to future GDP growth. As briefly discussed in the introduction, the logic behind this method is as follows. Investors forecast changes in future economic conditions based on the information diffused over the financial market. The dynamic expectation of investors about changes in economy is reflect by the movements of prices of financial assets. We can create an explicit linear relationship between future macroeconomic activity and financial market in the regression of the macroeconomic variable of interest on a set of financial asset returns (base assets). The base assets should be

Table 5.8: Summary statistics for number of stocks in each of the 25 size and book-to-market sorted portfolios

Size	BE/ME	Obs.	Mean	Min	Max
Small	Low	246	41	1	73
	2	258	40	1	88
	3	270	41	1	111
	4	258	41	1	116
	High	258	23	1	66
2	Low	257	29	1	60
	2	258	40	1	104
	3	246	43	1	109
	4	270	38	1	102
	High	270	37	1	99
3	Low	246	32	3	73
	2	258	37	1	100
	3	270	35	1	93
	4	270	40	1	98
	High	258	41	1	92
4	Low	270	38	1	109
	2	258	36	1	98
	3	246	38	1	89
	4	246	37	1	90
	High	258	42	1	88
Big	Low	258	48	1	140
	2	270	33	1	78
	3	270	29	1	68
	4	246	32	1	61
	High	258	45	2	123

This table reports the numbers of observations of 25 portfolios returns and the summary statistics for the number of stocks contained in each of the 25 size and book-to-market sorted portfolios. Sample period is July 1992 to December 2014, with 270 months. The numbers, reported in the Column “Obs.”, less than 270 suggest the data are missing, which happens from July 1992 to June 1995.

Table 5.9: Summary statistics for returns on the 25 portfolios

Size	BE/ME	Mean	Std. Dev.	T-value	Min	Max
Small	Low	2.1215	11.0970	2.92	-26.19	45.60
	2	2.6179	10.8162	3.70	-28.03	38.66
	3	2.5332	11.0428	3.51	-33.17	48.92
	4	2.3167	10.3730	3.42	-26.86	33.52
	High	1.9802	10.3448	2.93	-25.06	40.28
2	Low	1.6221	10.2346	2.42	-27.52	45.26
	2	2.1297	10.8133	3.01	-29.25	42.86
	3	2.2090	10.9609	3.08	-37.43	50.17
	4	2.2412	11.3982	3.01	-33.05	52.27
	High	1.8508	10.2596	2.76	-28.92	41.35
3	Low	1.3580	9.7978	2.12	-27.25	32.38
	2	1.8345	10.2202	2.75	-27.43	46.79
	3	1.8097	10.7389	2.58	-33.10	46.44
	4	2.0970	10.4283	3.08	-25.44	43.63
	High	1.5097	9.9563	2.32	-25.65	40.22
4	Low	1.1791	9.2600	1.95	-25.89	27.65
	2	1.4783	10.0132	2.26	-26.98	42.18
	3	1.6693	10.1874	2.51	-28.24	39.67
	4	1.6459	10.2374	2.46	-25.68	38.89
	High	1.7677	10.6790	2.53	-30.16	48.75
Big	Low	0.9605	9.0260	1.63	-25.72	48.63
	2	0.9953	9.6309	1.58	-29.69	39.27
	3	1.3848	9.9516	2.13	-26.61	48.20
	4	1.6037	9.8707	2.49	-27.13	54.22
	High	1.3586	9.5343	2.18	-28.43	35.00

This table reports the summary statistics for returns on the 25 portfolios. Sample period is July 1995 to December 2014, with 234 monthly observations.

good representatives for the financial market. As proposed in Breeden et al. (1989)[14], the fitted value from the regression can be used to construct a mimicking portfolio which contains the same information as the macroeconomic variable, but this information is expressed in terms of portfolio returns. The advantage of this method is that if the mimicking portfolio is rewarded a risk premium, then the variable linked to the mimicking portfolio can reveal the identities of state variables which determine expected returns. A further discussion on the economic interpretation for the mimicking portfolio of macroeconomic variables can be referred to Cochrane (1999).

As for the choice of base assets, Lamont (2001) suggests that all available assets whose unexpected returns are maximally correlated with unexpected components of the economic variable should be involved. However, it is impractical to include all such assets in the regression. Vassalou (2003) and Nguyen et al. (2009) choose six equity portfolios and two bonds portfolios. The six equity portfolios are the excess returns on the six portfolios of building blocks for the Fama-French factors; while the two bond portfolios are TERM and DEF. These assets are good representatives for both equity and fixed-income markets. The six equity base assets are the building blocks for the size and value factors which can largely explain the equity returns in the market. The stocks in the six Fama-French portfolios are the same stocks in the overall equity market portfolio, thus it is not necessary to add the market portfolio as a base asset.[15] The term spread and default spread are both important variables in the fixed income market, and have the verified ability to explain the financial asset returns. Our choices of base assets in the Chinese market follow the two studies[16] except that we include an additional momentum portfolio here because we find the momentum factor can better predict China's future GDP growth than the size and value factors in the previous chapter.

However, in the context of asset pricing, only innovations in variables should be rewarded

[14]Breeden et al. (1989) form a maximum correlation portfolio (MCP) highly correlated with real consumption growth rate. Lamont (2001) extended this argument to an "economic tracking portfolio" whose returns track economic variables such as inflation, expected output or future GDP growth.

[15]In fact, including the market portfolio slightly decreases the adjusted R^2 of the regressions of future GDP growth on equity base assets.

[16]Although we use indices as proxies for the building blocks due to unavailability of relevant data, we believe the results would not change too much.

returns or risk premiums in asset returns. Therefore, the expected component of information contained in the mimicking portfolios should be “filtered” so that the component left in the GDP mimicking portfolio is the innovation in GDP growth (a.k.a. news related to future GDP growth) unexpected by investors based on the information of financial market. Lamont (2001) proposes a method to achieve this aim by further including control variables into the right-hand-side of the above regression. The chosen control variables are considered to be able to predict the returns on base assets. Thus, by including the control variables that are able to predict the returns on base assets in the regression of future GDP growth on the base assets, it effectively extracts the information component for forecasting future GDP growth from the base assets. By running the regression of future GDP growth rate on base assets together with control variables, we construct the GDP factor by summing up the productions of coefficient estimates related to base assets and base asset returns, but dropping the terms related to control variables. Thus, the information contained in the GDP factor is filtered by the control variables, and the component left in the GDP factor is the innovative information related to future GDP growth. The specification of this method is presented in the rest of this section. We first examine the ability of the base assets to predict future GDP growth in the absence and presence of control variables, and then construct the GDP mimicking portfolio (the GDP factor) based on the regression results.

5.3.1 The ability of the base assets to predict future GDP growth

Before constructing the mimicking portfolio, we examine whether the nine base assets can predict future GDP growth in this subsection by running the following regression:

$$GDPGR_{t,t+4} = a + cB_{t-1,t} + e_{t,t+4} \tag{5.1}$$

where $GDPGR_{t,t+4}$ is the four quarters ahead GDP growth rate between t and $t+4$, $B_{t-1,t}$ is the vector of returns on the nine base assets between $t-1$ and t. The seven equity portfolio returns are excess returns over the risk-free rate (RF). Therefore, all nine base assets represent the zero-investment portfolios. Standard errors are corrected for serial correlation up to three lags and White's (1980) heteroskedasticity using the Newey-West (1987) estimator.

The first columns of Panels 1 and 2 of Table 5.10 present the results from the regressions of four quarters ahead nominal and real GDP growth rates on the nine base assets, respectively. It is difficult to interpret the individual coefficient estimates because of the existence of multicollinearity among regressors, as suggested by Vassalou (2003). In fact, we only care about the extent to which the base assets together can predict future GDP growth, thus we focus on the adjusted R^2 here. The adjusted R^2 is 44.4% for the regression of future nominal GDP growth rate, suggesting that a significant proportion of variation in future nominal GDP growth is forecast by the nine base assets. But in the case of future real GDP growth, the adjusted R^2 is only 8.58%. The results of F-test reject the null hypothesis that the coefficients of nine base assets are jointly equal to zero at the 10% level in both cases.

Next, we examine the extent to which the nine base assets contain important incremental information related to future GDP growth by running the following regression including the control variables:

$$GDPGR_{t,t+4} = a + cB_{t-1,t} + k_1 TERM_{t-2,t-1} + k_2 DEF_{t-2,t-1} + k_3 RF_{t-1,t} + e_{t,t+4} \tag{5.2}$$

where $TERM_{t-2,t-1}$, $DEF_{t-2,t-1}$ and $RF_{t-1,t}$ act as the control variables in the regression, and their subscripts suggest that the values of control variables are known at each period. The results are reported in the second columns of Panels 1 and 2 of Table 5.10 for the cases of future nominal and real GDP growth, respectively. The values of adjusted R^2 increase significantly to 71.09% and 51.36%, respectively, for both cases. The F-tests reject the null hypothesis that the coefficients of nine base assets are jointly equal to zero in the case of nominal GDP growth, but cannot reject in the real GDP growth case.

The results presented here are consistent with those we find in the previous chapter that the size and value factors are not able to predict future GDP and the momentum factor contains more information related to future nominal than real GDP growth.

Table 5.10: The ability of base assets to predict future GDP growth

		Panel 1		Panel 2	
		GDPGR-N		GDPGR-R	
Base Assets	BG	0.3663	0.0720	0.1292	0.0151
		(0.68)	(0.29)	0.80	(0.21)
	BV	-0.4549	-0.1936	-0.1403	-0.0457
		(-1.73)	(-0.95)	(-1.79)	(-0.72)
	MG	-0.1848	-0.1774	-0.1160	-0.1196
		(-0.28)	(-0.44)	(-0.59)	(-0.92)
	MV	0.3055	0.2546	0.1476	0.1177
		(0.23)	(0.44)	(0.37)	(0.70)
	SG	-0.3845	0.0229	-0.0684	0.0452
		(-0.41)	(0.05)	(-0.23)	(0.36)
	SV	0.4203	-0.0416	0.0638	-0.0567
		(0.29)	(-0.07)	(0.15)	(-0.38)
	MOM	-0.0874	0.0662	-0.0338	0.0353
		(-0.77)	(0.76)	(-0.92)	(1.08)
	TERM	0.3047	41.2112	-0.5272	8.9152
		(0.09)	(3.17)	(-0.47)	(2.24)
	DEF	-21.3597	-31.2826	-3.3011	-3.8151
		(-3.57)	(-4.21)	(-1.82)	(-1.61)
Control Variables	Constant	27.7494	38.4657	12.0888	16.9799
		(5.05)	(7.33)	(6.61)	(8.38)
	TERM-lag		-27.6675		-5.2673
			(-2.69)		(-1.66)
	DEF-lag		3.5753		-2.9294
			(0.38)		(-0.89)
	RF		-16.7210		-6.0919
			(-6.14)		(-6.52)
	Adj. R^2	0.4440	0.7109	0.0858	0.5136
	F-test	3.15	3.83	1.72	1.13
	p-value	0.0028	0.0005	0.0989	0.3553

This table reports the results of regressions of nominal and real GDP growth rates on the base assets with and without the control variables.

$$GDPGR_{t,t+4} = a + cB_{t-1,t} + k_1 TERM_{t-2,t-1} + k_2 DEF_{t-2,t-1} + k_3 RF_{t-1,t} + e_{t,t+4}$$

where $GDPGR_{t,t+4}$ is the four quarters ahead GDP growth rate between t and $t+4$; $B_{t-1,t}$ is the vector of returns on the nine base assets between $t-1$ and t; $TERM_{t-2,t-1}$, $DEF_{t-2,t-1}$ and $RF_{t-1,t}$ are the control variables in the regression, and their subscripts suggests the values are known. The estimated coefficients of nine base assets in second columns of both panels will be used to construct the GDP factors. Sample period of GDPGR-N is from Q1 1993 to Q4 2014, while period of GDPGR-R from Q3 1992 to Q4 2014.

5.3.2 Construction of the mimicking portfolio

Following Lamont (2001), Vassalou (2003) and Nguyen et al. (2009), the GDP mimicking portfolio is constructed by first running the following regression model:

$$GDPGR_{t,t+4} = cB_{t-1,t} + kZ_{t-2,t-1} + e_{t,t+4} \tag{5.3}$$

where $GDPGR_{t,t+4}$ is future GDP growth rate; $B_{t-1,t}$ is the vector of the nine equity and fixed-income base assets, $Z_{t-2,t-1}$ denotes a set of control variables, including a constant, $TERM_{t-2,t-1}$, $DEF_{t-2,t-1}$ and $RF_{t-1,t}$, which are able to predict the returns on the base assets.[17] We then obtain the estimated coefficients related to base assets, $\hat{c}$, from the above regression. Next, we construct the return on the mimicking portfolio, or the GDP factor (GDP^F), by summing up the productions of the returns on the base assets and their corresponding coefficient estimates and dropping the vector of control variables, $Z_{t-2,t-1}$. Thus, the GDP factor, GDP^F, is then equal to

$$GDP^F_{t-1,t} = \hat{c}B_{t-1,t} \tag{5.4}$$

As GDP is reported quarterly, equations (5.2) or (5.3) is estimated using quarterly data. Thus, $GDPGR_{t,t+4}$ is quarterly observed year-on-year growth in GDP. Once the vector of coefficients $\hat{c}$ in equation (5.3) is estimated, we can construct either monthly or quarterly returns on the mimicking portfolio from either monthly or quarterly base asset returns, respectively. To construct monthly returns on the mimicking portfolio, we need an assumption that the estimated value of parameter c is not influenced by the change in data frequency. As in Nguyen et al. (2009), we will conduct the following analysis at the monthly frequency. Thus, we plug monthly data into equation (5.4) to calculate the monthly returns on the mimicking portfolio.

Summary statistics for monthly returns on the GDP factors constructed from nominal and real GDP growth rates, i.e., GDP^F-N and GDP^F-R, are presented in Table 5.11. The mean returns on the two mimicking portfolios are positive and significantly different from zero. This

[17] Vassalou also includes CAY of Lettau and Ludvigson (2001a). We do not have data for it, thus exclude it from our analysis.

Table 5.11: Summary statistics for the monthly returns on the mimicking portfolio

	GDP^F-N	GDP^F-R
Mean	1.5402	0.9081
Median	0.3474	0.6233
Std. Dev.	6.5093	1.1983
T-value	3.89	12.45
Min	-6.7534	-0.6401
Max	20.5313	3.7103
Skewness	0.9945	0.9334
Kurtosis	3.1835	2.8661
Correlation with SMB	0.1817	0.1092
Correlation with HML	-0.1866	-0.1270
Correlation with WML	-0.0595	-0.0292

This table reports the summary statistics for the monthly returns on the two mimicking portfolio, GDP^F-N and GDP^F-R. Sample period is July 1992 to December 2014, with 270 observations. Notice that the sample size of both GDP factors reported in this table is larger than that to be explored in regression analysis in Table 5.12.

means that if news related to future GDP growth is a factor that can explain a certain portion of cross-sectional variation in returns, then its associated risk premium should be positive and statistically significant during this period from July 1992 to December 2014. Notice that the sample size of both GDP factors reported in this table is larger than that to be explored in regression analysis and reported in Table 5.12 where the period is from July 1995 to December 2014, with 36 fewer observations. We will see that during the latter period, the mean of returns on the nominal GDP factor, GDP^F-N, is negative but insignificant, while for the real GDP factor, GDP^F-R, the mean is still significant and positive. We should also notice that even during the former period, although it is statistically significant, the t-statistics for GDP^F-N is much smaller than that for GDP^F-R. A more detailed discussion about these phenomena will be provided in Section 5.5.1.

5.4 Major questions and methods

Because the conclusion of Nguyen et al. (2009) disagrees Vassalou's (2003) argument that the size and value factors are state variables proxying for news related to future GDP growth, we mainly follows the methods of Nguyen et al. (2009) in order to compare our results with theirs.

Specifically, we address the following two major questions in the context of the Chinese stock market followed by the methods designed to answer the corresponding questions.

- Are the common pricing factors, i.e., MKT, SMB, HML, and/or WML, the state variables that mainly capture the news (risk) related to future GDP growth?
- Is news related to future GDP growth the systematic risk priced in the cross section of equity returns?

Answers to these questions are helpful for understanding Vassalou's (2003) macroeconomic risk-based explanation for the success of the Fama-French model, i.e. whether the Fama-French factors are state variables that capture the risk related to future GDP growth. If the answers to both questions are "yes", then the supportive evidence for Vassalou's (2003) argument is met.

As to the first question, Fama and French (1993) claim that their size and book-to-market equity factors (SMB and HML) proxy for state variables in the context of Merton's (1973) ICAPM. However, which fundamental variables the factors proxy for are unclear based on the model. If Vassalou's (2003) argument holds, then the factors should at least largely capture the same risk as that contained in the mimicking portfolio (GDP factor) which, by construction, only captures the innovation in future GDP growth. We can answer this question based on the following individual regressions of a GDP augmented four-factor pricing model which included the GDP factor as an additional factor:

$$\begin{aligned} R_{pt} - RF = \quad & a_p + b_p \left(R_{mt} - RF \right) + s_p SMB_t + h_p HML_t \\ & + w_p WML_t + g_p GDP_t^F + e_{pt} \end{aligned} \tag{5.5}$$

where $p = 1, 2, \cdots, N$; RF is the risk-free rate; $R_{pt} - RF$ is the portfolio return in excess of the risk-free rate; $R_{mt} - RF$ is the market portfolio return in excess of the risk free rate (or equivalently, the market factor MKT_t); SMB_t is the return on the mimicking portfolio for the size factor; HML_t is the return on the mimicking portfolio for the book-to-market factor; WML_t is the return on the mimicking portfolio for the momentum factor, and GDP_t^F is the return on the mimicking portfolio for the GDP factor. In this setup, if Vassalou's (2003) argument is applied in China, then we should expect a significant estimated slope coefficient related to the GDP factor. Meanwhile, if the factors mainly capture the news related to future GDP growth, then adding a better proxy for the innovation in GDP growth should dramatically decrease the significance of the slopes related to other factors. On the other hand, if the significance of the factor loadings do not change atfer the inclusion of the GDP factor, then there is no supportive evidence for that the factors mainly capture the information about macroeconomic risk.

As GDP is an aggregate macroeconomic variable, it is quite possible that the factors together with each other capture the innovative information about future GDP growth, thus this information may be concealed rather than clearly uncovered in the empirical time-series regressions of the pricing models, i.e., adding the GDP factor may not change the significance of slopes of other factors. We address the second question to further examine whether news related to future GDP growth is the systematic risk priced in the cross-section of equity returns no matter whether the factors do or do not *mainly* capture news related to future GDP growth. No matter which is the case, the second question is important for the contention whether the Fama-French factors are state variables mainly capturing news related to future GDP growth. First, if it is the former case, i.e., the factors mainly capture news related to future GDP growth, then we can further examine Vassalou's (2003) argument by answering the second question. Because in the standard pricing model, the valid factors are priced in the cross-section of equity returns, i.e., the estimated risk factor premiums are significant in the cross-sectional regressions. If the factors *mainly* capture the fundamental risk in future GDP growth, then adding the GDP factor should dramatically decrease the significance of those factor premiums and the GDP factor premium then became significant, because the GDP factor is a better proxy for the innovation in GDP growth by construction. In this case, we confirm that Vassalou's (2003) argument is applied in China. Second, even if the latter is the case, we cannot rule out the possibility that the factors may contain some innovative information about future GDP growth. Thus, adding the GDP factor may not change the factor premiums to be insignificant, but it could significantly decrease the magnitudes of the estimated factor premiums and the GDP factor premium can still be significant. In this case, we conclude that the pricing factors may not mainly capture the risk in future GDP growth, but they do contain some innovative information of GDP growth that are systematic and priced in equity returns. Last, if we do not find any evidence supporting for affirmative answers to both questions, then the results are consistent with Nguyen et al. (2009) in Australia.

To answer the second question, we follow a one-step method adapted from Cochrane (2001) which is based on Hansen's (1982) Generalized Methods of Moments (GMM). As an alternative to the standard Fama-MacBeth two-stage regression, this setup allows us to estimate the pricing models and the factor risk premiums simultaneously in one step. More importantly, the GMM approach allows for conditional heteroscedasticity, serial correlation, and non-normal distribution. Thus, it can correct for the generated-regressors problem in standard errors in the context of Fama and MacBeth's (1973) approach, while the Fama-MacBeth regressions provide standard errors corrected only for cross-sectional correlation but not for time-series autocorrelation. This makes the GMM setup a superior method over the Fama-MacBeth two-stage regressions.

In order to find evidence for the second question, we compare the results of the GMM system regression of three standard asset pricing models and their corresponding GDP augmented counterparts. For convenience of description, we only present the specification of the GMM system for the GDP augmented four-factor model here.[18] If the GDP augmented four-factor model is correct by assumption, then we expect that the following equation holds.

$$\begin{aligned} E(R_p) - RF = \quad & b_p \left[E(R_m) - RF\right] + s_p E(SMB) + h_p E(HML) \\ & + w_p E(WML) + g_p E\left(GDP^F\right) \end{aligned} \tag{5.6}$$

The empirical counterpart for this model is given by equation (5.5). By applying expec-

[18] The other models are nested versions of this one which we do not present to save space.

tations to equation (5.5) and comparing with equation (5.6), we could find that if the pricing model is correct, then the intercept term should be equal to zero. Because we will examine the validity of the system as a whole, we impose the restriction of zero intercepts to the empirical regression equation to obtain the first equation (5.7) in the system. Besides, we introduce five mean-adjusted transformations to the independent variables in equation (5.7) into the system, which are equations (5.8), (5.9), (5.10), (5.11) and (5.12). By combining those equations together, this setup of system allows us to simultaneously estimate the five risk loadings on individual factors through regression equation (5.7) of pricing model and the five risk premiums through equations (5.8) to (5.12).

$$R_{pt} - RF = b_p\left(R_{mt} - RF\right) + s_p SMB_t + h_p HML_t + w_p WML_t + g_p GDP_t^F + e_{pt} \tag{5.7}$$

$$R_{mt} - RF = \lambda_m + e_{mt} \tag{5.8}$$

$$SMB_t = \lambda_s + e_{st} \tag{5.9}$$

$$HML_t = \lambda_h + e_{ht} \tag{5.10}$$

$$WML_t = \lambda_w + e_{wt} \tag{5.11}$$

$$GDP_t^F = \lambda_g + e_{gt} \tag{5.12}$$

Following Faff (2001) and Nguyen et al. (2009), the GMM approach is chosen to perform the asset pricing tests. The system of equations (5.7) to (5.12) contains $6N+5$ sample moment equations (i.e., $1/T \sum_{t=1}^{T} e_{pt}$, $1/T \sum_{t=1}^{T} e_{pt}(R_{mt} - R_f)$, $1/T \sum_{t=1}^{T} e_{pt} SMB_t$, $1/T \sum_{t=1}^{T} e_{pt} HML_t$, $1/T \sum_{t=1}^{T}$

$e_{pt}WML_t$, $1/T \sum_{t=1}^{T} e_{pt}GDP_t^F$ for $p = 1, 2, \ldots, N$, and $1/T \sum_{t=1}^{T} e_{jt}$ for $j = m, s, h, w, g$), with $5N + 5$ unknown parameters (i.e., $\Phi = (b_1, b_2, \ldots, b_N,\ s_1, s_2, \ldots, s_N,\ h_1, h_2, \ldots, h_N,\ w_1, w_2, \ldots, w_N,\ g_1, g_2, \ldots, g_N,\ \lambda_m, \lambda_s, \lambda_h, \lambda_w, \lambda_g)$. Hence, the test has N $(= (6N+5) - (5N+5))$ overidentifying restrictions and is given by:

$$GMM = (T - N - 1) \cdot g_T\left(\hat{\phi}\right)' \cdot S_T^{-1} \cdot g_T\left(\hat{\phi}\right) \tag{5.13}$$

where $g_T\left(\hat{\phi}\right) = \frac{1}{T}\Sigma_{t=1}^{T} f_t\left(\hat{\phi}\right)$ is the empirical moment condition vector; and the GMM statistic represents the small-sample adjusted version of Hansen's (1982) J-statistic following MacKinlay and Richardson (1991), which is asymptotically distributed as a chi-square statistic with N degrees of freedom.

We then run GMM system regressions on the following models: the CAPM, the Fama–French model and the four-factor model, as well as their corresponding GDP augmented counterparts. Three outcomes of results are possible to show up corresponding to our previous discussion of answers to the second question. First, if news related to future GDP growth is systematic and priced in equity returns then we should expect the estimated premium on the GDP factor (λ_g) in the GDP augmented system regressions to be significantly positive. If news related to future GDP growth is the fundamental risk that induces the success of the MKT, SMB, HML and/or WML factors in explaining equity returns, then in the system regression of the GDP augmented models, the factor premiums on MKT (λ_m), SMB (λ_s), HML (λ_h) and/or WML (λ_w) will tend to be insignificant. This indicates that in the presence of the GDP factor, MKT, SMB, HML and/or WML lose their ability to explain the cross-section of equity returns and thus confirms Vassalou's (2003) argument. Second, if news related to future GDP growth is not the main fundamental risk that contributes to the ability of priced factors to explain equity returns but the risk in future GDP is systematic, then the premiums on priced factors and the GDP factor could be all significant, but the magnitudes of estimated factor premiums would decrease compared to those in standard pricing models. Last, if the priced factors are not proxies for news related to future GDP growth and GDP growth is not the systematic risk priced in equity returns, then the significance and magnitudes of factors premiums would not change even in the presence of the GDP factor in the system regressions.

5.5 Main results

This section discusses the results of our main analysis in this study. We first discuss in more detail the factors (especially, the nominal GDP factor) as the independent variables in our individual regressions, and then analyze the individual regressions of the GDP augmented four-factor model on 25 size and book-to-market sorted portfolios in the following subsection. The results provide evidence for our first major question. The results of GMM system regression are examined in Section 5.5.3 and answer our second major question.

5.5.1 More detailed discussion on the factors

Panel A of Table 5.12 reports the basic descriptive statistics for independent variables, i.e., MKT, SMB, HML, WML and two GDP factors constructed using nominal and real GDP growth rates, and Panel B reports the correlations of independent variables used in the analysis. We recall that the sample period in this table is from July 1995 to December 2014, which is 36 months shorter than that reported in Table 5.11. This is the reason why the means of the two GDP factors are different from those in Table 5.11. Particularly, in Table 5.12 the mean of GDP^F-R is still positive and significantly different from zero, but the mean of GDP^F-N is negative but insignificant. At the very first glance, one may attribute this dramatic change in the mean of nominal GDP factor to the following facts. We construct the GDP^F-N factor (as well as the GDP^F-R factor) using the data of GDP growth rates spanning the period from July 1992 to December 2014. In the first 36 months of this period, the nominal GDP growth rates are extremely high and the high inflation rates contribute to the high nominal GDP growth

Table 5.12: Basic descriptive statistics for and correlations of independent variables

	MKT	SMB	HML	WML	GDP^F-N	GDP^F-R
Panel A: Summary statistics for all factors						
Mean	1.1853	0.6962	0.3017	0.2900	-0.3903	0.5717
Median	0.735	0.960	0.265	0.005	-0.3035	0.5326
Std. Dev.	8.9551	4.1967	2.9943	4.1753	4.4048	0.8718
t-statistic	2.02**	2.54**	1.54	1.06	-1.36	10.03***
Minimum	-26.85	-17.3	-11.8	-9.71	-6.7534	-0.6401
Maximum	36.29	14.17	14.84	23.94	14.6364	3.3952
Skewness	0.4476	-0.4317	-0.0720	1.0572	0.9242	1.1684
Kurtosis	4.6798	4.5027	6.6197	7.4941	4.3563	4.6659
Panel B: Correlation between factors						
MKT	1					
SMB	0.0924	1				
HML	0.0291	-0.2695	1			
WML	0.2071	-0.1401	-0.0921	1		
GDP^F-N	0.1601	-0.0256	0.0059	0.0315	1	
GDP^F-R	0.1725	-0.1094	0.0744	0.0680	0.9811	1

This table reports the descriptive statistics for and correlations of the monthly returns on all the factors. Sample period is July 1995 to December 2014, with 234 monthly observations. Notice that the sample size reported in this table is smaller than that reported in Table 5.11.

rates.[19] These lead the values of GDP^F-N to be also very high in the first 36 months. But here in Table 5.12 we report the mean of GDP^F-N excluding the extremely high values of GDP^F-N in the 36 months before July 1995 from the sample. Thus, the mean of GDP^F-N becomes negative only because of the way of constructing the GDP^F-N factor and the exclusion of large positive values from the sample.

However, the consistently positive and significant mean of GDP^F-R makes us suspect the aforesaid explanation. In fact, we also construct the GDP^F-N factor using the data of GDP growth rates only covering the period from July 1995 to December 2014, and the mean of GDP^F-N is still negative and insignificant. We provide an economic explanation for these facts as follows. The nominal GDP growth is determined by the real GDP growth and the inflation. So the returns on the mimicking portfolio of the nominal GDP factor are also influenced by real GDP growth and inflation. For a normal equity asset, real GDP growth and inflation have offsetting effects on its returns. That is, higher real GDP growth will increase the asset's

[19] See Section 5.2.3 of this study for a more detailed description of GDP growth and inflation rates.

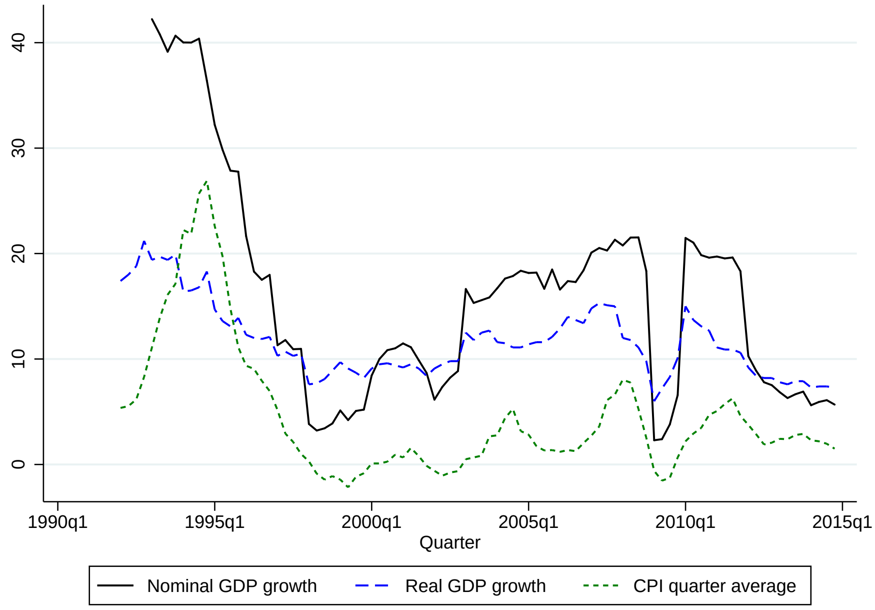

Figure 5.3: Trends of nominal and real GDP growth rates and CPI

returns, but higher inflation will decrease the returns. Similar to the normal assets, the returns on GDP^F-N will be determined by the offsetting effects of both real GDP growth and inflation. If the effect of real GDP growth is greater than that of inflation then the return on GDP^F-N will be positive. If the effect of real GDP growth is less than that of inflation then the return on GDP^F-N will be negative. This explanation is consistent with the fact that the median of GDP^F-N is less than that of GDP^F-R in both Tables 5.11 and 5.12. It also suggests that although the inflation was high during the period from July 1992 to July 1995, the impact of China's real economy growth on equity returns beat the negative impact of high inflation, which is consistent with the high real GDP growth rates during the same period. To have a visual picture of this, we draw the trends of nominal and real GDP growth rates as well as the Consumer Price Index (CPI) in Figure 5.3. CPI is monthly observed, and its value in the figure is quarterly averaged after subtracting 100 from the index series.

Except for the negative mean of the nominal GDP factor, all other factors have positive means. But, only MKT, SMB, and GDP^F-R have significant means being different from zero at the 5% level. The correlation coefficients between the Fama-French factors and the market factor are both small. The relationship between SMB and HML is negative. WML has a positive relationship with MKT, but negative relationships with the Fama-French factors. The two GDP factors are positively related to the market excess return, book-to-market equity, and momentum factors, but negatively with the size factor. It is not surprising that we find an extremely high correlation between the two GDP factors.

5.5.2 Individual regressions of the GDP augmented four-factor model

Tables 5.13 and 5.14 report the results of individual regressions of a GDP augmented four-factor model on the 25 size and book-to-market sorted portfolios in the cases of nominal and real GDP growth, respectively. The results will be compared with those from individual regressions of a four-factor model on the same 25 portfolios reported in Table 5.15.

In both cases, the coefficients on the market, size, book-to-market equity, and momentum factors are very similar to those of four-factor regression presented in Table 5.15. Specifically, the market factor is statistically significant and positive for all 25 portfolios with the values around one. SMB is statistically significant for 24 portfolios and positive except for the five big size portfolios. HML is statistically significant at the 10% level for 15 high and low book-to-market portfolios while the middle book-to-market portfolios do not present significant coefficients related to HML. The coefficient estimators related to the momentum factor are less statistically significant, and only six portfolios present the significance of coefficients. The results here suggest that in addition to the market factor, the size factor is the most important factor in explaining the individual returns on the 25 size-value sorted portfolios. The value factor is less important. As shown in Chapters 2 and 3, even over the longer period in this chapter, the momentum factor is least able to explain the 25 portfolio returns in individual time-series regressions.

There are five coefficient estimators related to the GDP factor statistically significant at least at the 10% level for the same portfolios in both cases of nominal and real GDP factors as shown in Tables 5.13 and 5.14. The magnitudes of coefficients are generally greater in the real GDP growth case. This is due to the fact that the inflation-adjusted GDP growth rate is less than the nominal GDP growth rate as long as the inflation rate is positive.

The adjusted R^2 for 25 regressions in the case of nominal GDP growth ranges from 82.9% to 95.53%, with the average of 91.66%, while for the real GDP growth case, the adjusted R^2 has a range of 82.9% to 95.54% with an average value of 91.67%. Compared to the adjusted R^2 from the four-factor regressions which spans from 82.93% to 95.52%, with the average of 91.63%, we can see only a slight increase in the average value of the adjusted R^2 for the GDP augmented four-factor models.

In addition, the regression intercepts are significant at the 10% level for 6 and 3 (out of 25) portfolios in the nominal and real GDP growth cases, in comparison with 5 portfolios in the four-factor regressions. The results of individual regressions, therefore, indicate that the performance of GDP augmented four-factor model and the standard four-factor model is very similar and any improvement from including the GDP factor into the four-factor model is only marginal. The results might suggest either that there is no information contained in both nominal and real GDP factors helpful for explaining the equity returns, or that the information related to equity returns contained in both nominal and real GDP factors is also embodied in the priced factors. However, which one should be the case cannot be answered at the moment. But the second stage of our main analysis will provide evidence for it.

Moreover, compared with the findings of Nguyen et al. (2009) in the Australian market, we find a greater adjusted R^2 from the GDP augmented four-factor model and a slightly more significant GDP factor in the Chinese market.[20]

As we discuss before, if news related to future GDP growth induces the significance of the MKT, SMB, HML and/or WML factors, if any, then in the presence of a superior proxy for news related to future GDP growth, i.e., the GDP factor here, those factors should lose their ability to explain equity returns. If this is the case, the factor loadings on MKT, SMB, HML and/or WML will become insignificant while the loading on the GDP factor should be significant and positive. However, we do not observe this according to the above regression

[20] Vassalou (2003) does not report the individual regressions in her study, thus we cannot make a comparison with the US market.

analysis in the Chinese market. When news related to future GDP growth is introduced as a risk factor into the four-factor asset pricing model, the explanatory powers of the four factors are unaffected while the loading on GDP^F is generally insignificant. Overall, the results of our individual regressions suggest that there is no significant evidence here to support the argument that MKT, SMB, HML and/or WML are proxies for the risk associated with future GDP growth.

Table 5.13: Individual regressions of a GDP^F-N augmented four-factor model on 25 size and book-to-market sorted portfolios, period July 1995 to December 2012, obs. 234

Size	BE/ME	b_p	s_p	h_p	w_p	g_p	a_p	Adj. R^2
Small	Low	0.9560***	1.0972***	-0.1779	-0.1853	0.0544	0.3533	0.8290
		(17.24)	(8.20)	(-1.05)	(-1.30)	(0.75)	(1.27)	
	2	0.9766***	1.1223***	-0.0380	-0.0266	0.0278	0.7091***	0.9094
		(36.14)	(11.03)	(-0.35)	(-0.29)	(0.51)	(3.16)	
	3	0.9958***	1.1950***	-0.0765	0.1192*	0.0623	0.5339***	0.9476
		(40.54)	(18.96)	(-0.96)	(1.96)	(1.43)	(2.74)	
	4	0.9470***	1.1047***	0.1367	-0.0473	0.0163	0.4041**	0.9262
		(36.57)	(10.45)	(1.58)	(-0.40)	(0.44)	(2.19)	
	High	0.9311***	1.1898***	0.2459***	0.0631	-0.0764*	-0.0740	0.9420
		(32.80)	(20.81)	(2.92)	(1.13)	(-1.74)	(-0.43)	
2	Low	0.9839***	0.7836***	-0.5414***	-0.1144**	-0.0108	0.1027	0.9258
		(27.22)	(12.09)	(-5.90)	(-2.12)	(-0.26)	(0.59)	
	2	1.0476***	0.9301***	-0.1800***	0.0072	0.0441*	0.3100**	0.9553
		(45.63)	(19.82)	(-2.81)	(0.20)	(1.77)	(2.38)	
	3	1.0195***	1.0822***	0.0334	0.1429**	0.0609	0.2195	0.9487
		(39.59)	(15.61)	(0.47)	(2.02)	(1.47)	(1.24)	
	4	1.0486***	1.1350***	0.1136	0.1479*	0.0048	0.1328	0.9263
		(31.25)	(13.76)	(1.29)	(1.85)	(0.08)	(0.72)	
	High	0.9664***	0.9341***	0.4188***	0.0383	-0.0130	-0.0874	0.9134
		(33.46)	(15.53)	(4.06)	(0.67)	(-0.19)	(-0.47)	
3	Low	0.9484***	0.7311***	-0.3083***	0.0177	0.0179	-0.1801	0.9224
		(34.72)	(15.01)	(-4.61)	(0.42)	(0.56)	(-1.11)	
	2	0.9853***	0.7324***	-0.0981	0.1809**	0.0649	0.1593	0.9203
		(43.40)	(8.46)	(-1.07)	(2.42)	(1.57)	(0.83)	
	3	1.0257***	0.8609***	0.0711	0.0365	0.0943	-0.0005	0.9090
		(32.15)	(10.62)	(0.73)	(0.39)	(1.43)	(0.00)	
	4	1.0306***	0.7677***	0.2258**	-0.0345	0.1225***	0.3306*	0.9376
		(40.20)	(13.06)	(2.54)	(-0.79)	(3.14)	(1.70)	
	High	0.9987***	0.6478***	0.3839***	0.0334	-0.0423	-0.2669	0.9268
		(41.74)	(10.03)	(4.76)	(0.61)	(-1.08)	(-1.47)	

Table 5.13: (Continued)

Size	BE/ME	b_p	s_p	h_p	w_p	g_p	a_p	Adj. R^2
4	Low	0.9119***	0.4337***	-0.3924***	0.0343	-0.0091	-0.0989	0.8730
		(26.93)	(6.45)	(-3.51)	(0.52)	(-0.18)	(-0.54)	
	2	1.0179***	0.4508***	-0.2216**	-0.0179	0.0591	0.0532	0.9073
		(35.55)	(5.36)	(-2.17)	(-0.22)	(1.14)	(0.23)	
	3	1.0628***	0.2792**	0.0678	-0.1002	0.0703	0.2513	0.9010
		(36.12)	(2.53)	(0.70)	(-0.93)	(1.28)	(0.97)	
	4	1.0553***	0.4335***	0.1760**	-0.0090	0.0808*	0.0743	0.9230
		(36.00)	(6.24)	(1.99)	(-0.12)	(1.75)	(0.38)	
	High	1.0595***	0.4634***	0.7111***	0.1165	0.0637	-0.0342	0.9038
		(32.80)	(4.23)	(7.56)	(0.91)	(1.18)	(-0.15)	
Big	Low	0.9425***	-0.2022***	-0.8362***	0.0719**	0.0164	0.2220	0.9324
		(34.59)	(-3.79)	(-12.09)	(2.05)	(0.39)	(1.35)	
	2	1.0315***	-0.0922	-0.4031***	0.0530	-0.0279	-0.0678	0.9255
		(43.20)	(-1.47)	(-4.74)	(0.97)	(-0.81)	(-0.35)	
	3	1.0663***	-0.2967***	0.0623	-0.0277	-0.0203	0.3088	0.9073
		(32.03)	(-3.24)	(0.58)	(-0.29)	(-0.40)	(1.45)	
	4	1.0190***	-0.2342**	0.3501***	0.0189	0.1507***	0.5066**	0.8962
		(24.73)	(-2.53)	(3.05)	(0.22)	(2.66)	(2.34)	
	High	0.9521***	-0.2838***	0.7975***	0.0881	0.0411	0.1776	0.9067
		(36.17)	(-2.81)	(8.78)	(0.66)	(0.96)	(1.00)	

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level
This table reports the results of individual regressions of a GDP^F-N augmented four-factor model on the 25 size and book-to-market sorted portfolios.

$$R_{pt} - RF = a_p + b_p\left(R_{mt} - RF\right) + s_p SMB_t + h_p HML_t + w_p WML_t + g_p GDP_t^F + e_{pt}$$

where MKT_t is the market excess return over the risk-free rate; SMB_t is the return on size factor; HML_t is the return on book-to-market factor; WML_t is the return on momentum factor; GDP_t^F is substituted by the return on GDP^F-N; and ε_{it} is the error term.
Standard errors are corrected for serial correlation up to three lags and White's (1980) heteroskedasticity using the Newey-West (1987) estimator to calculate the t-statistics.
Sample period is July 1992 to December 2014, with 270 observations.

Table 5.14: Individual regressions of a GDP^F-R augmented four-factor model on 25 size and book-to-market sorted portfolios, period July 1995 to December 2012, obs. 234

Size	BE/ME	b_p	s_p	h_p	w_p	g_p	a_p	Adj. R^2
Small	Low	0.9555***	1.1012***	-0.1822	-0.1871	0.2838	0.1693	0.8290
		(17.48)	(8.05)	(-1.08)	(-1.32)	(0.65)	(0.57)	
	2	0.9758***	1.1252***	-0.0407	-0.0276	0.1794	0.5958***	0.9094
		(36.81)	(11.23)	(-0.37)	(-0.30)	(0.60)	(2.70)	
	3	0.9947***	1.2004***	-0.0820	0.1171*	0.3609	0.3031	0.9478
		(40.84)	(19.53)	(-1.05)	(1.94)	(1.50)	(1.52)	
	4	0.9470***	1.1058***	0.1355	-0.0478	0.0793	0.3522*	0.9262
		(36.69)	(10.46)	(1.56)	(-0.41)	(0.39)	(1.91)	
	High	0.9326***	1.1830***	0.2526***	0.0658	-0.4476**	0.2118	0.9423
		(33.29)	(20.69)	(3.01)	(1.20)	(-2.00)	(1.08)	
2	Low	0.9845***	0.7821***	-0.5402***	-0.1139**	-0.0863	0.1561	0.9258
		(27.41)	(12.04)	(-5.92)	(-2.11)	(-0.38)	(0.81)	
	2	1.0470***	0.9337***	-0.1837***	0.0057	0.2475*	0.1511	0.9554
		(45.69)	(19.99)	(-2.86)	(0.16)	(1.96)	(1.07)	
	3	1.0186***	1.0873***	0.0282	0.1408**	0.3434	-0.0009	0.9488
		(39.50)	(15.73)	(0.40)	(2.01)	(1.45)	(-0.01)	
	4	1.0472***	1.1373***	0.1123	0.1474*	0.1088	0.0695	0.9264
		(31.66)	(13.49)	(1.28)	(1.86)	(0.33)	(0.34)	
	High	0.9679***	0.9313***	0.4208***	0.0391	-0.1463	0.0006	0.9135
		(33.98)	(15.39)	(4.15)	(0.68)	(-0.42)	(0.00)	
3	Low	0.9488***	0.7316***	-0.3093***	0.0173	0.0602	-0.2220	0.9223
		(34.53)	(14.91)	(-4.63)	(0.41)	(0.34)	(-1.30)	
	2	0.9837***	0.7386***	-0.1040	0.1785**	0.3990*	-0.0941	0.9207
		(42.96)	(8.71)	(-1.14)	(2.43)	(1.70)	(-0.48)	
	3	1.0256***	0.8669***	0.0640	0.0336	0.4510	-0.2962	0.9088
		(32.76)	(10.54)	(0.67)	(0.36)	(1.20)	(-1.35)	
	4	1.0289***	0.7777***	0.2156**	-0.0386	0.6795***	-0.1062	0.9381
		(40.74)	(13.28)	(2.42)	(-0.88)	(3.26)	(-0.54)	
	High	1.0004***	0.6429***	0.3882***	0.0350	-0.2981	-0.0803	0.9271
		(42.07)	(10.00)	(4.78)	(0.65)	(-1.45)	(-0.42)	

Table 5.14: (Continued)

Size	BE/ME	b_p	s_p	h_p	w_p	g_p	a_p	Adj. R^2
4	Low	0.9125***	0.4324***	-0.3913***	0.0348	-0.0768	-0.0516	0.8730
		(27.07)	(6.33)	(-3.47)	(0.53)	(-0.28)	(-0.26)	
	2	1.0175***	0.4550***	-0.2262**	-0.0198	0.3023	-0.1433	0.9073
		(35.63)	(5.44)	(-2.21)	(-0.24)	(1.05)	(-0.62)	
	3	1.0632***	0.2831***	0.0629	-0.1022	0.3099	0.0456	0.9008
		(36.68)	(2.60)	(0.65)	(-0.94)	(0.94)	(0.22)	
	4	1.0549***	0.4392***	0.1697*	-0.0115	0.4081	-0.1913	0.9230
		(36.30)	(6.39)	(1.94)	(-0.15)	(1.58)	(-0.99)	
	High	1.0581***	0.4693***	0.7054***	0.1142	0.3836	-0.2783	0.9040
		(33.80)	(4.29)	(7.65)	(0.90)	(1.23)	(-1.19)	
Big	Low	0.9426***	-0.2013***	-0.8374***	0.0714**	0.0747	0.1726	0.9324
		(34.75)	(-3.82)	(-12.01)	(2.01)	(0.34)	(0.99)	
	2	1.0328***	-0.0956	-0.4001***	0.0542	-0.2051	0.0601	0.9256
		(43.33)	(-1.53)	(-4.66)	(0.99)	(-1.02)	(0.31)	
	3	1.0662***	-0.2978***	0.0637	-0.0271	-0.0872	0.3669*	0.9073
		(32.13)	(-3.27)	(0.59)	(-0.29)	(-0.31)	(1.71)	
	4	1.0180***	-0.2234**	0.3382***	0.0141	0.7735***	0.0043	0.8962
		(24.71)	(-2.40)	(2.92)	(0.16)	(2.61)	(0.02)	
	High	0.9521***	-0.2812***	0.7944***	0.0868	0.1984	0.0476	0.9067
		(36.22)	(-2.76)	(8.79)	(0.66)	(0.86)	(0.23)	

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level
This table reports the results of individual regressions of a GDP^F-R augmented four-factor model on the 25 size and book-to-market sorted portfolios.

$$R_{pt} - RF = a_p + b_p\left(R_{mt} - RF\right) + s_p SMB_t + h_p HML_t + w_p WML_t + g_p GDP_t^F + e_{pt}$$

where MKT_t is the market excess return over the risk-free rate; SMB_t is the return on size factor; HML_t is the return on book-to-market factor; WML_t is the return on momentum factor; GDP_t^F is substituted by the return on GDP^F-R; and ε_{it} is the error term.
Standard errors are corrected for serial correlation up to three lags and White's (1980) heteroskedasticity using the Newey-West (1987) estimator to calculate the t-statistics.
Sample period is July 1992 to December 2014, with 270 observations.

Table 5.15: Individual regressions of a four-factor model on 25 size and book-to-market sorted portfolios, period July 1995 to December 2012, obs. 234

Size	BE/ME	b_p	s_p	h_p	w_p	a_p	Adj. R^2
Small	Low	0.9604***	1.0946***	-0.1789	-0.1859	0.3290	0.8293
		(16.61)	(8.25)	(-1.06)	(-1.33)	(1.21)	
	2	0.9789***	1.1210***	-0.0385	-0.0269	0.6967***	0.9096
		(34.40)	(10.90)	(-0.35)	(-0.29)	(3.21)	
	3	1.0009***	1.1920***	-0.0777	0.1186*	0.5062***	0.9473
		(38.92)	(18.28)	(-1.00)	(1.92)	(2.65)	
	4	0.9483***	1.1039***	0.1364	-0.0475	0.3968**	0.9264
		(37.26)	(10.46)	(1.58)	(-0.41)	(2.20)	
	High	0.9248***	1.1935***	0.2473***	0.0640	-0.0400	0.9412
		(29.03)	(20.68)	(2.88)	(1.11)	(-0.24)	
2	Low	0.9830***	0.7841***	-0.5412***	-0.1143**	0.1075	0.9261
		(26.40)	(12.09)	(-5.91)	(-2.11)	(0.63)	
	2	1.0512***	0.9279***	-0.1808***	0.0067	0.2903**	0.9552
		(43.58)	(19.38)	(-2.80)	(0.19)	(2.25)	
	3	1.0245***	1.0793***	0.0323	0.1422**	0.1923	0.9483
		(38.94)	(15.26)	(0.45)	(1.97)	(1.12)	
	4	1.0490***	1.1347***	0.1136	0.1479*	0.1307	0.9266
		(29.52)	(13.90)	(1.29)	(1.85)	(0.75)	
	High	0.9654***	0.9347***	0.4191***	0.0385	-0.0817	0.9137
		(31.34)	(15.56)	(4.10)	(0.66)	(-0.47)	
3	Low	0.9498***	0.7302***	-0.3086***	0.0175	-0.1881	0.9226
		(35.32)	(14.96)	(-4.60)	(0.42)	(-1.18)	
	2	0.9906***	0.7293***	-0.0992	0.1801**	0.1304	0.9199
		(44.54)	(8.14)	(-1.08)	(2.37)	(0.68)	
	3	1.0334***	0.8564***	0.0694	0.0354	-0.0424	0.9079
		(29.19)	(10.58)	(0.71)	(0.37)	(-0.20)	
	4	1.0406***	0.7618***	0.2236**	-0.0359	0.2761	0.9353
		(36.28)	(12.76)	(2.43)	(-0.79)	(1.43)	
	High	0.9952***	0.6498***	0.3847***	0.0338	-0.2481	0.9268
		(39.48)	(9.96)	(4.70)	(0.62)	(-1.37)	

Table 5.15: (Continued)

Size	BE/ME	b_p	s_p	h_p	w_p	a_p	Adj. R^2
4	Low	0.9112***	0.4342***	-0.3922***	0.0345	-0.0948	0.8735
		(25.74)	(6.52)	(-3.52)	(0.53)	(-0.54)	
	2	1.0227***	0.4480***	-0.2226**	-0.0186	0.0268	0.9071
		(34.88)	(5.27)	(-2.18)	(-0.23)	(0.12)	
	3	1.0686***	0.2758**	0.0665	-0.1010	0.2200	0.9005
		(34.40)	(2.45)	(0.69)	(-0.95)	(0.89)	
	4	1.0620***	0.4297***	0.1746*	-0.0099	0.0383	0.9221
		(33.16)	(5.96)	(1.96)	(-0.13)	(0.20)	
	High	1.0647***	0.4603***	0.7100***	0.1158	-0.0625	0.9035
		(29.56)	(4.14)	(7.48)	(0.89)	(-0.29)	
Big	Low	0.9439***	-0.2030***	-0.8365***	0.0717**	0.2147	0.9327
		(33.17)	(-3.79)	(-12.10)	(2.05)	(1.35)	
	2	1.0292***	-0.0909	-0.4026***	0.0533	-0.0554	0.9256
		(42.24)	(-1.44)	(-4.74)	(0.98)	(-0.29)	
	3	1.0647***	-0.2957***	0.0627	-0.0274	0.3178	0.9076
		(31.68)	(-3.23)	(0.58)	(-0.29)	(1.56)	
	4	1.0314***	-0.2414***	0.3474***	0.0172	0.4395**	0.8922
		(22.06)	(-2.60)	(2.92)	(0.20)	(1.97)	
	High	0.9555***	-0.2858***	0.7968***	0.0876	0.1593	0.9068
		(36.84)	(-2.83)	(8.71)	(0.66)	(0.92)	

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level
This table reports the results of individual regressions of a four-factor model on the 25 size and book-to-market sorted portfolios.

$$R_{pt} - RF = a_p + b_p\,(R_{mt} - RF) + s_p SMB_t + h_p HML_t + w_p WML_t + e_{pt}$$

where MKT_t is the market excess return over the risk-free rate; SMB_t is the return on size factor; HML_t is the return on book-to-market factor; WML_t is the return on momentum factor; and ε_{it} is the error term.
Standard errors are corrected for serial correlation up to three lags and White's (1980) heteroskedasticity using the Newey-West (1987) estimator to calculate the t-statistics.
Sample period is July 1992 to December 2014, with 270 observations.

5.5.3 Results of restricted system based on GMM estimation

Table 5.16 and 5.17 present the estimated risk loadings on the four pricing factors in addition to the nominal and real GDP factors, respectively. We find that the risk loadings on the market, size, and book-to-market factors are very similar to those from individual regressions in the previous subsection for both cases of nominal and real GDP factors. The magnitudes of risk loadings on the momentum factor and the nominal and real GDP factors are different from those estimated based on individual regressions. This is not a surprise because we impose a restriction of zero intercepts in the GMM system regressions. However, the significance of these risk loadings is similar to those obtained form individual regressions in the previous subsection. Thus, this GMM setup in which we can estimate the pricing model and factor premiums simultaneously also confirms the robustness of findings in the previous subsection. Because the risk loadings are not the focus of this subsection, in the next we pay our attention to the estimations of factor risk premiums.

Table 5.16: GMM regressions of a GDP^F-N augmented four-factor model on 25 size and book-to-market sorted portfolios

Size	BE/ME	b_p	s_p	h_p	w_p	g_p
Small	Low	0.9399***	1.2035***	-0.0463	-0.1189	0.0994**
		(28.97)	(13.53)	(-0.46)	(-1.23)	(2.14)
	2	0.9739***	1.2680***	-0.0199	-0.0060	0.0100
		(48.01)	(21.89)	(-0.31)	(-0.09)	(0.23)
	3	0.9796***	1.2374***	-0.0907	0.0793*	0.0155
		(53.11)	(29.43)	(-1.49)	(1.72)	(0.49)
	4	0.9517***	1.2369***	0.1462***	0.0629	0.0144
		(54.83)	(16.94)	(2.58)	(0.66)	(0.54)
	High	0.9414***	1.2343***	0.2276***	0.0936**	-0.0422
		(48.60)	(31.61)	(3.53)	(1.96)	(-1.29)
2	Low	1.0018***	0.8109***	-0.4749***	-0.0937**	0.0304
		(49.33)	(16.68)	(-6.94)	(-2.21)	(0.98)
	2	1.0216***	0.9742***	-0.1751***	0.0389	0.0236
		(55.63)	(31.00)	(-3.33)	(1.51)	(0.84)
	3	1.0018***	1.0750***	-0.0342	0.1004*	0.0328
		(61.87)	(24.16)	(-0.73)	(1.94)	(1.16)
	4	1.0249***	1.0861***	0.0761	0.1329**	-0.0258
		(49.28)	(18.45)	(1.42)	(2.11)	(-0.82)
	High	0.9784***	0.9859***	0.4087***	-0.0006	0.0335
		(49.51)	(22.24)	(5.06)	(-0.01)	(0.71)
3	Low	0.9657***	0.7374***	-0.3377***	0.0224	0.0484*
		(42.24)	(19.90)	(-6.77)	(0.63)	(1.87)
	2	0.9895***	0.8027***	-0.1518**	0.0872	0.0220
		(58.65)	(15.94)	(-2.15)	(1.47)	(0.70)
	3	1.0072***	0.7979***	-0.0381	0.0164	0.0596
		(50.27)	(12.96)	(-0.54)	(0.24)	(1.55)
	4	1.0086***	0.7993***	0.2238***	-0.0063	0.0892**
		(53.54)	(18.66)	(3.29)	(-0.18)	(2.58)
	High	1.0144***	0.6253***	0.3140***	0.0553	-0.0147
		(48.06)	(11.73)	(4.77)	(1.21)	(-0.55)

Table 5.16: (Continued)

Size	BE/ME	b_p	s_p	h_p	w_p	g_p
4	Low	0.9125***	0.5113***	-0.4320***	0.0575	0.0408
		(36.00)	(10.57)	(-5.49)	(1.01)	(1.11)
	2	1.0047***	0.5287***	-0.2821***	0.0574	0.0275
		(44.31)	(9.38)	(-3.89)	(0.86)	(0.80)
	3	1.0752***	0.4232***	0.0175	0.0001	0.0231
		(51.35)	(6.11)	(0.26)	(0.00)	(0.62)
	4	1.0397***	0.4359***	0.1448**	-0.0129	0.0444
		(53.87)	(9.75)	(1.99)	(-0.20)	(1.37)
	High	1.0323***	0.4217***	0.6048***	0.0067	0.0322
		(58.15)	(5.49)	(9.27)	(0.07)	(0.96)
Big	Low	0.9426***	-0.1841***	-0.8520***	0.0886***	0.0354
		(54.86)	(-5.36)	(-17.52)	(3.38)	(1.30)
	2	1.0287***	-0.1104**	-0.4464***	0.0315	-0.0246
		(47.78)	(-2.30)	(-7.11)	(0.73)	(-1.03)
	3	1.0614***	-0.1757***	0.0197	0.0414	-0.0283
		(38.94)	(-3.16)	(0.25)	(0.55)	(-0.81)
	4	1.0020***	-0.1636***	0.4029***	0.1004	0.1040***
		(37.77)	(-2.85)	(5.32)	(1.58)	(2.71)
	High	0.9592***	-0.3010***	0.7884***	-0.0716	0.0599
		(52.00)	(-4.04)	(9.96)	(-0.66)	(1.64)

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level
This table reports the estimated risk loadings on the four factors in addition to GDP^F-N factor from the restricted system based on GMM estimation. The test of the GDP augmented four-factor model is based on the following system:

$$R_{pt} - RF = b_p\left(R_{mt} - RF\right) + s_p SMB_t + h_p HML_t + w_p WML_t + g_p GDP_t^F + e_{pt}$$

$$R_{mt} - RF = \lambda_m + e_{mt}$$

$$SMB_t = \lambda_s + e_{st}$$

$$HML_t = \lambda_h + e_{ht}$$

$$WML_t = \lambda_w + e_{wt}$$

$$GDP_t^F = \lambda_g + e_{gt}$$

The associated z-statistic for the factor risk loadings is contained in parentheses below the coefficient estimate. Sample period is July 1995 to December 2014, with 234 observations.

Table 5.17: GMM regressions of a GDP^F-R augmented four-factor model on 25 size and book-to-market sorted portfolios

Size	BE/ME	b_p	s_p	h_p	w_p	g_p
Small	Low	0.9581***	1.2410***	-0.1189	-0.0899	0.7051**
		(27.00)	(15.51)	(-1.33)	(-1.15)	(2.24)
	2	0.9664***	1.2099***	-0.0565	0.0655	0.7027***
		(46.56)	(20.31)	(-0.70)	(1.00)	(3.00)
	3	0.9650***	1.2171***	-0.0691	0.0649*	0.6099***
		(51.16)	(27.89)	(-1.14)	(1.77)	(3.17)
	4	0.9514***	1.2291***	0.1279**	0.0472	0.3527**
		(53.61)	(23.00)	(2.25)	(0.80)	(2.36)
	High	0.9285***	1.2490***	0.2521***	0.1277***	-0.1425
		(49.17)	(30.80)	(3.93)	(3.13)	(-0.87)
2	Low	0.9872***	0.8597***	-0.4508***	-0.0949**	0.1015
		(41.84)	(17.40)	(-6.34)	(-2.23)	(0.67)
	2	1.0286***	0.9650***	-0.1833***	0.0318	0.2843***
		(54.58)	(28.79)	(-3.53)	(1.05)	(2.73)
	3	1.0036***	1.0665***	-0.0152	0.0826**	0.3340**
		(52.24)	(27.73)	(-0.30)	(2.31)	(1.98)
	4	1.0379***	1.0977***	0.0718	0.1102**	0.1888
		(46.26)	(24.55)	(1.26)	(2.23)	(0.90)
	High	0.9625***	1.0188***	0.3751***	0.0498	-0.1748
		(44.19)	(26.69)	(5.55)	(1.17)	(-0.75)
3	Low	0.9587***	0.7524***	-0.3592***	0.0314	0.0458
		(42.58)	(19.82)	(-6.85)	(0.86)	(0.33)
	2	0.9946***	0.7715***	-0.1713**	0.1020**	0.2648
		(56.84)	(12.05)	(-2.45)	(2.32)	(1.42)
	3	1.0121***	0.8559***	0.0004	-0.0222	0.1308
		(46.13)	(17.61)	(0.01)	(-0.39)	(0.46)
	4	1.0115***	0.8213***	0.1599**	-0.0160	0.5631***
		(50.61)	(20.59)	(2.40)	(-0.46)	(3.47)
	High	1.0071***	0.6876***	0.3530***	0.0548	-0.2466*
		(53.05)	(14.51)	(5.65)	(1.29)	(-1.69)

Table 5.17: (Continued)

Size	BE/ME	b_p	s_p	h_p	w_p	g_p
4	Low	0.9380***	0.5060***	-0.4749***	0.0663	-0.0542
		(35.76)	(11.50)	(-5.53)	(1.39)	(-0.27)
	2	1.0103***	0.5320***	-0.2803***	0.0611	0.2237
		(43.03)	(9.95)	(-3.71)	(1.13)	(1.08)
	3	1.0726***	0.4240***	0.0185	0.0097	0.2864
		(47.12)	(6.56)	(0.28)	(0.16)	(0.96)
	4	1.0462***	0.4485***	0.1097*	-0.0048	0.1750
		(51.65)	(8.82)	(1.84)	(-0.09)	(0.77)
	High	1.0484***	0.4242***	0.6074***	0.0021	0.1725
		(42.79)	(6.54)	(9.66)	(0.03)	(0.73)
Big	Low	0.9338***	-0.1585***	-0.8766***	0.0974***	0.2202
		(48.25)	(-4.00)	(-18.59)	(3.65)	(1.38)
	2	1.0184***	-0.0633	-0.4186***	0.0373	-0.2106
		(51.39)	(-1.41)	(-6.62)	(0.97)	(-1.21)
	3	1.0624***	-0.1797***	0.0593	0.0381	0.2961
		(38.82)	(-3.50)	(0.71)	(0.67)	(1.35)
	4	1.0106***	-0.1611***	0.2899***	0.1063*	0.8460***
		(30.63)	(-2.64)	(3.24)	(1.79)	(3.73)
	High	0.9497***	-0.2997***	0.7985***	-0.0630	0.1708
		(42.40)	(-4.96)	(11.87)	(-0.99)	(1.04)

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level
This table reports the estimated risk loadings on the four factors in addition to GDP^F-R factor from the restricted system based on GMM estimation. The test of the GDP augmented four-factor model is based on the following system:

$$R_{pt} - RF = b_p\,(R_{mt} - RF) + s_p SMB_t + h_p HML_t + w_p WML_t + g_p GDP_t^F + e_{pt}$$

$$R_{mt} - RF = \lambda_m + e_{mt}$$

$$SMB_t = \lambda_s + e_{st}$$

$$HML_t = \lambda_h + e_{ht}$$

$$WML_t = \lambda_w + e_{wt}$$

$$GDP_t^F = \lambda_g + e_{gt}$$

The associated z-statistic for the factor risk loadings is contained in parentheses below the coefficient estimate. Sample period is July 1995 to December 2014, with 234 observations.

Table 5.18 presents the estimated risk premiums on the factors from the restricted system based on the GMM estimation and the overidentifying tests of the cross-equation restrictions on the CAPM, the GDP augmented CAPM, the standard Fama–French model, the GDP augmented Fama-French model, the four-factor model including the momentum factor and the GDP augmented four-factor model. In the system regressions, we focus our analysis on the GMM statistics and the estimated factor premiums. Here, the GMM statistics represents the small-sample adjusted Hansen's (1982) J-statistic, which is used to perform the tests of overidentifying restrictions of the models and is subject to a chi-square distribution. If the factor premium is positive and significant, we conclude that the factor is systematic and priced in the cross section of equity returns. On the other hand, if the factor premium is insignificant, or negative and significant, we conclude that the factor is not priced and therefore not systematic.[21]

Panel A presents the results of GMM regressions for three standard models, that is, the CAPM, Fama-French, and four-factor models. The first system regression reported is for the CAPM. The GMM statistic is insignificant implying that the model cannot be rejected. The market premium (λ_m) is positive and statistically significant at the 5% level with the z-statistic being 2.34. The second system regression reported is for the Fama–French model. The GMM statistic is insignificant which supports the overall favorability of the model. The market premium (λ_m) is still positive but statistically significant at the 10% level with the z-statistic being 1.87. The size premium (λ_s) and the book-to-market equity premium (λ_h) are both positive and statistically significant at the 5% level. The third system regression reported is for the four-factor model. The GMM statistic is again insignificant; therefore, we cannot reject the model. The magnitude and significance of the market, size, and book-to-market equity risk premiums are very similar to those in the Fama-French model. The added momentum factor presents an insignificant risk premium in the system regression. The main finding of Panel

[21]As in Nguyen et al. (2009), the five factors "employed in this analysis have been constructed such that a positive premium is consistent with a risk-based explanation for the factors. Thus, an observed negative premium would be inconsistent with a risk-based explanation and would therefore indicate that the factor is not priced and is not systematic."

Table 5.18: Generalised method of moments (GMM) system tests of the asset pricing models specified

	GMM	λ_m	λ_s	λ_h	λ_w	λ_g
Panel A: standard models						
CAPM	32.682	1.3451**				
	[0.1392]	(2.34)				
FF	27.1858	1.0487*	0.7474***	0.4239**		
	[0.3467]	(1.87)	(3.07)	(2.28)		
Four-factor	26.5003	0.9672*	0.7829***	0.4036**	0.1153	
	[0.3813]	(1.73)	(3.21)	(2.20)	(0.52)	
Panel B: Nominal GDP augmented models						
CAPM GDP_N	35.0639*	1.2611**				-0.8451
	[0.0871]	(2.20)				(-1.62)
FF GDP_N	32.0213	1.0157*	0.7720***	0.3794**		-0.7273
	[0.1574]	(1.80)	(3.22)	(2.02)		(-1.39)
Four-factor GDP_N	31.3453	0.9395*	0.8257***	0.3597*	0.0573	-0.7653
	[0.1778]	(1.68)	(3.46)	(1.94)	(0.26)	(-1.45)
Panel C: Real GDP augmented models						
CAPM GDP_R	25.4057	0.8813*				0.6199***
	[0.4398]	(1.69)				(8.78)
FF GDP_R	20.777	0.8704*	0.7062***	0.3422*		0.5480***
	[0.7050]	(1.66)	(2.98)	(1.87)		(7.60)
Four-factor GDP_R	19.4246	0.9239*	0.7048***	0.3549**	0.1849	0.5548***
	[0.7763]	(1.77)	(2.98)	(1.97)	(0.87)	(7.61)

*** Statistically significant at the 1% level
** Statistically significant at the 5% level
* Statistically significant at the 10% level
This table reports the estimated risk premiums on the four factors in addition to GDP^F-N and GDP^F-R factors, respectively, from the restricted system based on GMM estimation. The test of the GDP augmented four-factor model is based on the following system:

$$R_{pt} - RF = b_p (R_{mt} - RF) + s_p SMB_t + h_p HML_t + w_p WML_t + g_p GDP_t^F + e_{pt}$$

$$R_{mt} - RF = \lambda_m + e_{mt}$$

$$SMB_t = \lambda_s + e_{st}$$

$$HML_t = \lambda_h + e_{ht}$$

$$WML_t = \lambda_w + e_{wt}$$

$$GDP_t^F = \lambda_g + e_{gt}$$

The generalised method of moments test statistic (GMM), testing that the asset pricing models hold, is distributed as a chi-square with N degrees of freedom. The associated p-value is contained in square brackets below the GMM statistic.
The associated z-statistic for the factor premium is contained in parentheses below the coefficient estimate.
Sample period is July 1995 to December 2014, with 234 observations.

A is that the market, size and value effects are systematic risks which are priced in the cross section of equity returns, but the momentum effect is not.

It is interesting to notice that the mean return on the value factor is insignificant in Table 5.12, while the estimated risk premium on the value factor is significant from the GMM system regressions. The results from the GMM regressions suggest the existence of value premium in the Chinese market although it is generally weaker than the size premium over the sample period. We attribute the difference in significance of value premium to that the GMM approach should be better to estimate the risk premiums (which acts as the cross-sectional stage of the famous Fama-MacBeth regression) than the arithmetic time-series average factor returns. The arithmetic time-series average factor returns as the risk premium estimates only consider the premiums individually but do not consider their roles in the context of asset pricing. The risk premiums estimated from the cross-sectional regressions, for example, the Fama-MacBeth approach or the GMM approach employed in this study, are obtained by examining the ability of factors to explain the cross-sectional variation in equity returns.

Panel B reports the GMM system tests on the three corresponding augmented models which include the nominal GDP factor. The GMM statistics and the estimates of risk premiums on the four common pricing factors are very similar to those found in Panel A, except for an important difference that the GMM statistics for the nominal GDP augmented CAPM is statistically significant at the 10% level thus it is not in favor of the model. The results of these system regressions reported in this panel indicate that the estimated nominal GDP risk premiums are negative and statistically insignificant in all three augmented models. This suggests that the nominal GDP factor is not priced in all three augmented models presented in this panel, thus it is not a systematic risk.

The results of GMM system tests on the three real GDP growth augmented models are reported in Panel C. All GMM statistics for the three models are statistically insignificant, thus the null hypothesis of over-identifying restrictions for all three real GDP augmented models cannot be rejected. The significance of all risk premiums on MKT, SMB, HML and WML is the same as that reported in Panel B. Recall that the aim of the system regression analysis is to test whether the risk associated with future GDP growth is priced in equity returns. If this is the case, then in the presence of the GDP factor, the risk premiums on the four common factors should decrease in magnitudes to some extent or even turn to be insignificant, while the premium on the GDP factor should be significantly positive. In fact, the momentum premium is always insignificant in all cases, thus we focus our analysis on the risk premiums on MKT, SMB, HML, and GDP^F-R. Compared to the three standard models, the magnitudes of risk premiums on MKT, SMB and HML decrease significantly, while the risk premium on the real GDP factor is significantly positive in all three real GDP enhanced models. Besides, the magnitude of the GDP^F-R premium is very close to the mean of returns on the real GDP mimicking portfolio reported in Table 5.12. Specifically, if we compare the risk premiums on the market, size and value factors in both the standard and the corresponding real GDP augmented Fama-French models, we can conclude that adding the real GDP factor into the Fama-French dramatically decreases the magnitudes of the estimated premiums on MKT, SMB and HML.

Another interesting finding from the GMM system regression is that whether the news related to future GDP growth is priced in the cross-sectional equity portfolio returns depends on the way to construct the GDP factor. Here, the GDP factor constructed using real GDP growth rate is priced in the cross section of equity returns. On the other hand, it is not the case for nominal GDP growth. Unlike real GDP growth which has a positive effect on equity returns, inflation has an opposite effect on equity returns. Because the nominal GDP factor involves the offsetting effects of both real GDP growth and inflation on equity returns, including inflation together with real GDP growth contaminates the information captured by the nominal GDP factor. As a result, the nominal GDP factor has no explanatory power for equity returns and is not priced in equity returns. However, the ability of the real GDP factor to explain the cross-sectional portfolio returns seems to be limited according to our empirical results, even though the estimated premium on it is statistically significant. We provide a potential explanation for this interesting finding in the concluding section.

5.6 Conclusion

Finding the macroeconomic variables behind the priced factors is an important task of asset pricing. Vassalou (2003) argues that news related to future GDP growth can primarily explain the anomalies related to size and book-to-market effects in the US context. However, Nguyen et al. (2009) find opposite evidence in the Australian market. Inspired by their contradictory conclusions and by the reasonably good performance of the Fama-French model in the Chinese equity market, we investigate which argument is appropriate in the Chinese context. Specifically, we examine whether the ability of the market, size, book-to-market equity and momentum factors, if any, to explain the Chinese equity returns comes from the fact that they are proxies for news related to future nominal and/or real GDP growth. We also overcome the difficulty of lacking recorded data, which commonly exists in young emerging markets, by calibrating data to enlarge the sample size on which our analysis is based.

To examine the above contention in the context of Chinese market, we address two questions: “Are the common pricing factors, i.e., MKT, SMB, HML, and/or WML, the state variables that mainly capture the news (risk) related to future GDP growth?” and “Is news related to future GDP growth the systematic risk priced in the cross section of equity returns?”. If answers to the two questions are both “yes”, then Vassalou’s (2003) argument is applied. If both answers are “not yes”, then the conclusion is consistent with Nguyen et al. (2009). Interestingly, our empirical results suggest “not yes” to the first question and “yes” to the second. Thus, the Chinese evidence reconcile the contradictory conclusions of the two existing studies.

Specifically, to answer the first question, we run the formal individual regressions of 25 size-value sorted portfolio returns on four common pricing factors, i.e., MKT, SMB, HML and WML, with and without an additional GDP factor capturing news related to future GDP growth. The results indicate that the momentum factor is not helpful for explaining the 25 portfolio returns and that adding the nominal or real GDP factor can only marginally improves the performance of the standard Fama-French model. Thus, we do not find evidence positively supportive for the first question from the individual regressions.

The most interesting findings are from the results of GMM system regressions designed to answer the second question. The system includes the regression equations of pricing model with zero-intercept restriction imposed, as well as the mean-adjustment processes to each of the pricing factors in the model. This setup allows us to estimate the risk loadings and the risk premiums on factors simultaneously and is a superior over the two-stage Fama-MacBeth approach because it can correct for the generated-regressors problem when estimating the risk premiums. We use this setup to examine the CAPM, the Fama-French model and a four-factor model as well as their corresponding GDP augmented counterparts. The results of risk premium estimates provide hints for the answer to our second question. We find that the market, size and value factors are all systematic risks that priced in the cross section of equity returns with or without a GDP factor included, while both the momentum factor and the nominal GDP factor are not systematic. Interestingly, we find that the estimated premiums on the real GDP factor that captures news related to future real GDP growth is statistically significant. This suggests that the risk in real GDP growth is systematic risk priced in equity returns. Moreover, adding the real GDP factor significantly decreases the magnitudes of risk premiums on the market, size and value factors. This suggests that the market, size and value factors together contain some overlapping innovative information about future real GDP growth. Thus, our results here partially support Vassalou's (2003) argument.

A potential explanation for our above findings is given as follows. Because the GDP growth is an aggregate macroeconomic variable, it assembles different specific macroeconomic risks together. Thus, it may conceal the specific information related to certain macroeconomic variables. On the contrary, each priced factor may capture certain specific but currently unknown fundamental risk(s). As a result, adding the GDP factor cannot improve the standard pricing models in the individual regressions. On the other hand, based on the results of GMM system regression, the significantly positive risk premium on the real GDP factor, together with the decreased but still significantly positive risk premiums on the market, size and book-to-market factors, implies that the equity returns contain the aggregate information about the future real GDP growth. Thus, the news related to future real GDP growth is priced in the equity returns. This suggests that real GDP growth, to some extent, is the fundamental risk that drives the movement of equity prices, and also that the equity returns reflect the investors' concern about the uncertainty of the future economy, particularly, the risk in future real GDP growth.

The negative and insignificant risk premium on the nominal GDP factor suggests that news related to future nominal GDP growth is not priced in equity returns. We attributed this to the fact that the nominal GDP factor contains information about both real GDP growth and inflation. The effects of real GDP growth and inflation on equity returns are offsetting, thus, including inflation contaminates the information contained in the nominal GDP factor. As a result, the nominal GDP factor is not priced or systematic in equity returns.

Our findings in the Chinese market partially support the argument of Vassalou (2003) that the reasonably good performance of the Fama-French model is because SMB and HML capture the fundamental risk associated with future GDP growth, thus reconciling the previous findings in the US and Australian markets. This suggests that the information of fundamental macroeconomic risks captured by the priced factors could be country-specific. However, an appropriate explanation for the country-specific evidence is difficult to achieve. Because the Chinese evidence is right in the middle of the evidences from the US and Australian markets, the contradictory conclusions from the US and Australia make it difficult for an appropriate explanation. Therefore, the sources of fundamental risks that drive the movement of asset prices behind the priced factors are still not clear across different countries. This leaves an interesting topic for further investigating on this explanation. For future research, some other macroeconomic variables, for example, consumption, asset wealth and labor income of Lettau and Ludvigson (2001a), or financial variables, such as aggregate dividend yield and term spread, default spread, and Treasury-bill yield of Petkova (2006), are worthy of being examined. Whether the state variables that drive the asset prices are generally applied across countries or just country-specific is an important question to be answered. Whether a consistent economic interpretation for priced factors exists across different countries is also worthy of exploring. Once the specific state variables are figured out, a theoretical pricing model based on these variables would be more convincing.

Chapter 6

Concluding Remarks

This thesis conducts four separate but closely relevant empirical studies on asset pricing in the Chinese stock market. The Chinese stock market is a fast developing emerging market and deserves wide attentions from researchers for its increasing importance and specific features different from many other developed markets. Within the area of asset pricing, the factor pricing models are one of the focused branches for studying.

We first examine the performance of the CAPM, the Fama-French model and its extension that includes the additional momentum and liquidity factors in explaining the 25 size and book-to-market sorted portfolio returns in the Chinese Main-board stock market at the level of daily returns with a quarterly frequency for factor reforming. By comparing the empirical performance of different pricing models, the Fama-French model performs very well compared to the CAPM, even the factors are constructed with a higher than yearly frequency and the models are examined in the different market trends. Adding the additional factors seems to only marginally improve the explanatory power of the Fama-French three-factor model.

The traditional methods to construct the Fama-French factors are complex and time-consuming, which limits the examination and application of the Fama-French model in both academic research and financial practice. This motivates us to examine a simplified method to construct the factors by using the Chinese style indexes. The results support this simpler method as an alternative to the traditional methods. This simplified method makes it much easier to examine and apply the Fama-French model and, more importantly, makes it much cheaper and feasible to hedge risks related to the size and value effects based on the Fama-French model.

The last two chapters make efforts to examine whether GDP growth is the macroeconomic variable hidden behind the priced factors but driving the movement of asset prices in the context of the Chinese market. The first attempt is in Chapter 4 to examine whether the common pricing factors, i.e., the market, size, value and momentum factors, are correlated with current and future GDP growth in China. We find that only the momentum factor can largely capture and predict both current and future GDP growth, but the other three factors cannot. The second attempt is in Chapter 5 to examine whether the priced factors, that is, the market, size and value factors in the Chinese case here, proxy for the risk associated with future GDP growth as the state variable in the context of Merton's (1973) ICAPM, such that the Fama-French model can successfully explain the cross-sectional variation in the equity returns. The findings here and in the previous studies suggest that the information about macroeconomic risks captured by the priced factors, such that priced in the equity returns, is country-specific. Therefore, the sources of macroeconomic risks that fundamentally drive the movement of asset prices are still not clear. Thus, it leaves the following questions open for future research. For example, which specific macroeconomic variables the priced factors proxy for; whether those variables determine the success of factor pricing models in explaining equity returns; whether the variables are generally applied across countries or just country-specific; and what are the economic interpretations for the sources of fundamental risk that drive asset prices.

In the rest we briefly discuss several possible extensions of the studies of this thesis for future research. Although we have shown that the Fama-French three-factor model performs very well in explaining the cross-sectional variation in equity returns in the Chinese context, the portion left unexplained by the model cannot be neglected. Thus, a direct extension of this thesis is to examine the variants of the Fama-French three-factor model that includes additional factors in the Chinese stock market. Among the very recent literature of factor pricing models, Fama and French (2015a) add the profitability and investment factors into their three-factor model to make it a five-factor model.[1] Adding these two factors is inspired by the dividend discount model of Miller and Modigliani (1961). The model suggests that the book-to-market equity is a noisy proxy for expected return for the fact that the market equity also responds to forecasts of earnings (profitability) and investment. Many studies also find that much of the unexplained average returns by the Fama-French three-factor model is related to profitability and investment.[2]

In Fama and French (2015a), the profitability and investment factors are defined similarly

[1] Hou et al. (2015) is a paper closest to Fama and French (2015a), in which they examine a four-factor model that excludes the book-to-market (HML) factor. But they do not provide explanation for why HML is excluded in their model.

[2] For example, see Titman et al. (2004), Novy-Marx (2013) and Aharoni et al. (2013).

as the size and value factors. The profitability factor (RMW) is the difference between the returns of firms with robust and weak operating profitability; and the investment factor (CMA) is the difference between the returns of firms that invest conservatively and firms that invest aggressively. They find that the five-factor model performs better than the three-factor model, but it fails to capture the low average returns on portfolios of small size firms that invest a lot despite low profitability. Moreover, when the new factors are included in the model, the value factor (HML) seems to become redundant in the sense that the average return on HML is captured by the exposures of HML to other factors. But, they also warn that this result may be sample specific and more tests should be conducted by using different data in other periods or for other countries outside the US. As we find in the previous chapters that the value effect is generally weak compared to the size effect, this might be explained by the model of Miller and Modigliani (1961) in the sense that the book-to-market equity is a noisy proxy for expected return. Thus, the Chinese market provides an appropriate test field to investigate whether the profitability and investment factors would be better able to explain expected returns than the value factor as suggested by Miller and Modigliani (1961).

On the other hand, as we have discussed before, due to the specific features of the Chinese stock market which distinguish itself from other developed markets, the conclusions found from the developed markets could be inapplicable in the Chinese context.[3] Specifically, the profitability and investment factors, as well as the value factor, are firm-specific factors constructed based on the characteristics of company's fundamentals. These characteristics to some extent reflect the firm's long-term performance. In most of the developed markets, institutional investors are the dominant market players. For example, according to Celik and Isaksson (2013), institutional investors held around 60% of all publicly listed stocks in the US and 82% in Japan in 2011, and the portion of public equity held by institutional investors

[3]For example, Wang and Xu (2004) find that before the shareholder structure reform commenced in 2005, the value factor has no power in explaining the cross-sectional variation in the Chinese equity returns. Instead, they identify another factor based on the free float ratio and argue that their free-float factor should take the place of the value factor in the Fama-French three-factor model because it "serves as a signal of the quality of corporate governance in a Chinese company", thus, "is a better proxy for a Chinese company's growth potential and investment opportunity" than the value factor due to the specific features of the Chinese market.

was around 89% in the UK in 2012. The professionally-managed institutions are much better at analyzing fundamentals, focus more on long-term investment, and attempt to avoid being motivated by fear and irrational exuberance. Unlike those mature markets, most trades on the Chinese stock market are made by individual retail investors, rather than institutional investors. According to the China Securities Regulatory Commission (CSRC) 2014 Annual Report[4], the investors are classified into three types: professional institutions, natural persons and general institutions. The former are commonly called institutional investors that invest on behalf of other people or organizations, while the latter two are the commonly called retail investors who buy and sell securities for their personal account. As of the end of 2014, professional institutions held 14.22% of A shares by market capitalization, while natural persons and general institutions held 25.03% and 60.75%, respectively. In the same year, natural persons accounted for 85.37% of the annual turnover, while professional and general institutions accounted for only 12.15% and 2.37%, respectively.[5] The extremely high proportion of annual turnover triggered by the natural persons suggests that the majority of Chinese individual investors care less about the company's fundamentals and long-term prospects, seek for short-term gains, and trade speculatively and irrationally with strong herding behavior in buying winners and selling losers. The investment philosophy of the general public in the Chinese market makes it interesting to examine whether the profitability and investment factors are also helpful for explaining the portion of equity returns that cannot be explained by the Fama-French three-factor model in the Chinese context.

Although the Fama-French three-factor model can explain up to more than 90% cross-sectional variation in equity returns, a large body of literature documents that the Fama-French three-factor model performs poorly when we consider some return anomalies. Suggested by Lewellen et al. (2010), some prominent anomalies are accruals (Sloan, 1996), net share issues (Ikenberry et al., 1995, Loughran and Ritter, 1995), momentum (Jegadeesh and Titman, 1993), and volatility (Ang et al., 2006).[6] These are known to cause problems for the Fama-French

[4]Referred to the CSRC website at http://www.csrc.gov.cn/pub/newsite/zjhjs/zjhnb/.

[5]The Shanghai-Hong Kong Stock Connect accounted for the rest 0.11%.

[6]For a wider array of anomalies, see also Chan et al. (1996), Daniel and Titman (2006), Campbell et al.

three-factor model. Thus, a further extension of this thesis is to examine whether the new extensions of the Fama-French three-factor model improve the performance when we consider those anomalies. Hou et al. (2015) examine the performance of their four-factor model dropping the value factor (*HML*) in accounting for 6 categories of nearly 80 proxies for anomalies in the US: momentum, value-versus-growth, investment, profitability, intangibles and trading frictions. Fama and French (2016) contribute to the relevant literature by examining the Fama-French five-factor model when face with the anomalies not targeted by the model in the US market. Both papers find that the models including the market, size, investment and profitability factors outperform the Fama-French three-factor model in capturing many (but not all) of the anomalies. It could be of interest to find evidence in the Chinese market. Besides, when we consider variants of the Fama-French three-factor model and their performance in explaining equity returns or the return anomalies, it could be also interesting if we narrow down the sample period such that each period only includes one bull or one crash and examine and compare the performance of the models in different market trends. Both the US and Chinese data contain events like the 1929 US crash and recent tech bubble and financial crisis. More experience of equity markets will allow for more analysis.

As their earlier claim about the three-factor model that *SMB* and *HML* capture the risk associated with an unidentified state variable in Merton's (1973) Intertemporal CAPM, Fama and French (2015a) also suggest an interpretation for their five-factor model as another version of the Intertemporal CAPM in which the risk premiums that are not captured by the market factor are introduced by up to four unidentified state variables. In this sense, the size, value, profitability and investment factors are not state variable mimicking portfolios, but just diversified portfolios that capture the effects of state variables on expected returns without identifying the unknown state variables. Thus, the extension to the last task of this thesis, i.e., to find the fundamental macroeconomic risks that drive the movement of asset prices, could be to examine whether the Fama and French's (2015a) new claim is true by examining the possible macroeconomic variables using a similar framework of Chapter 5. Broadly speaking,

(2008), Cooper et al. (2008), Hafzalla et al. (2011), Novy-Marx (2013), and Aharoni et al. (2013).

the task is to find evidence supportive for the risk-based explanation behind the factors. For the most part, the task can be done by linking the factors to macroeconomic variables and business cycle fluctuations as in Liew and Vassalou (2000), Lettau and Ludvigson (2001a,b) and Vassalou (2003). On the other hand, as Campbell (1996) points out that empirical implementations of the ICAPM model should not rely on choosing important macroeconomic variables, we should also consider to relate the factors in the model to innovations in state variables that forecast future investment opportunities.[7] As a result, we should also be cautious of choosing the state variables before conducting the study.

[7] For example, Petkova (2006) does not choose the macroeconomic variables, such as GDP or consumption growth, as the state variables. Instead, he includes the short-term T-bill, term spread, aggregate dividend yield, and default spread into the set of relevant state variables.

References

Aharoni, G., Grundy, B., and Zeng, Q. (2013). Stock returns and the miller modigliani valuation formula: Revisiting the fama french analysis. *Journal of Financial Economics*, 110(2):347 – 357.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, 5(1):31 – 56.

Amihud, Y. and Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17(2):223 – 249.

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1):259–299.

Aylward, A. and Glen, J. (2000). Some international evidence on stock prices as leading indicators of economic activity. *Applied Financial Economics*, 10(1):1–14.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1):3 – 18.

Bernanke, B. S. and Blinder, A. S. (1992). The federal funds rate and the channels of monetary transmission. *The American Economic Review*, pages 901–921.

Black, F. (1972). Capital market equilibrium with restricted borrowing. *The Journal of Business*, 45(3):444–55.

Black, F. (1993). Beta and return. *The journal of portfolio management*, 20(1):8–18.

Breeden, D. T., Gibbons, M. R., and Litzenberger, R. H. (1989). Empirical tests of the consumption-oriented capm. *The Journal of Finance*, 44(2):231–262.

Bremer, M. and Sweeney, R. J. (1991). The reversal of large stock-price decreases. *The Journal of Finance*, 46(2):747–754.

Brennan, M. J., Chordia, T., and Subrahmanyam, A. (1998). Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics*, 49(3):345 – 373.

Brennan, M. J., Wang, A. W., and Xia, Y. (2004). Estimation and test of a simple model of intertemporal capital asset pricing. *The Journal of Finance*, 59(4):1743–1776.

Cakici, N., Chan, K., and Topyan, K. (2015). Cross-sectional stock return predictability in china. *The European Journal of Finance*, pages 1–25.

Campbell, J. Y. (1996). Understanding risk and return. *Journal of Political Economy*, 104(2):298–345.

Campbell, J. Y., Hilscher, J., and Szilagyi, J. (2008). In search of distress risk. *The Journal of Finance*, 63(6):2899–2939.

Capaul, C., Rowley, I., and Sharpe, W. F. (1993). International value and growth stock returns. *Financial Analysts Journal*, pages 27–36.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82.

Carpenter, J. N., Lu, F., and Whitelaw, R. F. (2015). The real value of china's stock market. Technical report, National Bureau of Economic Research.

Celik, S. and Isaksson, M. (2013). Institutional investors as owners: who are they and what do they do? *OECD Corporate Governance Working Papers*, (11):1.

Chan, L. K., Hamao, Y., and Lakonishok, J. (1991). Fundamentals and stock returns in japan. *The Journal of Finance*, 46(5):1739–1764.

Chan, L. K., Jegadeesh, N., and Lakonishok, J. (1996). Momentum strategies. *The Journal of Finance*, 51(5):1681–1713.

Chen, C., Hu, X., Shao, Y., and Wang, J. (2015). Fama-french in china: Size and value factors in chinese stock returns. *University of Hong Kong Working Paper*.

Chen, X., Kim, K. A., Yao, T., and Yu, T. (2010). On the predictability of chinese stock returns. *Pacific-Basin Finance Journal*, 18(4):403–425.

Chen, Z.-H. (2004). Cross-sectional variations and three factors asset pricing model: Empirical evidences from china a share market. *Chinese Journal of Management Science*, 12(6):12–17. in Chinese.

Cochrane, J. H. (1999). New facts in finance. Working Paper 7169, National Bureau of Economic Research.

Cochrane, J. H. (2001). *Asset Pricing*. Princeton University Press.

Cooper, M. J., Gulen, H., and Schill, M. J. (2008). Asset growth and the cross-section of stock returns. *The Journal of Finance*, 63(4):1609–1651.

Cremers, M., Petajisto, A., and Zitzewitz, E. (2012). Should benchmark indices have alpha? revisiting performance evaluation. Technical report, National Bureau of Economic Research.

Daniel, K. and Titman, S. (1997). Evidence on the characteristics of cross sectional variation in stock returns. *The Journal of Finance*, 52(1):1–33.

Daniel, K. and Titman, S. (2006). Market reactions to tangible and intangible information. *The Journal of Finance*, 61(4):1605–1643.

Daniel, K., Titman, S., and Wei, K. (2001). Explaining the cross-section of stock returns in japan: Factors or characteristics? *The Journal of Finance*, 56(2):743–766.

Datar, V. T., Naik, N. Y., and Radcliffe, R. (1998). Liquidity and stock returns: An alternative test. *Journal of Financial Markets*, 1(2):203 – 219.

Davis, J. L. (1994). The cross-section of realized stock returns: The pre-compustat evidence. *The Journal of Finance*, 49(5):1579–1593.

Davis, J. L., Fama, E. F., and French, K. R. (2000). Characteristics, covariances, and average returns: 1929 to 1997. *The Journal of Finance*, 55(1):389–406.

DeBondt, W. F. and Thaler, R. H. (1987). Further evidence on investor overreaction and stock market seasonality. *Journal of Finance*, pages 557–581.

Dimson, E. and Marsh, P. (1998). Murphy's law and market anomalies. *Available at SSRN 135681*.

Faff, R. W. (2001). An examination of the fama and french three-factor model using commercially available factors. *Australian Journal of Management*, 26(1):1–17.

Faff, R. W. (2003). Creating fama and french factors with style. *Financial Review*, 38(2):311–322.

Faff, R. W. (2004). A simple test of the fama and french model using daily data: Australian evidence. *Applied Financial Economics*, 14(2):83–92.

Fama, E. F. (1981). Stock returns, real activity, inflation, and money. *The American Economic Review*, 71(4):545–565.

Fama, E. F. and French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of Political Economy*, 96(2):246–273.

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Fama, E. F. and French, K. R. (1995). Size and book-to-market factors in earnings and returns. *The Journal of Finance*, 50(1):131–155.

Fama, E. F. and French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *The journal of Finance*, 51(1):55–84.

Fama, E. F. and French, K. R. (1998). Value versus growth: The international evidence. *The journal of Finance*, 53(6):1975–1999.

Fama, E. F. and French, K. R. (2012). Size, value, and momentum in international stock returns. *Journal of Financial Economics*, 105(3):457–472.

Fama, E. F. and French, K. R. (2015a). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1 – 22.

Fama, E. F. and French, K. R. (2015b). International tests of a five-factor asset pricing model. *Fama-Miller Working Paper*.

Fama, E. F. and French, K. R. (2016). Dissecting anomalies with a five-factor model. *Review of Financial Studies*, 29(1):69–103.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):pp. 607–636.

Ferson, W. E. and Foerster, S. R. (1994). Finite sample properties of the generalized method of moments in tests of conditional asset pricing models. *Journal of Financial Economics*, 36(1):29–55.

Ferson, W. E. and Harvey, C. R. (1994). Sources of risk and expected returns in global equity markets. *Journal of Banking & Finance*, 18(4):775–803.

Friedman, B. M. and Kuttner, K. N. (1992). Money, income, prices, and interest rates. *The American Economic Review*, pages 472–492.

Gibbons, M. R., Ross, S. A., and Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica: Journal of the Econometric Society*, pages 1121–1152.

Gompers, P. A. and Metrick, A. (1998). Institutional investors and equity prices. Working Paper 6723, National Bureau of Economic Research.

Gregory, A., Tharyan, R., and Christidis, A. (2009). The fama-french and momentum portfolios and factors in the uk. *University of Exeter Business School, Xfi Centre for Finance and Investment Paper*, (09/05).

Gregory, A., Tharyan, R., and Christidis, A. (2013). Constructing and testing alternative versions of the fama–french and carhart models in the uk. *Journal of Business Finance & Accounting*, 40(1-2):172–214.

Grinold, R. C. (1993). Is beta dead again? *Financial Analysts Journal*, 49(4):28–34.

Gustafson, K. E. and Miller, J. D. (1999). Where has the small-stock premium gone? *The Journal of Investing*, 8(3):45–53.

Hafzalla, N., Lundholm, R., and Van Winkle, E. M. (2011). Percent accruals. *The Accounting Review*, 86(1):209–236.

Hansen, B. E. (1992). Testing for parameter instability in linear models. *Journal of policy Modeling*, 14(4):517–533.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–1054.

He, J. and Ng, L. K. (1994). Economic forces, fundamental variables, and equity returns. *The Journal of Business*, 67(4):599–609.

Hodrick, R. J. and Zhang, X. (2001). Evaluating the specification errors of asset pricing models. *Journal of Financial Economics*, 62(2):327 – 376.

Hou, K., Xue, C., and Zhang, L. (2015). Digesting anomalies: An investment approach. *Review of Financial Studies*, 28(3):650–705.

Ikenberry, D., Lakonishok, J., and Vermaelen, T. (1995). Market underreaction to open market share repurchases. *Journal of Financial Economics*, 39(2):181–208.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1):65–91.

Kothari, S. P., Shanken, J., and Sloan, R. G. (1995). Another look at the cross-section of expected stock returns. *The Journal of Finance*, 50(1):185–224.

Lakonishok, J., Shleifer, A., and Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *The Journal of Finance*, 49(5):1541–1578.

Lamont, O. A. (2001). Economic tracking portfolios. *Journal of Econometrics*, 105(1):161 – 184.

Lettau, M. and Ludvigson, S. (2001a). Consumption, aggregate wealth, and expected stock returns. *The Journal of Finance*, 56(3):815–849.

Lettau, M. and Ludvigson, S. (2001b). Resurrecting the (c)capm: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy*, 109(6):1238–1287.

Lewellen, J., Nagel, S., and Shanken, J. (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial economics*, 96(2):175–194.

Liao, L. and Shen, H.-b. (2008). Fama-french three factors model and the effect of the split-share structure reform [j]. *The Journal of Quantitative & Technical Economics*, 9:117–125. in Chinese.

Liew, J. and Vassalou, M. (2000). Can book-to-market, size and momentum be risk factors that predict economic growth? *Journal of Financial Economics*, 57(2):221–245.

Lintner, J. (1965a). Security prices, risk, and maximal gains from diversification. *The Journal of Finance*, 20(4):587–615.

Lintner, J. (1965b). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, pages 13–37.

Liu, W. (2004). Liquidity premium and a two-factor model. *EFA 2004 Maastricht Meetings Paper No. 2678*.

Loughran, T. and Ritter, J. R. (1995). The new issues puzzle. *The Journal of Finance*, 50(1):23–51.

MacKinlay, A. (1987). On multivariate tests of the capm. *Journal of Financial Economics*, 18(2):341 – 371.

MacKinlay, A. C. (1995). Multifactor models do not explain deviations from the capm. *Journal of Financial Economics*, 38(1):3–28.

MacKinlay, A. C. and Richardson, M. P. (1991). Using generalized method of moments to test mean-variance efficiency. *The Journal of Finance*, 46(2):511–527.

Mao, X., Chen, M., and Yang, Y. (2008). Long-run return performance following listed rights issue: Based on the improved three-factor model. *Journal of Financial Research*, 5:013. in Chinese.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica*, 41(5):867–887.

Miller, M. H. and Modigliani, F. (1961). Dividend policy, growth, and the valuation of shares. *The Journal of Business*, 34(4):411–433.

Morck, R., Yeung, B., and Yu, W. (2000). The information content of stock markets: why do emerging markets have synchronous stock price movements? *Journal of Financial Economics*, 58(1-2):215 – 260. Special Issue on International Corporate Governance.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica: Journal of the Econometric Society*, pages 768–783.

Narayan, P. K. and Zheng, X. (2010). Market liquidity risk factor and financial market anomalies: Evidence from the chinese stock market. *Pacific-Basin Finance Journal*, 18(5):509–520.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–08.

Nguyen, A., Faff, R., and Gharghori, P. (2009). Are the fama-french factors proxying news related to gdp growth? the australian evidence. *Review of Quantitative Finance and Accounting*, 33(2):141–158.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1):1 – 28.

Nyblom, J. (1989). Testing for the constancy of parameters over time. *Journal of the American Statistical Association*, 84(405):223–230.

Pastor, L. and Stambaugh, R. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, 111(3):642–685.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies*, 22(1):435–480.

Petkova, R. (2006). Do the fama-french factors proxy for innovations in predictive variables? *The Journal of Finance*, 61(2):581–612.

Pham, V. T. L. (2007). Constructing fama-french factors from style indexes: Japanese evidence. *Economics Bulletin*, 7(7):1–10.

Rahim, R. A. and Nor, A. H. S. M. (2006). A comparison between fama and french model and liquidity-based three-factor models in predicting the portfolio returns. *Asian Academy of Management Journal of Accounting and Finance*, 2(2):43–60.

Rosenberg, B., Reid, K., and Lanstein, R. (1985). Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, 11(3):9–16.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3):341–360.

Ross, S. A. (1977). *Risk and return in finance*, chapter Return, risk and arbitrage. Ballinger, Cambridge, MA.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review*, 71(3):289–315.

Stock, J. H. and Watson, M. W. (1989). New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual*, 4:351–394.

Stock, J. H. and Watson, M. W. (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature*, 41(3):788–829.

Stock, J. H. and Watson, M. W. (2007). *Introduction To Econometrics*. Boston: Pearson/Addison Wesley.

Titman, S., Wei, K. J., and Xie, F. (2004). Capital investments and stock returns. *Journal of financial and Quantitative Analysis*, 39(04):677–700.

Treynor, J. L. (1961). Market value, time, and risk. *Unpublished manuscript*, pages 95–209.

Treynor, J. L. (1962). Toward a theory of market value of risky assets. *Unpublished manuscript. A final version was published in 1999, in Asset Pricing and Portfolio Performance: Models, Strategy and Performance Metrics. Robert A. Korajczyk (editor) London: Risk Books, pp. 15-22.*

Vassalou, M. (2003). News related to future gdp growth as a risk factor in equity returns. *Journal of Financial Economics*, 68(1):47 – 73.

Vassalou, M. and Xing, Y. (2004). Default risk in equity returns. *The Journal of Finance*, 59(2):831–868.

Wang, F. and Xu, Y. (2004). What determines chinese stock returns? *Financial Analysts Journal*, 60(6):65–77.

Wheelock, D. C. and Wohar, M. E. (2009). Can the term spread predict output growth and recessions? a survey of the literature. *Federal Reserve Bank of St. Louis Review*, 91(5 Part 1):419–440.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4):817–838.

Xu, J. and Zhang, S. (2014). The fama-french three factors in the chinese stock market. *China Accounting and Finance Review*, 16(2).