The Role of ETFs in Asset Pricing, Mutual Fund Performance, and Market Prediction

LI, JUNQI

The Role of ETFs in Asset Pricing, Mutual Fund Performance, and Market Prediction

Junqi Li

A thesis submitted in fulfillment of the requirements for the Degree of Doctor of Philosophy

Department of Economics and Finance

Durham Business School

Durham University

October 2017

Acknowledgements

I never expected that I will be a PhD student in a world top 100 university one day. Due to the financial distress, my father had no opportunities to receive a higher education. However, he never yields to any difficulty and becomes a manager by keeping learning and improving. My parents always provide me the best support whenever I need in the past four years.

I would like to say many thanks my supervisor team. I cannot make it today without their guidance, encouragement and support. Through my PhD study, Dr Anurag Banerjee is always there whenever I need any help. He teaches me so many research skills, such as econometrics, Matlab, Scientific Workplace, LaTex, with great patience. He spent thousands of hours on cultivating me and introducing me to his friends. He made me way much better than I was.

Dr Xing Wang is also an amazing supervisor. He helps me to obtain research data, develop research topics and establish models. He is always willing to help me anytime and provide me construct suggestions. When I was down, he always encourage and support to overcome difficulties. In addition, he shares so many his books, paper, and codes with me, which benefits a lot.

I would like to thank Daniel Hung for helping my first chapter. He is very innovative and inspires me always. He is strict but very supportive. Moreover, he is very keen to take me to academic conference and introduce me to other scholars. Without his help, my research would be much harder.

I am using this opportunity to express my gratitude to them throughout my PhD study. I will always keep learning and improving. Day as a teacher father for life. They will always have my respect. No matter where I will be and what I will do, I will try my best to make them pound for having me as their PhD student.

Abstract

This thesis investigates the various roles that the information provided by Exchange Traded Funds (ETFs) could play in asset pricing and market prediction. The empirical analysis contains three parts: The first part extracts information from the US ETFs market and constructs explanatory returns to price the Fama-French portfolios. It aims to provide a parsimonious model (the ETF-factor model) that is able to compete with the five-factor model of Fama and French (2015) and the q-factor model of Hou, Xue, and Zhang (2015). The second part applies the ETF-factor model, along with other conventional pricing models, to measure US equity fund performance. In addition, it attempts to develop relative pricing models as passive benchmarks for measuring US fixed-income fund performance by using information from bond ETFs. The purpose of the third part is to develop a new measure of Chinese investor behaviour that has predictive power for the Chinese market by using the information provided by respective ETFs.

The results suggest that ETFs deserve more attention in academic research. In line with conventional financial theory, ETFs' market dramatically increases the investment universe and securitizes illiquid assets. It comes as no surprise that the risk factors developed from ETFs have explanatory power for a cross-section of stock returns. In addition, a proxy for the bond market can be developed from bond ETFs. This avoids the subjective selection of the bond index as a passive benchmark and can provide a unique pricing model for bonds. Furthermore, research on ETFs contributes to the behavioral finance literature. Investor sentiment is a very important concept in behavioral finance. This thesis finds evidence that the investor behaviour that uses information from ETFs explains and predicts the Chinese market. In addition, it could lead to a profitable high-frequency trading strategy in actual trading.

Overall, this thesis researches ETFs from a new perspective. It does not view the ETFs as an investment vehicle, but consider ETFs as a type of fundamental asset in the economy. The findings of this thesis contribute to the literature of asset pricing, behavioural finance, and market prediction, and identifies new areas for future research.

Declaration & Statement of Copyright

The work described in this thesis is entirely my own, except where I have acknowledged help from, or collaboration with, a named person, or given reference to a published source or thesis.

The material contained in this thesis has not been submitted in support of an application for another degree or qualification in this or any other university.

Junqi Li (2017)

Contents

List of Tables

List of Figures

1. Introduction

1.1. Background

The Exchange Traded Funds (ETFs) are normally tradable securities that passively track the performance of stocks, bonds, and many other commodities in the US market. SPDR S&P 500 ETF was introduced in January 1993 and became the largest ETF in the world. Shortly afterwards, ETFs became very popular with institutional investors and financial advisors. The innovation in types of ETFs continues in the 21st century. The first bond ETF was introduced in 2002, and the first commodity ETF was introduced in 2004. In 2007, strong market demand produced a large growth in ETFs. At the end of the last financial crisis, the first actively managed ETF was introduced to the US market. At the end of 2010, assets managed by US ETFs reached $1 trillion. Currently, the ETFs market continues its strong growth, and $2.54 trillion in assets were being managed at the end of 2016 and $3.42 trillion in assets were being managed at the end of 2017.

The Chinese ETFs market has only gained attention in recent years, but it has experienced rapid growth. In October 2016, over $24 billion in assets were managed by ETFs in the Chinese market. Compared to the size of the US ETFs market, the Chinese ETFs market is relatively small. However, it is expected to become one of the major ETFs markets in the world. So far, the biggest Chinese ETF is the China 50 ETF that manages assets of approximately $4 billion.

Thanks to their unique characteristics, ETFs dramatically extend the investment universe. Investors can invest in non-tradable or illiquid assets more easily by buying respective ETFs at the cost of a trivial tracking error. Compared to conventional mutual funds, ETFs give investors more trading flexibility. Conventional mutual fund shares can only be traded at the end of the trading day, while ETFs can be traded during the day on the stock exchange. Moreover, ETF trading incurs lower cost than mutual fund trading. Investment strategies that were too expensive to implement have become affordable due to the innovation in ETFs. Thus, it become easier to invest on specific indexes, sectors, countries via ETFs.

Previous ETF studies often focused on ETF performance and compared ETFs with conventional mutual funds (see, e.g., Harper *et al.* (2006), Guedj and Huang (2009), Agapova (2011)). Gastineau

(2004) thinks ETFs and conventional index mutual funds are competitors and studied the lagged pre-tax performance of ETFs. Further, Ben-David *et al.* (2014) say that ETFs attract investors from mutual funds and have an impact on the price and liquidity of underlying securities. Recently, Meziani (2016) thoroughly introduced ETFs as investment vehicles.

With the dramatic growth of the ETF market, recent ETF studies pay more attention to the economic consequences of ETF trading. Ben-David *et al.* (2014) find that ETF trading increases the volatility of underlying securities. Similar results are provided by Madhavan and Sobczyk (2015) and Krause *et al.* (2014). Glosten *et al.* (2016) show that ETF trading increases the informational efficiency of underlying stocks. However, Israeli *et al.* (2017) argue that an increase in ETF ownership can lead to weaker price efficiency, which is mainly due to higher trading costs.

On the other hand, pricing anomalies are still at the center of asset pricing literature after the 2007-08 financial crisis (see, e.g., Fama and French (2008, 2016), Hou *et al.* (2015), Cederburg and O'Doherty (2015)). The mainstream asset pricing models are Sharpe's (1964) Capital Asset Pricing Model (CAPM), Lintner's addition to the CAPM (Sharpe-Lintner CAPM) (1965), and Fama and French's models (1993, 2015). The five-factor model of Fama and French (2015) is the latest version of the multi-factor pricing model, but it faces a challenge from Foye's (2017) finding that the empirical performance of this five-factor model is relatively poor in the UK market.

Compared to a multi-factor model, CAPM is based on the solid theory of return and risk. The first statement of Roll's (1977) critique says that CAPM is valid only if the market is mean-variance efficient considering all assets. However, the second statement of Roll's (1977) critique says that it is unrealistic to observe all investment opportunities. In other words, the theoretical market portfolio is unobservable. Previous studies considered investment opportunities in the stock market but failed to consider non-stock assets. As mentioned earlier about the features of ETFs, the ETF market not only considers stock investment opportunities but also other investments such as bond and commodity opportunities. According to Roll's (1977) critique, the ETF market is a better proxy for a theoretical market portfolio than the stock market.

However, the foundation of the efficient market hypothesis is questionable because of documented pricing anomalies, such as size premium, value premium, liquidity premium, which lead to the birth of behavioral finance. Some scholars accept the fact that market participants do not always behave rationally and study the psychology of investors. Market sentiment is a typical example of crowd psychology. Some recent studies have shown that pricing anomalies are related to investor sentiment (see, e.g., Baker and Wurgler (2006), Ho and Hung (2009), Stambaugh *et al.* (2012)). These studies show that investor sentiment contributes to the explanation of a cross-section of stock returns. Additionally, Huang *et al.* (2014) show that investor sentiment can predict stock returns.

Despite the rapid growth of studies on ETFs, there are very few scholars who extend ETF studies to fields such as asset pricing and market prediction. First, the ETF market provides additional information about underlying asset markets. Some previous studies[1] have investigated how stock ETFs affect the stock market, but no studies have ever used ETFs to price assets, such as stocks and bonds. Second, the ETF market provides more information about investor behavior. For example, an index ETF makes the underlying index tradable and liquid. Thus, it is possible to observe trading information, such as the bid/ask price and trading volume. Serious investors can perform a technical analysis of an index by using the trading information provided by the respective ETF.

1.2. Objectives

This thesis intends to extend the studies on ETFs to the fields of: 1). asset pricing, 2). mutual fund performance, 3). investor sentiment and market prediction. Despite many researchers' focus on these fields, very few previous studies have developed any asset pricing or investor sentiment model by using the information provided by the ETF market. Further, this thesis attempts to adopt a new pricing model as a mutual fund performance measurement technique and to use investor sentiment to predict the stock market.

[1] Krause *et al.* (2014) find that trading ETFs increase volatility in their component stocks, which leads to an increase in liquidity of stocks held by the ETF. Israeli *et al.* (2017) find that ETF trading could increase trading costs and decline analysts covering the stocks. Thus, ETF trading could result in less informative security prices for underlying stocks.

According to Lintner's (1965) CAPM, security returns, on average, are proportional to market return. However, in the 1980s, the explanatory power of market return faced a challenge from security characteristics, such as the earning-to-price ratio and market capitalization. By extracting unique information from various pricing anomalies, Fama and French (FF) find that size factor (SMB) and value factor (HML) provide the greatest extra explanatory power for a cross-section of stock returns (see, e.g., Jaffe *et al.* (1989), Fama and French (1992, 1993)).[2] Many studies have confirmed that the FF three-factor model is superior to CAPM.

Due to the lack of solid theoretical foundation, much research seeks a better understanding of size and value effects. Some studies argue that the size factor is the premium for bearing liquidity risk because, normally, the small stocks are less liquid than the big stocks. The best evidence is provided by Liu (2006). He proposes a two-factor model that includes the market and liquidity factors. This two-factor model is superior to the FF three-factor model in explaining a cross-section of stock returns. In addition, this model successfully describes anomalies related to size, book-to-market ratio, investment, and so on. On the other hand, some studies argue that the value factor is the premium for bearing financial distress risk. Fama and French (2015) find that adding the investment and profitability factors makes the value factor redundant. This finding is confirmed by the q-factor model of Hou *et al.* (2015).

After the latest financial crisis, the conventional pricing models faced challenges from various cross-sectional anomalies. Conventional models focus on the sources of risk that are restricted to the stock market. Alternative pricing models were developed to explain the new anomalies, but no consensus was achieved. According to Roll's (1977) critique, the theoretical market portfolio is unobservable. The proxy for the theoretical market portfolio cannot capture the systematic risk completely in empirical research. Thus, many researchers devote themselves to identifying other risk factors that capture the rest of the systematic risk.

---

[2] The Fama and French factor model is designed to describe stock returns in asset pricing and portfolio management. According to Fama and French (1992, 1993), the three factors are 1). Proxy for the market risk, 2). SMB (Small Minus Big), the average return on the three small portfolios minus the average return on the three big portfolios, 3). HML (High Minus Low), the average return on the two value portfolios minus the average return on the two growth portfolios.

One purpose of this thesis is to recognize the sources of risk, not from the stock market, but from the ETF market. According to Roll's (1977) critique, this approach is more reasonable than other approaches in current research. Another purpose of this thesis is to apply the new pricing model to measure the US mutual fund performance.

The final objective of this thesis is to show that information from the ETF market can contribute to actual trading. Much research has shown that investor sentiment affects a cross-section of stock returns.[3] Recent research shows that investor sentiment can also predict the stock market.[4] Due to the ETFs' stock-like feature, some specific ETFs can show the attitudes of investors toward a stock market. For example, the bid/ask volume of SPDR S&P 500 ETF trading reflects whether or not investors are willing to long or short the S&P 500 index.

1.3. Gap in the Literature

The area that has not been sufficiently explored in ETF literature is how to take advantage of useful information uncovered by the ETF market. In the exsiting ETF literature, most studies treat the ETFs as innovative investment vehicles and focus on comparing their performance with mutual fund performance. Compared to conventional mutual funds, ETFs become attractive because of lower trading costs and tax efficiency. But the most important benefit of an ETF is the stock-like features offered. Investors can carry out the same types of trades that they can with a stock. Therefore, the ETF market provides more useful information about stocks in the aspect of portfolio management. Unfortunately, very few ETF research capitalizes the useful information reflected by the ETF market.

This thesis attempts to apply the information in ETF market to other research fields. In the field of asset pricing, hundrends of risk factors have been proposed to describe the cross-section of stock returns (see, e.g., Harvey, Liu, and Zhu (2015), McLean and Pontiff (2016)). From a portfolio perspective, DeMiguel *et al.* (2017) find only a small number of firm-specific characteristics are

[3] For example, Baker and Wurgler (2006) have documented the impact of investor sentiment on the cross-section of stock returns.

[4] For example, Schmeling (2009) that sentiment negatively forecasts aggregate stock market returns on average across countries. Huang *et al*. (2015) show that investor sentiment is a powerful predictor of stock returns.

significant.[5] Characteristics are related to risk factors because the return of a long-short portfolio based on a characteristic can be used as a proxy for unkown risk factor. Their results are provide more empirical evidence to suport the five-factor model of Fama and French (2015). Because of using the information from the ETF market, this thesis contributes to the asset pricing literature by providing alternative proxies for unkown risk factors.

In the field of mutual fund performance, the most used measurement models are CAPM (Sharpe (1964), Lintner (1965)), the three-factor model of Fama and French (1993) and the four-factor model of Carhart (1997). However, Berk and Van Binsbergen (2015) argue that these models are not appropriate choice of benchmarks. Those mutual fund researchers interpret the factors in factor models as investment opportunities available to investors, but actually the long-short portfolios are not available to investors. If the risk factors are produced by using ETFs, this problem is solved. Investors, particularly retail investors, are able to take long or short positions in underlying assets by trading ETFs. This thesis contributes to the mutual fund performance literature by providing multi-factor models in which the factors are investment opportunities available to investors in practice.

In the fields of investor sentiment and market prediction, many researchers successfully predict the market by using different measures of investor sentiment such as the sentiment index of Baker and Wurgler (2006), the CFA Institute's Global Market Sentiment Survey and put-call ratio which is widely used in industry. Due to the ETFs' stock-like feature, it becomes possible to measure the investor sentiment by observing trading behaviour in the ETF market. Because of herding behaviour in Chinese market, this new measure of investor sentiment may predict the movement of stocks. Thus, this thesis contributes to the literature of investor sentiment and Chinese market prediction by constructing new measure of investor sentiment and providing new predictive models.

[5] DeMiguel *et al.* (2017) consider more than 50 firm-specific characteristics and find that only beta, unexpected quarterly earnings, return volatility, asset growth, 1-month momentum and gross profitability are significant. Moreover, they find that 12-month momentum and book-to-market are not significant.

1.4. Research Questions

Previous literature has provided numerous findings on ETFs, but the understanding of the ETF market is still poor. Due to their stock-like features, ETFs can be seen as traded indexes. The ETFs are able to provide much more information than other funds Thus, there are plenty of questions that are interesting to answer because previous studies always treat ETFs as new investment vehicles, not fundamental assets such as stocks and bonds. For example, is the ETF market a better candidate as the proxy for the theoretical market portfolio in CAPM? Does the ETF market produce new risk factors that describe the market anomalies? Is it possible to develop passive benchmarks from the ETF market to measure fixed-income fund performance? Can the stock market be predicted based on the information provided by specific ETFs?

The first two questions challenge the conventional approach adopted by previous asset pricing research. In conventional pricing models, the risk factors are developed from the stock market. In other words, previous research assumes that all sources of risk are from the stock market. Obviously, this assumption conflicts with Roll's (1977) critique. More investment opportunities in non-stock markets need to be considered. First, this study defines the value-weight ETFs market portfolio as the new proxy for the theoretical market portfolio, which provides a new market factor. Along with this new market factor, the size factor exhibits no significant explanatory power. Second, this study develops another risk factor that is defined as the difference between the return on a portfolio of commodity ETFs and the return on a portfolio of bond ETFs. Henriksen and Kværner (2017) show that long-run profitability is related to commodity price change. This indicates that anomalies related to profitability may be the compensation for commodity price change.

The second question focuses on an area that attracts very little attention. Few studies discuss the pricing models for fixed-income securities. In addition, there is no consensus on the passive benchmark with which to measure fixed-income fund performance. Fama and French (1993) propose the term and default factors to explain a cross-section of bond returns. Some research finds that the liquidity risk is priced in US corporate bond returns. This thesis aims to develop a more general pricing model that can be used to measure the performance of all types of fixed-income funds. Due to the existence of bond ETFs, it is reasonable to construct a bond market factor that plays the role of the market factor in CAPM. Combined with other factors discussed in previous research, it is possible to develop a pricing model that describes the abnormal performance of fixed-income funds. This empirical pricing model contributes to the literature of bond valuation.

The third question explores the use of the information provided by ETFs. The purpose of an ETF is to replicate the performance of an index. Because ETFs are traded on the stock exchange, the bid/ask volumes and prices listed in the order book are observable. The bid/ask volume reflects the investors' willingness to buy and sell. This study constructs a new measure of investor behaviour using the bid/ask volumes of respective ETFs. This measure of investor behaviour varies in range from -1 to 1. Similar to the put-call ratio, it reflects the overall mood of the market. Huang *et al.* (2014) construct a measure of investor sentiment and successfully use it to predict the aggregate stock market. Motivated by their research, this thesis aims to show the predictive power of this new measure of investor behaviour. This study is the first to construct a measure of investor behaviour by taking advantage of the ETF market. In addition, it predicts the stock market based on the measure of investor behaviour. Thus, this thesis provides more evidence to support the findings of Huang *et al.* (2014) and extends the field of ETF research.

1.5. Structure of the Thesis

Chapter 2 reviews the literature that is related to the research objectives of this thesis. This chapter starts by reviewing the price efficiency of ETFs, and ETFs' impact on underlying markets. Then it provides an overview of documented financial anomalies and relevant traditional asset pricing models. Further, it discusses the existence of skilled fund managers and summarizes the common risk factors in bonds. Finally, it briefly reviews the impact of investor sentiment on the markets and empirical evidence found in the Chinese market.

Chapter 3 proposes and tests a parsimonious asset pricing model (an ETF-factor model) with two factors comprised of ETFs. This chapter starts by reviewing the asset pricing literature and exploring the sources of risk in the ETF market. Then it shows that the risk factor, which is the difference between the return on a portfolio of commodity ETFs and the return on a portfolio of bond ETFs, is priced in a cross-section of stock returns after the 2007-08 financial crisis. In detail, this factor successfully describes the anomalies related to profitability, size, and liquidity. Moreover, this chapter constructs an alternative market factor defined as the excess return on the value-weight ETFs index. This alternative market factor describes the size and value premiums in the sample period. Finally, this chapter shows that the ETF-factor model beats the FF five-factor model (Fama and French, 2015) and the q-factor model (Hou *et al.* 2015) in the time period after the latest crisis.

Chapter 4 applies the conventional pricing models and ETF-factor model (discussed in Chapter 3) to the US equity mutual fund performance. It finds that the value-weight portfolio of US domestic equity funds performs close to the market portfolio. By comparing the results to those measured by conventional pricing models, it aims to demonstrate the validity of the ETF-factor model as a performance measurement technique. In addition, it develops relative pricing models as passive benchmarks for measuring fixed-income fund performance. The proxy for the bond market is the most important explanatory return, and the slope factor provides extra explanatory power. Finally, this chapter employs the simulation approach of Fama and French (2010) to provide more evidence.

Chapter 5 develops a new measure of investor behaviour for the Chinese stock market by using the information provided by the corresponding Chinese ETF. It shows that investor behaviour has explanatory power on the contemporaneous half-daily index return either in a linear model or non-parametric model. In addition, this chapter finds continuity of investor behaviour in the same day and propose a two-step prediction process. The in-sample and out-of-sample forecasting ability of the market sentiment indicate that this study can lead to a high-frequency trading strategy in actual trading.

Chapter 6 concludes the findings of this thesis and provides suggestions for further research.

2. Review of Literature

2.1. Introduction

This chapter reviews the literature of ETFs, asset pricing, mutual fund performance and investor sentiment. The primary purpose of this thesis is to study whether information from the ETF market helps to explain the cross-section of stock returns, measure the mutual fund performance and predict the stock market. Indeed, the literature of these fields is highly correlated. The well-constructed asset pricing models such as CAPM and the three-factor model of Fama and French (1993) are widely used by mutual fund researchers to measure fund performance. Further, some researchers add investor sentiment to existing asset pricing models to predict the stock returns (see, Baker and Wurgler (2006)).[6]

This chapter includes four sections. The first section provides the definition of ETFs, reviews the price efficiency of ETFs and briefly summarises the impact of ETFs on the stock market. The second section reviews several well documented financial anomalies, then present traditional asset pricing models that are closely related to those anomalies. The third section briefly reviews how mutual fund researchers apply the traditional asset pricing models to measure the equity mutual fund performance. In addition, this section present some widely used risk factors in bond market. The fourth section reviews the literature of investor sentiment, particularly in the Chinese stock market.

2.2. Exchange Traded Funds

The US Securities and Exchange Commission (SEC) defines ETFs as 'a type of exchange-traded investment product that must register with the SEC under the 1940 Act as either an open-end investment company (generally known as "funds") or a unit investment trust.' Initially, ETFs are all passively managed funds tracking the performance of specific US equity indexes. But now some newer ETFs also track indexes of fixed-income instruments, foreign securities, and commodities.

[6] Baker and Wurgler (2006) add a lagged sentiment proxy to the four-factor model of Carhart (1997) and perform predictive regressions. They find that the cross-section of future stock returns is conditional upon their measure of investor sentiment.

According to Pauline Skypala, US ETFs were managing $2.54 trillion in assets at the end of 2016 (Financial Times, Feb 2017). In addition, about 90% of funds flow into passively managed products.

Unlike mutual funds, ETFs are traded on stock exchanges and at market prices. Thanks to liquidity, many ETFs are traded at market prices close to NAV (net asset value), rather than at discounts or premiums. An ETF's NAV is the residual of the ETF's assets minus its liabilities. There are several reasons why ETFs are so popular: first, ETFs provide much higher liquidity compared to open-ended mutual funds; second, the transaction cost is relatively very low; and, third, ETFs cover nearly all the areas in the market that are also covered by traditional open-ended funds, which means ETFs provide an extended investment universe compared to the stock market.

According to Vanguard, there are five ETF structures: open-end funds, unit investment trusts, grantor trusts, exchange-traded notes, and partnerships. Each type of structure has its own pros and cons. Most ETFs are structured as open-end funds, which aim to provide exposure to stock and bond asset classes. Some ETFs that track broad asset classes are structured as unit investment trusts. For example, commodity ETFs are typically structured as trusts. Those ETFs investing solely in physical commodities or currencies are normally structured as grantor trusts. Exchange-traded notes are prepaid forward contracts, so they are excluded from ETFs in this thesis. Very few ETFs are partnership ETFs, but these can accommodate a wide range of types of investments.

Currently, all ETFs are subject to the Securities Act of 1933 (1933 Act) and the Securities Exchange Act of 1934 (1934 Act). In addition, ETFs are structured as open-end funds or unit investment trusts and are also ruled by the 1940 Act. ETFs are regulated like conventional mutual funds, but some exemptions are widely supported by the industry because of ETFs' unique structural and operational features. In March 2008, the SEC unanimously proposed two new rules under the Investment Company Act of 1940: Rules 6c-11 and 12d1-4. Fuller and Rosenberger (2008) claim that ' new proposed Rule 6c-11 permits certain index-based and actively managed ETFs that have fully transparent portfolios to launch without obtaining individualized exemptive orders from the SEC, as is currently required, essentially codifying the relief in these orders.' In other words, unnecessary regulatory burdens are to be removed and innovations are encouraged. In addition, the proposed Rule 12d1-4 allows significant investments (mutual funds and other types of investment companies) in ETFs to a greater extent. However, the SEC failed to adopt Rule 6c-11 because of concerns about leveraged and inverse ETFs. Thomas Conner (2009) says one concern could be the use of derivatives by those ETFs, which deeply concerns the government. But, generally, the assets under active ETFs are trivial compared to those managed by the passive ETFs.

2.2.1. Pricing Efficiency

Many studies have investigated pricing efficiency of ETFs in the US ETF market. The pricing efficiency of ETFs is that market price of an ETF should be kept in line with the fund's net asset values through the fund's creation and redemption process. Chu and Hsieh (2002) find that the SPDRs, ETFs that track the performance of the S&P 500 index, are priced efficiently. If there is the presence of mispricing signals, the market corrects the prices immediately. Ackert and Tian (2008) find that SPDRs are traded at a very small discount or premium. Further, Engle and Sarkar (2006) state that domestic ETFs are priced efficiently in the market. If there are premiums or discounts, they are generally small and corrected in the next several minutes. The international ETFs may be traded at relatively larger premiums or discounts. But compared to closed-end funds, ETFs have much smaller, and less persistent, premiums or discounts. On the other hand, Marshall *et al.* (2013) and Petajisto (2017) show that there exist arbitrary opportunities in the ETF market, indicating mispricing remain.

Some research is also conducted to examine price efficiency of ETFs in emerging markets. Jiang *et al.* (2010) find that the price and net asset value (NAV) of the Shanghai 50 ETF (SSE 50) are co-integrated, generally, but that the ETF price may deviate from the NAV if the Chinese market becomes volatile. The empirical results of Lin *et al.* (2006) indicate that the first Taiwanese ETF, the Taiwan Top 50 Tracker Fund (TTT), is price efficient and produces almost the same returns to the underlying index. Charteris (2013) focuses on seven South African ETFs and finds that five are traded at premiums or discounts, but that mispricing will be corrected in no more than two trading days. In general, compared to other index trackers, ETFs are relatively efficiently priced in the market.

2.2.2. Impact on the Stock Market

Recent studies suggest that ETFs play an important role in the price discovery process of the underlying indexes or stocks (see, e.g., Chen and Strother (2008), Ivanov *et al.* (2013), Glosten *et al.* (2016)). Chen and Strother (2008) show that investment of ETFs helps determine the price of the underlying assets in the Chinese market through the interaction of buyers and sellers. Ivanov *et al.* (2013) achieves the same conclusion in the US market by analysing the data of S&P 500 ETFs, Russell 2000 ETFs etc.

Glosten *et al.* (2016) find that ETF trading increases price efficiency for underlying securities and pushes the prices to reflect more systematic fundamental information. However, Israeli *et al.* (2017) argue that ETF ownership can increase transaction costs and decrease price efficiency. Increasing ETF ownership means decreasing traders of the underlying securities in the market, and, thus, probably fewer analysts will be willing to cover these securities.

Moreover, some recent research has started to investigate the impact of ETF trading on the volatility of underlying securities. Bradley and Litan (2010) show that ETF trading poses systematic risks when the market declines. Ben-David *et al.* (2014) find that ETF ownership has a strong and positive impact on the volatility of underlying securities because noise traders cause mispricing via ETF trading. Krause *et al.* (2014) discovered volatility spill-overs from ETFs to underlying stocks. The level of volatility spill-overs is related to ETF liquidity and the weight of component stocks. A recent study by Bhattacharya and O'Hara (2016) agrees that ETF trading can increase the volatility of underlying security prices and lead to persistent distortions from fundamentals at the individual asset level. However, they show that ETF trading helps underlying prices converge to intrinsic values at the aggregate level.

Finally, ETF trading has an inevitable impact on the liquidity of underlying securities. Hamm (2014) finds that ETFs reduce the liquidity of individual stocks because uninformed investors exit the stock market in favour of ETFs. Agarwal *et al.* (2016) find that stocks with higher ETF ownership have greater commonality in liquidity. ETF ownership pushes the co-movement in the liquidity of underlying stocks in the basket. As ETFs continue gaining higher ownership of stocks, more studies on how ETF trading affects the underlying markets are needed.

2.3. Asset Pricing

One goal of this thesis is to propose a new asset pricing model based on information from the ETF market. In this section, a brief overview of the financial anomalies that this thesis attempts to describe is provided, and a concise review of asset pricing models is given. An anomaly often occurs with respect to asset pricing models, particularly in the CAPM. The existence of anomalies is often interpreted as that securities or security portfolios perform inconsistently to the notion of efficient markets. However, this argument may be inappropriate because the anomalies may be the consequence of an incorrect equilibrium model. For example,

Previous studies have identified many anomalies such as size and value effects.[7] The number of documented anomalies is still growing. This thesis focuses just on the cross-sectional return patterns. According to the CAPM, if the market is efficient and the pricing model is correct, the average security returns should conform to a linear relationship. In the late 1970s, anomalies related to firm characteristics, such as earning-to-price ratio and market capitalization, questioned the explanatory power of the market factor. The following section summarizes the major anomalies and presents relevant asset pricing models.

2.3.1. Financial Anomalies

The size effect is one of the first anomalies to be uncovered by previous studies. Banz (1981) was the first to find that the small firms listed on the New York Stock Exchange (NYSE) tend to produce higher average returns. A recent study by Van Dijk (2011) finds that the size premium in the US stock market is positive and large over the last decade. The size effect has been documented in most major stock markets around the world.

[7] The size effect is the negative relation between the expected return and the market value of the stock. Banz (1981) was the first to document this phenomenon for U.S. stocks. The value effect refers to the positive relation between the expected return and the book-to-market ratio (a ratio that compares the book value of a firm to its market value).

The value effect is another common anomaly that has been widely discussed. Basu (1977) finds that firms with high earning-to-price ratios tend to produce positive alphas.[8] Further, many studies document the significant relationship between a cross-section of stock returns and the book-to-price ratio (see, e.g., Rosenberg *et al.* (1985), De Bondt and Thaler (1987)). The value effect is also well documented in previous and recent studies.

The momentum effect means that the stocks that outperform peers during the last 3 to 12 months will continue to do so in the near future (Jegadeesh & Titman 1993).[9] Moskowitz and Grinblatt (1999) confirm the momentum effect at the level of industry. Griffin *et al.* (2005) show that the momentum effect exists in a world-wide range.

The liquidity effect says that investors require a liquidity premium in a cross-section of stock returns. Amihud and Mendelson (1986, 1989) show that illiquid stocks produce significantly higher returns and use bid-ask spreads to explain returns of stocks listed on the NYSE. The liquidity effect has been confirmed by many asset pricing studies (see, e.g., Brennan and Subrahmanyam (1996), Amihud (2002), Acharya and Pedersen (2005)).

The investment effect has received much attention recently.[10] Thomas and Zhang (2002) show that firms with low inventory growth rates perform better than those with high inventory growth rates. Titman *et al.* (2004) find that stocks with higher capital investments tend to produce negative alphas. Some recent research has shown the relationship between the investment and book-to-market ratio.

The profitability effect says that there is a link between the stock return and profitability. It has been discussed by Haugen and Baker (1996) and is confirmed by some recent papers (see, e.g., Li *et al.* (2009), Balakrishnan *et al.* (2010)). Similar to the investment effect, profitability has been shown to be related to the book-to-market value.

[8] In asset pricing literature, alpha typically means the intercept in an asset pricing model.
[9] The momentum effect refers the tendency for rising asset prices to rise further, and falling prices to keep falling.
[10] The investment effect refers to the negative relation between the expected return and the firm's investment.

2.3.2. Traditional Asset Pricing Models

A large number of documented pricing anomalies (size anomaly, book-to-market anomaly etc.) at the firm level indicate that CAPM fails the portfolio-link tests. Much research has been done to develop a multi-factor model that describe most documented anomalies. This section reviews the traditional asset pricing models employed in this thesis and presents some empirical questions that remain unresolved.

The Sharpe-Lintner CAPM (Sharpe (1964), Lintner (1965)) says that the market portfolio is the tangency portfolio and the risk premium of an asset is determined by the market risk premium and beta.[11] The early time-series tests of Gibbons (1982) and Stambaugh (1982) show that the market beta can explain the difference in average returns across securities. Several widely aceepted models are presented as following:

1). The Fama and French three-factor model (Fama & French (1993, 1996)) is a multifactor model that includes the market, size, and value factors. Fama and French find that the average size and value factors are positive, indicating that investors require premium from the exposures to small firms and firms with a high book-to-market ratio. A number of studies have provided evidence that the Fama and French three-factor model improves the asset pricing ability of the CAPM.

2). The Carhart four-factor model (Carhart 1997) is an extension of the Fama and French three-factor model that includes a momentum factor. The motivation for inclusion of momentum strategies is that investors incorporate the 52-week price high/low in their investment decisions. The momentum factor is a self-financing portfolio that takes long positions on winners and negative positions on losers. After the momentum anomaly was documented, Jegadeesh and Titman (2001) also found that it is persisted throughout the 1990s, which indicates that the momentum anomaly is not from temporary mispricing.

3). The Pástor and Stambaugh four-factor model (Pástor & Stambaugh 2003) is an extension of the Fama and French three-factor model through including a liquidity factor. This liquidity factor is

[11] In asset pricing literature, beta usually refers to the estimated coefficient on the market factor.

a measure of market-wide liquidity. They find that the liquidity risk describes half of the returns generated by a momentum strategy.

4). The Fama and French five-factor model (Fama & French 2015) is an extension of their three-factor model that includes investment and profitability factors. In this model, the value factor becomes redundant because it is explained completely by the other four factors. Although this model fails to describe the portfolios made up of small firms that have low profitability but are aggressive in investment, it improves the explanatory power of a cross-section of stock returns relative to the three-factor model.

5). The q-factor model (Hou *et al.* 2015) includes the market factor, a size factor, an investment factor, and a profitability factor. Their empirical results confirm that the value factor is redundant when using the investment and profitability factors. Based on the GRS tests[12], the q-factor model is superior for capturing anomalies relative to the Fama and French three-factor model and the Carhart four-factor model.

However, the biggest challenge these models are facing is that the risk factors, apart from the market factor, are related to specific firm-level anomalies. As Cederburg and O'Doherty (2015) point out, the anomaly-based evidence against the CAPM is exaggerated because most firm-level anomalies are related to each other and do not provide unique information. One example is that the value factor is completely explained by the investment and profitability factors (Fama & French 2015). In addition, those anomalies due to market mispricing probably weaken or disappear after they are first uncovered. For example, some research finds that the size effect appears to diminish after the success of Fama and French's three-factor model (see, e.g., Horowitz *et al.* (2000), Schwert (2003), Hou and Van Dijk (2014)). Thus, it may be not wise to construct the risk factors under the guidance of documented anomalies. There is most likely an alternative approach for developing asset pricing models.

[12] The GRS test is a statistical test of the hypothesis that a group of α are jointly indistinguishable from zero (Gibbons, Ross, and Shanken, 1989). Equivalently, it is a test that some linear combination of the factor portfolios is mean-variance efficient.

In summary, the traditional asset pricing models are constructed guided by the documented pricing anomalies. Empirical asset pricing researchers aim to eliminate the appearance of pricing errors in practice. One important application of the pricing models is the evaluation of mutual fund performance. In the next section, key literature on mutual fund performance will be reviewed and some common risk factors in bond markets will be presented.

2.4. Mutual Fund Performance

Another goal of this thesis is to investigate the existence of skilled managers in the US mutual fund industry after the 2007-08 financial crisis. Evaluation of mutual fund performance is a very important application of the asset pricing models. On the other hand, some widely used benchmarks for mutual funds enrich the literature of asset pricing. The validity of an asset pricing model will be enhanced if it is widely accepted and used by the mutual fund researchers. For example, Carhart (1997) add the momentum factor to the three-factor pricing model (Fama and French, 1993) and investigate the persistence of fund manager skill. Due to its success in measuring mutual fund performance, the Carhart four-factor model has become an important multi-factor asset pricing model. The following section will review several pricing models that are widely used as passive benchmarks for equity mutual funds. Besides, several documented common risk factors in bond markets will be discussed because they can help measure the fixed-income fund performance.

2.4.1. Equity Mutual Fund Performance

Many early studies find that, in aggregate, the mutual fund industry does not outperform its benchmarks (see, e.g., Jensen (1968), Carhart (1997), Gruber (1996), Daniel *et al.* (1997)). The performance measurement techniques in the academic literature include the CAPM, Fama and French three-factor model, and Carhart four-factor model. However, the recent paper by Barber *et al.* (2016) found that the CAPM is the best performance measurement model because most investors are unaware of other risk premiums and only sophisticated investors consider exposure to other risk factors. Berk and Van Binsbergen (2016) confirm this conclusion by showing that investors tend to use CAPM even although the CAPM does not explain cross-sectional variation well.

Apart from focusing on the alphas, some recent research develops new techniques to identify skilled managers.[13] Barras *et al.* (2010) developed the "False Discovery Rate" approach to estimating the proportion of skilled managers. They found that, in the past 20 years, the skilled managers tend to disappear, while the unskilled managers increase substantially. Cuthbertson *et al.* (2008) applied bootstrap simulations to distinguish the skilled or lucky managers in the U.K. equity fund industry. They found that only very few outperforming managers have stock picking abilities, while most underperforming managers lack such skills. In other words, superior performance is probably due to luck, and inferior performance is due to poor skill. A similar conclusion is reached by Fama and French (2008) when they apply bootstrap simulations to infer the existence of skilled manager from actively managed US equity mutual funds. These recent studies all conclude that the skilled fund managers (not the lucky fund managers) are diminishing.

2.4.2. Common Risk Factors in the Bond Market

Apart from focusing on ETF performance, this thesis also measures fixed-income mutual fund performance. Although the bond market is very important, previous research has paid little attention to bond pricing models. Therefore, the consensus is yet to be reached regarding this matter. Litterman and Scheinkman (1991) find that most of the variation in bond returns can be explained by attributes of the yield curve: level, steepness, and curvature. Fama and French (1992) propose two common factors in the bond market: the term and default factors. Elton *et al.* (1995) employ an aggregate bond index, an aggregate stock index, two indices capturing default and option effects, and two macro variables that incorporate the unexpected changes in inflation and real GNP in the model.

---

[13] In literature of mutual fund performance, alpha normally refers to the item of intercept in the measurement model. The significant alpha indicates the superior or inferior kill of the fund manager.

Recent studies continue investigating the common risk factors in the bond market, particularly in the corporate bond market. The default premium is priced in corporate bonds (see, e.g., Driessen (2004), Berndt *et al.* (2005), Longstaff *et al.* (2005)). In addition, some research shows that liquidity risk is an important determinant of expected corporate bond returns (see, e.g., Lin *et al.* (2011), De Jong and Driessen (2012), Acharya *et al.* (2013)). Meanwhile, these studies also confirm the validity of term and default factors in corporate bonds. Further, Hsu *et al.* (2015) find that the default risk is related to a firm's innovation, which plays a role in bond pricing. In summary, most previous research agrees that the term, default, and liquidity risks are systematic risks in the bond market, but have yet to reach a consensus on other risks.

In summary, evaluation of mutual fund performance is an important application of asset pricing models. The existing literature of equity mutual funds employ a number of pricing models such as the three-factor model of Fama and French (1993) and the four-factor model of Carhart (1997). Among the literature of fixed-income funds, the default risk factor and the term factor are widely used to measure the fund performance. The next section will review the literature of investor sentiment, which is an important application of behavioural finance in asset pricing research.

2.5. Investor Sentiment

A growing body of research shows that some pricing anomalies cannot be fully described by the traditional asset pricing theories (reviewed in section 2.3 and 2.4). However, researchers find that some anomalies can be explained by the theories of behaviour finance. Behavioural finance is a field of finance that proposes psychology-based theories to investigate the phenomenon in stock market. Richard Thaler, who is awarded the Nobel Memorial Prize in Economic Sciences for his contributions to behavioural economics, believes that investors will capitalize on cognitive biases such as the endowment effect and loss aversion.[14] The connection between the behaviour finance and the asset pricing research is through the concept 'investor sentiment (investor

[14] Cognitive biases are systematic patterns of deviation from norm or rationality in judgment, and are often studied in psychology and behavioural economics. Endowment effect refers the hypothesis that people tend to overestimate things merely because they own them. Loss aversion refers to people's tendency to prefer avoiding losses to acquiring equivalent gains.

attention)'. A number of literature connects the changes of investor sentiment on financial markets to fundamental principles of asset pricing and market prediction (see, e.g., Barberis *et al.* (1998), Barberis and Thaler (2003), and Baker & Wurgler (2007)). The following section reviews the literature on measurements of investor sentiment in US and Chinese markets.

2.5.1. Measurements and Impacts in the Stock Market

Market sentiment, typically explained by crowd psychology, reflects investors' opinions and emotions concerning the market, as a whole. Investor sentiment is an approach to measuring the market sentiment. There are various methods of measuring investor sentiment. Brown and Cliff (2005) use a direct survey measure of investor sentiment. Some groups, such as the CFA Institute's Global Market Sentiment Survey, conduct these surveys to proxy the investor sentiment. Instead of surveying investor sentiment, Baker and Wurgler's (2006) develop a sentiment index using variables such as size, profitability, volatility, etc. In industry, financial analysts prefer to use the put-call ratio to measure market sentiment. The put-call ratio is calculated by dividing the number of traded put options by the number of traded call options.

Much recent research shows that investor sentiment has a significant impact on the stock market. Brown and Cliff (2005) show that investor sentiment is one of the determinants of asset values. Their measure of investor sentiment has a forecasting ability for market returns over several years. Other studies also find that their measures of investor sentiment can predict stock returns (see, e.g., Lemmon and Portniaguina (2006), Kaplanski *et al.* (2015), Massa and Yadav (2015)). In particular, Baker and Wurgler (2006) show that the cross-section of future stock returns is conditional upon the current measure of investor sentiment. Small, volatile, and unprofitable stocks, especially, generate relatively lower (higher) future returns when investor sentiment is high (low). Further, Huang *et al.* (2015) develop a new investor sentiment index and show its strong in-sample and out-of-sample forecasting ability on the aggregate stock market. In short, numerous recent studies show that investor sentiment plays an important role in determining and predicting stock returns. The above literature is about the US market. As the Chinese market is also a very important stock market globally, the next section will briefly review the relevant literature.

2.5.2. Investor Sentiment in the Chinese Stock Market

Although there is ample previous research that investigates the relationship between investor sentiment and stock returns in the US market, researchers have only begun to look at the Chinese market seriously. The Chinese stock market is relatively young compared to the US stock market. Thus, Chinese investors, particularly individual investors, are less rational compared to their counterparts invested in the US stock market. Chen *et al.* (2003) and Wang *et al.* (2006) state that most Chinese investors lack investment knowledge and skills. They exhibit behavioural biases and more easily make trading mistakes. Further, Li *et al.* (2016) find that, in the Chinese stock market, individual investors are heavily influenced by market sentiment and important events, while institutional investors are less influenced.

Some recent studies have found empirical evidence that investor sentiment plays an important role in the Chinese stock market. Burdekin and Redfern (2009) find that stock market sentiment has a negative impact on Chinese savings account growth from 2003 to 2007. Chi *et al.* (2012) use mutual fund flows to measure investor sentiment for different stocks and reveal the tremendous impact of investor sentiment on stock returns in the Chinese market. Ni *et al.* (2015) show that the nonlinearity of investor sentiment on monthly stock returns in the Chinese A-share stock market is significant. The effect is positive and strong in the short term, then it becomes negative and weak in the long term. Nguyen and Bhatti (2015) find the existence of a negative idiosyncratic volatility effect in the Chinese stock market. They suggest investor sentiment as a source of risk in the Chinese stock market.

2.6. Conclusion

Now it is time to discuss the relations among the four sections of literature review. In finance, asset pricing models are used to determine a theoretically appropriate expected return of an asset. CAPM is the first single factor model that describes the relationship between systematic risk and expected return for assets. However, those documented financial anomalies in section 2.3.1 demonstrate that CAPM fails in practice. To describe the documented financial anomalies, many researchers add other risk factors to the CAPM and develop multi-factor asset pricing models. The measurement of mutual fund performance is an important application of these asset pricing models. The most widely used pricing models in mutual fund industry are CAPM, Fama and French three-factor model and Carhart four-factor model. Several other risk factors discussed in asset pricing literature, such as default risk factor, are used to measure the bond mutual fund performance. Investor sentiment is another important concept in finance literature. As the literature in section 2.5 shows, adding lagged proxy for investor sentiment to asset pricing model helps to predict the market movement.

The literature of ETFs is quite independent from the literature of asset pricing, mutual fund performance and investor sentiment. The most important contribution of this thesis is to build up the bridge between the ETF research and other research fields. The rest of the thesis is organized as follows: chapter 3 construct ETF-factors using data of ETFs and proposes a new asset pricing model; chapter 4 employs this new asset pricing model to measure the mutual fund performance; chapter 5 constructs a new measure of investor sentiment using data of ETFs and proposes a predictive model.

3. A Parsimonious Asset Pricing Model with Exchange Traded Funds

3.1. Introduction

Financial economists doing asset pricing research face several challenges such as newly documented anomalies and poor empirical practice. From the traditional asset pricing approaches, an asset in a complete market is priced relative to a finite number of random states of the world. The price of an asset can be rewritten as linear combinations of the states. Unfortunately, the experience of the era after the 2007 has indicated that the effectiveness of these models in pricing stocks is very poor. For example, conditional CAPM alphas are related to 9 asset pricing anomalies: size, book-to-market, momentum, reversal, profitability, asset growth, net stock issues, accruals, and financial distress (Dittmar and Lundblad, 2017). To improve the empirical performance, the newly anomalies related to the profitability and investment of firms are further constructed as explainable variables in the pricing model (Fama and French, 2015).

Foremost of all is the identification of an appropriate empirical measure of the unobservable theoretical construction of the market portfolio in the CAPM that should capture a variety of wealth capital, including non-tradable or illiquid assets (Roll (1977)) and human capital (Jagannathan & Wang 1996). Recent empirical research proposes several pricing factors that capture the dimensions of risks other than the market, including the size and book-to-market factors (see, Fama and French (2006)), the liquidity risk factor (see, e.g., Amihud and Mendelson (1986), Pástor and Stambaugh (2003), Acharya and Pedersen (2005), Liu (2006)), the investment and profitability factors (see, e.g., Fama and French (2015)), and the q-factor model (Hou *et al.* 2015).[15]

In this chapter, this study proposes and tests a parsimonious asset pricing model with two factors comprised of ETFs. To distinguish the model from others, it will be known as the ETF-factor model. One of the motivations for this model’s development is Roll's critique (1977) that the true

[15] Ironically, however, many more factors are being advocated that attempt to explain a cross-section of stock returns (Harvey, Liu, and Zhu (2016)). The evidence of pricing anomalies presents yet another major hurdle for asset pricing models (see, e.g., Fama and French (2008) for a summary). Fama and French (2015), for example, show that their five-factor model fails to explain the returns on small stocks with high investment, but low profitability.

market portfolio should include every asset, including commodities, collectibles, and anything with any worth. CAPM only validates if the market is mean-variance efficient considering all investment opportunities. By now, it is clear that there are always anomalies because the true market portfolio is unobservable. Therefore, ignorance of other asset classes, such as bonds, commodities, and gold, won't lead to an appropriate CAPM. The other motivation is the evolution and growth of the ETF industry. This growth is forecast to continue. Robin Wigglesworth says, 'ETFs now account for about 30 percent of all US trading by value, and 23 percent by share volume' (Financial Times, 2017). The shift from actively managed mutual funds to passive funds, such as ETFs, continues. In 2015, less than 1% of over $2 trillion in assets managed under US ETFs were tied to actively managed ETFs. In other words, most ETFs replicated and continue to replicate indexes. Moreover, ETFs cover many areas of the market, including stock indexes, stock market sectors, commodities, currencies, bonds, and so on.

There are several advantages to using ETFs in asset pricing. First, the ETF market extends the investment universe relative to the stock market by including more investment opportunities, which enhances the mean-variance efficiency. Second, ETFs increase the commonality in the liquidity of underlying stocks (Agarwal *et al.* 2016). Third, risk factors are easier to construct via the ETF market, and ETF portfolios are rebalanced more frequently than Fama and French (FF) portfolios. In the ETF-factor model, the expected return of an asset is described by two ETF-factors: the ETFs market excess return (ETF_V) and the difference between the return on a portfolio of commodity ETFs and the return on a portfolio of bond ETFs (CMB_V):

$$\mathrm{E}(R_i) - R_f = b_i\mathrm{E}(ETF_V) + c_i\mathrm{E}(CMB_V) \tag{3.1}$$

where $\mathrm{E}(R_i)$ is expected return on asset i, R_f is the risk-free rate, $\mathrm{E}(ETF_V)$and $\mathrm{E}(CMB_V)$ are factor premiums, and $b_i$and c_i are factor loadings. To evaluate the performance of the ETF-factor model, a GRS test was performed on seven sets of 25 double-sort portfolios during the period from April 2009 to December 2016. The ETF-factor model is rejected by the GRS test at the 5% level for just one set of 25 double-sort portfolios. In contrast, the FF five-factor model (Fama & French, 2015) is rejected in two, and the q-factor model (Hou, Xue, and Zhang, 2015) is rejected in four sets of 25 double-sort portfolios. Apart from the expected market premium, all other expected premiums proposed by Fama and French (2015), Hou *et al.* (2015), Pástor and Stambaugh (2003), and Carhart (1997) are described by the new factor CMB_V. In the regressions using CMB_V to explain these factors, the alphas are small and insignificant. In addition, the CMB_V loadings for both the profitability factors in Fama and French (2015) and Hou *et al.* (2015), the traded liquidity factor in Pástor and Stambaugh (2003), and the momentum factor in Carhart (1997) are statistically significant at the 0.05 level. Moreover, evidence suggests that value and size anomalies are priced by the new market factor ETF_V, which measures the expected market premium.

The construction of the new factor, $CMB_{V,}$ is inspired by recent research. Normally, fixed-income securities have exposures to interest rate and credit risks, and commodities have exposures to equity risk and momentum. Cochrane and Piazzesi (2002) construct a single, tent-shaped linear combination of forward rates (hereafter, CP factor), which predicts bond returns. Further, Lustig *et al.* (2013) argue that the CP factor also signals investment opportunities in the stock market. Finally, Koijen *et al.* (2017) find that differential exposure to innovations in the CP factor explains the anomalies caused by the value effect. Commodity risk also plays a role in cross-sectional stock returns. Gokmenoglu and Fazlollahi (2015) find that investors react to the volatility of oil and gold prices in the long-run. Hou and Szymanowska (2013) construct commodities-based tracking portfolios that capture consumption growth and find that they are able to explain return driven by consumption growth. In addition, Brooks *et al.* (2016) find that sensitivity to backwardation or contango has an impact on the risk premium. Rising commodities prices are interpreted as bad economic news, and investors demand higher premiums for bearing commodity price risk. Recently, Henriksen and Kværner (2017) found that the overall expected profitability in the US stock market is negatively related to a shock to oil prices. This effect is stable and statistically significant in the long run. These findings demonstrate why the commodity bond return, CMB_V, exhibits strong explanatory power for the profitability factors in the FF five-factor and q-factor models.

The contribution of this thesis is to provide a parsimonious asset pricing model that can compete with the FF five-factor and q-factor models. Diverging from the conventional approach in asset pricing, this study creatively constructs risk factors using ETFs. This research is the first to exploit the link between ETFs and risk premiums. The rest of the chapter is organized as follows: Section 2 fully describes the definition of ETFs and the data. Section 3 reports the key empirical results and relevant analysis. Section 4 concludes the paper.

3.2. Data and Summary Statistics

3.2.1. Data Sample

This thesis' ETFs database covers the five types of ETFs that were traded in the US market monthly from January 1998 to December 2016. Monthly data on total return index, NAV, and outstanding shares was obtained from the Bloomberg Terminal. Monthly ETF returns are computed from the respective total return indices. In terms of asset classes, Bloomberg classifies ETFs as equity ETFs, fixed-income ETFs, commodity ETFs, alternative ETFs, specialty ETFs, and so on. Speciality ETFs, such as covered call ETFs and hedge fund ETFs, can be very complex. Because of their complex investment structures, it is impossible to determine which dimension of risk they capture. Therefore, specialty ETFs have been excluded. Further, the gold ETFs have been separated from the other commodity ETFs. Although returns on gold may be correlated with other commodities, Lawrence (2003) argues the gold is different due to its high liquidity and quick response to price changes.

Most ETFs trading in the marketplace are index-based ETFs, but some newer actively managed ETFs are leveraged or inverse ETFs. Leveraged ETFs deliver multiples of the performance of their tracking benchmark, and inverse ETFs deliver the performance opposite their benchmark. There is no doubt that leveraged and inverse ETFs are welcomed by some speculators and hedgers, but they are not in the mainstream. According to Ben-David *et al.* (2016), the US equity ETFs manage $1.33 trillion in assets passively, while only $24 billion in assets is managed actively. The active ETFs represent only 1.8% of the assets under management in the US equity ETFs market. This pattern is also observed in fixed-income ETFs. The sample for this study includes 1395 equity ETFs, 321 fixed-income ETFs, 130 commodity ETFs, 17 gold ETFs, and 30 alternative ETFs for the period January 1998 to December 2016.

3.2.2. Construction of the ETF Factors

As described earlier, this thesis proposes to construct a new proxy for the theoretical market portfolio based on the ETF market. A theoretical market portfolio includes every type of asset available in the financial market, with each asset weighted in proportion to the total market. In terms of the ETFs market, each investor holds a certain positive amount of each ETF in equilibrium. Because every ETF is a set of assets, investors actually hold a certain amount of each risky asset in the investment universe created by the ETFs market. In the framework of CAPM, the value-weighted factor ETF_V is constructed as:

$$ETF_V = \sum_{i=1}^{n} w_i\,(r_i - r_f),\ w_i = {}^{v_i}\!/_{\sum_{i=1}^{n} v_i} \text{ and } \sum_{i=1}^{n} w_i = 1 \quad (3.2)$$

where r_i is return of the ith ETF, v_i is market value of the ith ETF, i = 1, 2, ... n, w_i is the weight for the ith ETF and r_f is the risk-free rate.

An asset class is a group of securities that have similar characteristics. Across asset classes, the types and degrees of various risk exposures can be distinct. In the world of Fama and French, the investment universe is the entire stock market. Following the concept of asset class, Fama and French obtain the spread between two groups of stocks. For example, to construct the size factor, stocks are allocated to two groups: small and big stocks; to construct the size factor, stocks are allocated to three groups: value, neutral, and growth stocks. As Fama and French (1995) mention, firms with a high book-to-market ratio tend to be persistently distressed. Clearly, the value stocks can be viewed as an asset class, given that the investment universe is the stock market.

However, the ETFs market not only covers the stock market, but also the bond market, commodity market, and so on. Because ETFs are investment funds, anomalies do not exist in the ETFs market. Thus, there are no anomalies leading to sources of risk. However, based on the types of underlying assets, ETFs can be easily classified as equity ETFs, bond ETFs, commodity ETFs, and so on. Therefore, these types of ETFs can be employed to construct proxies for the excess returns on the stock market, bond market, commodity market, and so on. Constructing these explanatory returns is illustrated in (3.2). Thus, the five ETF-factors are: the equity ETF factor (EE_V), bond ETF factor (EB_V), commodity ETF factor (EC_V), gold ETF factor (EG_V), and alternative ETF factor (EA_V).

3.2.3. Summary Statistics of ETF-Factors

[Insert table 3.1]

Table 3.1 reports summary statistics of the time-series of ETF-factors in different periods. The average EE_V return is 1.21% per month and its standard deviation is 4.08% during the period April 2009 to December 2016 (Panel A of Table 3.1). Targeting fixed-income securities, EB_V generates a lower average monthly return of 0.36% but also a lower monthly volatility of 0.87%. Similarly, the average EA_V return is 0.10% and the volatility is 1.19% per month. The similarity of the characteristics of EB_V and EA_V is due to the significant investment in mortgage-backed securities by alternative ETFs. Apparently, but with no surprise, EC_V and EG_V reflect much higher monthly volatility: 6.24% and 9.00%, respectively. The average EC_V is -0.21% per month, which is in line with the declining oil price during this period. Contrary to the negative return of EC_V, EG_V generates a positive mean: 0.21% per month. In the last column of Panel A, the ETF market factor ETF_V has a mean of 1.06% per month, which is slightly lower than the average EE_V return. This is due to the lower average returns of non-equity ETFs. The standard deviation of ETF_V is 3.45% per month, suggesting that ETF_V is less volatile than EE_V.

Panel B reports the descriptive statistics of conventional factors during the period of April 2009 to December 2016. The mean market factor is slightly higher than that of EE_V. This is because equity ETFs can invest in foreign markets, and they put different weights on the same underlying assets than the stock market. The traded liquidity factor, VWF, has the average return: -0.07% per month with t-statistic = 0.25. The momentum factor has the smallest average return: -0.35% per month with t-statistic = 0.67. FF factors have average returns varying between 0.09% and 0.23% per month.

In Panel C, the low correlations of EE_V and EB_V, and EC_V and EG_V show that different ETF factors capture various other risks apart from equity risk. Although some people argue that gold acts more like a currency, EC_V and EG_V are highly correlated with a correlation of 0.96 in this case. EB_V has relatively low correlations with the other four ETF factors. Like EB_V, EA_V show less co-movement with other ETF factors, with the exception of EE_V.

Panel D reports the correlation matrix between ETF factors and traditional factors during the period April 2009 to December 2016. First of all, the correlation between EE_V and the market factor is 0.97, which demonstrates that, as a whole, equity ETFs successfully replicate the US stock market. SMB correlates to EE_V, EC_V and EG_V with values of 0.41, 0.17 and 0.17, respectively, indicating that size premium not only has exposure to equity risk, but also commodity risk. On the other hand, HML only correlates to EE_V with a value of 0.24 and is almost independent of other ETF factors. According to Fama and French (2015), RMW is the equity spread between stocks with robust profitability and stocks with weak profitability, and CMA is the equity spread between stocks with low investment and stocks with high investment.

Interestingly, RMW is negatively correlated with both EC_V and EG_V with the same value of -0.19. Fama and French (2015) define profitability as revenues minus cost of goods sold, interest expense and selling, and general and administrative expenses per dollar. Fluctuating commodity prices can have a tremendous impact on the public companies' cost of goods sold, and, hence, the profitability. Titman *et al.* (2004) finds a negative relationship between a firm's capital investment and stock performance. Fama and French (2015) construct the factor CMA to capture this anomaly. However, CMA exhibits very low correlations with these five ETF factors. The momentum factor, MOM, negatively correlates to these ETF factors with values between -0.19 and -0.37. Pástor and Stambaugh (2003) described a traded liquidity factor, which is derived from ten portfolios based on sensitivity to the non-traded liquidity innovation factor. The traded liquidity factor, VWF, shows a positive correlation with both EC_V and EG_V, suggesting that liquidity in the stock market can be linked to commodity risk.

Panels E and F report the descriptive statistics of conventional factors and the ETF market factor, $ETF_{V,}$ during the period January 1998 to December 2016 and the period January 1998 to March 2009, respectively. Combining the Panels B, E and F, it is found that the ETFs market factor, $ETF_{V,}$

has a higher average return than $R_i - R_f$ prior to April 2009, but a lower average return afterwards. The profitability factor, RMW, has a mean of 0.43% per month before April 2009, but its mean value reduces dramatically afterwards. Before April 2009, the momentum factor, MOM, has the largest mean, with a value of 0.84% (t-statistic = 1.67). Afterwards, this value reduces dramatically to -0.35%. The mean of the traded liquidity factor, VWF, also declines dramatically after April 2009, from 1.05% to -0.07%. The average returns of SMB, HML, and CMA, are relatively similar before and after April 2009.

[Insert table 3.2]

Fama and French (2015) argue that HML is droppable in the five-factor model because the average HML return can be described by other factors. In the spirit of Ross (1976), the excess return of an asset is the linear combination of the factor risk premia. In line with Huberman and Kandel (1987), adding a factor that can be explained by existing factors does not provide improvements to the mean-variance efficiency of the tangent portfolio. Therefore, each risk factor is supposed to capture a unique source of risk. Table 3.2 shows the results of using four ETF factors to describe the average return on the fifth factor during the period April 2009 to December 2016. It was found that each ETF factor is not completely explained by the others. In the EE_V regression, the alpha is 0.9% per month (t-statistic = 3.66) and the R-squared coefficient is only 0.61. This shows that EE_V, a proxy for the stock market, is not replaceable. As for the description of EB_V, the alpha is 0.3% per month (t-statistic = 3.45) and the R-squared coefficient is 0.37. Considering the average EB_V return is 0.36%, about 83% of average EB_V return is not explained by other ETF factors. Besides, the very low R-squared coefficient indicates that most of the uncertainty of EB_V is not captured.

In the regressions to explain EC_V and EG_V, the R-squared coefficients are about 0.9 due to the high correlation between EC_V and EG_V. However, the alphas are statistically significant with values of - 0.5% (t-statistic = 2.78) and 0.5% (t-statistic = 2.10), respectively. Although EC_V and EG_V are highly correlated, they cannot substitute for each other. Besides, the average EA_V return is not completely explained in the regression, although the loadings on EE_V and EB_V are statistically significant with values of 0.17 (t-statistic = 5.90) and 0.32 (t-statistic = 3.22).

[Insert table 3.3]

Fama and French (1996) show that size and book-to-market portfolios are proxies for multifactor-minimum-variance (MMV) portfolios. Indeed, SMB and HML are linear combinations of MMV portfolios. Both explanatory variables in ICAPM and risk factors in APT must be the expected excess returns on MMV portfolios. If ETF factors can describe the average returns of FF factors, then the linear span of ETF factors at least covers that of FF factors. Table 3.3 reports the results of regressions that use ETF factors to explain conventional factors. The market factor has been dropped from the conventional factors because EE_V is already the proxy for the stock market. The results show that FF factors and other conventional factors are explained by ETF factors because intercepts in all regressions are statistically insignificant.

In the SMB regression, the alpha is indistinguishable from zero with a value of 0.01% per month (t-statistic = 0.03). The R-squared coefficient is 0.25, which is quite large relative to Fama and French (1996). Interestingly, the size premium, SMB, is negatively related to EB_V. The slope coefficient estimate is -0.80 (t-statistic = 2.43), which says that big stocks are more bond-like than small stocks. The regression intercept for HML is 0.10% per month (t-statistic = 0.33). Interestingly, the loading on EC_V is positive, while the loading on EG_V is negative. This pattern is also consistent in the regressions used to explain RMW, CMA, and VWF. Except for RMW, conventional factors have negative loadings on EB_V. Generally, ETF factors do a good job of describing average returns of conventional factors, at least on US data after the 2007-08 financial crisis.

3.3. Model Performance

3.3.1. A Parsimonious Model using ETF Factors

Traditional asset pricing approaches to financial markets have focused on the completeness argument, which provides mathematically elegant models, such as CAPM. With very simple extensions, FF models make very precise predictions of cross-sectional stock returns. However, the effectiveness of asset pricing models has been very poor since the 2007-08 financial crisis. According to Fama and French (2015), GRS tests reject all models that combine FF factors for six sets of left-hand-side portfolio, during the period July 1963 to December 2013. The shortcoming of FF factors is the lack of theoretical perspective. Hahn and Lee (2006) find that changes in the default spread and term spread can be proxies for SMB and HML. Their results provide more economic meaning for SMB and HML, but the concepts of small and value stocks are still subjective. Although the construction of the ETF factors also lacks theoretical fundamentals, the selection criteria, based on the types of ETFs, are objectively available on the market.

This section discusses whether or not ETF factors describe the expected returns of seven sets of double-sort portfolios. If a model completely explains the expected returns, the alphas in the regressions should be indistinguishable from zero. Gibbons, Ross, and Shanken's (1989) GRS statistic tests the mean-variance efficiency of the market portfolio. The GRS test statistic follows an *F* distribution that requires the error terms are normal as well as uncorrelated and homoscedastic. These strong statistical assumptions are violated easily in practice. Thus, only one test does not provide solid evidence. In section 3.3.4 'Cross-Sectional Regression Based Tests', more empirical tests of asset pricing models will be performed.

Table 3.4 summarises the GRS statistics and mean absolute[16] for tests of conventional models and ETF-factor models during the period April 2009 to December 2016. Six conventional asset pricing models are considered: CAPM, the FF three-factor model, the FF five-factor model, the Carhart four-factor model, the FF three-factor model with combined VWFs, and the q-factor model (Hou *et al.* 2015).

[Insert Table 3.4]

Panel A of Table 3.4 shows that the FF five-factor model is superior to other conventional models in describing the expected returns on double-sort portfolios. Except for the tests on the 25 B/M-Inv portfolios and the 25 Size-OP portfolios, the FF five-factor and Carhart four-factor models completely capture expected returns of five sets of double-sort portfolios. Although both produce similar GRS test results, the FF five-factor model provides minor improvements on the GRS statistics. In addition, improvements in the average absolute intercept in terms of explaining expected returns

[16] Mean absolute is the average absolute value of the intercepts for a set of double-sort portfolios.

are observed for each set of double-sort portfolios. The results provided the evidence to Fama and French (2015) that the FF five-factor model is better than FF three-factor model in terms of describing the average stock returns.

Now, the ETF-factor model performance will be investigated. As the equity ETF factor EE_V is an alternative proxy for market excess return, it must appear in all ETF-factor models. Panel B of Table 3.4 compares the performance of ETF-factor models that combine EE_V with any combination of EB_V, EC_V, and EG_V. Based on the GRS statistics and mean absolute, using EE_V alone as the explanatory variable fails to describe the expected returns of five sets of double-sort portfolios, indicating poor model performance. Even CAPM in Panel A completely explains four sets of double-sort portfolio returns. The two-factor models that combine EE_V with EB_V or EG_V provide very limited improvements on GRS statistics and mean absolute. Adding EB_V as an explanatory variable in the regression even produces larger average absolute intercepts.

However, the two-factor model that combines EE_V with EC_V completely explains the expected returns of 25 OP-Inv portfolios and 25 Size-Inv portfolios, while other two-factor models fail to do so. For all seven sets of double-sort portfolios, the two-factor model that combines EE_V with EC_V produces lower GRS statistics and smaller average absolute intercepts than the single-factor model that employs EE_V and the other two-factor models. Evidently, then, EE_V and EC_V must be included in the ETF-factor model. In addition, the three-factor model that combines EE_V and EC_V with EG_V completely describes the expected returns of 25 Size-Prior portfolios. The GRS test says the three-factor model that combines EE_V and EC_V with EG_V is as good as the FF five-factor model, at least in the given sample period. Interestingly, if the bond ETF-factor, EB_V, is added to the regressions, the four-factor model exhibits very poor performance, based on both the GRS statistics and mean absolute.

3.3.2. The New Market Portfolio

The version of the CAPM developed by Sharpe (1964) and Lintner (1965) is based on Markowitz's (1952) mean-variance analysis and portfolio selection. However, early tests strongly reject the

Sharpe-Lintner CAPM and demonstrate its poor empirical performance. Roll (1977) argues that it is empirically impossible to construct a market portfolio that includes every single possible available asset, including real estate, precious metals, stamp collections, jewelry, and anything else with any worth. Consequently, it is not possible to test the CAPM. However, if a proxy can be found that explains the expected returns, CAPM still contributes significantly to the empirical research. In line with Roll's (1977) critique, Stambaugh (1982) constructs market portfolios that include returns for bonds, real estate, consumer durables, and common stocks. But he finds that adding non-equity assets does not improve the mean-variance efficiency because the volatility of the stock returns dominates the volatility of the expanded market returns. In addition, Fama and French (1998) find that the international version of CAPM fails to explain the value anomaly.

This thesis proposes a new proxy for the theoretical market portfolio using the rapidly growing ETFs market. Apart from equity ETFs, bond ETFs, commodity ETFs, gold ETFs and alternative ETFs are included in the market portfolio. As I describe in the data section, this ETFs market portfolio is a value-weight index. The ETF-factor ETF_V, described in Table 3.1, is the excess return on the ETF market portfolio. The traditional market factor $R_m - R_f$ is the excess return on the market. It is calculated as the value-weight return on all NYSE, AMEX, and NASDAQ stocks (from CRSP) minus the monthly risk-free rate.

[Insert Table 3.5]

Table 3.5 compares the performance of ETF_V and $R_m - R_f$ as the explanatory variable in CAPM. Panel A of Table 3.5 reports the GRS statistics and mean absolute during the period April 2009 to December 2016. Although the traditional CAPM produces slightly smaller GRS statistics and mean absolute, both ETF_V and $R_m - R_f$ completely explain expected returns of four double-sort portfolios. Panel B of Table 3.5 reports the results during the period January 1998 to March 2009. In this period, ETF_V is a better proxy for the excess return on the market portfolio. It explains five sets of portfolios returns, while $R_m - R_f$ only explains four. Generally, using ETF_V instead $R_m - R_f$ provides improvements on both GRS statistics and mean absolute. Panel C of Table 3.5 shows that ETF_V and $R_m - R_f$ perform quite similarly during the period January 1998 to December 2016.

Generally, $R_m - R_f$ produces a slightly smaller mean absolute, while ETF_V provides some improvements on GRS statistics.

3.3.3. The ETF-factor Model

Panel B of Table 3.4 shows that EC_V contributes to explaining returns, while EB_V drags down the model performance. A model that includes EB_V as an explanatory variable only explains the expected returns of 25 B/M-OP portfolios and Size-B/M portfolios, but a model that includes EC_V and no EB_V explains expected returns of at least four sets of double-sort portfolios. In other words, a long position in EC_V helps explain returns, but a long position in EB_V does not. Therefore, this study is interested in the spread between the commodity ETF-factor (EC_V) and the bond ETF-factor (EB_V). This commodity-bond return will be designated as CMB_V (commodity minus bond). Specifically, the excess return on asset i is:

$$\mathrm{E}(R_i) - R_f = b_i\mathrm{E}(ETF_V) + c_i\mathrm{E}(CMB_V) \tag{3.3}$$

Table 3.6 shows that the ETF-factor model is the superior model relative to the FF five-factor model during the period April 2009 to December 2016. The GRS tests in Panel A support the inference that the ETF-factor model has little trouble explaining the expected returns on the double-sort portfolios. However, the expected returns on 25 Size-OP portfolios are not completely explained by the ETF-factor model. In addition, the ETF-factor model and FF five-factor model produce very similar average absolute intercepts. Thus, if CMB_V contributes to the description of the anomalies, then conventional factors, excluding $R_m - R_f$, are described by CMB_V.

[Insert Table 3.6]

Panel B of Table 3.6 tests this inference with regressions that use CMB_V to explain the conventional factors and q-factors (Hou *et al.* 2015). The results are consistent with this inference. Except for the $R_m - R_f$ and market factors in the q-factor model, other expected premiums are all described by CMB_V. In the regressions used to explain RMW, MOM and VWF, the intercepts are, respectively, 0.05% (t-statistic = 0.33), -0.44% (t -statistic = 0.80), and 0.00% (t-statistic = 0.00). In addition, their loadings on CMB_V are all statistically significant at an alpha level of 0.05. Specifically, the profitability factor, RMW, significantly relates to CMB_V with a factor loading of -0.05 (t-statistic = 2.01). RMW is the return on a zero-investment portfolio that purchases stocks with robust profitability and sells stocks with weak profitability. The negative CMB_V loading indicates that a firm's profitability is negatively affected by the commodity price change.

The oil market is the world's largest commodity market and oil, itself, is a typical commodity. Thus, the commodity ETFs market must be highly correlated with the oil market and this fact provides an economic link between the profitability factor, RMW, and the new factor, CMB_V. In addition, the explanatory power of CMB_V on the profitability factor is confirmed in the regression to explain ROE, the profitability factor constructed by Hou *et al.* 2015. In the regression, the intercept is -0.14% (t-statistic = 0.55), and the CMB_V loading is -0.11 (t-statistic = 2.40). These results indicate the robustness of this finding.

[Insert Table 3.6]

Surprisingly, the traded liquidity factor, VWF, is absolutely explained by CMB_V. In the sample period, the average VWF is -0.07% per month (Panel B of Table 3.1). However, in the VWF regression, the intercept (average return unexplained by CMB_V) is 0.00% per month and the CMB_V loading is approximately -0.17 (t-statistic = 2.00). Although the volatility transmission mechanism between stock, bond, and commodity markets exists (Mensi *et al.* 2013), no relevant paper that investigate how these markets affect the liquidity in stock market has been found.

3.3.4. Time-Series Regression Details

[Insert Table 3.7]

In this section, more regression results are reported for the tests on 25 B/M-Inv portfolios and 25 Size-OP portfolios, whose returns are not completely explained by either the FF five-factor model or the q-factor model. However, the ETF-factor model passes the GRS test on 25 B/M-Inv portfolios. The GRS statistic is 1.58, and the p-value is 0.073. Table 3.7 summarises the alphas in the FF five-factor model, the q-factor model, and the ETF-factor model for both the 25 B/M-Inv portfolios and the 25 Size-OP portfolios. The common problem of these models is the strong negative intercept for the portfolio in the largest B/M quintile and highest Inv quintile. The FF five-factor model produces the biggest t-statistic, with a value of 2.84, and the ETF-factor model produces the smallest t-statistic, with a value of 2.03. However, the portfolio in the third B/M quintile and the fourth Inv quintile is likely to show anomalies. In this case, the FF five-factor explains the portfolio return, but the other two models fail. The alpha in the ETF-factor model is -0.37% per month (t-statistic = 1.97), and the alpha in the q-factor model is -0.41 % per month (t-statistic = 2.80). Fama and French (2015) say that high investment is associated with growth (low B/M), and low investment is associated with value (high B/M). The strong negative alpha for the portfolio in the biggest B/M quintile and highest Inv quintile indicates that aggressive investment is not a wise strategy for companies that lack growth potential.

[Insert Table 3.8]

Panel A of Table 3.8 shows the FF five-factor slopes for $R_m - R_f$, SMB, HML, RMW, and CMA in the regressions for 25 B/M-Inv portfolios. The slopes for $R_m - R_f$ are all significantly close to 1. The HML slopes are strongly positive for stocks with a high B/M ratio and slightly negative for stocks with a low B/M ratio. The CMA slopes are positive for stocks with low investment and negative for stocks with high investment. These results confirm that HML and CMA do play important roles in explaining returns on 25 B/M-Inv portfolios.

Panel B of Table 3.8 shows the two-factor slopes for ETF_V and CMB_V in the regressions for 25 B/M-Inv portfolios. Interestingly, the slopes for ETF_V, an alternative proxy for the excess return on market portfolio, are not always close to 1. It is observed that stocks with higher B/M ratio have larger ETF_V loadings, indicating that they bear more systematic risk. Particularly, the ETF_V slopes are all larger than 1.5 for the five portfolios in the largest B/M quintile. This evidence suggests that value anomalies can be priced by the ETF market factor, $ETF_{V,}$ in the given sample period. The CMB_V slopes are significantly negative for stocks with low investment. For 25 B/M-Inv portfolios, 60% of portfolio returns in the 1st, 2nd, and 3rd Inv quintiles have significantly negative CMB_V loadings, while only 10% of portfolios returns in the 4th and 5th Inv quintiles have significantly negative CMB_V loadings.

The GRS test in Tables 3.5 and 3.6 says that the FF five-factor model, the two-factor model that includes ETF_V and CMB_V, and the q-factor model all fail to describe the expected returns on 25 Size-OP portfolios during the period April 2009 to December 2016. But in the ETF-factor model, the alphas across 25 Size-OP portfolios are statistically insignificant at the 5% level (Table 3.7). On the other hand, the FF five-factor model produces two significant alphas, and the q-factor model produces four significant alphas. For the FF five-factor model, the anomalies are associated with the portfolio in the smallest Size quintile and 2nd OP quintile and the portfolio in the 4th Size quintile and 2nd OP quintile. This result is partly in line with Fama and French's (2015) observation that small stocks in the lowest profitability show significantly negative alphas, during the period July 1963 to December 2013.

Panel C of Table 3.8 shows the FF five-factor slopes for $R_m - R_f$, SMB, HML, RMW, and CMA in the regressions for 25 Size-OP portfolios. The slopes for $R_m - R_f$ are almost close to 1. The SMB slopes are significantly positive for small stocks and negative for the big stocks. The RMW slopes are significantly negative for the portfolios in the lowest profitability quintile and become positive with increasing profitability.

Panel D of Table 3.8 shows the two-factor slopes for ETF_V and CMB_V in the regressions for 25 Size-OP portfolios. Similar to the pattern in Panel B of Table 3.8, the ETF_V slopes can be significantly different from 1. Across the smallest Size quintile, the ETF_V loadings vary from 1.27 to 1.54, all of which are significant. Across the biggest Size quintile, the ETF_V loadings vary from 0.9866 to 1.3779 and four of them are very close to 1. This evidence supports that small stocks bear more systematic risk than big stocks. In addition, 7 out of 10 portfolios in the 4th and 5th Size quintile have significantly negative exposure to CMB_V loadings, indicating that big stocks are very sensitive to the spread between commodity and bond. For portfolios in the 1st, 2nd, and 3rd Size quintile, their absolute CMB_V loadings increase with size and tend to become statistically significant. This evidence suggests that small stocks with higher profitability are negatively sensitive to CMB_V.

These results do not deny the explanatory power of SMB, HML, RMW, and CMA. It has been demonstrated that these factors help describe the relevant anomalies during the period April 2009 to December 2016. However, it is also shown that size and value anomalies are described by the ETFs market factor, ETF_V. This evidence shows that ETF_V is a better market risk factor compared to that in CAPM. In addition, the new factor CMB_V also has explanatory power on anomalies related to size, investment, and profitability. It is a great supplement to ETF_V and makes the ETF-factor model highly competitive.

3.3.5. Cross-Sectional Regression Based Tests

Tests of asset pricing models usually use the cross-sectional regression approach of Fama and MacBeth (1973) and the time-series regression approach of Gibbons, Ross and Shanken (1989). In the time-series regression approach, GRS tests whether the factors are multifactor-mean-variance-efficient (MMVE). While the cross-section regression assumes those factors are multifactor-minimum-variance (MMV) and only test whether they are also MMVE.

Section 3 has applied GRS test (based on time-series regressions) to investigate whether some linear combination of the ETF-factor portfolios are on the minimum variance boundary. The results show that the ETF-factors EB_V, EC_V and CMB_V contribute to describe the cross section of expected FF portfolio returns. To complete the evaluation of asset pricing models, this subsection considers GLS cross-section test of linear factor model. To make results more comparable to those in asset pricing literature, specific index returns that have longer historical data are used as proxies for ETF-factors.[17] The Bloomberg Barclays US Government/Credit Bond Index is used as proxy for the bond ETF market and the S&P GSCI Total Return Index in USD is used as proxy for the commodity ETF market.[18] Finally, the simulation techniques of Lewellen, Nagel and Shanken (2010) are used to provide confidence intervals for test statistics.

---

[17] The construction of proxies for ETF-factors are exactly the same as the way to construct ETF-factors in section 3.2.2. Instead of using aggregate ETF market returns, this section uses specific index returns because they can provide longer historical data.

[18] The Bloomberg Barclays US Government/Credit Bond Index is a broad-based flagship benchmark that measures the non-securitized component of the US Aggregate Index. It includes investment grade, US dollar-denominated, fixed-rate Treasuries, government-related and corporate securities. The S&P GSCI Total Return Index in USD is widely recognized as the leading measure of general commodity price movements and inflation in the world economy. Index is calculated primarily on a world production weighted basis, comprised of the principal physical commodities futures contracts.

There are two steps of running the cross-section regressions: first, the factor loadings are estimated in time-series regression (for example, equation (3.1)); second, estimate the relation between expected returns and the factor loadings estimated in the first step. Using the equation (3.1) as the time-series regression, the cross-section regression is:

$$\mathrm{E}_t\left(R_{i,t}\right) - R_f = \mu + \beta_1 \mathrm{b} + \beta_2 \mathrm{c} + \alpha \tag{3.4}$$

where $\mathrm{E}_t\left(R_{i,t}\right)$ is a N×1 vector of expected returns on FF portfolios and N is the number of test assets, R_f is a N×1 vector of average risk-free rate, μ is the vector of estimated intercepts, b is a N×1 vector of factor loadings on ETF_V (estimated in first step), c is a N×1 vector of factor loadings on CMB_V (estimated in first step), β_1 and β_2 are the slopes, α is the vector of true pricing errors.

In line with Lewellen, Nagel and Shanken (2010), two new test statistics are considered in the cross-section regression. Shanken's (1985) T^2 test in cross-section regression is similar to the GRS test in time-series regression. It focuses on pricing errors (α in equation (3.4)) in the cross-section regression of expected returns on factor loadings. The T^2 statistic tests whether the pricing errors are jointly zero. Another statistic is the quadratic q that measures the difference between the maximum generalized squared Sharpe ratio and that attainable from a model's mimicking portfolios (if the model completely describes the cross section of expected returns, the q is indistinguishable from zero). Clearly, the smaller the T^2 statistic and q statistic are, the better the tested pricing model is.

Four models discussed in the previous subsections are selected to be investigated. The models include: 1). The five-factor model of Fama and French (2015); 2). The ETF-factor model that includes three variables constructed from the ETFs: ETF_v, EB_v and EC_v.[19] ETF_v is the proxy for the market factor, EB_v is the proxy for the bond factor and EC_v is the proxy for the commodity factor; 3). The ETF-factor model presented in equation (3.1); 4). Sharpe-Lintner CAPM. Because cross-section tests

[19] The t-test on whether the coefficient on EB_v equals the negative coefficient on EC_v is performed. In the period of April 2009 to December 2016, the t-statistics is 0.93. But in the period of February 1973 to October 2017, the t-statistics is 3.62. These results indicate that Model1 and Model2 (both in table 3.9) are statistically no difference in the period of April 2009 to December 2016.

in original research mainly focus on FF Size-BM portfolios, the goal here is to test how well a model describes the cross section of expected returns on 25 FF Size-BM portfolios.

[Insert Table 3.9]

Table 3.9 reports cross-section regressions for all four models in two periods of April 2009 to December 2016 and February 1973 to October 2017. The expected returns of 25 FF Size-BM portfolios used by the tests are the same as those in previous subsections, but with longer length of historical data. This table reports the GLS R^2, the cross-section T^2 statistic, q statistic described earlier and the GRS statistic.[20] Confidence intervals for the true values of q, the GLS R^2 and GRS statistic are obtained using the simulation technique of Lewellen, Nagel and Shanken (2010).[21]

Table 3.9 shows several key results. First, the FF five-factor model performs superior in terms of slope estimates, cross-section GLS R^2, T^2 statistic, q statistic and GRS statistic. Using the data in the period of April 2009 to December 2016, the T^2, GLS R^2, q and GRS are 24.39, 0.00, 0.26 and 1.12 for CAPM while these statistics are 18.71, 0.21, 0.22 and 1.15 for the FF five-factor model. The FF five-factor model produces smaller T^2 and q statistics, indicating that it does better job of explaining the cross section of 25 Size-B/M portfolio returns. The FF five-factor model produces higher GLS R^2 (0.21), which means this model explains the maximum Sharpe ratio available on the 25 Size-B/M portfolios better than the CAPM. Interestingly, the two ETF-factor models do not provide much improvements over the CAPM when the performance is measured by the T^2 or q statistic.

Second, results change dramatically from time to time. The T^2 and GRS statistics are insignificant in tests for all models in the period of April 2009 to December 2016 but reject all models in the period of February 1973 to October 2017. However, the GLS R^2 and q statistic are

[20] GLS R^2 is more rigorous than OLS R^2. A high OLS R^2 can be produced even though the factor mimicking portfolios are far from mean-variance efficient, but a high GLS R^2 can only be produced if the model can explain the maximum Sharpe ratio available on the test assets.

[21] For the simulation of T^2 statistic, q statistic and GLS R^2, the true factor loadings are fixed the same. The expected returns (e.g., $\mathrm{E}_t(R_{i,t})$ in equation (3.4)) are simulated using the estimated factor loadings, estimated cross-section slopes, a scalar constant and a random variable drawn from normal distribution. Lewellen, Nagel and Shanken (2010) explain the simulation technique in detail.

improved for all models compared to their counter parts in the period of April 2009 to December 2016. In line with the results of GRS test in tables 3.4 to 3.6, most models are not rejected by GRS test in the period of April 2009 to December 2016.[22] Although the FF five-factor model improves test statistics, these improvements are not significant.

Third, the two ETF-factor models provide limited improvements compared to CAPM. The T^2 and q statistics for the ETF-factor models are almost the same as those for the CAPM, either in the period of April 2009 to December 2016 or February 1973 to October 2017. The ETF-factor models produce smaller the GRS statistic. The confidence intervals for GRS statistic show that there is no significant improvement provided by ETF-factor models. For example, in the period of April 2009 to December 2016, a 95% confidence interval for the true GRS statistic is from 1.33 to 5.07 for Model 2 and is from 1.36 to 5.13 for CAPM.

Finally, in the spirit of taking the estimated coefficients seriously (Lewellen, Nagel and Shanken, 2010), table 3.9 shows that none of the models describes the level of expected returns because the intercepts in the cross-section regression are all significantly greater than zero. According to Lewellen, Nagel and Shanken (2010), the intercepts can be interpreted as the estimated zero-beta rates over the risk-free rate. Using excess quarterly returns, they find that the annualized zero-beta rates range from 7.8% to 14.3% over the risk-free rate. Considering both sample periods in table 3.9, the annualized zero-beta rates range from 16.08% to 21.12% over the risk-free rate. This result confirms the claim of Lewellen, Nagel and Shanken (2010) that the difference in lending versus borrowing costs is not the reason of big zero-beta rates.

3.4. Conclusion

Previous work identifies asset-pricing anomalies related to size, book-to-market, profitability, investment, and liquidity. This chapter investigates the performance of the FF five-factor model, the q-factor model, and the ETF-factor models derived from ETFs based on the GRS statistic of

[22] The GRS statistic for CAPM in table 3.9 is slightly different from its counterpart in table 3.5, but this difference is not substantial and does not change the conclusion. The reason is that the data of 25 FF Size-BM portfolios used in table 3.5 is updated using the 201704 CRSP database, while the data used in table 3.9 is updated the 201712 CRSP database. In addition, the construction of FF portfolios is updated in 201712. More details are available on the website of French Data Library.

Gibbons, Ross and Shanken (1989) and the GLS R^2, T^2 statistic and q statistic of Lewellen, Nagel and Shanken (2010).

This chapter offers four contributions to the existing literature of asset pricing. First, it proposes several ETF-factors as proxies for the unknown risk. The ETF-factors are different from the existing risk factors for two reasons: 1). ETF-factors are tradable in the market while the risk factors of Fama and French (1993, 2015), Carhart (1997), Pastor and Stambaugh (2003) are not available for investors; 2). ETF-factors cover more types of underlying assets and extend the universe of investment. The interaction between asset classes reflects future investment opportunity that contributes to price assets (Chen, 2002). Second, it provide solid empirical evidence for the five-factor model of Fama and French (2015), q-factor model of Hou et al. (2015) by using latest data and standard empirical asset pricing tests. Third, it enriches the literature of ETFs by extending the application of ETF research to the asset pricing field. Fourth, it provides an appropriate model to measure the mutual fund performance. This application will be fully investigated in the chapter 4.

The empirical results of time-series regression suggest that, after the 2007-08 financial crisis, the ETF-factor models outperform other conventional models, considering the GRS tests for seven sets of double-sort portfolios. Specifically, size and value anomalies are captured by the ETFs market factor, ETF_V. The ETF-factor CMB_V, which is the proxy for the spread between commodity and bond, significantly captures the anomalies related to profitability, liquidity, and momentum. The time-series and cross-sectional regression results highlight an alternative approach in asset pricing because very few previous studies construct proxies for unknown risk factors from the ETF market. The empirical results of cross-section regression show that multi-factor models do not provide significant improvements on the basic CAPM, including the ETF-factor models. But the FF five-factor model can produce smaller T^2 statistic and q statistic and greater GLS R^2.

Although the ETF-factor models do not perform superior in the cross-sectional regression based tests, they still deserve further investments in future research. Most importantly, the new factor CMB_V has not been persuasively defined. Moreover, there is no answer to how the spread between commodity and bond reshapes risk premiums. Ang *et al.* (2006) find that aggregate volatility risk is embedded in cross-sectional returns. In addition, Mensi *et al.* (2013) find significant volatility transmission across commodity and equity markets. These results suggest that cross-sectional returns may react to the volatility in the commodity market. In addition, Henriksen and Kværner (2017) demonstrates the negative relationship between oil price change and the overall expected profitability in the US stock market, which is stable and significant. However, direct evidence is still missing. Besides, this study shows that size and value anomalies are priced by the new ETFs market factor, ETF_V, after the latest financial crisis. More time is needed to decide whether or not this is a new normal.

Table 3-1 Summary Statistics for Explanatory Returns for Pricing Models

R_m is the value-weight return on all NYSE, AMEX, and NASDAQ stocks and R_f is the one-month Treasury bill rate. The construction of *SMB*, *HML*, *RMW,* and *CMA* follows Fama and French (2015). *SMB* is the size factor, *HML* is the value factor, *RMW* is the profitability factor and *CMA* is the investment factor. The construction of the momentum factor *MOM* is in line with Carhart (1997). The traded liquidity factor follows Pastor and Stambaugh (2003) and is downloaded from their data library. The equity ETF-factor EE_v is the excess return on a value-weight portfolio of equity ETFs. The bond ETF-factor EB_v is the excess return on a value-weight portfolio of fixed-income ETFs. The commodity ETF-factor EC_v is the excess return on a value-weight portfolio of commodity ETFs. The gold ETF-factor EG_v is the excess return on a value-weight portfolio of gold ETFs. The alternative ETF-factor EA_v is the excess return on a value-weight portfolio of alternative investment ETFs. ETF_v is the excess return on a value-weight portfolio of US ETFs. The table shows the average return, the standard deviation, t-statistics for the average return, cross-correlations. The period of Panel A, B, C, D is April 2009 through December 2016, the period of Panel E is January 1998 through December 2016 and the period of Panel F is January 1998 to March 2009.

Panel A: Summary Statistics for Monthly ETF-factors; April 2009 to December 2016

	EE_v	EB_v	EC_v	EG_v	EA_v	ETF_v
Average Return	1.21	0.36	-0.21	0.21	0.10	1.06
Standard Deviation	4.08	0.87	6.24	9.00	1.19	3.45
t-statistics	2.86	4.03	-0.32	0.22	0.82	2.95

Panel B: Summary Statistics for Monthly Conventional Factors; April 2009 to December 2016

	$R_m - R_f$	*SMB*	*HML*	*RMW*	*CMA*	*MOM*	*VWF*
Average Return	1.38	0.23	0.20	0.09	0.19	-0.35	-0.07
Standard Deviation	3.88	2.52	2.43	1.58	1.39	5.02	2.75
t-statistics	3.43	0.88	0.78	0.52	1.32	-0.67	-0.25

Panel C: Cross-Correlations for Monthly ETF-factors; April 2009 to December 2016

	EE_v	EB_v	EC_v	EG_v	EA_v
EE_v	1.00				
EB_v	0.53	1.00			
EC_v	0.54	0.40	1.00		
EG_v	0.49	0.39	0.96	1.00	
EA_v	0.75	0.58	0.50	0.45	1.00

Panel D: Cross-Correlation between ETF-factors and Conventional Factors; April 2009 to December 2016

	EE_v	EB_v	EC_v	EG_v	EA_v
$R_m - R_f$	0.97	0.40	0.46	0.40	0.70
SMB	0.41	0.00	0.17	0.17	0.19
HML	0.24	-0.08	0.04	0.00	0.04
RMW	-0.33	0.01	-0.19	-0.19	-0.15
CMA	0.07	-0.06	-0.02	-0.06	-0.02
MOM	-0.37	-0.26	-0.22	-0.19	-0.24
VWF	0.13	0.06	0.26	0.24	0.15

Panel E: Summary Statistics for Conventional Factors and ETF_v ; January 1998 to December 2016								
	$R_m - R_f$	*SMB*	*HML*	*RMW*	*CMA*	*MOM*	*VWF*	ETF_v
Average Return	0.49	0.30	0.21	0.29	0.30	0.35	0.59	0.45
Standard Deviation	4.57	3.25	3.27	3.14	2.22	5.53	3.67	4.74
t-statistics	1.64	1.39	0.96	1.40	2.06	0.97	2.45	1.43

Panel F: Summary Statistics for Conventional Factors and ETF_v ; January 1998 to March 2009								
	$R_m - R_f$	*SMB*	*HML*	*RMW*	*CMA*	*MOM*	*VWF*	ETF_v
Average Return	-0.11	0.35	0.22	0.43	0.38	0.84	1.05	0.03
Standard Deviation	4.91	3.68	3.75	3.87	2.65	5.83	4.14	5.43
t-statistics	-0.27	1.09	0.67	1.30	1.66	1.67	2.96	0.07

Table 3-2 Using Four ETF-factors in Regressions to Explain the Fifth, April 2009 to December 2016

The equity ETF-factor EE_v is the excess return on a value-weight portfolio of equity ETFs. The bond ETF-factor EB_v is the excess return on a value-weight portfolio of fixed-income ETFs. The commodity ETF-factor EC_v is the excess return on a value-weight portfolio of commodity ETFs. The gold ETF-factor EG_v is the excess return on a value-weight portfolio of gold ETFs. The alternative ETF-factor EA_v is the excess return on a value-weight portfolio of alternative investment ETFs. Int is the regression intercept and R^2 is the R-squared, adjusted for degrees of freedom.

	Int	EE_v	EB_v	EC_v	EG_v	EA_v	R^2
EB_v	0.0028	0.0403		-0.0221	0.0264	0.2838	0.37
t-stat	3.45	1.18		-0.57	0.96	2.48	
EC_v	-0.0047	0.1298	-0.1324		0.6283	0.2006	0.93
t-stat	-2.78	1.85	-0.61		24.22	0.97	
EG_v	0.0052	-0.1115	0.3576	1.4168		-0.1659	0.92
t-stat	2.10	-1.13	1.16	20.31		-0.51	
EA_v	-0.0021	0.1671	0.3219	0.0379	-0.0139		0.62
t-stat	-2.82	5.90	3.22	0.89	-0.46		

Table 3-3 Using Five ETF-factors in Regressions to Explain the Conventional Factors, April 2009 to December 2016

SMB is the size factor, HML is the value factor, RMW is the profitability factor and CMA is the investment factor. They are well explained by Fama and French (2015). The momentum factor MOM is defined by Carhart (1997). VWF is the traded liquidity factor constructed by Pastor and Stambaugh (2003). The ETF-factors EE_v, EB_v, EC_v, EG_v and EA_v are the same as those in Table 2. Int is the regression intercept and R^2 is the R-squared.

	Int	EE_v	EB_v	EC_v	EG_v	EA_v	R^2
SMB	-0.0001	0.4217	-0.7982	-0.1361	0.0923	-0.2969	0.25
t-stat	-0.03	4.46	-2.43	-0.86	0.99	-1.16	
HML	0.0010	0.3290	-0.5789	0.0984	-0.0867	-0.4896	0.15
t-stat	0.33	2.73	-1.36	0.62	-0.89	-1.57	
RMW	0.0018	-0.2061	0.4256	0.0550	-0.0520	0.1902	0.17
t-stat	0.94	-3.03	2.26	0.61	-0.95	1.08	
CMA	0.0019	0.0773	-0.1454	0.0752	-0.0624	-0.1407	0.04
t-stat	1.08	1.16	-0.53	0.78	-1.05	-0.78	
MOM	0.0036	-0.4713	-0.6911	-0.1662	0.0993	0.5849	0.15
t-stat	0.97	-1.36	-0.75	-0.43	0.51	1.79	
VWF	0.0009	-0.0322	-0.2278	0.2008	-0.0570	0.1927	0.08
t-stat	0.27	-0.16	-0.61	1.16	-0.49	0.31	

Table 3-4 Summary Statistics for Tests of Conventional Asset Pricing Models and ETF-factor Models, April 2009 to December 2016

The sets of dependent excess returns include the 25 *B/M-Inv* portfolios, the 25 *BM-OP* portfolios, 25 *OP-Inv* portfolios, 25 *Size-B/M* portfolios, 25 *Size-Inv* portfolios, 25 *Size-Mom* portfolios and 25 *Size-OP* portfolios. In Panel A, the conventional models include CAPM, FF3 for the three-factor model (Fama and French, 1993), FF5 for the five-factor model (Fama and French, 2015), C4 for the Carhart four-factor model (Carhart, 1997), PS for Pastor-Stambaugh model (Pastor and Stambaugh, 2003) and the q-factor model (Hou, Xue and Zhang, 2015). ETF_v represents the one-factor model in which the explanatory variable is the excess return on ETFs market. In Panel B, the ETF-factor models include the one-factor model in which EE_v is the explanatory variable, the two-factor models that combine EE_v (abbreviated to *E*) with EB_v (abbreviated to *B*), EC_v (abbreviated to *C*) or EG_v (abbreviated to *G*) , the three-factor models that combine EE_v with EB_v and EC_v, EB_v and EG_v or EC_v and EG_v, and the four-factor model in which the explanatory variables are EE_v, EB_v, EC_v and EG_v. GRS is the F-statistics of Gibbons, Ross, and Shanken (1989) that test whether the regression intercepts for a set of portfolios are jointly zeros. p(GRS) is the p-value of GRS. Ave$|\alpha|$ is the average absolute value of the intercepts for a set of double-sort portfolios.

Panel A	CAPM	FF3	FF5	C4	PS	ETF_v	q-factor
B/M-Inv Portfolio							
GRS	1.97	2.03	1.85	2.00	2.07	1.88	2.28
p(GRS)	0.0150	0.0118	0.0259	0.0135	0.0103	0.0215	0.0043
Ave$\vert\alpha\vert$	0.0023	0.0022	0.0019	0.0021	0.0022	0.0034	0.0021
BM-OP Portfolio							
GRS	1.03	1.00	0.86	0.98	0.98	1.17	1.20
p(GRS)	0.4443	0.4856	0.6590	0.5092	0.5009	0.2990	0.2745
Ave$\vert\alpha\vert$	0.0032	0.0029	0.0025	0.0028	0.0029	0.0032	0.0024
OP-Inv Portfolio							
GRS	1.41	1.36	1.20	1.38	1.40	1.63	1.49
p(GRS)	0.1327	0.1594	0.2760	0.1527	0.1419	0.0584	0.1034
Ave$\vert\alpha\vert$	0.0025	0.0024	0.0018	0.0023	0.0023	0.0029	0.0020
Size-B/M Portfolio							
GRS	0.94	1.13	0.97	1.11	1.20	1.19	1.31
p(GRS)	0.5562	0.3360	0.5138	0.3591	0.2787	0.2818	0.1938
Ave$\vert\alpha\vert$	0.0017	0.0012	0.0010	0.0012	0.0012	0.0024	0.0015
Size-Inv Portfolio							
GRS	1.47	1.59	1.56	1.64	1.58	1.65	1.74
p(GRS)	0.1090	0.0689	0.0796	0.0585	0.0720	0.0540	0.0395
Ave$\vert\alpha\vert$	0.0018	0.0016	0.0013	0.0016	0.0016	0.0025	0.0015
Size-Mom Portfolio							
GRS	1.71	1.68	1.61	1.65	1.69	1.86	1.78
p(GRS)	0.0434	0.0493	0.0670	0.0560	0.0482	0.0231	0.0340
Ave$\vert\alpha\vert$	0.0022	0.0021	0.0019	0.0020	0.0022	0.0030	0.0023
Size-OP Portfolio							
GRS	1.92	1.89	1.76	1.94	1.90	2.14	2.24
p(GRS)	0.0180	0.0213	0.0374	0.0179	0.0208	0.0071	0.0052

Ave|α|	0.0018	0.0017	0.0012	0.0017	0.0017	0.0026	0.0017	
Panel B	*E*	*E,B*	*E,C*	*E,G*	*E,B,C*	*E,B,G*	*E,C,G*	*E,B,C,G*
B/M-Inv Portfolio								
GRS	2.06	2.59	1.83	1.95	2.23	2.39	1.91	2.32
p(GRS)	0.0101	0.0011	0.0263	0.0164	0.0052	0.0027	0.0194	0.0036
Ave|α|	0.0035	0.0050	0.0029	0.0031	0.0043	0.0045	0.0030	0.0043
BM-OP Portfolio								
GRS	1.35	1.69	1.14	1.27	1.39	1.53	1.08	1.31
p(GRS)	0.1669	0.0470	0.3268	0.2176	0.1448	0.0878	0.3843	0.1908
Ave|α|	0.0032	0.0041	0.0027	0.0029	0.0036	0.0037	0.0025	0.0036
OP-Inv Portfolio								
GRS	1.84	2.68	1.61	1.73	2.31	2.47	1.53	2.19
p(GRS)	0.0254	0.0007	0.0633	0.0392	0.0037	0.0019	0.0886	0.0062
Ave|α|	0.0030	0.0046	0.0025	0.0026	0.0038	0.0041	0.0027	0.0040
Size-B/M Portfolio								
GRS	1.34	1.66	1.27	1.27	1.40	1.50	1.46	1.48
p(GRS)	0.1698	0.0531	0.2197	0.2201	0.1416	0.0987	0.1143	0.1049
Ave|α|	0.0025	0.0051	0.0018	0.0021	0.0044	0.0047	0.0018	0.0042
Size-Inv Portfolio								
GRS	1.83	2.70	1.59	1.75	2.40	2.65	1.57	2.29
p(GRS)	0.0264	0.0007	0.0687	0.0373	0.0025	0.0009	0.0756	0.0041
Ave|α|	0.0026	0.0053	0.0019	0.0022	0.0045	0.0049	0.0021	0.0043
Size-Mom Portfolio								
GRS	2.02	2.73	1.78	1.91	2.36	2.52	1.66	2.29
p(GRS)	0.0120	0.0006	0.0329	0.0191	0.0029	0.0015	0.0536	0.0042
Ave|α|	0.0030	0.0051	0.0023	0.0026	0.0043	0.0046	0.0023	0.0042
Size-OP Portfolio								
GRS	2.32	3.31	2.16	2.20	2.96	3.07	2.51	3.29
p(GRS)	0.0033	0.0001	0.0067	0.0056	0.0002	0.0001	0.0016	0.0001
Ave|α|	0.0027	0.0052	0.0022	0.0024	0.0045	0.0048	0.0022	0.0043

Table 3-5 Summary Statistics for Tests of CAPM and the Single ETF-factor Model in Different Period

This table compares the model performance of CAPM and the Single ETF-factor Model in three periods: April 2009 to December 2016, January 1998 to March 2009 and January 1998 to December 2016. The sets of dependent excess returns are the same as those in Table 4. The pricing models are CAPM and the single ETF-factor model that is the one-factor model in which ETF_v is the explanatory variable. GRS is the F-statistics of Gibbons, Ross, and Shanken (1989) that test whether the regression intercepts for a set of portfolios are jointly zeros. p(GRS) is the p-value of GRS. Ave$|\alpha|$ is the average absolute value of the intercepts for a set of double-sort portfolios.

	April 2009 to December 2016		January 1998 to March 2009		January 1998 to December 2016	
Models	CAPM	ETF_v	CAPM	ETF_v	CAPM	ETF_v
B/M-Inv Portfolio						
GRS	1.97	1.88	1.73	1.72	1.90	1.80
p(GRS)	0.0150	0.0215	0.0287	0.0298	0.0085	0.0144
Ave$\|\alpha\|$	0.0023	0.0034	0.0026	0.0022	0.0023	0.0030
BM-OP Portfolio						
GRS	1.03	1.17	1.19	0.86	1.51	1.34
p(GRS)	0.4443	0.2990	0.2690	0.6539	0.0650	0.1372
Ave$\|\alpha\|$	0.0032	0.0032	0.0026	0.0021	0.0023	0.0027
OP-Inv Portfolio						
GRS	1.41	1.63	1.36	1.54	1.37	1.38
p(GRS)	0.1327	0.0584	0.1404	0.0678	0.1198	0.1169
Ave$\|\alpha\|$	0.0025	0.0029	0.0028	0.0027	0.0024	0.0028
Size-B/M Portfolio						
GRS	0.94	1.19	2.69	2.28	2.98	2.69
p(GRS)	0.5562	0.2818	0.0002	0.0019	0.0000	0.0001
Ave$\|\alpha\|$	0.0017	0.0024	0.0033	0.0025	0.0026	0.0033
Size-Inv Portfolio						
GRS	1.47	1.65	1.67	1.57	2.23	1.96
p(GRS)	0.1090	0.0540	0.0374	0.0591	0.0012	0.0059
Ave$\|\alpha\|$	0.0018	0.0025	0.0035	0.0027	0.0030	0.0036
Size-Mom Portfolio						
GRS	1.71	1.86	1.32	1.40	1.76	1.75
p(GRS)	0.0434	0.0231	0.1668	0.1211	0.0183	0.0189
Ave$\|\alpha\|$	0.0022	0.0030	0.0038	0.0035	0.0033	0.0038
Size-OP Portfolio						
GRS	1.92	2.14	1.56	1.21	1.73	1.63
p(GRS)	0.0180	0.0071	0.0620	0.2438	0.0209	0.0354
Ave$\|\alpha\|$	0.0018	0.0026	0.0035	0.0028	0.0029	0.0034

Table 3-6 Summary Statistics for Tests of the ETF-factor Model

This table shows the statistics for tests of ETF-factor model and investigates the explanatory power of the commodity-bond factor (CMB_v). Panel A tests the ability of the ETF-factor model to explain monthly excess returns on *B/M-Inv* portfolios, the 25 *BM-OP* portfolios, 25 *OP-Inv* portfolios, 25 *Size-B/M* portfolios, 25 *Size-Inv* portfolios, 25 *Size-Mom* portfolios and 25 *Size-OP* portfolios. Panel B shows the results of using the commodity-bond factor (CMB_v) to explain the risk premiums in the five-factor model of Fama and French (2015) and the q-factor model of Hou, Xue and Zhang (2015), except for the excess market return $R_m - R_f$.

Panel A

Models	FF5	ETF-factor
B/M-Inv Portfolio		
GRS	1.85	1.58
p(GRS)	0.0259	0.0728
Ave$\lvert\alpha\rvert$	0.0019	0.0026
BM-OP Portfolio		
GRS	0.86	0.86
p(GRS)	0.6590	0.6480
Ave$\lvert\alpha\rvert$	0.0025	0.0025
OP-Inv Portfolio		
GRS	1.20	1.34
p(GRS)	0.2760	0.1718
Ave$\lvert\alpha\rvert$	0.0018	0.0023
Size-B/M Portfolio		
GRS	0.97	1.06
p(GRS)	0.5138	0.4073
Ave$\lvert\alpha\rvert$	0.0010	0.0015
Size-Inv Portfolio		
GRS	1.56	1.30
p(GRS)	0.0796	0.1964
Ave$\lvert\alpha\rvert$	0.0013	0.0015
Size-Mom Portfolio		
GRS	1.61	1.54
p(GRS)	0.0670	0.0848
Ave$\lvert\alpha\rvert$	0.0019	0.0019
Size-OP Portfolio		
GRS	1.76	1.96
p(GRS)	0.0374	0.0156
Ave$\lvert\alpha\rvert$	0.0012	0.0018

Panel B

	Int	CMB_v	R^2
SMB	0.0027	0.0752	0.03
t-stat	1.28	1.55	
HML	0.0021	0.0223	0.00
t-stat	0.81	0.59	
RMW	0.0005	-0.0532	0.04
t-stat	0.33	-2.01	
CMA	0.0019	-0.0036	0.00
t-stat	1.29	-0.13	
MOM	-0.0044	-0.1658	0.04
t-stat	-0.80	-2.00	
MOM	0.0000	0.1241	0.07
t-stat	0.00	2.55	
$q\text{-}SMB$	0.0029	0.0704	0.03
t-stat	1.37	1.37	
$q\text{-}I/A$	0.0021	-0.0208	0.01
t-stat	1.30	-0.70	
$q\text{-}ROE$	-0.0014	-0.1113	0.07
t-stat	-0.55	-2.40	

Table 3-7 Intercepts in regressions for 25 B/M-Inv portfolios and 25 Size-OP portfolios

According to Fama and French (2015), the 25 *B/M-Inv* portfolios are the intersections of 5 portfolios formed on the ratio of book equity to market equity (BE/ME) and 5 portfolios formed on investment (Inv), and the 25 *Size-OP* portfolios are the intersections of 5 portfolios formed on size (market equity, ME) and 5 portfolios formed on profitability (OP). The Model FF 5 is the five-factor model of Fama and French (2015) and the ETF-factor model is the two-factor model that contains the ETFs market factor and the commodity-bond factor. Panel A shows the intercepts and their t-statistics in regressions for 25 *B/M-Inv* portfolios. Panel B shows the intercepts and their t-statistics in regressions for 25 *Size-OP* portfolios.

Panel A: *B/M-Inv Portfolio*

	Low Inv	2	3	4	High Inv	Low Inv	2	3	4	High Inv
Model: FF5										
			Alpha					t(Alpha)		
Low B/M	0.0007	-0.0001	0.0008	-0.0025	0.0019	0.37	-0.05	0.56	-1.49	0.98
2	0.0012	0.0009	-0.0004	0.0004	0.0016	0.57	0.76	-0.29	0.29	0.82
3	0.0019	0.0029	-0.0017	0.0028	-0.0039	0.84	1.21	-1.09	1.81	-1.52
4	-0.0015	-0.0004	0.0025	0.0037	-0.0012	-0.64	-0.19	1.24	1.47	-0.48
High B/M	1.0518	1.2344	1.0606	1.1771	1.0952	0.29	-1.47	1.33	0.09	-2.84
Model: ETF-factor										
			Alpha					t(Alpha)		
Low B/M	0.0025	0.0017	0.0029	-0.0020	0.0021	0.93	0.99	1.69	-0.98	0.68
2	0.0037	0.0031	0.0011	0.0012	0.0017	1.26	1.68	0.54	0.64	0.66
3	0.0040	0.0046	-0.0015	0.0037	-0.0027	1.45	1.85	-0.68	1.97	-1.04
4	0.0009	0.0007	0.0038	0.0041	-0.0009	0.27	0.23	1.69	1.90	-0.33
High B/M	-0.0004	-0.0042	0.0012	-0.0010	-0.0080	-0.08	-0.96	0.24	-0.26	-2.03
Model: q-factor										
			Alpha					t(Alpha)		
Low B/M	0.0009	0.0005	0.0017	-0.0018	0.0036	0.51	0.39	1.16	-1.09	1.60
2	0.0026	0.0014	0.0004	0.0008	0.0020	1.05	0.99	0.29	0.57	0.96
3	0.0037	0.0031	-0.0010	0.0041	-0.0032	1.44	1.22	-0.66	2.80	-1.28
4	0.0000	0.0003	0.0037	0.0043	-0.0006	0.02	0.13	1.77	1.99	-0.24

High B/M	0.0010	-0.0032	0.0021	-0.0001	-0.0072	0.31	-1.07	0.61	-0.02	-2.15

Panel B: *Size-OP Portfolio*

	Low Inv	2	3	4	High Inv	Low Inv	2	3	4	High Inv
Model: FF5										
			Alpha					t(Alpha)		
Low B/M	-0.0018	0.0041	0.0005	0.0022	-0.0027	-1.14	3.19	0.39	1.27	-1.63
2	0.0005	-0.0002	0.0016	0.0015	-0.0006	0.34	-0.18	1.49	1.23	-0.35
3	-0.0013	0.0010	0.0004	-0.0004	0.0002	-0.75	0.87	0.37	-0.27	0.10
4	-0.0021	0.0028	0.0016	0.0021	-0.0001	-1.05	2.19	1.13	1.37	-0.04
High B/M	-0.0018	0.0041	0.0005	0.0022	-0.0027	-0.34	-0.34	0.26	0.98	-0.57
Model: ETF-factor										
			Alpha					t(Alpha)		
Low B/M	-0.0028	0.0042	0.0004	0.0029	-0.0020	-0.65	1.18	0.12	0.77	-0.52
2	-0.0017	0.0000	0.0017	0.0012	-0.0002	-0.39	0.00	0.56	0.39	-0.06
3	-0.0029	0.0019	0.0012	0.0008	0.0017	-0.71	0.68	0.46	0.29	0.63
4	-0.0040	0.0028	0.0019	0.0031	0.0011	-1.15	1.24	0.96	1.63	0.48
High B/M	-0.0015	-0.0001	0.0010	0.0020	0.0018	-0.47	-0.05	0.75	1.56	1.04
Model: q-factor										
			Alpha					t(Alpha)		
Low B/M	-0.0011	0.0046	0.0010	0.0034	-0.0007	-0.81	2.84	0.60	1.71	-0.31
2	0.0007	0.0003	0.0026	0.0030	0.0015	0.35	0.23	1.77	2.08	0.77
3	-0.0008	0.0013	0.0015	0.0014	0.0016	-0.41	0.98	1.30	0.97	0.92
4	-0.0018	0.0032	0.0023	0.0032	0.0012	-0.86	2.50	1.67	2.45	0.80
High B/M	-0.0012	-0.0009	0.0009	0.0012	0.0009	-0.46	-0.69	0.98	1.79	0.63

Table 3-8 Slopes in regressions for 25 B/M-Inv portfolios and 25 Size-OP portfolio

According to Fama and French (2015), the 25 *B/M-Inv* portfolios are the intersections of 5 portfolios formed on the ratio of book equity to market equity (BE/ME) and 5 portfolios formed on investment (Inv), and the 25 *Size-OP* portfolios are the intersections of 5 portfolios formed on size (market equity, ME) and 5 portfolios formed on profitability (OP). The Model FF 5 is the five-factor model of Fama and French (2015) and the ETF-factor model is the two-factor model that contains the ETFs market factor and the commodity-bond factor. In the FF5 model, $R_m - R_f$ is the excess market return, SMB is the size factor, HML is the value factor, RMW is the profitability factor and CMA is the investment factor. In the ETF-factor model, ETF_v is the excess ETFs market return and CMB_v is the commodity-bond return. Panel A shows the slopes and their t-statistics in regressions for 25 *B/M-Inv* portfolios. Panel B shows the slopes and their t-statistics in regressions for 25 *Size-OP* portfolios.

Panel A: *B/M-Inv Portfolio*

	Low Inv	2	3	4	High Inv	Low Inv	2	3	4	High Inv
Model: FF5										
			$R_m - R_f$					t($R_m - R_f$)		
Low B/M	1.1519	0.9206	0.9319	1.0178	1.0954	21.80	26.14	22.95	23.18	14.86
2	1.0193	0.8905	0.9723	0.9915	1.0856	12.56	19.70	16.73	23.40	15.73
3	0.9954	0.9247	1.0435	0.9773	1.1232	10.76	11.84	19.18	18.57	15.99
4	1.0205	0.9815	0.8999	0.9572	0.8893	15.79	14.90	15.77	10.73	11.66
High B/M	1.0518	1.2344	1.0606	1.1771	1.0952	14.50	22.16	12.73	15.72	12.00
			SMB					t(SMB)		
Low B/M	0.0443	-0.1469	-0.1573	0.0449	0.0391	0.63	-1.99	-2.56	0.58	0.74
2	0.0724	-0.0272	-0.0266	-0.0542	-0.0698	0.74	-0.39	-0.37	-0.70	-0.73
3	0.0620	0.0486	0.1257	0.0674	0.1067	0.66	0.54	1.49	0.76	1.06
4	0.2825	-0.0915	0.1115	0.1399	0.1518	2.66	-1.00	1.16	1.52	1.39
High B/M	0.2633	0.2802	0.3642	-0.0130	0.2945	2.02	2.00	3.04	-0.10	1.86
			HML					t(HML)		
Low B/M	-0.3712	-0.3679	-0.2234	-0.2056	-0.2129	-4.09	-5.28	-3.00	-2.25	-2.32
2	-0.1638	-0.2464	-0.1499	-0.0738	0.1956	-1.09	-3.14	-1.69	-0.98	1.68
3	-0.0165	-0.2980	0.0900	0.2327	0.3463	-0.11	-2.29	0.93	1.56	4.90
4	0.3849	0.0253	0.1883	0.7628	0.5392	2.95	0.21	1.91	4.27	3.70

High B/M	1.0142	0.8751	1.0419	1.0725	0.7052	6.76	6.57	7.39	6.39	4.19
			RMW					t(RMW)		
Low B/M	-0.2109	0.1462	0.1110	0.1916	0.0635	-2.14	1.76	1.18	1.39	0.42
2	0.1323	0.0176	0.1789	0.0042	-0.0779	1.08	0.19	2.41	0.04	-0.47
3	0.2306	-0.2334	0.0078	0.2351	-0.0606	1.25	-1.36	0.06	1.62	-0.49
4	0.1491	-0.1545	0.3006	0.1842	-0.0506	0.96	-1.15	2.61	1.22	-0.36
High B/M	-0.3358	-0.3565	-0.5470	-0.4392	-0.3757	-1.84	-1.81	-3.69	-2.22	-1.46
			CMA					t(CMA)		
Low B/M	0.5544	0.3277	0.0477	-0.2555	-0.8686	2.98	2.67	0.35	-1.60	-5.43
2	0.7995	0.6972	0.4352	-0.1995	-0.5673	3.44	5.90	2.88	-1.47	-2.53
3	0.5795	0.5616	0.3498	-0.3912	-0.7372	1.99	3.09	2.24	-1.38	-5.05
4	0.5578	0.8109	0.0825	-0.4659	-0.5942	1.89	3.73	0.42	-1.14	-2.54
High B/M	-0.1878	-0.6415	-0.4604	-0.7570	-0.4741	-0.84	-2.85	-1.61	-2.99	-1.80
Model: ETF-factor										
			ETF_v					t(ETF_v)		
Low B/M	1.3223	0.9619	0.9482	1.1630	1.1862	15.13	16.05	13.45	12.55	12.15
2	1.1825	0.9962	1.1183	1.1104	1.2735	10.79	11.85	18.18	15.69	16.23
3	1.1701	1.0673	1.3664	1.1596	1.3171	10.05	9.27	19.93	10.62	14.66
4	1.3176	1.2127	1.1011	1.2755	1.1426	10.31	12.69	15.84	7.79	11.73
High B/M	1.5626	1.7084	1.6094	1.6009	1.5227	14.10	14.42	12.89	11.41	12.41
			CMB_v					t(CMB_v)		
Low B/M	-0.0737	-0.0774	-0.0326	-0.1264	-0.0690	-1.53	-2.18	-1.03	-2.25	-1.30
2	-0.0944	-0.0652	-0.1258	-0.0870	-0.0912	-1.85	-1.91	-3.16	-1.95	-1.76
3	-0.1089	-0.0400	-0.1640	-0.0729	0.0183	-2.03	-0.87	-3.38	-1.48	0.38
4	-0.0728	-0.1541	-0.0821	-0.0804	-0.0303	-0.78	-2.51	-2.24	-1.44	-0.61
High B/M	-0.1896	-0.1598	-0.1901	-0.1434	-0.0683	-2.04	-1.98	-2.05	-1.81	-0.73

Panel B: ***Size-OP Portfolio***

	Low OP	2	3	4	High OP	Low OP	2	3	4	High OP
Model: FF5										
			$R_m - R_f$					t($R_m - R_f$)		
Small	0.9797	0.8183	0.8437	0.9177	1.0058	20.97	25.06	21.78	17.78	20.03
2	1.0469	0.9630	0.9366	0.9489	1.1078	27.51	27.04	39.83	21.43	21.66
3	1.0707	1.0076	1.0086	1.0376	1.1381	25.18	27.77	32.69	33.50	27.14
4	1.1510	0.9943	0.9548	1.0095	1.0999	19.56	31.93	21.34	24.41	27.90
Big	1.1407	0.9983	0.9881	0.9824	0.9719	17.78	26.35	36.19	57.67	41.05
			SMB					t(SMB)		
Small	1.0956	0.9906	0.9456	0.9313	1.1193	13.10	17.96	16.50	10.85	14.95
2	0.9206	0.8680	0.8770	1.0069	1.0939	14.21	14.91	17.04	13.89	12.47
3	0.8239	0.6566	0.6919	0.6771	0.5883	11.15	12.59	12.45	13.44	8.16
4	0.3754	0.3341	0.3226	0.3529	0.3199	4.06	5.79	4.73	6.33	4.98
Big	-0.2408	-0.2255	-0.1214	-0.1933	-0.1531	-3.52	-4.19	-2.24	-5.27	-3.79
			HML					t(HML)		
Small	-0.0733	0.3588	0.4319	0.3155	0.5012	-0.81	3.89	4.90	3.10	5.55
2	-0.4409	0.1574	0.1894	0.2524	0.2562	-6.77	1.85	3.54	3.09	3.47
3	-0.2549	0.0359	0.0500	0.2477	0.0279	-3.51	0.42	0.86	3.01	0.34
4	-0.0677	0.1881	0.0667	0.0487	0.0898	-0.53	2.39	1.04	0.66	1.29
Big	0.2994	0.2061	-0.0230	-0.0336	-0.1415	2.75	3.80	-0.45	-0.82	-3.44
			RMW					t(RMW)		
Small	-0.6455	0.0750	0.1692	0.2003	0.4276	-6.24	0.84	2.01	1.68	3.45
2	-0.9989	0.0247	0.2033	0.3877	0.4951	-10.08	0.24	2.84	3.78	4.49
3	-0.7422	-0.1620	0.2006	0.4633	0.3828	-6.38	-1.74	2.61	5.38	4.10
4	-0.7687	-0.0846	0.0449	0.1405	0.2035	-5.73	-1.07	0.43	1.54	2.04
Big	-0.8785	-0.3592	-0.0059	-0.0069	0.4552	-6.56	-4.98	-0.08	-0.14	7.52
			CMA					t(CMA)		

Small	-0.0676	-0.1663	-0.0871	-0.0584	-0.0443	-0.47	-1.34	-0.70	-0.36	-0.28
2	0.0878	-0.0677	-0.0678	-0.3113	-0.3883	0.64	-0.60	-0.79	-2.68	-3.06
3	0.0212	0.0980	0.0453	-0.3109	-0.0045	0.15	0.78	0.52	-2.56	-0.04
4	-0.2536	-0.1370	-0.1635	-0.1211	-0.1269	-1.54	-1.09	-1.20	-1.12	-0.97
Big	-0.4910	0.0006	0.1634	0.0655	-0.0239	-2.39	0.01	1.76	1.29	-0.37

Model: ETF-factor

	ETF_v					t(ETF_v)				
Small	1.5100	1.2728	1.3298	1.3506	1.5447	13.56	12.71	12.53	14.21	11.40
2	1.5662	1.3911	1.3681	1.4241	1.5981	13.68	16.50	16.80	11.40	13.20
3	1.5625	1.3622	1.3695	1.3721	1.4514	14.71	18.41	18.07	15.25	18.68
4	1.5729	1.3132	1.2199	1.2642	1.3606	17.28	23.82	20.44	18.88	23.42
Big	1.3779	1.1803	1.1492	1.0769	0.9866	17.75	18.46	28.14	19.69	14.02
	CMB_v					t(CMB_v)				
Small	-0.0411	-0.0798	-0.1188	-0.0957	-0.1275	-0.55	-1.10	-1.66	-1.32	-1.32
2	-0.1095	-0.1005	-0.1266	-0.1473	-0.1312	-1.49	-1.58	-2.02	-1.70	-1.64
3	-0.1150	-0.0509	-0.1037	-0.0935	-0.0977	-1.52	-0.87	-1.48	-1.37	-1.67
4	-0.1371	-0.1184	-0.0923	-0.0643	-0.0911	-2.79	-2.64	-2.74	-1.34	-1.87
Big	-0.0492	-0.1099	-0.1163	-0.0734	-0.0758	-1.06	-3.33	-3.75	-3.48	-2.66

Table 3-9 Empirical Tests of Asset Pricing Models

This table shows the coefficients, Shanken t-statistics and other statistics from cross-sectional regressions of average excess returns on factor loadings for four models. The test assets are Fama and French's 25 Size-B/M portfolios. The cross-sectional T^2 statistic tests whether pricing errors in the cross-sectional regression are jointly zero, with p-values in parentheses. The R^2 is the GLS R^2 of Lewellen, Nagel and Shanken (2010), with simulated values in brackets. The sample estimate of q, the difference between the maximum generalized squared Sharpe ratio and that attainable from the mimicking portfolios, is reported and the simulated values are in brackets. The GRS statistic that tests whether the intercepts in time-series regressions are jointly zero is reported in the last several rows, with p-values in parentheses and simulated values in brackets. Ninety-five percent confidence intervals for the true R^2, q and GRS statistics are reported next to corresponding sample values. The models are estimated in two periods: April 2009 to December 2016 and February 1973 to October 2017. Model FF5 is the five-factor model of Fama and French (2015) that includes five variables. Model1 is an ETF-factor model that includes three variables: ETF_v, EB_v and EC_v. Model2 is the ETF-factor model that is presented in equation 3.1. CAPM is the single-factor model that includes the market factor as the variable.

Size-BM portfolio

	April 2009 to December 2016				February 1973 to October 2017			
	FF5	Model1	Model2	CAPM	FF5	Model1	Model2	CAPM
Coeff	0.0176	0.0134	0.0137	0.0145	0.0135	0.0150	0.0143	0.0140
	-0.0037	0.0016	0.0011	-0.0005	-0.0079	-0.0092	-0.0085	-0.0082
	0.0023	0.0011	0.0057		0.0025	-0.0017	0.0038	
	0.0016	0.0053			0.0034	-0.0027		
	0.0023				0.0031			
	-0.0005				0.0016			
t-stat	2.43	2.36	2.08	2.47	4.02	3.23	2.63	3.31
	-0.44	0.21	0.13	-0.07	-2.03	-1.85	-1.52	-1.78
	0.86	0.45	0.20		1.85	-0.07	0.15	
	0.64	0.18			2.63	-0.72		
	0.94				1.58			
	-0.23				0.73			
T^2	18.71	23.23	23.94	24.39	54.73	67.05	69.21	69.90
	(0.48)	(0.33)	(0.35)	(0.38)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.21	0.03	0.01	0.00	0.33	0.25	0.21	0.20
	[0.13,	[0.02,	[0.00,	[0.00,	[0.14,	[0.06,	[0.04,	[0.02,
	0.83]	0.58]	0.42]	0.15]	0.81]	0.78]	0.73]	0.69]
q	0.22	0.25	0.26	0.26	0.11	0.13	0.13	0.13
	[0.01,	[0.01,	[0.01,	[0.02,	[0.01,	[0.01,	[0.01,	[0.02,
	0.32]	0.32]	0.34]	0.36]	0.31]	0.31]	0.31]	0.32]
GRS	1.15	0.97	1.25	1.12	3.06	3.94	4.24	4.27
	(0.32)	(0.51)	(0.23)	(0.34)	(0.00)	(0.00)	(0.00)	(0.00)
	[1.48,	[1.33,	[1.65,	[1.36,	[2.69,	[3.74,	[3.84,	[3.92,
	5.65]	5.07]	5.62]	5.13]	5.84]	7.30]	7.56]	7.29]

4. The Performance of Equity and Fixed-Income Funds

4.1. Introduction

Mutual fund performance receives a lot of attention in financial literature. However, most previous work focused on just equity mutual fund performance. As a result, numerous well-developed and mathematically elegant performance measurement models for equity mutual funds are available (see, e.g., the five-factor model of Fama and French (2015), the q-factor model of Hou *et al.* (2015), and the four-factor model of Carhart (1997)). These models are constructed to capture directly the documented pricing anomalies related to size, book-to-market, momentum, and so on. The related literature shows that these models provide significantly improved explanatory power over the CAPM of Sharpe (1964) and Lintner (1965). However, the 2007-08 financial crisis changed the U.S economy and the global economy in significant ways. In the post-financial era, the effectiveness of these asset pricing models is facing challenges. Meanwhile, the investment vehicle innovations never stop. The fast-growing market for ETFs is a typical example. Chapter 3 exploits the unique characteristics of ETFs and construct the ETF-factor model, in which both factors are derived from ETFs. Thus, this study also employs the ETF-factor model as a performance measurement model for equity mutual funds.

Bonds are an important asset class, and there are a large literature on asset pricing for bonds. The expected corporate bond returns receive a lot of attentions from the asset pricing literature. Economists have done a lot of research to explain and forecast the term structure of government bond yields. So far, there has been no consensus on what models are best to price the expected returns on all types of bonds. According to previous work, two basic performance measurement models are frequently employed. The first is the three-factor model of Blake et al. (1993). This model and its extended models are widely used to measure bond performance (see, e.g., Elton *et al.* (1995), Gutierrez *et al.* (2009), Cici and Gibson (2012)). The three factors in the original model are the excess return on the overall bond market (the bond market factor), the excess return on a portfolio of high yield bonds (the default factor) and the excess return on a portfolio of mortgage-backed securities (the option premium). The mortgage-backed securities behave differently from other government securities due to their option features. On the other hand, Fama and French (1993) find that the term factor that captures the unexpected changes in interest rates and the default factor both have strong explanatory power on the expected bond returns. In most papers, these two factors are written as $TERM_t$ and DEF_t , respectively. Gebhardt *et al.* (2005) find that the term and default factors are successful in explaining a cross-section of corporate bond returns. These two factors are widely employed in studies of bond pricing (see, e.g., Houweling *et al.* (2005), Lin *et al.* (2011), Jostova *et al.* (2013)). In addition, these studies also find significant price momentum and liquidity risk in U.S corporate bond returns. However, this thesis intends to propose a more general performance measurement technique for all manner of fixed-income funds. Thus, the momentum impact in the corporate bond returns will be ignored for now, and the TED spread will be employed as the proxy for the liquidity risk.

This thesis' main results indicate that this study is important. While measured results are interesting for showing mutual fund performance, they can be more interesting because they provide an additional test of the relative pricing model. This study finds that the ETF-factor and q-factor models tend to produce similar decisions on the performance of US domestic equity mutual funds. If measured by the conventional models, more equity funds exhibit abnormal performance.[23] The bootstrap approach (Fama & French 2010) reveals that the ETF-factor and q-factor results are more reasonable if it is assumed that extremely superior and inferior managers are rare. The conventional models (CAPM, FF3, C4, and FF5)[24] tend to overestimate the number of inferior managers. Generally, all models agree that, in aggregate, US domestic equity mutual funds exhibit no abnormal performance relative to passive benchmarks, which is in line with Fama and French (2010).

On the other hand, this thesis proposes a new two-factor model that can explain the investment strategy followed by fixed-income fund managers. The two factors are the proxy for bond market ($EB_{V,t}$), i.e., the monthly excess return on a portfolio of bond ETFs; and the slope factor ($SLOPE_t$),

[23] The conventional models include the five-factor model of Fama and French (2015), the four-factor model of Carhart (1997), the three-factor model of Fama and French (1993), and CAPM.

[24] FF3 is short for Fama and French's three-factor model, C4 is short for Carhart's four-factor model and FF5 is short for Fama and French's five-factor model.

i.e., the spread between 10-year and 2-year Treasury bond monthly returns. These two factors dominate other explanatory returns in previous work, such as the term and default factors of Fama and French (1993). Compared to the three-factor model of Blake *et al.* (1993), another advantage of the current model is that $EB_{V,t}$ is explicitly defined as one that captures the systematic risk in the bond market. Thus, fixed-income fund managers avoid choosing the appropriate index. The proxy for the bond market ($EB_{V,t}$) alone captures most of volatility in the fixed-income fund market, which supports the validity of $EB_{V,t}$ as the bond-market factor. For corporate and government fixed-income bonds, the simulations indicate that both superior and inferior fixed-income fund managers are observed. But the wealth invested in the fixed-income funds does not produce any benefits or losses relative to the passive benchmark.

This chapter is organized as follows: Section 2 describes the fund sample and explanatory returns; Section 3 summarizes the benchmarks for evaluating equity and fixed-income funds; the aggregate portfolio performance and individual fund performance are discussed in Sections 4 and 5, respectively. The evidence produced by bootstrap simulations is shown in Section 6, and Section 7 concludes the chapter.

4.2. Data

The ETFs database used covers five types of US ETFs during the period April 2009 to December 2016. The relevant monthly data for total return index, net asset value, and outstanding shares are obtained from the Bloomberg Terminal. My mutual fund database covers all 58,710 mutual funds that are traded in the US market, which is from the CRSP (Centre for Research in Security Prices) database. The CRSP data start in 1962, but this thesis concentrates on the period after the latest financial crisis. In this mutual fund sample, only funds that invest primarily in the US and funds that invest primarily in fixed-income securities are used. In addition, only those funds that have at least 36 months of returns are included. Based on the CRSP Objective Code, the final sample contains monthly total returns and total net asset values data on 13,271 domestic equity funds and 5,700 fixed-income securities. The particular corporate fixed-income funds (564), government fixed-income funds (781) and money market fixed-income funds (1,654) are selected for further use.

As for the explanatory returns, the excess market return ($R_m - R_f$), the size factor (SMB), the value factor (HML), the momentum factor (MOM), and risk-free rate (one month treasury bill rate) from Wharton Research Data Services (WRDS) are used. The profitability factor (RMW) and the investment factor (CMA) are obtained from the Kenneth R. French Data Library. Hou, Xue, and Zhang (2015) kindly provided the data on their profitability factor (ROE) and investment factor (I/A). To construct the common factors in bond returns, the monthly 2-year, 5-year, and 10-year US Treasury bond returns were obtained from WRDS. The daily return on the BofA Merrill Lynch US Corporate A Index value and the daily TED spread are downloaded from the Federal Reserve Economic Data (FRED). Then they were converted from daily to monthly in frequency.

4.3. The Regression Framework

The main benchmarks used for evaluating equity fund performance are the ETF-factor model and the five-factor model of Fama and French (2015), but the q-factor model (Hou *et al.* 2015), Carhart's (1997) four-factor model (C4), and the CAPM are also considered. The ETF-factor model and the Fama and French five-factor model (FF5) are, respectively:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_1 ETF_V + \beta_2 CMB_V + \varepsilon_{i,t} \tag{4.1}$$

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_1 (R_m - R_f) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 RMW_t + \beta_4 CMA_t + \varepsilon_{i,t} \tag{4.2}$$

In the regressions, $R_{i,t}$ is the return on fund i for month t; $R_{f,t}$ is the risk-free rate for month t (the one-month US treasury bill rate); R_m is the value-weight return on all NYSE, AMEX, and NASDAQ stocks (from CRSP); ETF_V and CMB_V are the alternative excess market return and commodity-bond return in chapter 3; SMB_t, HML_t, RMW_t, and CMA_t are the size, value, profitability and investment factors of Fama and French (2015); α_i is the unexplained return left by the benchmark model or the abnormal return; and $\varepsilon_{i,t}$ is the error term. The explanatory returns describe the source of risk that the fund on the left may be exposed to. In other words, a linear combination of explanatory returns replicates the return on a comparable portfolio. A significantly positive intercept indicates good performance, and vice versa. The value-weight (hereafter, VW) aggregate of the US equity funds should be very close to the market portfolio. Thus, it is expected to have a zero intercept. The equal weight (hereafter, EW) aggregate of the US equity funds may deviate from the market portfolio because it puts more weights on small equity funds.

In order to evaluate fixed-income fund performance, six proxies for common risks in bond returns are considered: $EB_{V,t}$ and $EE_{V,t}$ from chapter 3; $TERM_t$ and DEF_t from Fama and French (1993); $SLOPE_t$ from Fang and Hung (2014); and and TED_t (TED spread). $EB_{V,t}$ and $EE_{V,t}$ are the value-weight excess returns on a portfolio of bond ETFs and a portfolio of equity ETFs, respectively. According to Fama and French (1993), the factor $TERM_t$ captures the risk of unexpected changes in interest rates, and the factor DEF_t captures the default risk. In this paper, $TERM_t$ is the difference between the monthly 5-year US Treasury bond return and the one-month Treasury bill rate, and DEF_t is the spread between the monthly return on the BofA Merrill Lynch US Corporate A Index value and the monthly 5-year US Treasury bond return. Fang and Hung (2014) define $SLOPE_t$ as the difference between the 10-year and 3-year Treasury bond rates, which is the proxy for the slope of the term structure. However, this thesis' proxy for the slope of the term structure is the difference between the 10-year and 2-year Treasury bond returns. TED_t is the monthly TED spread, i.e., the difference between the 3-month LIBOR based on US dollars and the 3-month Treasury bill rate. TED_t captures the liquidity premium in the bond market. It is found that the two-factor model using the value-weighted excess return on bond ETFs ($EB_{V,t}$) and the slope of the term structure ($SLOPE_t$) is superior for evaluating the fixed-income fund performance. The two-factor model with bond risk factors is:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_1 EB_{V,t} + \beta_2 SLOPE_t + \varepsilon_{i,t} \tag{4.3}$$

But five other models containing subsets of the six factors are also investigated. By comparing the results of the above two-factor model to these models, the impact of including different factors on the ability to explain the portfolio returns can be judged. The five models are:

1. A single-factor model based on the proxy for the bond market ($EB_{V,t}$).

2. A two-factor model employing the proxy for the bond market ($EB_{V,t}$) and the proxy for the stock market ($EE_{V,t}$).

3. A four-factor model including four non-market factors: the level factor ($TERM_t$), the default factor (DEF_t), the slope factor ($SLOPE_t$), and the liquidity factor (TED_t).

4. A five-factor that, in addition to the four non-market factors, incorporates the proxy for the bond market ($EB_{V,t}$).

5. A six-factor model that incorporates all the influences contained in the second and third models.

The first model is the simplest model and is analogous to CAPM. Besides, the proxy for the bond market ($EB_{V,t}$) is the most important factor when explaining bond returns. The second model, which includes two aggregate returns on bonds and stocks, tests the explanatory power of $EE_{V,t}$ on bond returns. The third model explores whether or not the four non-market factors are jointly enough to capture common risks in bond returns. The fourth model combines $EB_{V,t}$ with the four non-market factors. The idea is to see if the inclusion of the proxy for the bond market ($EB_{V,t}$) leads to better results. The fifth model includes all factors, which comprehensively investigates the explanatory power of each factor.

Table 4.1 shows summary statistics for the explanatory returns for the period April 2009 to December 2016. The statistics in Panel A show that the monthly market premium ($R_m - R_f$) has the highest average return, 1.38% per month (t-statistic = 3.43). The monthly ETFs market premium (ETF_V) has a very similar average return, which is 1.06% per month (t-statistic = 2.95). The commodity-bond return (CMB_V) has the smallest average return, -0.57% per month (t-statistic = -0.93). This result is in line with the fact that commodity prices have fallen to the lowest in this century after the 2007-08 financial crisis. Besides, it is observed that other risk premiums (SMB, HML, RMW, and CMA in the FF5 model) have relatively low average returns: respectively, 0.23% per month (t-statistic = 0.88), 0.20% per month (t-statistic = 0.78), 0.09% per month (t-statistic = 0.52), and 0.19% per month (t-statistic = 1.32). I/A and ROE are the investment and profitability factors used by Hou, Xue, and Zhang (2015), and they also have similar average returns: respectively, 0.22% per month (t-statistic = 1.46) and -0.08% per month (t-statistic = -0.30).

[Insert Table 4.1]

The statistics in Panel B provide more information about bond-market factors. $EE_{V,t}$ is the value-weight excess return on a portfolio of equity ETFs. It is an alternative market factor that captures the systematic risk in the stock market. Although it is not a bond market factor, it is used as a control variable in the regression. Apart from $EE_{V,t}$, $EB_{V,t}$ has the second highest average return: 0.36% per month (t-statistic = 4.03). In terms of bond returns, the average returns of $TERM_t$, $SLOPE_t$, and DEF_t are also large; they are, respectively, 0.21% per month (t-statistic = 1.87), 0.21% per month (t-statistic = 1.20), and 0.39% per month (t-statistic = 2.95). TED_t has the smallest average return: 0.02% per month (t-statistic = 20.96). Relative to $EB_{V,t}$, the higher volatility of $TERM_t$, $SLOPE_t$, and DEF_t suggests that they can explain sizeable time-series variation. Additionally, their large average returns indicate that they could explain much of cross-sectional returns on fixed-income portfolios.

The correlation matrix shows that $EB_{V,t}$ has low cross-correlations with other factors. The correlation between $EB_{V,t}$ and $EE_{V,t}$ is 0.53, but the correlation between $TERM_t$ and $EE_{V,t}$ is -0.28. This indicates that $EB_{V,t}$ has stock exposure, while $TERM_t$ captures the flight-to-quality effect. This pattern can also be found in the negative correlation between $SLOPE_t$ and $EE_{V,t}$. Unlike $TERM_t$ and $SLOPE_t$, DEF_t is positively correlated with $EE_{V,t}$, which is in line with Chava and Purnanandam (2010) that default risk is positively related to stock returns economically and statistically. $TERM_t$ and $SLOPE_t$ are highly correlated, which indicates that one of them could be redundant. The liquidity premium, TED_t, is almost not correlated to other factors, which indicates that the liquidity risk is not captured by others or it is trivial.

4.4. The Performance of Aggregate Portfolios

Table 4.2 shows estimates of regressions (4.1), (4.2), and (4.3) in the period April 2009 to December 2016 on EW and VW portfolios. In the VW portfolios, funds are weighted by monthly total net asset value (TNA). In the EW portfolios, funds are weighted equally every month. Panel A of Table 4.2 shows the regression results for equity EW and VW portfolios. The market factor loadings are close to 1.0, which is in line with expectations since the sample consists of domestic equity funds (funds invested primarily in US stocks). In the FF5 and q-factor models, the HML_t,

RMW_t, MOM_t, and ROE_t slopes are very close to zero. This pattern is not only found in the regression to explain the EW portfolio return, but also in the VW portfolio return. This indicates that, in aggregate, US domestic equity funds have little exposure to the value-growth, momentum, and profitability factors. This is similar to Fama and French's (2010) result that, in aggregate, active funds have little exposure to the value-growth and momentum factors for the period January 1984 to September 2006. The EW portfolio produces larger SMB_t slopes (approximately 0.15) than the VW portfolio (approximately 0.05), which indicates that small equity funds have more exposure to small stocks, but, in aggregate, US domestic equity funds invest like the market.

[Insert Table 4.2]

In the spirit of findings of chapter 3, $ETF_{V,t}$ is an alternative market factor that is the excess return on a more diversified market portfolio, and $CMB_{V,t}$ is a more general risk premium. In the regressions to explain the EW and VW portfolios, respectively, the $ETF_{V,t}$ loadings are 1.12 (t-statistic = 36.18) and 1.14 (t-statistic = 39.97), and the $CMB_{V,t}$ loadings are -0.06 (t-statistic = -3.44) and -0.07 (t-statistic = -4.05). Relative to the FF5 and other models, the factor loadings in the two-factor model are more statistically stable. The significantly negative loadings on $CMB_{V,t}$ mean that the aggregate return on US domestic equity funds have exposure to the commodity-bond spread.

The intercepts in the regressions summarize the performance of US domestic equity funds. Interestingly, the FF5, C4, and CAPM intercepts for the EW portfolio return are significantly large and negative; they are, respectively: -0.12% per month (t-statistic = -2.73), -0.13% per month (t-statistic = -3.09), and -0.16% per month (t-statistic = -2.79). In contrast, the ETF-factor and q-factor models tell us that the EW portfolio is not poor. The intercepts in these models are -0.05% per month (t-statistic = -0.50) and -0.05% per month (t-statistic = -1.75), both are trivial and statistically insignificant at the 0.05 level. In terms of VW portfolio performance, the conflicting result is eliminated. The FF5, C4, and CAPM intercepts for the VW portfolio return are, respectively, -0.05% per month (t-statistic = -1.28), -0.05% per month (t-statistic = -1.56), and -0.07% per month (t-statistic = -1.69). Compared to their counterparts for the EW portfolio return, the intercepts are closer to zero and statistically insignificant at the 0.05 level. This thesis' two-factor and the q-factor intercepts for the VW portfolio return are 0.04% per month (t-statistic = 0.44) and 0.03% per month (t-statistic = 1.26), both are trivial and statistically insignificant at the 0.05 level. All models say that the VW portfolio performance is no different from that of the passive benchmarks.

However, it is found that the intercepts generated by FF5, C4, and CAPM are robustly smaller than those generated by this thesis' two-factor model and the q-factor model. The difference could range from 0.08% to 0.11% per month or 0.96% to 1.32% per year, which is very large. This result is consistent no matter how the EW or VW portfolio performances are measured. Apart from this, the VW portfolio produce slopes that are close to zero for non-market explanatory returns, and the market factor $R_m - R_f$, alone, explains 99% of the variability in the VW portfolio monthly returns. In other words, the total wealth invested in US domestic equity funds generates market portfolio returns. The FF5, C4, and CAPM say that the EW portfolio is beaten by the passive benchmarks. The poor performance is attributed to the larger exposure to the size factor, SMB_t.

Panel B of Table 4.2 shows the regression results for fixed-income EW and VW portfolios. In the regressions used to explain the EW portfolio return, the $EB_{V,t}$ slopes are all statistically significant and vary from 0.56 to 0.83. $EB_{V,t,}$, alone, explains 74% of the variability in the EW portfolio monthly returns, indicating that $EB_{V,t}$ plays a role as the market factor in CAPM. It is found that adding the proxy for the stock market ($EE_{V,t}$) or other non-market factors increases the ability of the model to explain the time-series behaviour of the EW portfolio. In the second model, $EE_{V,t}$ shows some explanatory power along with $EB_{V,t}$. In the third model, the $TERM_t$ loading is 0.50 (t-statistic = 5.94), and the DEF_t loading is 0.34 (t-statistic = 9.25); both are statistically large. While the liquidity factor (TED_t) loading is large at 0.40 (t-statistic = 0.99), it is not statistically significant. And the EW portfolio return shows little exposure to the slope factor ($SLOPE_t$). These results are in line with work that shows that $TERM_t$ and DEF_t are the dominant variables in the common variation in bond returns (e.g., Keim and Stambaugh (1986), Fama and French (1989, 1993), Driessen (2004)). Although $TERM_t$ and DEF_t show strong explanatory power on the EW portfolio return, they explain less variability in the EW portfolio monthly returns than $EB_{V,t}$.

Interestingly, when looking at the fourth model, $EB_{V,t}$ and $SLOPE_t$ dominate other the factors, including $TERM_t$ and DEF_t. The $EB_{V,t}$ loading is 0.60 (t-statistic = 7.86), and the $SLOPE_t$ loading is 0.15 (t-statistic = 3.58), both are statistically significant. It is no surprise that $EB_{V,t}$ is the dominant variable, but $SLOPE_t$ shows significant explanatory power along with $EB_{V,t}$ (the $SLOPE_t$ loading is 0.15 (t-statistic = 3.58)). Because $TERM_t$ and $SLOPE_t$ are highly correlated (Table 4.1), the dominant variable $SLOPE_t$ should be kept, and the variable $TERM_t$ should be dropped to avoid multicollinearity. Turning to the regression results for the fifth model, it is found that adding the proxy for the stock market ($EE_{V,t}$) does not improve model performance. Among the six factors, only the $EB_{V,t}$ and $SLOPE_t$ loadings are large and statistically significant, while other factor loadings are trivial. If the four insignificant factors ($EE_{V,t}$, $TERM_t$, DEF_t and TED_t) are dropped, the regression results remain almost unchanged. The intercept in this thesis' two-factor model is zero, indicating that $EB_{V,t}$ and $SLOPE_t$ successfully describe the EW portfolio return. In addition, $EB_{V,t}$ and $SLOPE_t$ explains 82% of the variability in the EW portfolio monthly returns. All of these results provide solid evidence to support the validity of regression (4.3).

The findings above are consistent in explaining the VW portfolio return. For example, $EB_{V,t}$, alone, explains 79% of the variability in the VW portfolio monthly returns, and $EB_{V,t}$ and $SLOPE_t$, together, explain 89%. In the fifth model that includes all six factors, the VW portfolio has little exposure to the factors $EE_{V,t}$, $TERM_t$, DEF_t, and TED_t. The intercept in (4.3) for the VW portfolio return is zero, indicating that the VW portfolio return can be replicated by the passive benchmarks. All these results confirm the robustness of regression (4.3). Compared to the EW portfolio, the VW portfolio has less exposure to $EB_{V,t}$ and $SLOPE_t$. The loadings on these two factors are only half of their counterparts in the regression on the EW portfolio return. It is inferred that smaller fixed-income funds have more exposure to the factor $EB_{V,t}$.

4.5. Regression Results for Individual Funds

4.5.1. Equity Mutual Funds

In this section, the models will be employed to evaluate the performance of each fund. For the performance of equity funds, the focus will mainly be on the results produced by the ETF-factor model proposed in chapter 3 and the FF5 model. But results produced by the C4 model, the q-factor model, and CAPM will also be reported. For the performance of fixed-income funds, the focus will mainly be on the results produced by the regression (4.3), but the results produced by the five models discussed in the regression framework will also be reported. Due to the large number of funds, the average absolute value of the intercepts, the average adjusted R-squared, the percentage of significant intercepts at the 0.05 and 0.01 levels, and the percentages of significant risk factors at the 0.05 level will be reported for each model.

[Insert Table 4.3]

Table 4.3 summarizes the performance of equity funds and the relevant measurement models. The absolute value of intercepts produced by the ETF-factor model and the q-factor model are 0.17% and 0.18% per month, respectively, while the absolute value of intercepts produced by the FF5 model, the C4 model, and CAPM are 0.21%, 0.22% and 0.24% per month, respectively. At the 0.05 level, 8.49% of the ETF-factor model funds and 16.13% of the q-factor model funds have significantly abnormal returns. If measured by the FF5 model, the C4 model, and CAPM, these numbers are 23.98%, 28.06%, and 22.24%, respectively. The pattern that fewer funds have abnormal returns measured by the ETF-factor and q-factor models is consistent at the 0.01 level. For example, the ETF-factor model says that only 3.53% of funds have abnormal returns, while the FF5 model says this number is 11.66% at the 0.01 level.

Measured by the ETF-factor model, 97.80% of ETF_t slopes and 60.33% of CMB_t slopes are statistically significant at the 0.05 level. Almost all equity funds have significant loadings on ETF_t because it is the proxy for market premium. According to Anurag, Chi-Hsiou, and Junqi (2017), the commodity-bond return (CMB_t) well summarizes other risk premiums, excluding the market premium. Thus, it is not surprising that more than half of the equity funds have significant exposure to the factor CMB_t. Measured by the FF5, q-factor, C4, and CAPM models, approximately 98% of $R_m - R_f$ slopes are statistically significant. No matter which model is used, approximately 50% to 55% of SMB_t slopes and 40% to 50% of HML_t slopes are statistically significant. In the C4 model, 41.31% of funds have statistically significant loadings on the momentum premium (MOM_t). Considering the profitability factor loadings, 22.10% of RMW_t loadings are statistically significant, while 40.50% of ROE_t loadings are statistically significant. When looking at the investment factor loadings, 33.54% of CMA_t loadings are statistically significant, while 49.61% of I/A_t loadings are statistically significant. These results suggest that the profitability and investment factors of Hou *et al.* (2015) are more powerful in terms of explaining the expected equity fund monthly returns. As for the average adjusted R-squared, the ETF-factor model produces the smallest (R^2 = 0.81), and the q-factor produces the largest (R^2 = 0.87).

4.5.2. Fixed-Income Mutual Funds

Table 4.4 summarizes the percentage of sensitivities of the 5,700 fixed-income funds for the five testing models and the ETF-factor model. In addition, the percentages of sensitivities of funds grouped into corporate, government, and money market categories are examined. When the results across these models are examined, it is found that there are differences in the sensitivities of funds to the explanatory returns for each model. Panel B of Table 4.4 shows the measured results for 564 corporate fixed-income funds. Almost all corporate fixed-income funds have statistically significant sensitivity to the default factor (DEF_t) and the bond-market factor ($EB_{V,t}$). In the six-factor model, 82.45% of $EB_{V,t}$ slopes and 92.02% of DEF_t slopes are statistically significant at the 0.05 level. If the thesis' two-factor model is used to measure performance, all 564 corporate fixed-income fund returns have exposures to the bond-market factor ($EB_{V,t}$). The high percentages of sensitivities to the default factor and the bond-market factor are relatively constant and independent of the number of risk factors in the regressions.

[Insert Table 4.4]

Panel C shows the measured results for 781 government fixed-income funds. Compared to the results for the corporate fixed-income funds, there is more variation in the sensitivities. In the third model, 95.39% of $TERM_t$ slopes and 80.92% of DEF_t slopes are statistically significant at the 0.05 level. If the bond-market factor ($EB_{V,t}$) is included in the regressions, those numbers become 59.54% and 39.56%, respectively, while 50.96% of $EB_{V,t}$ slopes are statistically significant. However, if the term and the default factors are excluded from the regressions, more than 90% of government fixed-income funds have exposure to the bond-market factor. The percentages of sensitivities to the slope factor ($SLOPE_t$) remain reasonably constant in the third, fourth and fifth models, but increase to 90.40% in this thesis' two-factor model. These results indicate that the bond-market factor ($EB_{V,t}$) and the slope factor ($SLOPE_t$) act, in part, as the term factor and the default factor, respectively.

Turning to the performance of the 1,654 money market funds (Panel D of Table 4.4), the measured results indicate that all models fail to lead plausible conclusions. For example, the average R-squared produced by the two-factor model is only 0.04, which means that most of the variability in the monthly returns is not captured. While the fourth model produces the biggest average R-squared, its value is still very small: 0.14. For all models, the percentage of significant intercepts is in a range of 50.85% to 75.15% at the 0.05 level, and the percentages of sensitivities to all factors (except for the liquidity factor) are relatively low. Money market securities are short-term assets typically with a maturity of one year or less. For example, Treasury Bills (T-bills), commercial Paper, certificates of Deposit (CDs), and so on are typical types of money market securities. Investors in money market securities could face reinvestment risk (unexpected changes in interest rates), counterparty risk (default risk), and liquidity risk. But compared to the risk in bond returns, the reinvestment and counterparty risks in money market securities are trivial. Thus, the explanatory returns are too large to explain the returns on money market funds. However, more than half of the money market fund returns are sensitive to the liquidity premium (TED_t). This result is consistent in the models that include the liquidity premium (TED_t). For example, the six-factor model shows that 57.50% of money market funds have sensitivities to the liquidity premium (TED_t), while only a tiny fraction of them have exposures to other factors. The large number of significant intercepts and the small average R-squared suggest that low average returns on money market funds do not fit with the high average explanatory returns, except in the case of the liquidity premium.

Generally, the pattern of sensitivities is consistent with expectations. Almost all corporate fixed-income funds have statistically significant sensitivity to the default factor (DEF_t). Government fixed-income funds are more likely to have significant sensitivity to the term factor, $TERM_t$, than other categories. For almost half of money market funds, the sensitivity coefficient for the liquidity premium (TED_t) is statistically significant. Moreover, the six-factor model produces larger average R-squared than the two-factor model for funds in each category. Meanwhile, the six-factor model produces larger average absolute values of intercepts than the two-factor model. The performance measurement is based on comparison of mutual fund performance with the performance from a feasible strategy. Thus, the six-factor model provides investors with an investment strategy that captures more risk, while the two-factor model gives one that is superior to replicating the average return.

4.6. Cross-Sectional Bootstrap

Table 4.2 (Panel A) tells us that the total wealth invested in US domestic equity funds does not produce gross returns above or below those of passive benchmarks. But Table 4.3 says that some equity funds produce statistically significant abnormal returns relative to the same passive benchmarks. These results suggest that the performance of managers with superior skills is balanced by the performance of managers with inferior skills. In this section, the existence of superior and inferior managers is investigated using three bootstrap procedures in a world where true α is zero. The bootstrap statistical technique used is the cross-sectional bootstrap of Fama and French (2010). This study considers the average simulated t(α), rather than estimates of α, in its implementation because t(α) normalizes the estimates of α. Furthermore, t(α) can be interpreted as the abnormal return per risk. As Kosowski *et al.* (2006) mention, alphas for a fund having a short life or a risky fund tend to be outliers in the cross-section, but t(α) corrects the appearance of spurious outliers. This study is interested in the cross-section of true α, so it focuses on the cumulative distribution function (CDF) of t(α) estimates for actual fund returns and the average of

1000 simulation CDFs. Through comparing the percentiles of the t(α) estimates for actual returns with the respective average values from the simulations, an inference may be drawn about the existence of skilled managers. This procedure is briefly introduced in the following four steps:

1. A measurement model is used to compute the estimated alphas and the adjusted returns for all funds by subtracting the estimated alpha from a fund's monthly returns;

2. A pseudo time index is created through drawing 93 time periods from the original time sequence;

3. This study jointly resamples the adjusted and explanatory returns for this re-indexed time sequence across all funds;

4. Steps 3 and 4 are repeated for 1000 simulation runs.

The first step creates a world in which true alpha is zero. In other words, every manager only produces expected returns relative to the passive benchmarks. The second step breaks the autocorrelation of fund returns, but preserves the cross-correlation because the pseudo-time index is used for all funds. The third step retains the correlated heteroskedasticity of the explanatory returns and regression residuals because the disturbances of a benchmark model are a component of the adjusted returns. Following the fourth step, the distributions of t(α) are produced for the world in which true alpha is zero. Comparing the simulated distributions with the actual distribution of t(α), the question of whether or not there is a nonzero true alpha can be answered.

4.6.1. Equity Fund Returns

Table 4.5 summarizes the percentiles of t(α) estimates for actual and simulated US domestic equity fund returns during the period April 2009 to December 2016. When the FF5 model is employed as the performance measurement, all t(α) estimates for actual returns are below the respective average values from the simulations. For example, the 5th, 50th and 95th percentiles of the actual t(α) estimates are -3.08, -0.93 and 1.16, respectively, while their counterparts from the simulations are 1.61, -0.02 and 1.58, respectively. In addition, the simulations for all percentiles below the 70th produce a lower value of t(α) at the selected percentiles than those observed for actual returns. At the 99th percentile, only 27.30% of t(α) estimates from the simulations are lower than the t(α) estimates for actual returns. Figure 4.1 shows the actual and average simulated CDFs for the FF5 model. Clearly, the actual CDF is left to the simulated CDF at almost every position. The averages of the percentile values of FF5 t(α)'s from the simulations always beat their counterparts for actual returns. In other words, the funds that produce the positive abnormal expected returns are just lucky. However, the existence of inferior managers is positively based on the simulation tests.

[Insert Table 4.5]

[Insert Figure 4.1]

The ETF-factor results are similar to the FF5 results, but the difference is interesting. Using the ETF-factor model shrinks slightly the left tail of the cross-sections of t(α) estimates for actual returns. Moreover, all ETF-factor t(α) estimates for actual returns below the 99^{th} percentile are higher than their counterparts produced by the FF5 model. For example, the 5th, 50^{th}, and 95th percentiles of the actual FF5 t(α) estimates are -3.08, -0.93 and 1.16, respectively, while the ETF-factor counterparts are -2.21, -0.20, and 1.29, respectively. Compared to the FF5 results, more simulation runs are observed that produce lower values of t(α) than those for actual returns at the selected percentiles. Figure 2 shows the actual and average simulated CDFs for the ETF-factor model. Although the actual CDF is still left of the simulated CDF, the gap between them is smaller relative to the gap shown in figure 4.1.

[Insert Figure 4.2]

Interestingly, the q-factor results confirm the existence of superior managers. The left tail of the q-factor t(α) estimates for actual returns is still to the left of the simulated average, but the right tail of the q-factor t(α) estimates for actual returns moves to the right of the simulated average. For example, the 95th and 99th percentiles of the actual t(α) estimates are 1.65 and 2.54, respectively, while the respective average values from the simulations are 1.62 and 2.29. Apart from this, most values of the t(α) estimate produced by the simulations runs are smaller than those for actual returns at or above the 95th percentile. For example, 77.10% of simulation runs produce lower values of t(α) estimates than those for actual returns at the 99th percentile. Figure 3 shows the actual and average simulated CDFs for the q-factor model.

[Insert Figure 4.3]

Looking further at the percentiles of t(α) estimates for actual and simulated returns using CAPM, the FF3 model, and the C4 model as performance measurements, these models provide similar results to the FF5 results. Generally, they all conclude that there are inferior managers who produce negative true α relative to passive benchmarks and question whether or not there are superior managers who can really beat passive benchmarks. Across these three models, the CDF of t(α) estimates for actual returns is left of the average of the 1000 simulation CDFs at every selected percentile. Thus, the presence of skill is probably really luck. In addition, the tails of the FF5 t(α) estimates for actual returns are fatter than the tails of FF3 or CAPM t(α) estimates for actual returns. This suggests that adding other risk premiums pulls the alpha estimates away from zero.

Finally, the study found that the selection of the benchmark could be controversial. Based on the bootstrap simulations, the FF5, FF3, C4, and CAPM results all show the existence of inferior skill but uncover no evidence of superior skill. On the contrary, the q-factor results provide solid evidence of both superior and inferior skills. In addition, the ETF-factor results support the existence of inferior skill, but also show that the existence of superior skill is possible. Compared with the percentiles of t(α) estimates for actual returns produced by the FF5 model, using the ETF-factor model shrinks the left tail of the cross-sections of t(α) estimates. Compared with the percentiles of t(α) estimates for actual returns produced by the q-factor model, using the ETF-factor model shrinks the right tail. In other words, the ETF-factor alphas are less extreme than the FF5 and q-factor alphas.

4.6.2. Fixed-Income Fund Returns

Table 4.6 summarizes the percentiles of t(α) estimates for actual and simulated fixed-income fund returns during the period April 2009 to December 2016. The performance measurement techniques are the two-factor model described in (3), and the four- and six-factor models described in section 3. For the corporate fixed-income funds, the two-factor results point to the existence of superior managers in corporate fixed-income funds. For example, the 95th percentiles of the actual t(α) estimates are 3.78, while the average t(α) estimate from the simulations is only 1.22. In addition, the CDF of t(α) estimates for actual returns is always right of the simulated CDF at every selected percentile. Thus, these results suggest that when returns are measured by (3), corporate fixed-income fund managers easily produce positive true α relative to passive benchmarks.

[Insert Table 4.6]

However, the six-factor results for the corporate fixed-income fund returns show the evidence of inferior performance. The right tail of six-factor t(α) estimates for actual returns is still to the right of the average from the simulations, but the left tail of six-factor t(α) estimates for actual returns moves to the left of the average from the simulations. The four-factor results are similar to the six-factor results, but the right tail of the cross-sections of t(α) estimates for actual returns shrinks slightly. Therefore, the evidence of superior managers provided by the four-factor results becomes weaker.

Next, the government fixed-income fund performance is examined. The two-factor, six-factor, and four-factor results all show that there are inferior managers who produce negative true α. For instance, the 10th percentiles of the actual t(α) estimates, -2.06 (two-factor), -1.96 (six-factor), and -2.24 (four-factor), are much smaller than the average estimates from the simulation, which are -1.09, -1.23, and -1.19, respectively. On the other hand, inferences about the existence of superior managers vary when different models are used. In the tests that use the two-factor model, all percentiles of the t(α) estimate for actual funds above the 60^{th} percentile are always above the average values from the simulations. However, in the tests that use the six-factor or four-factor models, the percentiles of the t(α) estimates for actual returns are below the average values from the simulations at most selected percentiles. Only at the 98th percentile of the t(α) estimates for actual fund returns do they start to be above the simulated average. This result suggests that very few superior government fixed-income fund managers exist.

For the corporate fixed-income funds, the percentiles of the t(α) estimates for actual returns produced by the two-factor model are always to the right of the corresponding values produced by the six-factor model. For the government fixed-income funds, the left tail of the two-factor t(α) estimates for actual returns is close to the left tail of the six-factor t(α) estimates for actual returns, but the right tail of the two-factor t(α) estimates moves to the right, relatively. In brief, the two-factor model (4.3) tends to be optimistic about the corporate fixed-income fund performance, while the other two models are pessimistic about the government fixed-income fund performance.

4.7. Value Measure of Mutual Fund Performance

So far, this chapter uses the gross alpha (the intercept in the benchmark regression), a return measure, to evaluate the mutual fund performance. Unfortunately is not appropriate to measure the value of a mutual fund. For example, a manager who produces a gross alpha of 0.5% on a $5 billion fund adds more value than a manager who produces a gross alpha of 5% on a $5 million fund. In fact, the gross alpha is a good measure of manager skill only if all funds are the same size. Investors compete to find the managers who produce greater alpha and ultimately drive the alpha to zero. Therefore, the gross alpha is not the right measure of manager skill.

According to Berk and Van Binsbergen (2015), the dollar value of what the fund adds over the benchmark is the correct measure of manager skill. The realized value added is computed as the product of the risk adjusted gross return and the size of a fund. The size of a fund is defined as the total assets under management (AUM). For a fund at time t, the realized value added is between times t-1 and t. For a fund that exists for T period, the monthly average value (S_i) added can be estimated by simply averaging value added every month.25 Therefore, the monthly average value can be used to measure the manager skills of funds that exist in different periods and are different size.

[Insert Table 4.7]

[25] More details about the definition and calculation of the average value added (S_i) can be found in Berk and Van Binsbergen's (2015) paper "Measuring skill in the mutual fund industry". In line with their paper, this chapter uses the same mathematical symbol S_i to represent the average value added of fund i.

Table 4.7 estimates the cross-sectional sample distribution of average value added (S_i). Using the ETF-factors or q-factors to correct for risk, the average fund adds \$210,000 (t-statistic = 4.99) or \$140,000 (t-statistic = 5.01) per month. However, if the risk is adjusted by the FF5 model, the C4 model or the CAPM, the average fund adds -\$290,000 (t-statistic = -14.32), -\$290,000 (t-statistic = -12.85) or -\$380,000 (t-statistic = -16.29) per month. This result is very surprising for two reasons: first, it is very hard to believe that the mutual fund industry as a whole is losing money; second, the value added here is before management fee. If the cost of management is considered, the negative value added will become bigger.

Measured by the ETF-factor model, the fund at the 1st, 10th, 90th, and 99th percentile cutoffs respectively generate -\$4.09 million, -\$0.37 million, \$0.38 million and \$6.63 million per month. If measured by the q-factor model, these funds generate quite similar average value added: -\$4.57 million, -\$0.36 million, \$0.37 million and \$7.42 million per month. However, the results are quite different if measured by the models of FF5, C4 and CAPM. For example, measured by the FF5 model, the fund at the 1st, 10th, 90th, and 99th percentile cutoffs respectively generate -\$4.09 million, -\$0.37 million, \$0.38 million and \$6.63 million per month. The results measured by the C4 model and CAPM are very similar to the results measured by the FF5 model. Furthermore, measured by the ETF-factor or the q-factor model, around 62% funds produce negative average value added. If measured by the FF5, C4 and CAPM, around 80% funds produce negative average value added.

The cross-sectional weighted mean of the average value added (S_i) is reported in table 4.7. If measured by the ETF-factor model or the q-factor model, the weighted mean is greater than the arithmetic mean, indicating that surviving funds outperform. If measured by the FF5, C4 or CAPM, the weighted mean is smaller than the arithmetic mean, indicating that surviving funds underperform. Rohleder, Scholz and Wilkens (2010) have computed, tested and explained different definitions of the survivorship bias in mutual fund performance. They find clear evidence that non-survivor funds underperform significantly. Based on their research, the results produced by the ETF-factor model or the q-factor model seem more plausible.

In addition, table 4.7 estimates the cross-sectional sample distribution of average value added (S_i) for corporate fixed-income funds and government fixed-income funds. The pricing models employed are the two-factor model presented in equation (4.3) and the six-factor model discussed in the section 4.3. Simply judged by the results, the two-factor model (equation 4.3) highly likely overestimate the performance of corporate fixed-income funds due to the missing of default factor in the regression. Over 80% managers produce positive average value added (S_i). If measured by the six-factor model, around 50% managers produce positive average value added (S_i), which seems a plausible result. No matter which model is used to estimate the statistics for the government fixed-income funds, the results are quite similar: around half managers add value.

Berk and Van Binsbergen (2015) has mentioned that the t-statistics in table 4.7 because of two reasons: first, the value added are correlated across funds; second, the distribution of S_i has excess kurtosis. To solve this problem, they propose to test the strong form of the null hypothesis that manager skill is not persistent by using the skill ratio at time t: t-statistic of the average value added (S_i) over the history of the fund until time t.[26] At each time t, funds are sorted into 10 deciles based on the skill ratio. The fund's betas are estimated in the measurement horizon that has a length of 18 months. The horizons of the out-of-sample are: 3 years, 4 years, 5 years, 6 years and 6.75 years.[27] For each decile, the above statistics are computed as well as the mean and standard error of the time series. Further, the corresponding t-statistic is computed.

[Insert Table 4.8]

Table 4.8 reports the statistics about the out-of-sample performance. For the equity mutual funds, the performance persistence is not obvious. Using the ETF-factor model, the persistence is not observed with the sole exception of the four-year horizon. The similar conclusion can be drawn by using the models of FF5, C4 and CAPM. The q-factor specification provides some different results:

[26] More details about the definition of skill ratio and process of computing can be found in the six section of Berk and Van Binsbergen's (2015) paper "Measuring skill in the mutual fund industry". This chapter focuses on analysing the results, not restating the computing process.

[27] To avoid the selection bias, the first 18 observations in the measurement horizon will be dropped when computing the value added. At the end of the out-of-sample period, for example 3 years, funds are again sorted into deciles based on the skill ratio at that time for the next 36 monthly observations. The computing process will be repteated as long as the data allows.

the persistence of fund performance is observed in the three-year, four-year and five-year horizons. Although the estimates of average value added in the top decile are not statistically different from zero, they are almost greater than zero. For example, if measured by the ETF-factor model, the average value added of the top decile at each horizon are 0.72, 1.53, 1.18, 0.74 and 0.65 respectively.

On average, funds in the top decile beat those in the bottom decile frequently. For example, if measured by the ETF-factor model, on average the funds in the top decile have the chance in a range of 56% to 64% to beat the funds in the bottom decile. The difference between the average value added by the funds in top decile and the average value added by the funds in the bottom decile is statistically significant at 0.05 level, no matter in which out-of-sample horizon. But if measured by the FF5 model, no such significant difference is found. The results produced by the models of C4 and CAPM tend to support the significance of the difference, while the results produced by the q-factor model tend to decline the significance of the difference.

The last column reports the average fraction of total AUM in the top decile. On average, the top 10% funds manage about 10% assets. However, small difference exists if using different pricing models. Measured by the ETF-factor model or the q-factor model, the top 10% funds manage about 11.5% assets. Measured by the FF5 model or CAPM, the top funds manage about 9% or 7.5% assets. These results indicate that top funds are more likely to be big using the ETF-factor model or the q-factor model, but top funds are more likely to be small using the FF5 model or CAPM.

On the other hand, the performance of skilled fixed-income fund managers is found to be persistent. The null hypothesis that value added is not persistent is strongly rejected at all horizons. In addition, the top decile almost has a higher value added than the bottom decile and the difference between the average value added by the top and bottom is statistically significant. All these analysis are true for both corporate fixed-income funds and government fixed-income funds. The only difference is: the corporate fixed-income funds in the top decile are big while the government fixed-income funds in the top decile are small, based on the Fraction of Toal AUM reported in the last column.

4.8. Conclusion

This chapter expands the existing literature of mutual fund performance by documenting risk factors and using value measure. The ETF-factors are first used to evaluate the performance of equity mutual funds. Compared to existing risk factors, the ETF-factors are superior because they represent the opportunities on which investors can invest. In addition, this chapter uses the value measure proposed by Berk and Van Binsbergen (2015). The results contribute to the study of Fama and French (2010) on cross-section of mutual fund returns as extra evidence.

Three performance measurement models (the five-factor model of Fama and French (2015), the q-factor model of Hou *et al.* (2015), and the ETF-factor model proposed in chapter 3) were mainly employed to investigate the existence of skilled equity fund managers. These models conclude that the aggregate investment in US domestic equity funds produces no abnormal performance. Thus, if there are superior managers, their performance is balanced by that of inferior managers. In terms of the performance of individual funds, the FF5 results show that 23.98% of funds produce abnormal expected returns at the 0.05 level, while the ETF-factor and q-factor results show only 8.49% and 16.13% of funds, respectively, have significant alphas.

To investigate whether the abnormal performance in fund returns is skill or luck, the percentiles of the $t(\alpha)$ estimates for actual returns were compared with the average values from the cross-sectional bootstrap simulations that set true alpha to zero for every fund. Both the CDFs of the FF5 $t(\alpha)$ estimates and the ETF-factor $t(\alpha)$ estimates are always left of the corresponding CDFs from the simulations. However, the ETF-factor model shrinks the left tail of the $t(\alpha)$ estimates for actual returns, indicating that the ETF-factor results are less extreme. Conversely, the right tail of the q-factor $t(\alpha)$ indicates the existence of superior managers. These results are in line with the performance of the EW portfolio. The CAPM, C4, and FF5 intercepts are at least -0.12% per month, and all are statistically significant. The ETF-factor and q-factor intercepts are -0.05%, but statistically insignificant. Overall, the results indicate that smaller fund managers have insufficient skill to produce the passive returns.

The finding that small fund managers are inferior is in contrast with Chen, Hong, Huang, and Kubik (2004), who state that bigger funds performed worse than small funds during the years 1962 to 1999. The reason is the interaction of liquidity and organizational diseconomies. The most reasonable explanation for this controversy is the different sample period. The mutual fund data used by Chen, Hong, Huang, and Kubik (2004) is from the year 1962 to 1999. Fama and French (2010) select fund data in the period of January 1984 to September 2006. They skip the data during the period 1962 to 1983 because 15% of the funds report only annual returns. Their results provide more evidence of inferior skill of small fund managers, which is in line with the finding of this chapter.

Furthermore, this chapter has developed a two-factor model that successfully measure the performance of value-weight portfolios of fixed-income funds. The proxy for the bond market ($EB_{V,t}$) plays a role as the market factor in CAPM, and the slope factor ($SLOPE_t$) dominates other factors documented in Fama and French (1993), Elton *et al.* (1995), and Fang and Hung (2014). At the aggregate level, 89% of the variability in the monthly VW fixed-income fund returns is captured by the two explanatory returns. The addition of the default factor leads to an improvement in evaluating the performance of individual corporate fixed-income funds. To measure the performance of individual government fixed-income funds, adding the term factor helps capture more variability in monthly returns. When examining the existence of skilled managers for corporate and government fixed-income funds, the cross-sectional bootstrap simulations show that the six-factor results are more reasonable compared to the two-factor results. In the test that uses the six-factor model, the distribution of the t(α) estimates is roughly symmetric around zero.

Evaluating mutual fund performance provides an alternative test of the relevant pricing models. The ETF-factor results are plausible, which confirms the validity of the ETF-factor model. In addition, the two-factor and six-factor models for bonds are good candidates as benchmarks for measuring fixed-income fund performance for two reasons: first, most variation of expected returns of fixed-income funds is captured by the two-factor or six-factor model; second, measured abnormal returns of fixed-income funds, particularly corporate and government fixed-income funds, are in a plausible range.

Table 4-1 Summary Statistics for Monthly Explanatory Returns for the Relevant Pricing Models
R_m is the value-weight return on all NYSE, AMEX, and NASDAQ stocks and R_f is the one-month Treasury bill rate. The construction of SMB, HML, RMW, and CMA follows Fama and French (2015). SMB is the size factor, HML is the value factor, RMW is the profitability factor and CMA is the investment factor. The construction of the momentum factor MOM is in line with Carhart (1997). I/A is the investment factor and ROE is the profitability factor in the q-factor model of Hou, Xue, and Zhang (2015). The proxy for bond market EB_v is the excess return on a value-weight portfolio of bond ETFs. The proxy for bond market EE_v is the excess return on a value-weight portfolio of equity ETFs. $TERM$ is the difference between the monthly 5-year US Treasury bond return and the one-month Treasury bill rate. $SLOPE$ is the difference between the monthly the 10-year and 3-year Treasury bond rates. DEF is the spread between the monthly return on the BofA Merrill Lynch US Corporate A Index value and the monthly 5-year US Treasury bond return. TED is the difference between the 3-month LIBOR based on US dollars and the 3-month Treasury bill rate. This table shows the average monthly return, the corresponding standard deviation, and the t-statistics for the average monthly return in the period of April 2009 to December 2016.

Panel A: Explanatory Returns for Equity

	ETF_v	CMB_v	$R_m - R_f$	SMB	HML	RMW	CMA	I/A	ROE	MOM
Mean	1.06	-0.57	1.38	0.23	0.20	0.09	0.19	0.22	-0.08	-0.35
Std	3.45	5.95	3.88	2.52	2.43	1.58	1.39	1.43	2.50	5.02
t-stats	2.95	-0.93	3.43	0.88	0.78	0.52	1.32	1.46	-0.30	-0.67

Panel B: Explanatory Returns for Bond

	EB_v	EE_v	$TERM$	$SLOPE$	TED	DEF
Mean	0.36	1.21	0.21	0.21	0.24	0.39
Std	0.87	4.08	1.06	1.68	0.11	1.26
t-stats	4.03	2.86	1.87	1.20	20.96	2.95

Correlation Matrix

	EB_v	EE_v	$TERM$	$SLOPE$	TED	DEF
EB_v	1.00					
EE_v	0.53	1.00				
$TERM$	0.47	-0.28	1.00			
$SLOPE$	0.35	-0.45	0.87	1.00		
TED	0.21	0.14	-0.15	-0.12	1.00	
DEF	0.53	0.51	-0.31	-0.22	0.35	1.00

Table 4-2 Alphas and Betas in Regressions for EW and VW Portfolios of US Domestic Equity Funds and US Fixed-Income Funds

Panel A shows the intercepts and slopes for relevant versions of regressions estimated on EW and VW returns on US domestic equity funds. ETF represents the two-factor model derived from ETFs in chapter 3. FF5 represents the five-factor model of Fama and French (2015). The q-factor is the four-factor model of Hou, Xue and Zhang (2015). C4 represents the Carhart (1997) four-factor model and CAPM is the single-factor pricing model. The corresponding explanatory returns in each model is the same as those in Table 4.1. EW indicates that the dependent variable is equal-weight return, and VW indicates that the dependent variable is value-weight return.

Panel B shows the intercepts and slopes for relevant versions of regressions estimated on EW and VW returns on US fixed-income funds. Model 1 is a single-factor model that includes the bond market factor (EB_v). Model 2 is a two-factor model that employs the bond market factor (EB_v) and the alternative stock market factor (EE_v). Model 3 investigates the roles of term, slope, liquidity and default factors on evaluating the aggregate performance of US fixed-income funds. Model 4 not only incorporates the bond market factor (EB_v), but also term, slope, liquidity and default factors. In model 5, all the six factors are included. Eq (4.3) indicates the two-factor model that is explained as regression (4.3).

Apart from the intercepts and slopes, we also report the corresponding t-statistics and the adjusted R^2. On average, there are 13217 US domestic equity funds in our sample. The period is April 2009 through December 2016.

Panel A	EW					VW				
	ETF	FF5	q-factor	C4	CAPM	ETF	FF5	q-factor	C4	CAPM
Intercept	-0.05%	-0.12%	-0.05%	-0.13%	-0.16%	0.04%	-0.05%	0.03%	-0.05%	-0.07%
t-stats	-0.50	-2.73	-1.75	-3.09	-2.79	0.44	-1.28	1.26	-1.56	-1.69
ETF_v	1.12					1.14				
t-stats	36.18					39.97				
CMB_v	-0.06					-0.07				
t-stats	-3.44					-4.05				
$R_m - R_f$		0.93	0.91	0.92	0.97		0.96	0.95	0.96	0.97
t-stats		75.81	119.46	77.82	69.43		97.54	145.71	99.08	102.29
SMB		0.15	0.15	0.15			0.05	0.07	0.05	
t-stats		7.69	13.17	8.19			3.45	6.61	3.57	
HML		0.02		-0.05			0.00		-0.06	
t-stats		0.79		-2.80			0.08		-3.85	
RMW		0.00					0.01			
t-stats		0.14					0.25			

CMA		-0.11					-0.10					
t-stats		-2.89					-3.16					
I/A			-0.04					-0.05				
t-stats			-2.43					-3.37				
ROE			-0.02					0.01				
t-stats			-1.45					1.03				
MOM				-0.03					-0.03			
t-stats				-3.82					-3.71			
R^2	0.95	0.99	1.00	0.99	0.98	0.96	0.99	1.00	0.99	0.99		
Panel B	EW						VW					
	Model 1	Model 2	Model 3	Model 4	Model 5	Eq(4.3)	Model 1	Model 2	Model 3	Model 4	Model 5	Eq(4.3)
Intercept	-0.01%	0.00%	-0.09%	-0.03%	-0.05%	0.00%	0.00%	0.01%	0.00%	0.03%	0.03%	0.00%
t-stats	-0.23	0.13	-0.82	-0.42	-0.55	-0.13	0.06	0.55	0.08	1.01	0.89	0.27
EB_v	0.72	0.83		0.60	0.56	0.63	0.36	0.42		0.32	0.31	0.31
t-stats	16.41	17.80		7.86	4.87	16.13	18.79	21.66		10.85	7.04	20.72
EE_v		-0.05			0.01			-0.02			0.00	
t-stats		-4.57			0.51			-5.80			0.25	
TERM			0.50	-0.02	-0.01				0.27	-0.01	-0.01	
t-stats			5.94	-0.26	-0.12				7.00	-0.35	-0.27	
SLOPE			0.03	0.15	0.17	0.13			0.01	0.07	0.08	0.07
t-stats			0.56	3.58	3.37	6.32			0.27	4.39	3.96	8.76
TED			0.40	0.12	0.16				0.03	-0.12	-0.11	
t-stats			0.99	0.37	0.48				0.18	-0.98	-0.88	
DEF			0.34	0.03	0.04				0.17	0.00	0.00	
t-stats			9.25	0.60	0.72				9.83	-0.10	-0.02	
R^2	0.74	0.79	0.69	0.82	0.82	0.82	0.79	0.85	0.73	0.88	0.88	0.89

Table 4-3 Individual US Domestic Equity Fund Performance measured by ETF-factor, FF5, q-factor, C4 models, and CAPM

This table summarizes the individual equity fund performance and reports the percentage of significant factor loadings in each model. Ave|α|is the average absolute value of the intercept α. Per α (5%) and Per α (1%) are the percentages of the intercept α that are statistically significant at the 0.05 level and 0.01 level. Ave R^2 is the average adjusted R-squared across all regressions. The explanatory returns are the same as those in Table 1 and 2. We report the percentage of the regression slopes that are statistically significant at the 0.05 level in each model. In total, there are 13271 US domestic equity funds in our sample. The period is April 2009 through December 2016.

Models	ETF-factor	FF5	q-factor	C4	CAPM
Summary Statistics for Individual Fund Performance					
Ave|α|	0.17%	0.21%	0.18%	0.22%	0.24%
Per α (5%)	8.49%	23.98%	16.13%	28.06%	22.24%
Per α (1%)	3.53%	11.66%	6.85%	14.51%	10.06%
Ave R^2	0.81	0.86	0.87	0.86	0.82
Percentage of Significant Slopes at the 0.05 Level					
ETF_v	97.80%				
CMB_v	60.33%				
$R_m - R_f$		98.26%	98.25%	98.04%	98.43%
SMB		52.36%	55.68%	54.86%	
HML		42.79%		51.08%	
RMW		22.10%			
CMA		33.54%			
I/A			49.61%		
ROE			40.50%		
MOM				41.31%	

Table 4-4 Individual Fixed-Income Fund Performance measured by Relative Models on Bonds

This table summarizes the individual fixed-income fund performance and reports the percentage of significant factor loadings in each model. Panel A to D report the results for 5700 fixed-income funds, 564 corporate fixed-income funds, 781 government fixed-income funds and 1654 money market fixed-income funds respectively. Ave|α|is the average absolute value of the intercept α. Per α (5%) and Per α (1%) are the percentages of the intercept α that are statistically significant at the 0.05 level and 0.01 level. Ave R^2 is the average adjusted R-squared across all regressions. The explanatory returns are the same as those in Table 1 and 2. We report the percentage of the regression slopes that are statistically significant at the 0.05 level in each model. The period is April 2009 through December 2016.

Panel A: 5700 Fixed-Income Funds						
	Model 1	Model 2	Model 3	Model 4	Model 5	Eq(3)
Summary Statistics for Individual Fund Performance						
Ave|α|	0.10%	0.12%	0.18%	0.14%	0.16%	0.11%
Per α (5%)	33.39%	38.37%	19.88%	21.07%	20.39%	38.18%
Per α (1%)	26.32%	29.84%	13.60%	14.07%	13.18%	29.63%
Ave R^2	0.32	0.41	0.44	0.50	0.50	0.45
Percentage of Significant Slopes at the 0.05 Level						
EB_v	75.19%	72.49%		56.98%	38.96%	76.21%
EE_v		62.00%			5.58%	
$TERM$			55.00%	21.74%	21.18%	
$SLOPE$			46.74%	54.35%	46.81%	70.14%
TED			21.89%	21.84%	21.05%	
DEF			73.02%	26.40%	24.42%	
Panel B: 564 Corporate Fixed-Income Funds						
	Model 1	Model 2	Model 3	Model 4	Model 5	Eq(3)
Summary Statistics for Individual Fund Performance						
Ave|α|	0.10%	0.11%	0.14%	0.13%	0.13%	0.10%
Per α (5%)	21.45%	37.41%	12.41%	15.43%	12.77%	33.87%
Per α (1%)	10.11%	20.39%	4.08%	5.85%	4.96%	16.67%
Ave R^2	0.68	0.78	0.81	0.86	0.86	0.80
Percentage of Significant Slopes at the 0.05 Level						
EB_v	99.11%	99.82%		90.43%	82.45%	100.00%
EE_v		79.08%			9.22%	
$TERM$			97.52%	64.72%	64.36%	
$SLOPE$			52.66%	53.37%	46.81%	83.16%
TED			17.20%	16.13%	14.54%	
DEF			99.65%	90.96%	92.02%	
Panel C: 781 Government Fixed-Income Funds						
	Model 1	Model 2	Model 3	Model 4	Model 5	Eq(3)
Summary Statistics for Individual Fund Performance						
Ave|α|	0.01%	0.01%	0.02%	0.01%	0.01%	0.01%
Per α (5%)	75.15%	75.03%	52.54%	53.08%	50.85%	75.27%
Per α (1%)	66.63%	66.99%	40.69%	40.99%	38.51%	67.11%
Ave R^2	0.02	0.01	0.14	0.14	0.13	0.04

Percentage of Significant Slopes at the 0.05 Level						
EB_v	18.38%	9.37%		3.51%	1.93%	25.15%
EE_v		2.24%			0.79%	
$TERM$			5.93%	9.13%	8.46%	
$SLOPE$			9.07%	14.69%	10.70%	26.54%
TED			57.80%	57.50%	54.78%	
DEF			25.33%	19.95%	19.17%	
Panel D: 1654 Money Market Fixed-Income Funds						
	Model 1	Model 2	Model 3	Model 4	Model 5	Eq(3)
Summary Statistics for Individual Fund Performance						
Ave|α|	0.10%	0.11%	0.16%	0.14%	0.15%	0.09%
Per α (5%)	14.21%	21.77%	17.16%	16.52%	14.72%	24.20%
Per α (1%)	8.19%	12.16%	7.17%	7.68%	6.53%	13.44%
Ave R^2	0.37	0.60	0.72	0.76	0.76	0.68
Percentage of Significant Slopes at the 0.05 Level						
EB_v	97.18%	97.82%		50.96%	49.30%	91.42%
EE_v		89.12%			7.68%	
$TERM$			95.39%	59.54%	59.03%	
$SLOPE$			63.89%	76.44%	63.25%	90.40%
TED			13.44%	14.21%	13.96%	
DEF			80.92%	39.56%	34.96%	

Table 4-5 Percentiles of t(α) Estimates for Actual and Simulated U.S Domestic Equity Fund Return

This table compares the percentiles of t(α) estimates for actual equity fund returns (Act) with those for simulated equity fund returns at selected percentiles (Pct), and shows the percentage of 1000 t(α) estimates for simulated returns that are smaller than the corresponding t(α) estimate for actual returns at the selected percentiles (%<Act). Sim is the mean of the 1000 t(α) estimates for simulated returns at the selected percentiles. The relative models include the FF5, ETF-factor, q-factor, FF3, C4 models, and CAPM. The period is April 2009 through December 2016.

	FF5			ETF-factor			q-factor		
Pct	Sim	Act	%<Act	Sim	Act	%<Act	Sim	Act	%<Act
1	-2.29	-4.14	0.00%	-2.05	-3.20	1.40%	-2.26	-3.75	0.20%
2	-2.01	-3.72	0.00%	-1.80	-2.82	2.30%	-1.99	-3.31	0.30%
3	-1.84	-3.46	0.00%	-1.64	-2.59	3.10%	-1.82	-3.02	0.30%
4	-1.71	-3.24	0.00%	-1.52	-2.37	4.90%	-1.70	-2.80	0.40%
5	-1.61	-3.08	0.00%	-1.43	-2.21	5.90%	-1.60	-2.64	0.40%
10	-1.26	-2.62	0.00%	-1.10	-1.65	14.60%	-1.25	-2.14	0.50%
20	-0.83	-2.04	0.00%	-0.71	-1.02	24.40%	-0.82	-1.51	1.00%
30	-0.53	-1.64	0.00%	-0.43	-0.68	29.60%	-0.51	-1.07	1.30%
40	-0.27	-1.27	0.00%	-0.19	-0.42	31.60%	-0.24	-0.68	2.40%
50	-0.02	-0.93	0.00%	0.03	-0.20	32.30%	0.01	-0.31	6.30%
60	0.22	-0.58	0.00%	0.25	0.00	31.90%	0.26	0.01	11.10%
70	0.48	-0.22	0.00%	0.48	0.23	32.80%	0.53	0.36	23.20%
80	0.79	0.20	0.10%	0.76	0.51	32.70%	0.84	0.74	37.30%
90	1.22	0.71	0.80%	1.16	0.92	34.60%	1.27	1.21	46.40%
95	1.58	1.16	6.20%	1.49	1.29	37.30%	1.62	1.65	57.90%
96	1.68	1.28	7.40%	1.58	1.40	39.40%	1.73	1.79	62.00%
97	1.81	1.45	12.10%	1.70	1.52	38.50%	1.85	1.95	65.60%
98	1.98	1.69	18.90%	1.87	1.72	42.20%	2.02	2.14	66.60%
99	2.27	2.04	27.30%	2.13	2.01	44.20%	2.29	2.54	77.10%
	CAPM			FF3			C4		
Pct	Sim	Act	%<Act	Sim	Act	%<Act	Sim	Act	%<Act
1	-2.14	-3.90	0.10%	-2.23	-4.08	0.00%	-2.29	-4.19	0.00%
2	-1.89	-3.49	0.10%	-1.96	-3.65	0.00%	-2.01	-3.75	0.00%
3	-1.73	-3.27	0.10%	-1.79	-3.39	0.00%	-1.84	-3.47	0.00%
4	-1.62	-3.12	0.10%	-1.67	-3.23	0.00%	-1.71	-3.29	0.00%
5	-1.52	-2.98	0.10%	-1.57	-3.09	0.00%	-1.61	-3.15	0.00%
10	-1.19	-2.53	0.10%	-1.22	-2.63	0.00%	-1.25	-2.68	0.00%
20	-0.79	-1.99	0.10%	-0.80	-2.07	0.00%	-0.83	-2.11	0.00%
30	-0.49	-1.61	0.00%	-0.50	-1.67	0.10%	-0.52	-1.70	0.00%
40	-0.23	-1.29	0.00%	-0.24	-1.34	0.00%	-0.25	-1.36	0.00%
50	0.01	-0.99	0.00%	0.00	-1.01	0.10%	-0.01	-1.03	0.00%
60	0.25	-0.70	0.00%	0.24	-0.70	0.10%	0.24	-0.70	0.00%
70	0.51	-0.40	0.00%	0.50	-0.39	0.00%	0.51	-0.38	0.00%
80	0.80	-0.04	0.00%	0.80	0.00	0.00%	0.82	0.02	0.00%
90	1.21	0.54	0.20%	1.22	0.60	0.20%	1.25	0.62	0.30%

95	1.54	0.96	2.00%	1.57	1.03	2.40%	1.61	1.07	1.80%
96	1.63	1.11	4.20%	1.67	1.18	4.60%	1.71	1.22	4.40%
97	1.75	1.31	9.20%	1.80	1.36	7.20%	1.84	1.38	7.70%
98	1.91	1.52	15.30%	1.97	1.57	12.10%	2.02	1.60	11.20%
99	2.15	1.83	22.20%	2.24	1.88	16.50%	2.30	1.92	15.70%

Table 4-6 Percentiles of t(α) Estimates for Actual and Simulated US Corporate and Government Fixed-Income Fund Returns

This table compares the percentiles of t(α) estimates for actual corporate and government fund returns (Act) with those for simulated counterparts at selected percentiles (Pct), and shows the percentage of 1000 t(α) estimates for simulated returns that are smaller than the corresponding t(α) estimate for actual returns at the selected percentiles (%<Act). Sim is the mean of the 1000 t(α) estimates for simulated returns at the selected percentiles. The relative models are the Eq(3) that includes the bond market factor (EB_v) and the slope factor, the Model 4 that includes the term factor, the slope factor, the liquidity factor and the default factor, and the Model 5 that includes the six factors in the above two models. The period is April 2009 through December 2016.

Corporate Fixed-Income Fund Returns

	Eq(3)			Model 4			Model 5		
Pct	Sim	Act	%<Act	Sim	Act	%<Act	Sim	Act	%<Act
1	-1.83	-1.75	51.40%	-2.07	-2.91	8.60%	-1.76	-3.04	3.00%
2	-1.60	-1.38	65.10%	-1.82	-2.65	8.80%	-1.52	-2.84	2.90%
3	-1.46	-1.03	77.90%	-1.66	-2.58	7.20%	-1.37	-2.74	2.50%
4	-1.35	-0.80	84.00%	-1.54	-2.37	8.70%	-1.26	-2.51	3.30%
5	-1.26	-0.67	86.10%	-1.44	-2.17	10.70%	-1.18	-2.42	3.40%
10	-0.97	-0.14	94.70%	-1.11	-1.73	14.60%	-0.88	-1.90	7.40%
20	-0.63	0.48	97.30%	-0.71	-1.04	26.10%	-0.54	-1.18	17.60%
30	-0.39	0.84	98.00%	-0.42	-0.62	35.30%	-0.30	-0.85	21.40%
40	-0.20	1.13	98.60%	-0.17	-0.25	43.90%	-0.08	-0.53	28.40%
50	-0.01	1.40	98.80%	0.06	0.09	51.80%	0.12	-0.27	31.60%
60	0.16	1.73	99.30%	0.30	0.41	58.40%	0.33	-0.01	34.70%
70	0.35	2.01	99.70%	0.56	0.66	58.70%	0.56	0.22	36.50%
80	0.59	2.41	99.80%	0.87	0.98	60.30%	0.83	0.52	39.60%
90	0.92	2.94	99.80%	1.32	1.46	62.90%	1.23	1.07	50.80%
95	1.22	3.78	100.0%	1.71	1.94	67.20%	1.56	1.61	59.50%
96	1.31	3.91	100.0%	1.82	2.06	67.30%	1.66	1.75	61.70%
97	1.43	4.08	100.0%	1.97	2.26	69.20%	1.78	1.91	63.20%
98	1.58	4.29	100.0%	2.18	2.47	69.20%	1.97	2.03	62.10%
99	1.83	4.58	100.0%	2.51	2.79	67.30%	2.24	2.29	61.90%

Government Fixed-Income Fund Returns

	Eq(3)			Model 4			Model 5		
Pct	Sim	Act	%<Act	Sim	Act	%<Act	Sim	Act	%<Act
1	-2.00	-3.91	1.10%	-2.34	-3.86	2.10%	-2.24	-4.10	0.80%
2	-1.77	-3.50	1.40%	-1.99	-3.27	2.80%	-1.91	-3.47	1.30%
3	-1.63	-3.02	2.20%	-1.81	-2.98	3.90%	-1.74	-3.15	1.70%

4	-1.52	-2.80	3.30%	-1.67	-2.68	5.60%	-1.61	-2.91	2.10%
5	-1.42	-2.63	3.50%	-1.57	-2.50	6.80%	-1.52	-2.72	2.80%
10	-1.09	-2.06	5.90%	-1.23	-1.96	10.70%	-1.19	-2.24	3.80%
20	-0.73	-1.44	11.80%	-0.81	-1.38	14.60%	-0.79	-1.60	7.60%
30	-0.47	-1.11	12.00%	-0.48	-0.83	24.00%	-0.47	-1.08	11.90%
40	-0.23	-0.64	20.30%	-0.18	-0.49	26.70%	-0.19	-0.71	13.90%
50	0.01	-0.09	41.10%	0.11	-0.12	32.90%	0.08	-0.45	14.10%
60	0.24	0.48	68.30%	0.39	0.23	40.60%	0.35	-0.22	13.10%
70	0.48	0.93	80.70%	0.69	0.54	41.80%	0.63	0.08	15.60%
80	0.74	1.48	90.40%	1.03	0.87	42.00%	0.95	0.46	21.20%
90	1.09	2.21	96.60%	1.46	1.36	47.20%	1.37	1.07	35.60%
95	1.39	2.86	99.50%	1.82	1.72	49.30%	1.72	1.52	44.10%
96	1.49	3.10	99.90%	1.93	1.90	54.20%	1.82	1.61	44.40%
97	1.59	3.42	99.90%	2.07	2.04	54.10%	1.96	1.75	45.00%
98	1.73	3.91	100.0%	2.27	2.44	64.50%	2.14	2.26	61.10%
99	1.97	4.48	100.0%	2.63	2.79	62.70%	2.48	2.79	68.60%

Table 4-7 Value Measure of Equity, Corporate Fixed-Income and Government Fixed-Income Funds

This table estimates the monthly added value for every fund in the sample. The cross-sectional mean, standard error, t-statistic and percentiles are reported for the cross-sectional distribution. Percent with less than zero is the fraction of the distribution that has negative average value added. The cross-sectional weighted mean, standard error and t-statistic are weighted by the number of periods the fund exists. No. of Funds is the number of funds in the distribution. The numerical values are reported in $millions per month.

Equity represents domestic equity funds, Corporate FI represents corporate fixed-income funds and Government FI represents government fixed-income funds. ETF-factor is the two-facto model presented in equation (4.1), FF5 is the five-factor model of Fama and French (2015), q-factor is the q-factor model of Hou, Xue, and Zhang (2015), C4 is the four-factor model of Carhart (1997), CAPM is the single-factor model, Model 5 is the six-factor model discussed in the section 4.3 and Eq3 is the two-factor model presented in equation (4.3).

	Equity					*Corporate FI*		*Government FI*	
	ETF-factor	FF5	q-factor	C4	CAPM	Model 5	Eq3	Model 5	Eq3
Cross-Sectional Mean	0.21	-0.29	0.14	-0.29	-0.38	0.00	0.46	0.09	0.03
Standard Error of the Mean	0.04	0.02	0.03	0.02	0.02	0.11	0.09	0.07	0.05
t-statistic	4.99	-14.32	5.01	-12.85	-16.29	0.04	5.36	1.29	0.61
1st Percentile	-4.09	-7.20	-4.57	-7.37	-9.01	-6.87	-0.42	-4.48	-4.54
5th Percentile	-0.89	-1.78	-0.91	-1.84	-2.21	-1.37	-0.02	-1.31	-0.72
10th Percentile	-0.37	-0.82	-0.36	-0.83	-1.02	-0.62	0.00	-0.47	-0.27
50th Percentile	0.00	-0.03	0.00	-0.03	-0.04	0.00	0.04	0.00	0.00
90th Percentile	0.38	0.07	0.37	0.07	0.06	0.49	0.96	0.47	0.35
95th Percentile	1.13	0.31	1.09	0.33	0.29	1.33	1.96	1.13	1.05
99th Percentile	6.63	2.64	7.42	2.76	2.44	7.19	6.67	7.41	4.90
Percent with less than zero	61.77%	77.78%	61.89%	78.68%	80.75%	50.00%	17.55%	54.03%	49.30%
Cross-Sectional Weighted Mean	0.25	-0.32	0.19	-0.31	-0.41	-0.01	0.50	0.14	0.04
Se of the Weighted Mean	0.05	0.02	0.03	0.03	0.03	0.12	0.10	0.08	0.05
t-statistic	5.08	-13.65	5.53	-12.06	-15.45	-0.11	5.22	1.74	0.72
No. of Funds	13268	13268	13268	13268	13268	564	564	781	781

Table 4-8 Out-of-Sample Performance of Funds in the Top Decile

This table reports the out-of-sample statistics for the top decile. The first column reports the out-of-sample horizon. The second and third column labelled "Value Added" report the average value added of the top decile at each horizon and the associated p-value. The next two columns report the ratio of the time when the top decile generates a higher value added than the bottom decile. The last column reports the average fraction of total AUM in the top decile. All the p-values in this table are one tailed. The models are the same as those in table 4.7.

Horizon	Value Added		Top Outperforms Bottom		Fraction of Total AUM (%)
Years	*$ Mil*	*p-value(%)*	*Freq. (%)*	*p-value(%)*	
Equity Mutual Fund					
Model: ETF-factor					
3	0.72	18.12	63.89	3.03	11.29
4	1.53	2.44	64.58	0.34	11.59
5	1.18	5.51	61.67	1.07	11.73
6	0.74	13.59	58.33	3.36	11.82
6.75	0.65	15.36	56.79	3.37	11.81
Model: FF5					
3	0.49	2.26	61.11	8.09	8.81
4	0.23	15.93	52.08	14.02	9.05
5	0.23	14.99	50.00	11.95	9.20
6	-0.12	69.06	52.78	36.64	9.25
6.75	-0.22	83.54	48.15	44.61	9.20
Model: C4					
3	1.03	5.91	66.67	0.24	10.86
4	0.47	21.13	64.58	2.13	11.00
5	0.36	23.67	61.67	1.91	10.99
6	0.21	31.86	56.94	3.25	10.89
6.75	0.12	38.58	53.09	6.80	10.77
Model: q-factor					
3	1.62	0.55	61.11	0.52	11.45
4	1.01	2.95	50.00	9.26	11.51
5	0.77	4.19	53.33	16.86	11.46
6	0.51	10.21	54.17	28.60	11.34
6.75	0.54	8.05	56.79	19.91	11.23
Model: CAPM					
3	0.03	43.63	66.67	1.55	7.24
4	0.13	26.93	68.75	2.45	7.48
5	0.24	10.02	68.33	1.33	7.63
6	-0.12	70.22	62.50	9.36	7.69
6.75	-0.16	75.27	61.73	8.17	7.66
Corporate Fixed-Income Mutual Fund					
Model 5					
3	3.75	0.00	83.33	0.00	17.82
4	3.00	0.00	81.25	0.00	18.09
5	1.21	3.01	68.33	2.09	18.36

6	1.75	0.32	68.06	0.20	18.29
6.75	1.19	2.00	61.73	0.99	18.02
Government Fixed-Income Mutual Fund					
Model 5					
3	0.36	0.00	86.11	0.00	3.60
4	0.34	0.00	83.33	0.00	3.66
5	0.24	0.01	78.33	0.00	3.67
6	0.22	0.02	76.39	0.00	3.70
6.75	0.28	0.00	72.84	0.00	3.67

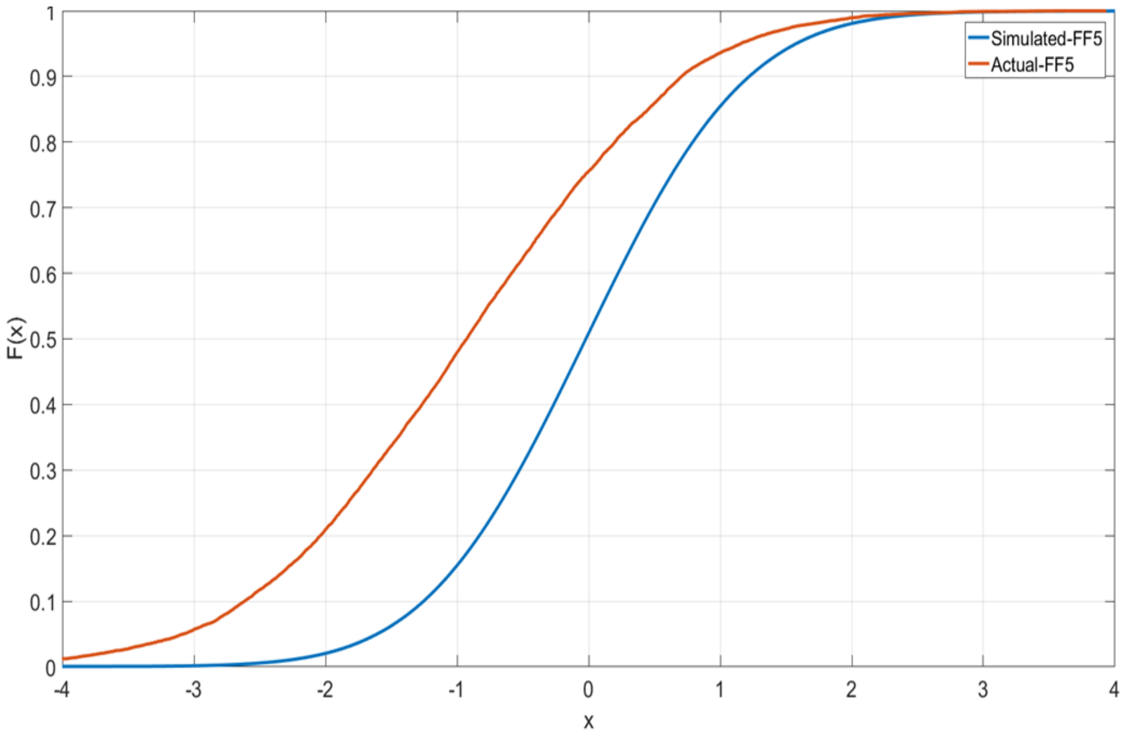

Figure 4.1 Simulated and actual CDF of FF5 t(α) estimates for US domestic equity fund returns

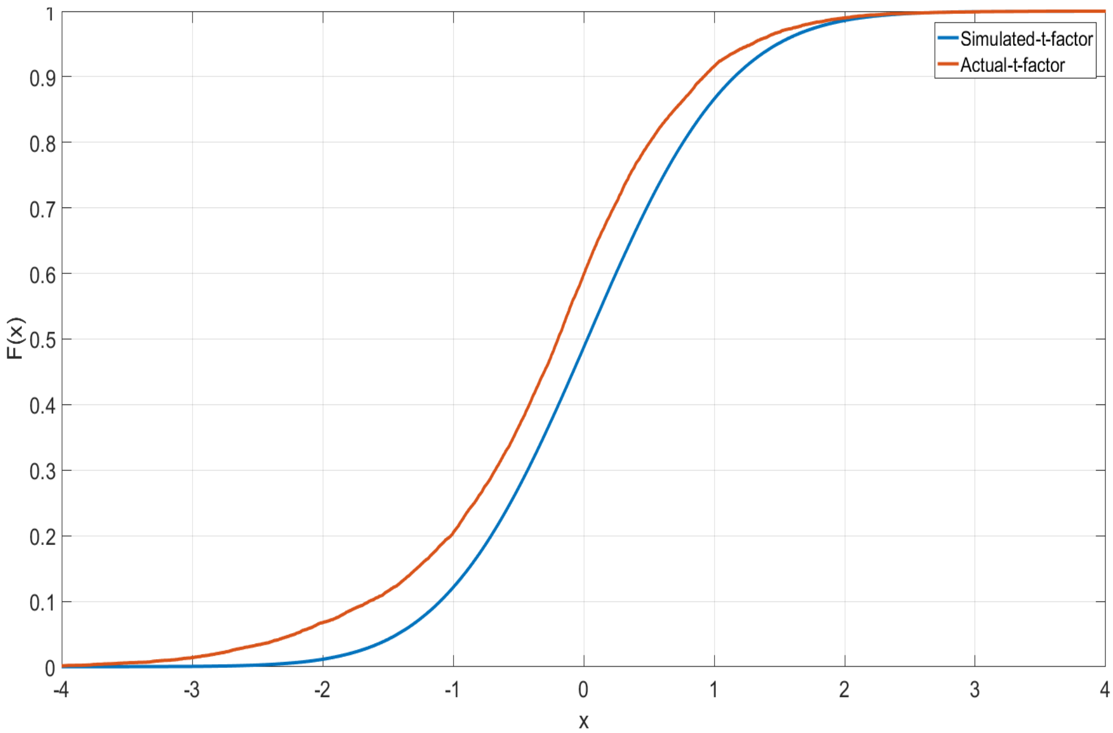

Figure 4.2 Simulated and actual CDF of ETF-factor t(α) estimates for US domestic equity fund returns

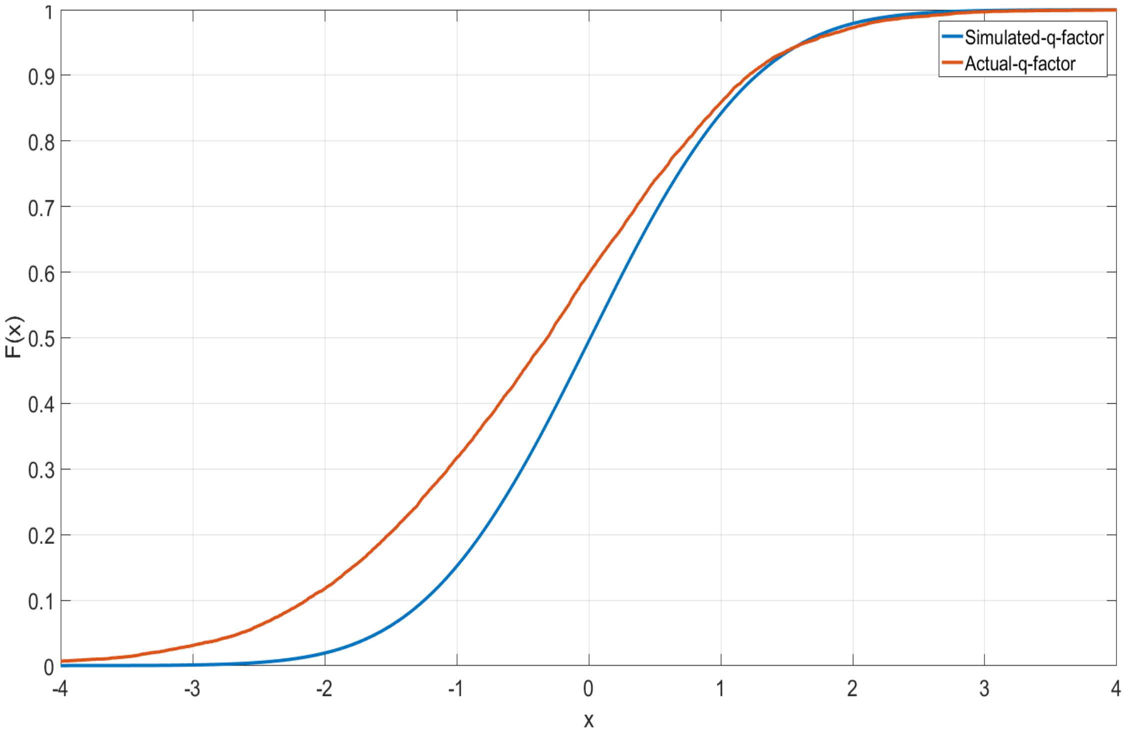

Figure 4.3 Simulated and actual CDF of q-factor t(α) estimates for US domestic equity fund return

5. Prediction of the Chinese Market with a New Measure of Investor Behaviour

5.1. Introduction

The Chinese stock market is the largest emerging market and the second largest market in the world, but it is widely accepted as an inefficient market. The latest evidence is the 2015-16 turbulence in the Chinese stock market. Nicholas Lardy said:' from June 2014 to June 2015, prices increased more than 150 percent on the Shanghai exchange... An unusually large part of this run-up was fuelled by retail investors who borrowed to buy equities. The market was priced way beyond perfection. Once prices fell even slightly, many of these investors found themselves needing to sell, leading to a sharp market correction' (26 Aug 2015, New York Times). One explanation for this market crash is the herd mentality: individual investors mimic the actions of the crowd of others in a rush to sell. A study by Yao *et al.* (2014) shows evidence of herding behaviour in the Chinese stock market, particularly in the B-share market (also see, e.g., Laurence *et al.* (1997), Groenewold *et al.* (2003), Li (2007)).

Empirically, recent studies suggest that investor sentiment has a strong impact on a cross-section of stock returns (see, e.g., Brown and Cliff (2005), Baker and Wurgler (2006, 2007), Stambaugh *et al.* (2012, 2014)). This finding can also be applied to the Chinese stock market. Gao and Kling (2008) show that Chinese institutional investors' sentiment can affect market volatility. Chi *et al.* (2012) find that investor sentiment has a significant impact on stock returns. Further, Ni *et al.* (2015) state that the influence of investor sentiment is persistent on monthly returns in the Chinese A-share market. A typical consequence is that Chinese investors are more likely to overreact in the short term. In addition, some studies provide evidence for the predictive ability of investor sentiment. Chen *et al.* (2014) proposes a measure of investor sentiment in China and show its good out-of-sample predictability. A study by Guo *et al.* (2017) shows that investor sentiment sometimes has predictive ability for Chinese stock prices.

In this chapter, the information from three index ETFs (ETF500, ETF300, and ETF50) is used to propose measure of investor behaviour. ETF500, ETF300, and ETF50 track the performance of the

CSI 500 index, CSI 300 Index, and SSE 50 Index, respectively, while minimizing tracking errors. The trade of ETFs can provide more past market data, price, and volume to predict the price movement of the underlying index. This study’s measure of investor behaviour uses the ask and bid volumes provided by the ETFs. The reason of calling it as measure of investor behaviour because the listed volumes in the order book reflect how the traders are willing to trade. This is very different from market trend that is defined as a perceived tendency of financial markets to move in a particular direction in a long-tern horizon. It is also different from market sentiment that is the feeling or tone of a market. The Investor behaviour in this chapter is simply an approximation of trading behaviour within a day.

Empirically, this measure of investor behaviour has strong explanatory power on the contemporaneous half-daily index return. Depending on the underlying index, the adjusted R^2 varies from 4.75% to 9.00% in the linear model. In addition, this measure of investor behaviour shows some in-sample and out-of-sample forecasting ability for market return. Apart from the linear model, this measure of investor behaviour shows explanatory power and predictive ability in two non-parametric models. The out-of-forecasting results provided by the coefficient varying model can be good, but they are very sensitive to the bandwidth selection. According to this thesis’ results, the linear model which uses the investor behaviour to directly explain or predict the market return performs the best. Hence, this study employs three measures of investor behaviour that are developed from the ETF500, ETF300, and ETF50 as the independent variables. The results show that these measures of investor behaviour together have stronger explanatory power on the market return.

Unfortunately, this study’s measure of investor behaviour has weak forecasting ability. To improve the out-of-sample forecasting results, a two-step prediction is developed: first, the investor behaviour at the next date is predicted; second, the estimated investor behaviour is used to predict the contemporaneous index return. The validity of the two-step prediction is supported by the evidence that the afternoon investor behaviour is predictable. For example, if the CSI 500 morning investor behaviour is used to predict the CSI 500 afternoon investor behaviour, the corresponding adjusted R^2 is 42.94%. This result indicates that approximately half of the afternoon volatility is predictable. Hence, two-step prediction provides some improvements in forecasting the CSI 500 index afternoon return.

This finding is a great complement to the previous studies (see, e.g., Laurence *et al.* (1997), Groenewold *et al.* (2003) and Yao *et al.* (2014)) which found that investor sentiment has a significant impact on the market and that the Chinese stock market is not efficient. Furthermore, this study investigates the interaction between the Chinese and US stock markets, and prices the information on the US stock market in the models. The results provide more empirical evidence to support the co-movements of Shanghai and New York stock prices (Chow, Liu and Niu, 2011). Finally, the continuity of investor behaviour is found within a day. It is possible to yield a number of applications of prediction models in high-frequency trading.

This Chapter is divided into six sections. In the second section, the measure of investor behaviour is constructed and its predictability tested. In the third section, the interaction between the Chinese and US stock markets is investigated. The econometric methodology and relevant results are presented in the fourth section. The fifth section explains the two-step prediction and shows its application on forecasting the Chinese market. The sixth section introduces the extended linear model, and the seventh section discusses a simple trading strategy as an example. The last section concludes the chapter.

5.2. The Measure of Investor Behaviour

5.2.1. Data

The SSE 50 Index selects the 50 largest stocks with good liquidity and representativeness from the Shanghai security market using a scientific and objective method. Next, the CSI 300 Index is designed for use as a performance benchmark and aims to reflect the price fluctuation and performance of the China A-share market. Then the CSI 500 Index (excluding the stocks either in the CSI 300 or ranking in the top 300 in the Shanghai and Shenzhen stock markets) aims to comprehensively reflect the price fluctuation and performance of the small-cap companies in the Shanghai and Shenzhen securities market. Tick data for the ETF50, ETF300, and ETF500 at a frequency of one second is obtained from the China Securities Depository and Clearing Corporation Limited.[28] These ETFs track the SSE 50 Index, CSI 300 Index, and CSI 500 Index, respectively. Tick data for the SSE 50 Index, CSI 300 Index, and CSI 500 Index at a frequency of one minute is also from the China Securities Index Company. Daily frequency data from the S&P500 is from the Yahoo Finance dataset. To construct the measure of investor behaviour, two types of volume are used: total bid/ask volume and the closest 10 bid/ask volume of below/above the corresponding price in the order book. In line with the research objective, all variables in the ETF dataset and the Index dataset are converted into half-daily frequency. The sample used covers the period of 16th April 2013 to 31st March 2016.

5.2.2. Construction of the Measure of Investor Behaviour

A put-call ratio is a popular tool specifically designed to help individual investors gauge the overall sentiment of the market. The ratio provides information about the trading volume of put options to call options. Normally, a higher ratio indicates a selling signal and a lower ratio indicates a buying signal because more put options traded in the market means traders are worried about prices falling and vice versa. The put-call ratio has been popularly used as a measure of investor sentiment in academic studies. Bandopadhyaya and Jones (2006) suggest that investor sentiment has explanatory power on short-term movements in asset prices. Moreover, Bandopadhyaya and Jones (2008) argue that the put-call ratio is superior to the volatility index as a measure of investor behaviour. Martikainen and Puttonen (1996) find that put-call ratio in the Finnish stock index options market has predictive power on the underlying returns in the stock market.

One way to construct a measure of investor behaviour is to use the volume of trade, which is the number of shares that change hands. The volume of trade is made up of buying volume and selling volume. Buying volume is the number of contracts that change hands at the ask price. Selling volume is the number of contracts that change hands at the bid price. The bid price is the highest current price someone is willing to buy shares at, and the ask price is the lowest current price

[28] I would like to thank Cheng Lu from China Securities Depository and Clearing Corporation Limited for providing me with high frequency ETF and index data.

someone willing to sell shares at. If contracts are traded at the bid price, that bid will disappear, and the new bid will be the lower price. Therefore, the number of contracts traded at bid price is selling volume because it has the potential to move the price down. Similarly, if contracts are traded at the ask price, that offer will disappear, and the new offer will be the higher price. Therefore, the number of contracts traded at ask price is buying volume because it has the potential to move the price up. Whether or not more transactions are occurring at the bid or offer gives traders short-term indications of price movement. However, one disadvantage of this measure is that there is no indication if no contracts change hands. This disadvantage is exaggerated in high-frequency trade. In this chapter, a measure of investor behaviour is constructed by dividing the difference of ask volume listed in the order book and bid volume in order book by their sum:

$$q \text{ (investor behaviour)} = \text{(ask volume - bid volume)/(ask volume + bid volume)} \qquad (5.1)$$

An order book lists the number of shares being bid or offered at each price point. Thus, alternatively, this ratio can be interpreted as excess sell orders over total listed orders. Although not all orders will be executed, excess ask volume in the order book suggests downward pressure on the stock price. There are two reasons: first, some investors may decrease the offered price to the best ask or the best bid to execute their orders, which increases selling volume that has the potential to move the market price down; second, even if no investors decrease the offered price, the excess ask volume forms resistance to stop the further rise in price. One advantage of using volume listed in the order book is that it provides pre-indications and retains predictable power.

In terms of volume in the order book, two types of the volume are available: total bid/ask volume and the closest 10 bid/ask volume of below/above the corresponding price. Total bid/ask volume includes all buying/selling volume listed in the book. Book depth refers to the number of price levels available at a given time in the book. A book can contain unlimited price levels, but sometimes the book is represented to a fixed depth, for example, 20 price levels. Therefore, total ask/bid volume includes all volume at these levels. However, orders with a price closer to the corresponding price have a higher probability of being executed, for example, the first 10 price levels. It would be interesting to know which type of volume is optimal, so two versions of the measure of investor behaviour will be used to find out. For convenience, the measure of investor behaviour constructed from total bid/ask volume will be defined as the first version, and the other will be defined as the second version. Eventually, the better of the two will be retained for further use.

Table 5.1 reports the summary statistics for the variables. The half-daily CSI 500 Index return has a mean of 0.11% per morning, while its counterparts of CSI 300 and SSE50 are smaller: 0.08% and 0.07%, respectively. Similarly, the half-daily CSI 500 Index return has a mean of 0.05% per afternoon, which is bigger than the average CSI 300 and SSE50 Index afternoon returns of 0.01% and 0.02%, respectively. These results show that the Chinese market generates a bigger average return in the morning, at least during the sample period. The CSI 500 Index return has higher volatility than the CSI 300 and SSE50 returns, both in the morning and afternoon. It is reasonable that the CSI 500 Index return generates the highest expected return and volatility because its components are smaller firms. In addition, it is observed that all index returns have negative skewness, both in the morning and afternoon, except for $r_{m,t}^{50}$ (SSE50 Index returns in the morning), indicating that the mode and median are larger than the mean. This observation is in line with the 2015-16 Chinese stock market turbulence, during which some extreme market movements heavily dragged down the average market performance.

[Insert Table 5.1]

As for the measures of investor behaviour, the statistics show that the measure of investor behaviour developed from the ETF500 is more volatile, both in the morning and afternoon. For example, the standard deviations of $q_{m,t}^{500}$, $q_{m,t}^{300}$ and $q_{m,t}^{50}$ are 24.84%, 21.64%, and 16.88%, respectively. These results indicate that trades associated with big firms could be more rational. In addition, there is difference between the two versions of investor behaviour. For example, the average $q_{m,t}^{500}$ and $q_{a,t}^{500}$ (first version) are -5.81% and -7.45%, while the average $q10_{m,t}^{500}$ and $q10_{a,t}^{500}$ (second version) are -6.18% and 0.80%. A similar pattern emerges from the statistics for the measures of investor behaviour developed from the ETF300 and ETF50 as well. Based on the first version of investor behaviours, Chinese investors are more willing to buy smaller stocks and sell big stocks. The evidence is that the investor behaviours developed from the ETF500 are negative, while those developed from the ETF300 and ETF50 are positive, indicating that Chinese investors are confident in smaller firms.

Furthermore, I report the correlation matrix for all variables in Table 5.1. The results indicate a negative correlation between the index return and the contemporaneous investor behaviour in the first version. For example, the correlations between $r_{m,t}^{500}$, $r_{m,t}^{300}$, $r_{m,t}^{50}$ and $q_{m,t}^{500}$, $q_{m,t}^{300}$, $q_{m,t}^{50}$ are -0.22, -0.26, and -0.29, respectively. But this negative correlation is trivial between the index return and the contemporaneous investor behaviour in the second version. In addition, there is a strong positive correlation between the morning and afternoon investor behaviours (both are first version). For example, the correlations between $q_{m,t}^{500}$, $q_{m,t}^{300}$, $q_{m,t}^{50}$ and $q_{a,t}^{500}$, $q_{a,t}^{300}$, $q_{a,t}^{50}$ are 0.66, 0.68, and 0.72, respectively. Compared to investor behaviours in the first version, those in the second version show weaker correlations with other variables. Lastly, the previous day's S&P500 return has relatively low correlations with today's Chinese market returns.

[Insert Figure 5.1]

[Insert Figure 5.2]

A plot of the measure of investor behaviour (developed from the ETF 500) during the sample period (16th April 2013 through 31st March 2016) appears in Figure 5.1, and the frequency distribution of the measure of investor behaviour is shown in Figure 5.2. Figure 5.1 shows that the measure of investor behaviour is more volatile during the 2015-16 Chinese stock market turbulence. This pattern is more obvious for the second version of investor behaviour. In addition, it is observed that the measure of investor behaviour is more volatile in the afternoon. Figure 5.2 shows that the distribution of the first version of investor behaviour is approximately symmetric, but that of the second version is left-skewed. For both versions of investor behaviour, the distributions in the afternoon are more spread out than those in the morning. This observation confirms that the afternoon investor behaviour is more volatile.

5.2.3. Regression Results: the Measure of Investor Behaviour on Return

In this section, the following two questions are investigated: first, whether or not this thesis' measure of investor behaviour explains or predicts the index return; second, which version is the better measure of investor behaviour. Specifically, consider the following estimates:

$$r_t = \beta_0 + \beta_1 q_t + \varepsilon_t \tag{5.2}$$

$$r_t = \beta_0 + \beta_1 q_{t-1} + \varepsilon_t \tag{5.3}$$

where r_t is the current return on the underlying index. In these regressions, the coefficient of q_t tells us whether or not the measure of investor behaviour explains the half-daily index return, and the coefficient of q_{t-1} tells us whether or not the measure of investor behaviour predicts the half-daily index return.

[Insert Table 5.2]

Table 5.2 summarises the time-series regression results. Panel A of Table 2 provides the estimates using information from the ETF500 and CSI 500 Index. Using the measure of investor behaviour constructed from total bid/ask volume (first version), results from the estimation of Equation (5.2) and (5.3) show that both q_t and q_{t-1} are significantly related to the half-daily CSI 500 Index return, indicating that both q_t and q_{t-1} play a role in short-term index movement. Their negative coefficients are as expected because the measure of investor behaviour is a contrarian measure. However, the second version of the measure loses explanatory power. Using the second version of the measure of investor behaviour, results from the estimation of Equations (5.2) and (5.3) show no significant relationship between q_t, q_{t-1} and the half-daily CSI 500 Index return.

Panel B of Table 5.2 presents the estimates using information from the ETF300 and CSI 300 Index. Using the first version of the measure of investor behaviour, results from the estimation of Equations (5.2) and (5.3) show that only q_t retains explanatory power on the half-daily CSI 300 Index return. This result indicates that the half-daily CSI 300 Index movement only relates to contemporaneous investor behaviour. In contrast to the results in Panel A, the second version of the measure of investor behaviour explains the half-daily CSI 300 Index return. From the estimation of Equation (5.2), the coefficient for q_t has a value of -0.0063 (t-statistic = -3.79). There is no evidence to support that q_{t-1} can predict the half-daily CSI 300 Index return index return.

Panel C of Table 5.2 presents the estimates using information from the ETF50 and SSE 50 Index. Using the measure of investor behaviour constructed from total bid/ask volume (first version), results from the estimation of Equations (5.2) and (5.3) show that the half-daily SSE50 Index return significantly relates to q_t, but not to q_{t-1}. Using the second version of the measure of investor behaviour, the results show that the SSE50 Index return is not significantly related to q_t and q_{t-1}. Regardless of statistical significance, there is a positive relationship between the half-daily SSE50 return and q_{t-1} in the second version. This paradoxical result indicates that the second version of the measure of investor behaviour may not be robust.

A comparison of the empirical results shows that the first version of the measure of investor behaviour that is constructed from total bid/ask volume is superior. It plays a significant role in explaining the contemporaneous return of the SSE 50 Index, CSI 300 Index, and CSI 500 Index, and retains the power of prediction for the CSI 500 return. Relatively, the second version of the measure of investor behaviour only explains the returns of the SSE 50 Index and CSI 300 Index and loses explanatory power on those of the CSI 500 Index, which reflects the performance of the small-cap companies in the Shanghai and Shenzhen securities market. In addition, the first version of the measure of investor behaviour provides consistent results and t-statistic improvements in regressions.

5.3. The Interaction between the Chinese and US Stock Markets

Bloomberg (2016) states that: 'The US and Chinese economies are increasingly intertwined, making their bilateral relationship key to global prospects.' Many statistics show a close relationship between the Chinese and US economies, but how close is the relationship between the two countries' stock markets? Jan Zilinsky examined the 120-day correlation between the S&P 500 and Shanghai Composite Index returns (2 Feb 2016, PIIE). He says: 'While the correlation has typically been positive over the last 10 years (global factors, after all, affect both China and the United States), there were many periods when the stock returns moved in the opposite direction.' Based on the chart he developed, the significant negative correlation mainly appears in the periods of 2014 to 2015.

Using time-varying regressions, Chow *et al.* (2011) find that the effect of a contemporaneous return for New York on that for Shanghai is persistent and significant in the period 2002 to 2010. Combined with other evidence, they point out that China's stock market has become more and more integrated into the world market. Based on their research, it is deduced that the information reflected in the current US market will affect the Chinese market that opens 13 hours later. The S&P 500 is adopted as the proxy for information from the US market because it is widely regarded as the best single gauge of large-cap US equities and considered as a leading indicator for the US economy. In this rapidly changing world, it is expected that Chinese investors can react to the information from the US market quickly. Therefore, the daily return of CSI 500 Index is divided into returns in the morning and afternoon. If the information from the US market is valuable, Chinese investors are expected to react in the morning session. Further, the impact of the Chinese stock market on the S&P 500 is investigated. The following linear regressions is carried to answer the questions:

$$r_{d,t} = \beta_0 + \beta_1 r_{sp,t-1} + \varepsilon_t \tag{5.4}$$

$$r_{m,t} = \beta_0 + \beta_1 r_{sp,t-1} + \varepsilon_t \tag{5.5}$$

$$r_{a,t} = \beta_0 + \beta_1 r_{sp,t-1} + \varepsilon_t \tag{5.6}$$

$$r_{sp,t} = \beta_0 + \beta_1 r_{d,t} + \varepsilon_t \tag{5.7}$$

where $r_{d,t}$ is the current daily return on the underlying index, $r_{m,t}$ is the half-daily return on the underlying index in the morning, $r_{a,t}$ is the half-daily return on underlying index in the afternoon, $r_{sp,t-1}$ is previous day's S&P500 return, and $r_{sp,t}$ is today's S&P500 return.

[Insert Table 5.3]

Table 5.3 summarises the time-series regression results. Panel A of Table 5.3 presents the estimates using the CSI 500 Index returns in the regressions. Results from the estimation of Equation (5.4) shows that the CSI 500 Index return is not sensitive to the previous day's S&P 500

Index return. Similarly, results from the estimation of Equations (5.5) and (5.6) shows that neither the CSI 500 Index return in the morning nor the afternoon is sensitive to the previous day's S&P 500 Index return. Although the t-statistics (-0.76 and 1.53 in Equations (5.5) and (5.6), respectively) indicate no significant relationship, the sign of the coefficients is interesting. The estimated coefficient in Equation (5.5) is negative, while that in Equation (5.6) is positive, indicating that the Chinese market reacts to the same information differently in the morning and afternoon. Daniel Egan says: ‘Traders who experience morning losses are about 16 percent more likely to assume above-average risk in the afternoon than traders with morning gains. This behavior has important short-term consequences for afternoon prices, as losing traders are prepared to purchase contracts at higher prices and sell contracts at lower prices than those that prevailed previously.’ Based on his research, behavioral biases in the Chinese market may be an answer.

Panel B of Table 5.3 presents the estimates using the CSI 300 Index returns as the dependent variable. Like the results in Panel A, there is no significant relationship between the CSI 300 Index return and the previous day's S&P 500 Index return at a daily frequency. Stepping back from the discussion of coefficient prominence, the sign of the coefficients estimated from Equations (5.5) and (5.6) are consistent with those in Panel A. However, there is an improvement in t-statistics from the estimation of Equation (5.5) using the CSI 300 Index returns. Panel C of Table 5.3 presents the estimates using the SSE 50 Index returns as the dependent variable. Like the patterns presented by Panels A and B, no significant relationship is found between the SSE 50 Index daily return and the previous day's S&P 500 Index return. However, results from the estimation of Equation (5.5) shows that the SSE 50 Index return in the morning is significantly sensitive to the previous day's S&P 500 Index return. The coefficient has a value of -0.1345 (t-statistic = -2.16). This negative relationship indicates a potential competition between top 50 Chinese companies and top 500 US companies. The pattern of coefficient sign from the estimation of Equations (5.5) and (5.6) is consistent, indicating different responses from the Chinese market in the morning and afternoon to the previous day's S&P 500 Index return.

Panel D of Table 5.3 summarizes the results in Equation (5.7) where the S&P 500 Index return is regressed on the CSI 500, CSI 300, or SSE 50 Index return. It is clear that the Chinese stock market return has a positive effect on the S&P 500 Index return in the same day. For example, the coefficients of q_t for the CSI 500, CSI 300, and SSE 50 are 0.0675 (t-statistic = 3.04), 0.0792 (t-statistic = 3.05), and 0.0706 (t-statistic = 2.69), respectively. All of them are statistically significant at the 0.01 level. These results indicate that US investors reacts to the information from the Chinese stock market in the same direction.

Overall, only the SSE 50 Index is sensitive to news from the US market. But the impact of the S&P 500 on the Chinese market should not be denied. One reason is that the sample period covers the 2015--16 Chinese stock market turbulence. Normally, the S&P 500 can be a good leading indicator for the Chinese market due to China-US economic and trade relations. However, it is an almost impossible task to predict a market that lost a third of the value of A-shares on the Shanghai Stock Exchange. Interestingly, there is spill-over from the Chinese stock market to the US market (proxied by the S&P 500 Index) through the 16th April 2013 to 31st March 2016.

5.4. Econometric Methodology

In this section, the parametric and non-parametric methods for in-sample estimation are provided. Further, predictive models using the contemporaneous relationship between the investor behaviour and index return are constructed.

5.4.1. Linear Models

First, assume that the current index return explained by the previous day's S&P500 return and investor behaviour with a linear relationship. The previous day's S&P500 return is included in the model because the growing Chinese market needs to react to the US market in a correct manner. To obtain how the Chinese market reacts to information from the US and investor behaviour, the following estimates are used:

$$r_{m,t} = \beta_0 + \beta_1 r_{sp,t-1} + \beta_2 q_{a,t-1} + \varepsilon_t \tag{5.8}$$

$$r_{m,t} = \beta_0 + \beta_1 r_{sp,t-1} + \beta_2 q_{m,t} + \varepsilon_t \tag{5.9}$$

$$r_{a,t} = \beta_0 + \beta_1 r_{sp,t-1} + \beta_2 q_{m,t} + \varepsilon_t \quad (5.10)$$

$$r_{a,t} = \beta_0 + \beta_1 r_{sp,t-1} + \beta_2 q_{a,t} + \varepsilon_t \quad (5.11)$$

where $r_{m,t}$ is the half-daily return on the underlying index in the morning, $r_{a,t}$ is the half-daily return on the underlying index in the afternoon, $r_{sp,t-1}$is the previous day's S&P500 return, $q_{a,t-1}$ is the previous day's investor behaviour in the afternoon, $q_{m,t}$ is the current investor behaviour in the morning, and $q_{a,t}$ is the current investor behaviour in the afternoon. This will provide detailed evidence for the important role of investor behaviour.

[Insert Table 5.4]

Table 5.4 reports the results of linear regressions. Panel A provides the estimated results for the CSI 500 Index in the period of 16th April 2013 through 31st March 2016. Consistent with the results in table 5.2, the half-daily index return is negatively related to investor behaviour. Statistical significance aside, the CSI 500 Index responds to the S&P500 in the opposite direction in the morning, but reverses the response with a larger magnitude in the afternoon. In the regressions of (5.9) and (5.10), the coefficients of $r_{sp,t-1}$ are -0.04 (t-statistic = -0.60) versus 0.15 (t-statistic = 1.96), which indicates that the CSI 500 Index moves in line with the S&P500 in the afternoon. Moreover, the magnitude of the coefficient and t-statistic are larger in the afternoon. The estimations of Equations (5.8) and (5.10) show that the previous day's afternoon investor behaviour and today's morning investor behaviour can barely forecast today's return in the morning and afternoon, respectively. Although β_2 in Equation (5.10) is statistically significant (t-statistic = -2.20) with a small value of -0.005, the adjusted R^2 is only 1.4%. However, the contemporaneous explanatory power of investor behaviour is statistically significant. In Equation (5.9), the morning investor behaviour, $q_{m,t}$, generates a regression slope of -0.012 (t-statistic = -4.62), and the adjusted R^2 is 4.75%. In Equation (5.11), the afternoon investor behaviour, $q_{a,t}$, generates a regression slope of -0.011 (t-statistic = -5.76) and an adjusted R^2 of 5.76%.

Panel B provides the estimated results for the CSI 300 Index in the same period. Similar to the results in Panel A, the predictive power of investor behaviour is relatively weak, but the contemporaneous explanatory power of investor behaviour is quite strong. In the regressions of

Equations (5.8) and (5.10), the coefficients of $q_{a,t-1}$ and $q_{m,t}$ are -0.002 (t-statistic = -0.7) and -0.004 (t-statistic = -0.15), respectively. However, in the regressions of Equation (5.9) and (5.11), the coefficients of $q_{m,t}$ and $q_{a,t}$ are -0.014 (t-statistic = -5.63) and -0.016 (t-statistic = -6.14), respectively. Moreover, the adjusted R^2s are 7.08% and 8.36%, respectively. The pattern of the CSI 300 Index's response to the S&P500 is the same. Statistical significance aside, the CSI 300 Index return negatively relates to the previous day's S&P500 return in the morning, but positively relates to it in the afternoon.

Panel C provides the estimated results for the SSE 50 Index in the sample period. The SSE 50 Index return relates negatively to the previous day's S&P500 return in the morning, but becomes positive in the afternoon. Unlike results from the CSI 300 and CSI 500, β_1 in Equation (5.8) is statistically significant (t-statistic = -1.98) with a value of -0.1350, which indicates a statistically significant impact from the S&P 500. But if $q_{a,t-1}$ is replaced by $q_{m,t}$, the significant impact from the S&P 500 disappears. In other words, the contemporaneous investor behaviour reflects information from the US market. Robustly, the contemporaneous explanatory power of investor behaviour is supported by the regression results. From the estimation of Equations (5.9) and (5.11), the coefficients of $q_{m,t}$ and $q_{a,t}$ are -0.020 (t-statistic = -6.54) and -0.021 (t-statistic = -5.82), respectively. Unsurprisingly, the in-sample adjusted R^2s are up to 8.21% and 9.00%, respectively.

In Huang *et al.* (2015), they construct an aligned investor sentiment index that has greater predictive power than other sentiment indices and popular macroeconomic variables. To argue the superior performance of the new investor sentiment index, in-sample and out-sample R^2 are used as the main measure to compare empirical models. To forecast market return using aligned investor sentiment index, the in-sample R^2 floats between 1% and 3% among all scenarios. The in-sample R^2 of this thesis' contemporaneous investor behaviour with ordinary least squares regression floats between 4.75% and 9.00%, substantially greater than theirs. However, this study cannot predict the return using a contemporaneous investor behaviour. Moreover, the investor behaviour has very weak forecasting power for the estimation of Equations (5.8) and (5.10). So, can the return still be predicted using the investor behaviour? Alternatively, a two-step predictive model that will be proposed in the next section.

5.4.2. Non-Parametric Models

Varying coefficient models are widely employed to capture dynamic patterns in finance and economics. In essence, the varying models are locally parametric models. Following the version of Fan and Zhang (2008), assume the form of the multivariate regression function as:

$$m(q, X) = X^T\beta(q) \tag{5.12}$$

for the unknown functional coefficient $X^T\beta(q) = (b_1(q), \dots, b_p(q))^T$, where the regression function is $m(q, X) = E(y|q, X)$. One prominent advantage over linear models is that nonlinear interactions are allowed between q and X. There are several approaches to estimate the $\beta(.)$ in Equation (4.13). Here, the kernel-local polynomial smoothing method of Fan and Zhang (2008) is adopted. To estimate the functional coefficient, the regression is:

$$y = X^T\beta(q) + \varepsilon \tag{5.13}$$

with $E(\varepsilon) = 0$ and $var(\varepsilon) = \sigma^2(q)$. In line with this thesis' research objective, consider four scenarios, as before:

Scenario 1: $q = (q_{a,0}, \dots, q_{a,t-1})^T, X = (X_0, \dots, X_{t-1})^T, y = (r_{m,1}, \dots, r_{m,t})$ (5.14)

Scenario 2: $q = (q_{m,1}, \dots, q_{m,t})^T, X = (X_0, \dots, X_{t-1})^T, y = (r_{m,1}, \dots, r_{m,t})$ (5.15)

Scenario 3: $q = (q_{m,1}, \dots, q_{m,t})^T, X = (X_0, \dots, X_{t-1})^T, y = (r_{a,1}, \dots, r_{a,t})$ (5.16)

Scenario 4: $q = (q_{a,1}, \dots, q_{a,t})^T, X = (X_0, \dots, X_{t-1})^T, y = (r_{a,1}, \dots, r_{a,t})$ (5.17)

where $r_{m,t}$ is the half-daily return on the underlying index in the morning, $r_{a,t}$ is the half-daily return on the underlying index in the afternoon, $q_{a,t-1}$ is the previous day's investor behaviour in the afternoon, $q_{m,t}$ is the current investor behaviour in the morning, $q_{a,t}$ is the current investor behaviour in the afternoon, and $X_{t-1} = \begin{bmatrix} 1 \\ r_{sp,t-1} \end{bmatrix}$ with $r_{sp,t-1}$ is the previous day's S&P500 return. Bandwidth selection is a key issue in the kernel smoothing approach. There are many methods for selecting the optimal bandwidth. For example, Fan and Zhang (1999) give a method to estimate the optimal bandwidth by minimizing the estimated MSE of estimated coefficients. Before estimating the optimal bandwidth, the pilot bandwidth for estimating bias and the covariance matrix needs to be chosen. Fan and Gijbels (1995) propose the residual squares criterion (RSC). But the focus of this research is to investigate the explanatory power and predictability of investor behaviour. Hence, this section will simply apply the varying coefficient models. To simply the problem, the bandwidth is set as 0.1 in the regressions.

Another nonparametric model proposed by Banerjee and Pitarakis (2014) is adopted as well. This method follows the piecewise local least squares principle. Basically, it relies on a straightforward disjoint binning principle and uses the Akaike Information Criteria to determine the number of bins. The functional coefficient regression model is summarised as:

$$y_t = f_0(q_{t-d}) + f_1(q_{t-d})x_t + u_t \tag{5.18}$$

$$x_t = x_{t-1} + v_t \tag{5.19}$$

where u_t and v_t are stationary residuals, and $f_0(q_{t-d})$ and $f_1(q_{t-d})$ are unknown functions of the random variables q_{t-d}, while x_t is assumed to follow a random walk process. One advantage of this model is that x_t can be a standard I(1) process, while y_t can be stationary because $f_0(q_{t-d})$ and $f_1(q_{t-d})$ are not constants. When $f_0(q_{t-d})$ and $f_1(q_{t-d})$ are constants, this model becomes a linear parametric model. Banerjee and Pitarakis (2014) finds the model useful for investigating the relationship between stock prices and dividends via a sentiment indicator. To apply this model, the previous four scenarios are still considered, but x_t becomes the previous day's S&P500 closing price. This assumption sounds reasonable because the US market is a relatively efficient market and efficient markets are random. Relative to linear models that directly employ the investor behaviour to explain or predict the market return, non-parametric models can provide more insights into the reaction to information from the US market. In other words, the reaction to the information from the US market relies on investor behaviour and the varying coefficients capture this impact.

[Insert Table 5.5]

Table 5.5 summarises the in-sample analysis for both non-parametric estimations. Because R^2 is not valid for the evaluation of the fit of nonparametric models, the MSE of the predictors will be reported. For comparison, the MSE generated from linear models (Equations (5.8) to (5.11) and Equations (5.5) to (5.6)) are added to Table 5.5. The estimated results of Equations (5.5) and (5.6) play a role as default results that simply reflect the impact of information from the US market. Generally, the varying coefficient model produces smaller MSEs of predictors than Banerjee's model and linear models that include an investor behaviour. However, the values of MSE using the coefficient model rely on the selection of bandwidth. The smaller the bandwidth is, the smaller the MSE and vice versa. Therefore, over-fitting may appear in this case. It is not surprising that the values of the MSE in Banerjee's model and linear models (Equations (5.8) to (5.11)) are quite similar. Banerjee's model is a linear model with constant coefficients in every bin that is decided by investor behaviour. Therefore, those distinguished betas capture the impact of investor behaviour.

Interestingly, both linear models that include the investor behaviour and non-parametric model provide some consistent results. First, they all improve the MSEs of predictors. Relative to default results, the estimated results for the varying coefficient model shows comprehensive improvements on the MSEs of predictors. Compared to lagged investor behaviour, using contemporaneous investor behaviour provides more improvements. For example, when regressing on the half-daily return on the CSI 500 in the morning by the previous day's S&P 500 return, the MSE value is 1.71×10^{-4}. But the MSE values are 1.66×10^{-4} and 1.55×10^{-4} when the functional coefficients rely on lagged and contemporaneous investor behaviours, respectively. However, Banerjee's model and linear models (Equations (5.8) to (5.11)) are sensitive to the selection of investor behaviour. Using the lagged investor behaviour, they barely provide improvements on the MSEs of predictors. In the same case of regressing on the half-daily return on the CSI 500 in the morning, the MSE for both Banerjee's model and the linear models is 1.71×10^{-4} using lagged investor behaviour. But the values of the MSE reduce to 1.62×10^{-4} and 1.63×10^{-4}, respectively when using a contemporaneous investor behaviour. Furthermore, these findings are consistent if regressed on half-daily returns on the CSI 500 in the afternoon or returns on the other two indices both in the morning and afternoon. This evidence again supports that contemporaneous investor behaviour contributes to the prediction of returns.

Through all models, the MSEs of predictors is relatively larger in the afternoon. This is in line with the summary statistics in Table 5.1 showing that the market is more volatile in the afternoon. For example, the CSI 300 Index return has a standard deviation of 1.25% in the afternoon, while it is 1.12% in the morning. Besides, the MSE of the estimated CSI 300 returns is the smallest compared to the MSEs of the other two. This is consistent with that results in Table 5.1 that the CSI 300 returns are the least volatile, both in the morning and afternoon.

5.5. Forecasting the Chinese Market

Looking at the summary statistics in Table 5.1, the market return correlates with the lagged and contemporaneous investor behaviours. The correlation between the market return and contemporaneous investor behaviour can be up to an absolute value of 0.29, in the case of the SSE 50 Index. Although the market return modestly correlates with the lagged investor behaviour, both the linear model and non-parametric model demonstrate that the lagged investor behaviour is an efficient input in the in-sample prediction. The contemporaneous investor behaviour is powerful in explaining the market return, but it cannot be employed to predict the market.

However, the correlation matrix in Table 5.1 provides some other useful information on investor behaviour. It is observed that the lagged investor behaviour highly correlates to the contemporaneous investor behaviour, especially within the same day. Using the correlation matrix of the CSI 500 Index returns as an example, $q_{a,t-1}$ (previous day's investor behaviour in the afternoon) correlates to $q_{m,t}$ (current investor behaviour in the morning) with a value of 0.34, and $q_{m,t}$ correlates to $q_{a,t}$ (current investor behaviour in the afternoon) with a value of 0.66. These significant correlations, especially the correlation between $q_{m,t}$ and $q_{a,t}$, encourage the idea of a two-step prediction. First, predict the investor behaviour for the next half day; second, employ the forecasted investor behaviour and the observed S&P500 to predict the market return. The core idea is that the market return is predicted indirectly through forecasting the investor behaviour, provided by the properties that the investor behaviour is predictable and the contemporaneous relationship between the investor behaviour and market return is significant. To legalise the first step of the prediction procedure stated above, consider the in-sample forecasting regression models:

$$q_{m,t} = \beta_0 + \beta_1 q_{a,t-1} + e_t \tag{5.20}$$

$$q_{a,t} = \beta_0 + \beta_1 q_{m,t} + \varepsilon_t \tag{5.21}$$

where $q_{a,t-1}$ is the previous day's investor behaviour in the afternoon, $q_{m,t}$ is the current investor behaviour in the morning, and $q_{a,t}$ is the current investor behaviour in the afternoon. The estimated results are displayed as the following:

CSI 500	$\widehat{q_{m,t}} = -0.0369+0.2827q_{a,t-1}$	$R^2 = 0.1156$
	$\widehat{q_{a,t}} = -0.0287+0.7885q_{m,t}$	$R^2 = 0.4294$
CSI 300	$\widehat{q_{m,t}} = 0.1409+0.2922q_{a,t-1}$	$R^2 = 0.0893$
	$\widehat{q_{a,t}} = -0.0457+0.6966q_{m,t}$	$R^2 = 0.4665$
SSE 50	$\widehat{q_{m,t}} = -0.1146+0.2077q_{a,t-1}$	$R^2 = 0.0548$
	$\widehat{q_{a,t}} = -0.0116+0.8074q_{m,t}$	$R^2 = 0.5144$

Interestingly, $q_{a,t}$ can be accurately predicted by $q_{m,t}$. Using the SSE 50 as an example, the beta is 0.8074 (t-statistic = 27.82), and the R^2 is 0.5144. This indicates a good fit. The significant predictive power of $q_{m,t}$ on $q_{a,t}$ is robust because evidence is also found in the cases of the CSI 500 and CSI 300. Generally, the relatively large beta indicates that $q_{a,t}$ moves proportionally with $q_{m,t}$, and the high R^2 shows that $q_{m,t}$ captures most of the volatility of $q_{a,t}$. On the other hand, $q_{a,t-1}$ has weak predictive power on $q_{m,t}$. Using the CSI 500 as an example, the beta is only 0.2827 (t-statistic = 7.12), and the R^2 is 0.1156. Compared to previous estimated results, it seems that $q_{a,t-1}$ is not a good predictor for $q_{m,t}$. In other words, within a day, the investor behaviour has a similar pattern after the lunch break but varies overnight. A good explanation can be provided: there is plenty of domestic and overseas information released that will shock the Chinese market, then crystallise the impacts on the investor behaviour, while investors need time to digest the new information and form the new investor behaviour.

Distinguished from normal predictive models, this thesis' predictive method requires knowing β_1 in Equations (5.20) and (5.21) in advance. In-sample prediction fails to satisfy this requirement, but out-of-sample prediction is able to solve this problem by providing the β_1 using the trained sample. In addition, out-of-sample prediction is applicable in practice and avoids the over-fitting issue. Campbell and Thompson's (2007) R_{OS}^2 is employed to measure the out-of-sample prediction performance. This statistic is computed as:

$$R_{OS}^2 = 1 - \frac{\sum_{t=1}^{T}(r_t - \widehat{r_t})^2}{\sum_{t=1}^{T}(r_t - \overline{r_t})^2} \tag{5.22}$$

where r_t is the forecasted return from the predictive model estimated through the period t-1, and r_t is the historical average return estimated through the period t-1. If R_{OS}^2 is a positive value, r_t is superior to the historical average return $\overline{r_t}$ in term of mean squared forecast error. For interest of comparison, both the linear model and non-parametric models will be considered in out-of-sample forecasting. If it is assumed that the window size is k (the fixed number of observations for sample training), there are T-k-1 rolling windows (out-of-sample evaluation periods). More specifically, r k = 300, 400, 500 will be used to test the robustness of the out-of-sample prediction. In the application of the linear model, the betas are estimated through the period f t-1, then predict the return at the date t. For the non-parametric models, the relationship between the investor behaviour and coefficients is obtained through the period t-1. Given the investor behaviour at t-1 or forecasted investor behaviour at t, the corresponding coefficients are found, then forecast the return at date t.

[Insert Table 5.6]

Panel A of Table 5.6 reports the out-of-sample forecasting results using $q_{a,t-1}$ and $q_{m,t}$ to predict $r_{m,t}$ and $r_{a,t}$, respectively. This predictive model is applicable in practice. Before the Chinese market opens on day t, the S&P500 return or closing price for day t-1 and the investor behaviour in the afternoon of day t-1 ($q_{a,t-1}$) are obtained. Thus, the return in the morning can be predicted using the predictive model estimated through the period t-1. Next, the investor behaviour in the morning of day t ($q_{m,t}$) at the lunch break is obtained. Then the afternoon return can be predicted using the predictive model estimated through the period until noon on day t. To set a control, assume the default model as the linear model with the only explanatory variable being the S&P500 return.

When forecasting the CSI 500 return in the morning ($r_{m,t}$), the default model delivers negative R^2_{OS}, indicating higher MSFE (mean square predictive error) than the historical average. In contrast, the linear model (specified in Equation (5.8)) delivers positive R^2_{OS} (varying from 0.14% to 0.35%), indicating lower MSFE (mean square predictive error). This result demonstrates the predictive ability of $q_{a,t-1}$ on $r_{m,t}$. In contrast, the performance of non-parametric models is controversial because both non-parametric models deliver negative R^2_{OS}. Besides, the performance of the varying coefficient model is highly sensitive to the bandwidth selection. The larger the bandwidth, the more improvement in R^2_{OS}. This model only delivers very modestly positive R^2_{OS} with a large bandwidth selection (here, h = 0.3 or 0.4) and long trained sample period (here, the window size is 500). Among all models, a longer trained sample period provides improvements on R^2_{OS}. However, this fails to find the predictive power of $q_{a,t-1}$ on $r_{m,t}$ using data from the CSI 300 and SSE 50. Among all models, all R^2_{OS} are negative regardless of the trained sample periods. In summary, when forecasting the CSI 500 return in the morning ($r_{m,t}$), $q_{a,t-1}$ presents modest out-of-sample predictive ability in the linear model.

When forecasting the CSI 500 return in the afternoon ($r_{a,t}$), in contrast, the default model delivers positive R^2_{OS}, indicating lower MSFE than the historical average. In other words, the S&P500 return on date t-1 presents some predictive power on the CSI 500 return in the afternoon on date t. If the morning investor behaviour $q_{m,t}$ is added to the model, the positive R^2_{OS} increases relative to those estimated from the default model. Clearly, this result suggests that $q_{m,t}$ has out-of-sample predictive power on $r_{a,t}$. As for the performance of the non-parametric models, the varying coefficient model generates positive R^2_{OS} if the window size is k = 400 or 500. Particularly, if the window size is k = 400 and bandwidth selection h = 0.1, the largest R^2_{OS} is obtained that has a value of 2.62%. In the case of predicting the CSI 300 afternoon return on date t, some positive R_{OS}^2 are observed. If the window size is set at k = 500, the default model and linear model deliver positive R_{OS}^2 of 0.78% and 0.54%, respectively. In addition, if the bandwidth selection is h = 0.4, the varying coefficient model improves R_{OS}^2 to 1.40%. Similar results are observed in the prediction of $r_{a,t}$ of the SSE 50. If the window size is set at k = 500, the default model and Banerjee's model deliver positive R_{OS}^2 of 0.36% and 0.34%, respectively. Surprisingly, if the bandwidth selection is h = 0.4, the varying coefficient model improves R_{OS}^2 to 4.08%. Strangely, the linear model generates negative R_{OS}^2 no matter the length of trained sample period. Generally, the predictive power of $q_{a,t-1}$ on $r_{m,t}$ is very poor since most R_{OS}^2 are negative, while $q_{m,t}$ shows some explanatory power on $r_{a,t}$ in the varying coefficient model.

Panel B of Table 5.6 reports the out-of-sample forecasting results using $\widehat{q_{m,t}}$ and $\widehat{q_{a,t}}$ to predict $r_{m,t}$ and $r_{a,t}$, respectively. The contemporaneous explanatory power of $q_{m,t}$ and $q_{a,t}$ on the corresponding index returns has been shown, but the market cannot be predicted. Instead, $q_{m,t}$ and $q_{a,t}$ are replaced by $\widehat{q_{m,t}}$ and $\widehat{q_{a,t}}$ in the models. As shown earlier, $\widehat{q_{m,t}}$ and $\widehat{q_{a,t}}$ are estimated in Equations (5.20) and (5.21). Thus, $\widehat{q_{m,t}}$ and $\widehat{q_{a,t}}$ can be used to predict the corresponding market returns at date t. When $\widehat{q_{m,t}}$ is employed to forecast the CSI 500 return in the morning ($r_{m,t}$), improvements are observed compared to the results in Panel A. Most R_{OS}^2 exceed their counterparts in Panel A. For example, if k = 500, the R_{OS}^2 produced by the linear model is 0.79%, which is larger than the corresponding R_{OS}^2 (0.35%) in Panel A. Apart from the linear model, both non-parametric models do not provide obvious improvements using $\widehat{q_{m,t}}$. Most R_{OS}^2 produced by the non-parametric models are slightly negative. If $\widehat{q_{a,t}}$ is used to forecast the CSI 500 return in the afternoon ($r_{a,t}$), the R_{OS}^2 of the linear model are 1.13%, 1.67%, and 1.93% for k = 300, 400, and 500, respectively. These results indicate that $\widehat{q_{a,t}}$ has stronger out-of-sample predictive ability than $q_{m,t}$ in the linear model. In both non-parametric models, $\widehat{q_{a,t}}$ does not exhibit superior predictive power relative to $q_{m,t}$. A similar, but slightly different pattern is found when the estimated contemporaneous investor behaviour is used to predict the CSI 300 or SSE 50 return. Neither $\widehat{q_{m,t}}$ nor $\widehat{q_{a,t}}$ exhibit out-of-sample predictive power since the majority of R_{OS}^2 are negative.

Furthermore, all of the models fail to beat the default model on the judgement of R^2_{OS}, indicating that the estimated contemporaneous investor behaviour does not provide any predictive power.

Overall, the S&P500 return at date t-1 shows limited predictive power on the morning return on date t ($r_{m,t}$), but it does have predictive power on the afternoon return on date t ($r_{a,t}$), particularly on $r_{a,t}$ of the CSI 500. Combining $r_{sp,t-1}$ with $q_{a,t-1}$ or $q_{m,t}$, the linear models consistently provide out-of-sample forecasting ability on $r_{m,t}$ and $r_{a,t}$ of the CSI 500. Compared to $q_{a,t-1}$ or $q_{m,t}$, $\widehat{q_{m,t}}$ and $\widehat{q_{a,t}}$ exhibit stronger out-of-sample predictive ability on $r_{m,t}$ and $r_{a,t}$ of the CSI 500. Both $q_{m,t}$ and $\widehat{q_{a,t}}$ show some predictive power on $r_{a,t}$ in the varying coefficient model, but the performance is very sensitive to bandwidth selection. In contrast, neither the lagged or estimated contemporaneous investor behaviour has out-of-sample predictive ability on the CSI 300 or SSE 50 return. Only the S&P500 return at date t-1 may help predict the afternoon return of the CSI 300 or SSE 50 at date t.

5.6. Combined Investor Behaviour

In order to further investigate the predictive power of investor behaviour, the linear model is extended by including three measures of investor behaviour. Similarly, estimate:

$$r_{m,t} = \beta_0 + \beta_1 r_{sp,t-1} + \beta_2 q^{500}_{a,t-1} + \beta_2 q^{300}_{a,t-1} + \beta_2 q^{50}_{a,t-1} + \varepsilon_t \tag{5.23}$$

$$r_{m,t} = \beta_0 + \beta_1 r_{sp,t-1} + \beta_2 q^{500}_{m,t} + \beta_2 q^{300}_{m,t} + \beta_2 q^{50}_{m,t} + \varepsilon_t \tag{5.24}$$

$$r_{a,t} = \beta_0 + \beta_1 r_{sp,t-1} + \beta_2 q^{500}_{m,t} + \beta_2 q^{300}_{m,t} + \beta_2 q^{50}_{m,t} + \varepsilon_t \tag{5.25}$$

$$r_{a,t} = \beta_0 + \beta_1 r_{sp,t-1} + \beta_2 q^{500}_{a,t} + \beta_2 q^{300}_{a,t} + \beta_2 q^{50}_{a,t} + \varepsilon_t \tag{5.26}$$

where $r_{m,t}$ is the half-daily return on the underlying index in the morning; $r_{a,t}$ is the half-daily return on the underlying index in the afternoon; $r_{sp,t-1}$ is the previous day's S&P500 return; $q^{500}_{a,t-1}$, $q^{300}_{a,t-1}$, and $q^{50}_{a,t-1}$ are the previous day's afternoon investor behaviours that developed from the ETF500, ETF300, and ETF50, respectively; $q^{500}_{m,t}$, $q^{300}_{m,t}$, and $q^{50}_{m,t}$ are today's morning investor behaviours that developed from the ETF500, ETF300, and ETF50, respectively; and $q^{500}_{a,t}$, $q^{300}_{a,t}$, and

$q_{a,t}^{50}$ are today's afternoon investor behaviours that developed from the ETF500, ETF300, and ETF50, respectively.

[Insert Table 5.7]

Panel A of Table 5.7 reports the in-sample estimation results for the Equations (5.23) to (5.26). Compared to the results in Table 5.4, it is found that the combination of three lagged measures of investor behaviour provides very limited improvements on adjusted R^2 and MSE. However, the combination of three contemporaneous measures of investor behaviour strengthens the explanatory power of the underlying index return. For example, if the underlying index is the CSI 500, the adjusted R^2 in Equations (5.23) and (5.25) are 0.06% and 1.22%, respectively, while the adjusted R^2 in Equations (5.24) and (5.26) are 6.62% and 11.55%, respectively. This pattern is robust and persistent if the underlying index becomes the CSI 300 or SSE50.

This finding is not surprising because the coefficients on additional measures of investor behaviour in (5.23) and (5.25) are trivial and statistically insignificant at the 0.05 level, while the coefficients on additional measures of investor behaviour in (5.24) and (5.26) are relatively large and statistically significant. For example, if the underlying index is the CSI 500, the coefficients on $q_{a,t-1}^{500}$, $q_{a,t-1}^{300}$, and $q_{a,t-1}^{50}$ in Equation (5.23) are -0.0031 (t-statistic = -1.63), -0.0004 (t-statistic = -0.14), and 0.0000 (t-statistic = 0.01), respectively. Clearly, the additional measures of afternoon investor behaviour developed from the ETF300 and ETF50 at date t-1 have very weak predictive power on the CSI 500 morning return at date t. In contrast, the coefficients on $q_{a,t}^{500}$, $q_{a,t}^{300}$, and $q_{a,t}^{50}$ in (5.26) are -0.0079 (t-statistic = -4.52), -0.0117 (t-statistic = -3.80), and -0.0075 (t-statistic = -1.98). This result indicates that the additional measures of afternoon investor behaviour provide extra explanatory power on the CSI 500 afternoon return. Similar evidence is also provided by the regressions on the CSI 300 or SSE50 returns.

Panel B of Table 5.7 reports the out-of-sample forecasting results for Equations (5.23) and (5.25). Unfortunately, it is found that the combination of measures of investor behaviour exhibits weak out-of-sample predictive ability for the Chinese market. For example, when predicting the CSI 500 morning return, the corresponding R_{OS}^2 are all negative, regardless of the length of window size.

When predicting the CSI 500 afternoon return, the R_{OS}^2 are smaller than the corresponding R_{OS}^2 in Table 5.6. If the CSI 300 or SSE50 returns are predicted, the R_{OS}^2 are smaller than the corresponding R_{OS}^2 in Table 5.6.

5.7. A Trading Strategy

This section proposes a simple trading strategy based on the empirical analysis of investor's behaviour. Because the morning return is very difficult to predict, this strategy only focuses on avoiding potential losses in the afternoon. The strategy defines three criteria and four scenarios. The three criteria are: market declines in the afternoon at t – 1; S&P 500 increases at t – 1; investors tend to buy more in the morning at t – 1. The four scenarios are defined as:

Scenario 1. None of the criteria is met. Scenario 2. Only one of the criteria is met.

Scenario 3. Two of the criteria is met. Scenario 4. All of the criteria is met.

During the lunch break at t, the strategy is to take a short position at scenario 2 and take a long position at other scenarios. The aim of this strategy is to keep potential gains and avoid potential losses. Without transaction cost, the total gain of this strategy on ETF500, ETF300 and ETF50 is 55.26%, 31.84% and 34.35% in the sample period. The total gain for the corresponding benchmark is 32.64%, 6.2% and 15.90% in the sample period. The daily risk adjusted return on ETF500, ETF300 and ETF50 is 0.0005 (t-statistics = 1.95), 0.0004 (t-statistics = 1.73) and 0.0004 (t-statistics = 1.54). The market beta is 0.68 (t-statistics = 38.28), 0.56 (t-statistics = 29.84) and 0.49 (t-statistics = 26.04) respectively.

Transaction cost in Chinese stock market includes two main parts: stamp duty (0.1%) and commission fee (0.3%). Taking a short position in stock market, it will trigger the stamp duty and commission twice for selling at the beginning of afternoon at t and buying back at the beginning of morning at t+1. In other words, every scenario 2 means a loss of 0.8%. Considering the transaction cost, the daily risk adjusted return on ETF500, ETF300 and ETF50 is -0.0024 (t-statistics = -8.48), -0.0029 (t-statistics = -10.56) and -0.0032 (t-statistics = -10.96). The significantly negative adjusted returns indicate that this strategy does not make profits in actual trading.

However, there are other approaches improving this strategy. First, the main part of the transaction cost is the commission fee. In fact, the commission fee can be negotiated to be down if the amount of investment is big. Second, it is more economically efficient to take a short position in derivative market and does not necessary to sell the underlying assets. For example, delta hedging is a good alternative option for taking a short position.[29]

5.8. Conclusion

In this chapter, a measure of the Chinese investor behaviour has been developed that uses the information of the corresponding ETF. With this new measure, it is shown that this thesis' measure of investor behaviour has strong explanatory power on the contemporaneous half-daily index return, in both the linear and non-parametric models. In addition, the measure of investor behaviour shows some forecasting ability for the Chinese market return, and two-step prediction improves the forecasting ability for the CSI 500 afternoon returns. Moreover, there is evidence that the Chinese stock market return has a significant impact on the US stock market return, at least during the sample period of 16th April 2013 to 31st March 2016.

The finding of this chapter provide more empirical evidence to the argument that the Chinese market is not efficient. The inefficiency of the Chinese market is well documented in existing literature (see, e.g., Yao *et al.* (2014)). In addition, the measure of the Chinese investor behaviour is a complement of investor behaviour, which enriches the literature of investor behaviour. Finally, the trading strategy discussed in this chapter may provide some inspirations of investors.

However, there are many open questions left by this study. First, the forecasting ability of the measure of sentiment is not strong enough to apply in practice. Can the forecasting ability increase if the data is hourly or even by the minute in frequency? Second, it is shown that the CSI 500 and CSI 300 returns are not sensitive to the previous day's S&P500 return. Thus, the Hang Seng Index or Singapore A50 be used? Finally, the put/call ratio is a widely accepted investor behaviour. If the put/call ratio is included in the models, can it contribute to predicting the Chinese stock market? If

[29] Delta hedging is an options strategy that aims to reduce, or hedge, the risk associated with price movements in the underlying asset, by offsetting long and short positions.

these questions are well answered, a high-frequency trading strategy may be developed that can make profits in actual trading.

Table 5-1 Summary Statistics for the Relevant Index Returns and Measures of Investor Behaviour

$r^{500}_{m,t}$, $r^{300}_{m,t}$ and $r^{50}_{m,t}$ are the morning returns on the CSI500, CSI300, and SSE50 Indices respectively at date t. $r^{500}_{a,t}$, $r^{300}_{a,t}$ and $r^{50}_{a,t}$ are the afternoon returns on the CSI500, CSI300, and SSE50 Indices respectively at date t. q is the first version of investor behaviour that uses the total ask/bid volumes, while q10 is the second version that only uses the closest 10 bid/ask volume of below/above the corresponding price in the order book. For the first version, $q^{500}_{m,t}$, $q^{300}_{m,t}$ and $q^{50}_{m,t}$ are the measures of morning investor behaviour that are developed from ETF500, ETF300, and ETF50 respectively at date t. $q^{500}_{a,t-1}$, $q^{300}_{a,t-1}$ and $q^{50}_{a,t-1}$ are the measures of afternoon investor behaviour that are developed from ETF500, ETF300, and ETF50 respectively at date t-1, and $q^{500}_{a,t}$, $q^{300}_{a,t}$ and $q^{50}_{a,t}$ are at date t. We also report all these variables above in the second version. For example, $q10^{500}_{m,t}$, $q10^{300}_{m,t}$ and $q10^{50}_{m,t}$ are also the measures of morning investor behaviour that are developed from ETF500, ETF300 and ETF50 respectively at date t, but they are the second version of investor behaviour. $r_{sp,t-1}$ is the daily *S&P500* return at the date t-1. For every variable, we report the time-series mean (Mean), standard deviation (Std), skewness (Skew), kurtosis (Kurt), minimum (Min) and maximum (Max). For variables developed from the same ETFs, we report the correlation matrix. The sample period is through 16th April 2013 through 31st March 2016.

	Summary Statistics						Correlation								
ETF 500	Mean	Std	Skew	Kurt	Min	Max	$r^{500}_{m,t}$	$r^{500}_{a,t}$	$r_{sp,t-1}$	$q^{500}_{a,t-1}$	$q^{500}_{m,t}$	$q^{500}_{a,t}$	$q10^{500}_{a,t-1}$	$q10^{500}_{m,t}$	$q10^{500}_{a,t}$
$r^{500}_{m,t}$	0.11%	1.31%	-1.40	11.97	-0.09	0.06	1.00	0.00	-0.04	-0.07	-0.22	-0.11	-0.01	-0.10	0.07
$r^{500}_{a,t}$	0.05%	1.39%	-0.39	8.21	-0.06	0.07	0.00	1.00	0.09	0.02	-0.09	-0.23	0.00	0.03	0.04
$r_{sp,t-1}$	0.05%	0.85%	-0.28	4.90	-0.04	0.04	-0.04	0.09	1.00	-0.07	0.03	0.02	0.01	0.03	-0.03
$q^{500}_{a,t-1}$	-7.48%	29.87%	0.00	3.28	-1.00	0.88	-0.07	0.02	-0.07	1.00	0.34	0.33	0.22	0.03	0.02
$q^{500}_{m,t}$	-5.81%	24.84%	0.01	3.14	-0.73	0.74	-0.22	-0.09	0.03	0.34	1.00	0.66	0.11	0.33	0.10
$q^{500}_{a,t}$	-7.45%	29.88%	-0.01	3.27	-1.00	0.88	-0.11	-0.23	0.02	0.33	0.66	1.00	0.04	0.09	0.22
$q10^{500}_{a,t-1}$	0.79%	33.63%	-0.43	3.51	-1.00	0.92	-0.01	0.00	0.01	0.22	0.11	0.04	1.00	0.29	0.27
$q10^{500}_{m,t}$	-6.18%	29.37%	-0.36	3.34	-0.92	0.79	-0.10	0.03	0.03	0.03	0.33	0.09	0.29	1.00	0.49
$q10^{500}_{a,t}$	0.80%	33.63%	-0.43	3.51	-1.00	0.92	0.07	0.04	-0.03	0.02	0.10	0.22	0.27	0.49	1.00
ETF 300	Mean	Std	Skew	Kurt	Min	Max	$r^{300}_{m,t}$	$r^{300}_{a,t}$	$r_{sp,t-1}$	$q^{300}_{a,t-1}$	$q^{300}_{m,t}$	$q^{300}_{a,t}$	$q10^{300}_{a,t-1}$	$q10^{300}_{m,t}$	$q10^{300}_{a,t}$
$r^{300}_{m,t}$	0.08%	1.12%	-0.37	7.96	-0.06	0.05	1.00	-0.08	-0.08	-0.03	-0.26	-0.25	0.00	-0.14	0.10
$r^{300}_{a,t}$	0.01%	1.25%	-0.69	8.29	-0.08	0.05	-0.08	1.00	0.07	0.05	0.00	-0.28	0.00	0.02	-0.13
$r_{sp,t-1}$	0.05%	0.85%	-0.28	4.91	-0.04	0.04	-0.08	0.07	1.00	-0.06	0.08	0.01	-0.03	0.02	-0.03
$q^{300}_{a,t-1}$	6.53%	22.13%	-0.32	3.57	-0.86	0.76	-0.03	0.05	-0.06	1.00	0.30	0.29	0.29	0.17	0.10

$q_{m,t}^{300}$	16.00%	21.64%	-0.48	3.31	-0.69	0.64	-0.26	0.00	0.08	0.30	1.00	0.68	0.17	0.39	0.17
$q_{a,t}^{300}$	6.58%	22.07%	-0.32	3.59	-0.86	0.76	-0.25	-0.28	0.01	0.29	0.68	1.00	0.10	0.19	0.29
$q10_{a,t-1}^{300}$	-2.98%	26.67%	0.05	3.35	-0.99	0.92	0.00	0.00	-0.03	0.29	0.17	0.10	1.00	0.30	0.22
$q10_{m,t}^{300}$	-1.51%	24.98%	-0.18	3.02	-0.93	0.72	-0.14	0.02	0.02	0.17	0.39	0.19	0.30	1.00	0.43
$q10_{a,t}^{300}$	-3.02%	26.70%	0.05	3.34	-0.99	0.92	0.10	-0.13	-0.03	0.10	0.17	0.29	0.22	0.43	1.00
ETF 50	Mean	Std	Skew	Kurt	Min	Max	$r_{m,t}^{50}$	$r_{a,t}^{50}$	$r_{sp,t-1}$	$q_{a,t-1}^{50}$	$q_{m,t}^{50}$	$q_{a,t}^{50}$	$q10_{a,t-1}^{50}$	$q10_{m,t}^{50}$	$q10_{a,t}^{50}$
$r_{m,t}^{50}$	0.07%	1.21%	0.31	7.23	-0.06	0.07	1.00	-0.11	-0.09	0.01	-0.29	-0.20	0.05	-0.12	-0.06
$r_{a,t}^{50}$	0.02%	1.30%	-0.37	9.59	-0.08	0.06	-0.11	1.00	0.06	0.01	-0.02	-0.29	-0.07	0.06	0.01
$r_{sp,t-1}$	0.05%	0.85%	-0.28	4.91	-0.04	0.04	-0.09	0.06	1.00	-0.10	0.15	0.12	-0.08	0.00	-0.03
$q_{a,t-1}^{50}$	9.63%	19.03%	-0.48	4.09	-0.71	0.79	0.01	0.01	-0.10	1.00	0.23	0.18	0.17	0.00	0.00
$q_{m,t}^{50}$	13.46%	16.88%	-0.60	4.32	-0.61	0.67	-0.29	-0.02	0.15	0.23	1.00	0.72	-0.02	0.21	0.16
$q_{a,t}^{50}$	9.71%	19.01%	-0.48	4.10	-0.71	0.79	-0.20	-0.29	0.12	0.18	0.72	1.00	-0.05	0.03	0.17
$q10_{a,t-1}^{50}$	-3.76%	17.69%	0.02	6.06	-0.87	0.94	0.05	-0.07	-0.08	0.17	-0.02	-0.05	1.00	0.16	0.14
$q10_{m,t}^{50}$	-1.17%	15.86%	-0.65	4.25	-0.70	0.43	-0.12	0.06	0.00	0.00	0.21	0.03	0.16	1.00	0.44
$q10_{a,t}^{50}$	-3.78%	17.71%	0.02	6.05	-0.87	0.94	-0.06	0.01	-0.03	0.00	0.16	0.17	0.14	0.44	1.00

Table 5-2 Regression of the Measure of Investor Behaviour on Corresponding Index Return

This table shows the estimates of regressions that use the measure of investor behaviour at date t or t-1 to estimate the corresponding index return. Panel A, B, and C report the relevant results for regressions on CSI500, CSI300 and SSE50 return respectively. For every regression, we report the number of observations. r_t^{500}, r_t^{300} and r_t^{50} are the half-daily returns on the CSI500, CSI300 and SSE50 Indices respectively at date t. q_t^{500}, q_t^{300} and q_t^{50} are the first version of investor behaviour, and $q10_t^{500}$, $q10_t^{300}$ and $q10_t^{50}$ are the second version. Y is the dependent variable, Int is the intercept, X is the independent variable, Obs is the number of observations. All variables are at half-daily frequency. The sample period is through 16th April 2013 to 31st March 2016.

	Y	Int	X	R^2	Obs	MSE
Panel A:	r_t^{500}		q_t^{500}			
ETF500	coefficient	0.0001	-0.0109	0.05	1390	1.727E-04
	t-statistics	0.25	-7.63			
	r_t^{500}		q_{t-1}^{500}			
	coefficient	0.0006	-0.0039	0.00	1389	1.818E-04
	t-statistics	1.49	-2.58			
	r_t^{500}		$q10_t^{500}$			
	coefficient	0.0008	-0.0013	0.00	1390	1.815E-04
	t-statistics	2.09	-1.05			
	r_t^{500}		$q10_{t-1}^{500}$			
	coefficient	0.0008	0.0006	0.00	1389	1.818E-04
	t-statistics	2.28	0.45			
Panel B:	r_t^{300}		q_t^{300}			
ETF300	coefficient	0.0020	-0.0139	0.07	1394	1.317E-04
	t-statistics	5.74	-8.59			
	r_t^{300}		q_{t-1}^{300}			
	coefficient	0.0006	-0.0011	0.00	1393	1.414E-04
	t-statistics	1.67	-0.68			
	r_t^{300}		$q10_t^{300}$			
	coefficient	0.0003	-0.0063	0.02	1394	1.387E-04
	t-statistics	1.05	-3.79			
	r_t^{300}		$q10_{t-1}^{300}$			
	coefficient	0.0005	0.0004	0.00	1393	1.415E-04
	t-statistics	1.52	0.24			
Panel C:	r_t^{50}		q_t^{50}			
ETF50	coefficient	0.0027	-0.0197	0.08	1394	1.446E-04
	t-statistics	6.12	-9.00			
	r_t^{50}		q_{t-1}^{50}			
	coefficient	0.0005	-0.0006	0.00	1393	1.575E-04
	t-statistics	1.20	-0.29			
	r_t^{50}		$q10_t^{50}$			
	coefficient	0.0004	-0.0036	0.00	1394	1.571E-04
	t-statistics	1.11	-1.53			
	r_t^{50}		$q10_{t-1}^{50}$			
	coefficient	0.0006	0.0041	0.00	1393	1.574E-04

t-statistics	1.67	1.95

Table 5-3 Using S&P500 Return to Forecast the Chinese Market

This table summarises the impact of *S&P500* return at date t-1 on the Chinese market at date t. All the variables have been introduced in Table 1 and Table 2. For every Chinese market index, we investigate its morning, afternoon and all day's reaction to previous day's *S&P500* return. Y is the dependent variable, Int is the intercept, X is the independent variables, R^2 is the adjusted R-squared coefficient, Obs is the number of observations for every regression, MSE is the mean squared error of predictor. The sample period is through 16th April 2013 to 31st March 2016.

	Y	Int	X	R^2	Obs	MSE
Panel A:	r_t^{500}		$r_{sp,t-1}$			
ETF500	coefficient	0.0016	0.0938	0.00	694	3.647E-04
	t-statistics	2.19	0.88			
	$r_{m,t}^{500}$		$r_{sp,t-1}$			
	coefficient	0.0012	-0.0541	0.00	694	1.714E-04
	t-statistics	2.30	-0.76			
	$r_{a,t}^{500}$		$r_{sp,t-1}$			
	coefficient	0.0004	0.1479	0.01	694	1.907E-04
	t-statistics	1.02	1.53			
Panel B:	r_t^{300}		$r_{sp,t-1}$			
ETF300	coefficient	0.0009	0.0055	0.00	696	2.613E-04
	t-statistics	1.46	0.05			
	$r_{m,t}^{300}$		$r_{sp,t-1}$			
	coefficient	0.0009	-0.1037	0.00	696	1.250E-04
	t-statistics	1.98	-1.71			
	$r_{a,t}^{300}$		$r_{sp,t-1}$			
	coefficient	0.0000	0.1092	0.00	696	1.565E-04
	t-statistics	0.09	1.20			
Panel C:	r_t^{50}		$r_{sp,t-1}$			
ETF50	coefficient	0.0009	-0.0404	0.00	696	2.811E-04
	t-statistics	1.32	-0.33			
	$r_{m,t}^{50}$		$r_{sp,t-1}$			
	coefficient	0.0007	-0.1345	0.01	696	1.441E-04
	t-statistics	1.55	-2.16			
	$r_{a,t}^{50}$		$r_{sp,t-1}$			
	coefficient	0.0002	0.0941	0.00	696	1.693E-04
	t-statistics	0.39	0.91			
Panel D:	$r_{sp,t}$		r_t^{500}			
S&P 500	coefficient	0.0004	0.0675	0.01	696	1.907E-04
	t-statistics	1.16	3.04			
	$r_{sp,t}$		r_t^{300}			
	coefficient	0.0004	0.0792	0.00	696	1.565E-04
	t-statistics	1.29	3.05			
	$r_{sp,t}$		r_t^{50}			
	coefficient	0.0004	0.0706	0.00	696	1.693E-04
	t-statistics	1.33	2.69			

Table 5-4 Forecasting and Explaining the Chinese Market Returns in Linear Models

This table shows the in-sample estimates of regression and slopes using linear models. Panel A, B, and C summarise the estimation results for CSI500, CSI300 and SSE50 indices respectively. Y is the dependent variable, Int is the intercept, X1 and X2 are the independent variables, R^2 is the adjusted R-squared coefficient, Obs is the number of observations for every regression, MSE is the mean squared error of predictor. For every regression, the variables are defined in Table 1 and 2. The sample period is through 16th April 2013 to 31st March 2016.

	Y	Int	X1	X2	R^2	Obs	MSE
Panel A	$r_{m,t}^{500}$		$\mathrm{r_{sp,t-1}}$	$\mathrm{q_{a,t-1}^{500}}$	0.0035	694	1.700E-04
	coeff	0.0009	-0.0620	-0.0031			
	t-stat	1.73	-0.86	-1.64			
	$r_{m,t}^{500}$		$\mathrm{r_{sp,t-1}}$	$\mathrm{q_{m,t}^{500}}$	0.0475	694	1.625E-04
	coeff	0.0005	-0.0423	-0.0117			
	t-stat	0.87	-0.60	-4.62			
	$r_{a,t}^{500}$		$\mathrm{r_{sp,t-1}}$	$\mathrm{q_{m,t}^{500}}$	0.0141	694	1.884E-04
	coeff	0.0001	0.1532	-0.0052			
	t-stat	0.21	1.71	-2.20			
	$r_{a,t}^{500}$		$\mathrm{r_{sp,t-1}}$	$\mathrm{q_{a,t}^{500}}$	0.0569	694	1.803E-04
	coeff	-0.0004	0.1547	-0.0105			
	t-stat	-0.79	1.93	-5.76			
Panel B	$r_{m,t}^{300}$		$\mathrm{r_{sp,t-1}}$	$\mathrm{q_{a,t-1}^{300}}$	0.0044	694	1.245E-04
	coeff	0.0010	-0.1066	-0.0017			
	t-stat	2.14	-1.59	-0.70			
	$r_{m,t}^{300}$		$\mathrm{r_{sp,t-1}}$	$\mathrm{q_{m,t}^{300}}$	0.0708	694	1.162E-04
	coeff	0.0030	-0.0769	-0.0135			
	t-stat	4.95	-1.22	-5.63			
	$r_{a,t}^{300}$		$\mathrm{r_{sp,t-1}}$	$\mathrm{q_{m,t}^{300}}$	0.0026	694	1.560E-04
	coeff	0.0001	0.1099	-0.0004			
	t-stat	0.19	1.28	-0.15			
	$r_{a,t}^{300}$		$\mathrm{r_{sp,t-1}}$	$\mathrm{q_{a,t}^{300}}$	0.0836	694	1.434E-04
	coeff	0.0011	0.1143	-0.0161			
	t-stat	2.53	1.47	-6.14			
Panel C	$r_{m,t}^{50}$		$\mathrm{r_{sp,t-1}}$	$\mathrm{q_{a,t-1}^{50}}$	0.0061	694	1.437E-04
	coeff	0.0008	-0.1350	-0.0002			
	t-stat	1.29	-1.98	-0.08			
	$r_{m,t}^{50}$	$\mathrm{r_{sp,t-1}}$	$\mathrm{r_{sp,t-1}}$	$\mathrm{q_{m,t}^{50}}$	0.0821	694	1.327E-04
	coeff	0.0034	-0.0746	-0.0199			
	t-stat	5.05	-1.11	-6.54			
	$r_{a,t}^{50}$		$\mathrm{r_{sp,t-1}}$	$\mathrm{q_{m,t}^{50}}$	0.0016	694	1.687E-04
	coeff	0.0005	0.1005	-0.0021			
	t-stat	0.64	1.24	-0.66			
	$r_{a,t}^{50}$	$\mathrm{r_{sp,t-1}}$	$\mathrm{r_{sp,t-1}}$	$\mathrm{q_{a,t}^{50}}$	0.0900	694	1.538E-04
	coeff	0.0022	0.1478	-0.0206			
	t-stat	3.42	1.59	-5.82			

Table 5-5 Forecasting and Explaining the Chinese Market Returns in Linear Models

This table compares the mean squared error of predictor between linear models and non-linear models. MC is the coefficient varying model of Fan and Zhang (2008). MB is the non-parametric model of Banerjee and Pitarakis (2014). ML is the linear model that includes previous day's *S&P500* return and the corresponding investor behaviour. MD is the default model that includes only previous day's *S&P500* return as the independent variable. Except for the default model, Y is the dependent variable, X1 and X2 are the independent variables. The corresponding mean squared error of predictor for each model is reported in the same row after the variables. All variables are defined in Table 1 and Table 2. The sample period is through 16th April 2013 to 31st March 2016.

	Y	X1	X2	MC	MB	ML	MD
Panel A	$r_{m,t}^{500}$	$r_{sp,t-1}$	$q_{a,t-1}^{500}$	1.661E-04	1.704E-04	1.700E-04	1.709E-04
	$r_{m,t}^{500}$	$r_{sp,t-1}$	$q_{m,t}^{500}$	1.549E-04	1.617E-04	1.625E-04	1.709E-04
	$r_{a,t}^{500}$	$r_{sp,t-1}$	$q_{m,t}^{500}$	1.834E-04	1.917E-04	1.884E-04	1.901E-04
	$r_{a,t}^{500}$	$r_{sp,t-1}$	$q_{a,t}^{500}$	1.740E-04	1.900E-04	1.803E-04	1.901E-04
Panel B	$r_{m,t}^{300}$	$r_{sp,t-1}$	$q_{a,t-1}^{300}$	1.170E-04	1.245E-04	1.245E-04	1.246E-04
	$r_{m,t}^{300}$	$r_{sp,t-1}$	$q_{m,t}^{300}$	1.104E-04	1.142E-04	1.162E-04	1.246E-04
	$r_{a,t}^{300}$	$r_{sp,t-1}$	$q_{m,t}^{300}$	1.498E-04	1.567E-04	1.560E-04	1.560E-04
	$r_{a,t}^{300}$	$r_{sp,t-1}$	$q_{a,t}^{300}$	1.394E-04	1.447E-04	1.434E-04	1.560E-04
Panel C	$r_{m,t}^{50}$	$r_{sp,t-1}$	$q_{a,t-1}^{50}$	1.398E-04	1.450E-04	1.437E-04	1.437E-04
	$r_{m,t}^{50}$	$r_{sp,t-1}$	$q_{m,t}^{50}$	1.253E-04	1.328E-04	1.327E-04	1.437E-04
	$r_{a,t}^{50}$	$r_{sp,t-1}$	$q_{m,t}^{50}$	1.603E-04	1.695E-04	1.687E-04	1.688E-04
	$r_{a,t}^{50}$	$r_{sp,t-1}$	$q_{a,t}^{50}$	1.427E-04	1.557E-04	1.538E-04	1.688E-04

Table 5-6 Out-of-Sample Forecasting Performance using Different Models

This table compares the out-of-sample forecasting performance of different models. The performance is judged by R_{OS}^2 of Campbell and Thompson (2008). Panel A reports the results using afternoon investor behaviour at date t-1 to predict the morning market return at date t. Panel B reports the results using morning investor behaviour at date t to predict the afternoon market return. Panel C reports the results using predicted morning market return to estimate the contemporaneous morning market return. Panel D reports the results using predicted afternoon market return to estimate the contemporaneous afternoon market return. MC is the coefficient varying model of Fan and Zhang (2008). MB is the non-parametric model of Banerjee and Pitarakis (2014). ML is the linear model that includes previous day's *S&P500* return and the corresponding investor behaviour. MD is the default model that includes only previous day's *S&P500* return as the independent variable. For the varying coefficient model, h means the selection bandwidth. To test the robustness of the forecasting results, we define three choices of window size: 300, 400 and 500. All R_{OS}^2 in this table are expressed in percentages.

	ETF500			ETF300			ETF50		
Panel A	Window Size			Window Size			Window Size		
$q_{a,t-1} \rightarrow r_{m,t}$	300	400	500	300	400	500	300	400	500
MD	-0.35	-0.19	-0.16	-0.14	0.11	0.01	-0.09	0.21	-0.01
ML	0.11	0.15	0.35	-0.56	-0.33	-0.38	-0.74	-0.34	-0.65
MB	-3.24	-2.57	-1.55	-3.29	-1.34	-2.27	-0.21	-0.49	0.24
MC (h = 0.1)	-12.20	-9.51	-6.05	-79.44	-68.86	-86.15	-96.37	-72.24	-83.63
MC (h = 0.2)	-2.64	-1.67	-0.08	-11.23	-7.51	-10.47	-18.53	-17.89	-20.03
MC (h = 0.3)	-1.29	-0.76	0.35	-3.09	-0.96	-1.38	-3.78	-4.09	-4.61
MC (h = 0.4)	-0.78	-0.48	0.44	-1.68	-0.23	-0.38	-1.72	-1.74	-1.83
Panel B	Window Size			Window Size			Window Size		
$q_{m,t} \rightarrow r_{a,t}$	300	400	500	300	400	500	300	400	500
MD	0.44	0.58	1.01	0.10	0.23	0.78	-0.33	-0.25	0.36
ML	0.76	1.07	1.43	-0.39	-0.14	0.54	-1.08	-0.68	-0.03
MB	-0.43	-0.10	-0.24	-3.02	-3.10	-2.14	-0.86	-0.78	0.34
MC (h = 0.1)	-0.01	2.62	1.59	-66.89	-76.60	-83.22	-59.08	-63.55	-65.31
MC (h = 0.2)	-0.20	1.81	0.97	-5.18	-5.57	-6.22	0.70	1.90	1.82
MC (h = 0.3)	-0.08	1.56	0.83	-0.17	0.24	0.78	1.45	2.68	3.66
MC (h = 0.4)	0.00	1.39	0.79	0.45	0.87	1.40	1.75	2.96	4.08
Panel C	Window Size			Window Size			Window Size		
$q^*_{m,t} \rightarrow r_{m,t}$	300	400	500	300	400	500	300	400	500

MD	-0.35	-0.19	-0.16	-0.14	0.11	0.01	-0.09	0.21	-0.01
ML	0.56	0.81	0.79	0.04	0.18	0.72	-0.56	-0.24	-0.40
MB	0.08	-0.16	-0.03	-3.17	-1.39	1.37	-3.15	-4.04	-1.89
MC (h = 0.1)	-1.56	-0.39	-0.31	-0.61	0.01	0.85	-0.94	-0.51	-0.67
MC (h = 0.2)	-0.92	-0.09	0.00	-0.18	0.08	0.64	-0.63	-0.39	-0.76
MC (h = 0.3)	-0.80	-0.13	-0.01	-0.12	-0.01	0.26	-0.67	-0.50	-0.96
MC (h = 0.4)	-0.77	-0.17	-0.03	-0.05	0.02	0.14	-0.66	-0.53	-1.02

Panel D	Window Size			Window Size			Window Size		
$q^*_{a,t} \rightarrow r_{a,t}$	300	400	500	300	400	500	300	400	500
MD	0.44	0.58	1.01	0.10	0.23	0.78	-0.33	-0.25	0.36
ML	1.13	1.67	1.93	-3.46	-2.29	-1.80	-4.60	-3.21	-2.89
MB	-0.36	-0.08	-0.13	-7.43	-5.24	-7.14	-4.58	-1.30	-6.13
MC (h = 0.1)	-0.92	1.05	0.80	-3.16	-1.21	-0.60	-5.93	-3.94	-3.80
MC (h = 0.2)	-0.75	0.83	0.58	-3.16	-1.18	-0.66	-3.67	-1.98	-1.76
MC (h = 0.3)	-0.51	0.88	0.64	-3.07	-1.38	-0.96	-2.98	-0.89	-0.62
MC (h = 0.4)	-0.30	0.97	0.76	-3.06	-1.54	-1.20	-2.56	-0.28	-0.12

Table 5-7 Forecasting Results Produced by Extended Linear Model

This table summarises the in-sample and out-of-sample forecasting results produced by the extended linear model. Panel A reports the estimated coefficients and corresponding t-statistics for every regression. Y is the dependent variable, Int is the intercept, X1, X2, X3, and X4 are the independent variables, R^2 is the adjusted R-squared coefficient, Obs is the number of observations for every regression, MSE is the mean squared error of predictor. Panel B reports R^2_{OS} of Campbell and Thompson (2008) for every out-of-sample forecasting. In every forecasting, the investor behaviour includes three sentiments that are developed from ETF500, ETF300, and ETF50 respectively. For example, $q_{a,t-1}$ includes $q^{500}_{a,t-1}$, $q^{300}_{a,t-1}$ and $q^{50}_{a,t-1}$. Similar to table 6, we define three choices of window size: 300, 400 and 500. All R^2_{OS} in this table are expressed in percentages.

Panel A	Y	Int	X1	X2	X3	X4	R^2	Obs	MSE
	$r^{500}_{m,t}$		$r_{sp,t-1}$	$q^{500}_{a,t-1}$	$q^{300}_{a,t-1}$	$q^{50}_{a,t-1}$			
Coeff		0.0010	-0.0624	-0.0031	-0.0004	0.0000	0.06%	694	1.700
t-stat		1.68	-0.86	-1.63	-0.14	0.01			
	$r^{500}_{m,t}$		$r_{sp,t-1}$	$q^{500}_{m,t}$	$q^{300}_{m,t}$	$q^{50}_{m,t}$			
Coeff		0.0018	-0.0315	-0.0105	-0.0097	0.0025	6.62%	694	1.588
t-stat		1.94	-0.44	-4.07	-3.32	0.53			
	$r^{500}_{a,t}$		$r_{sp,t-1}$	$q^{500}_{m,t}$	$q^{300}_{m,t}$	$q^{50}_{m,t}$			
Coeff		0.0003	0.1600	-0.0051	0.0010	-0.0030	1.22%	694	1.883
t-stat		0.54	1.84	-2.18	0.30	-0.81			
	$r^{500}_{a,t}$		$r_{sp,t-1}$	$q^{500}_{a,t}$	$q^{300}_{a,t}$	$q^{50}_{a,t}$			
Coeff		0.0013	0.1764	-0.0079	-0.0117	-0.0075	11.55%	694	1.686
t-stat		2.29	2.23	-4.52	-3.80	-1.98			
	$r^{300}_{m,t}$		$r_{sp,t-1}$	$q^{500}_{a,t-1}$	$q^{300}_{a,t-1}$	$q^{50}_{a,t-1}$			
Coeff		0.0009	-0.1098	-0.0015	-0.0012	-0.0001	0.29%	694	1.246
t-stat		1.62	-1.65	-0.86	-0.43	-0.05			
	$r^{300}_{m,t}$		$r_{sp,t-1}$	$q^{500}_{m,t}$	$q^{300}_{m,t}$	$q^{50}_{m,t}$			
Coeff		0.0032	-0.0545	-0.0049	-0.0094	-0.0084	9.34%	694	1.133
t-stat		4.30	-0.85	-2.32	-3.82	-2.41			
	$r^{300}_{a,t}$		$r_{sp,t-1}$	$q^{500}_{m,t}$	$q^{300}_{m,t}$	$q^{50}_{m,t}$			
Coeff		0.0001	0.1174	-0.0030	0.0014	-0.0031	0.45%	694	1.556
t-stat		0.09	1.43	-1.40	0.51	-0.87			
	$r^{300}_{a,t}$		$r_{sp,t-1}$	$q^{500}_{a,t}$	$q^{300}_{a,t}$	$q^{50}_{a,t}$			
Coeff		0.0013	0.1451	-0.0063	-0.0095	-0.0116	13.05%	694	1.359
t-stat		2.18	2.02	-3.36	-3.73	-3.19			
	$r^{50}_{m,t}$		$r_{sp,t-1}$	$q^{500}_{a,t-1}$	$q^{300}_{a,t-1}$	$q^{50}_{a,t-1}$			
Coeff		0.0006	-0.1371	-0.0013	-0.0013	0.0011	0.47%	694	1.439
t-stat		1.08	-1.89	-0.68	-0.46	0.33			
	$r^{50}_{m,t}$		$r_{sp,t-1}$	$q^{500}_{m,t}$	$q^{300}_{m,t}$	$q^{50}_{m,t}$			
Coeff		0.0039	-0.0719	-0.0017	-0.0088	-0.0140	10.13%	694	1.299
t-stat		5.43	-1.08	-0.88	-3.62	-4.62			
	$r^{50}_{a,t}$		$r_{sp,t-1}$	$q^{500}_{m,t}$	$q^{300}_{m,t}$	$q^{50}_{m,t}$			
Coeff		0.0002	0.1002	-0.0023	0.0014	-0.0025	0.07%	694	1.688
t-stat		0.24	1.17	-0.99	0.48	-0.61			

	$r_{a,t}^{50}$		$r_{sp,t-1}$	$q_{a,t}^{500}$	$q_{a,t}^{300}$	$q_{a,t}^{50}$			
Coeff		0.0016	0.1362	-0.0058	-0.0079	-0.0141	12.17%	694	1.483
t-stat		2.42	1.61	-2.55	-2.75	-3.52			
Panel B		ETF500			ETF300			ETF50	
Window Size	300	400	500	300	400	500	300	400	500
$q_{a,t-1} \rightarrow r_{m,t}$	-0.75%	-0.50%	-0.20%	-1.80%	-1.23%	-0.88%	-2.30%	-1.32%	-1.42%
$q_{m,t} \rightarrow r_{a,t}$	-0.53%	0.14%	0.68%	-1.47%	-0.71%	0.04%	-2.19%	-1.49%	-0.94%
$q_{m,t} \rightarrow r_{m,t}$	7.13%	7.58%	8.34%	8.57%	8.92%	10.93%	8.92%	9.93%	11.65%
$q_{a,t} \rightarrow r_{a,t}$	10.19%	10.40%	11.65%	9.67%	10.61%	11.35%	7.16%	8.46%	7.87%

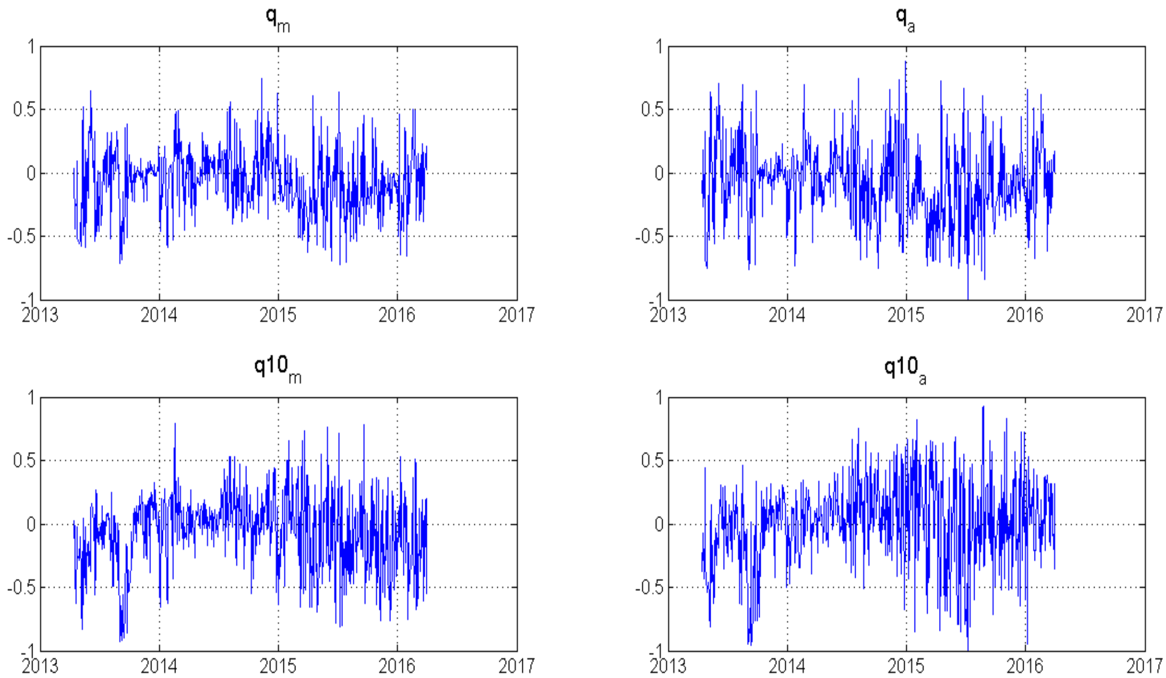

Figure 5.1 The Measures of Investor Behaviour Developed from ETF500

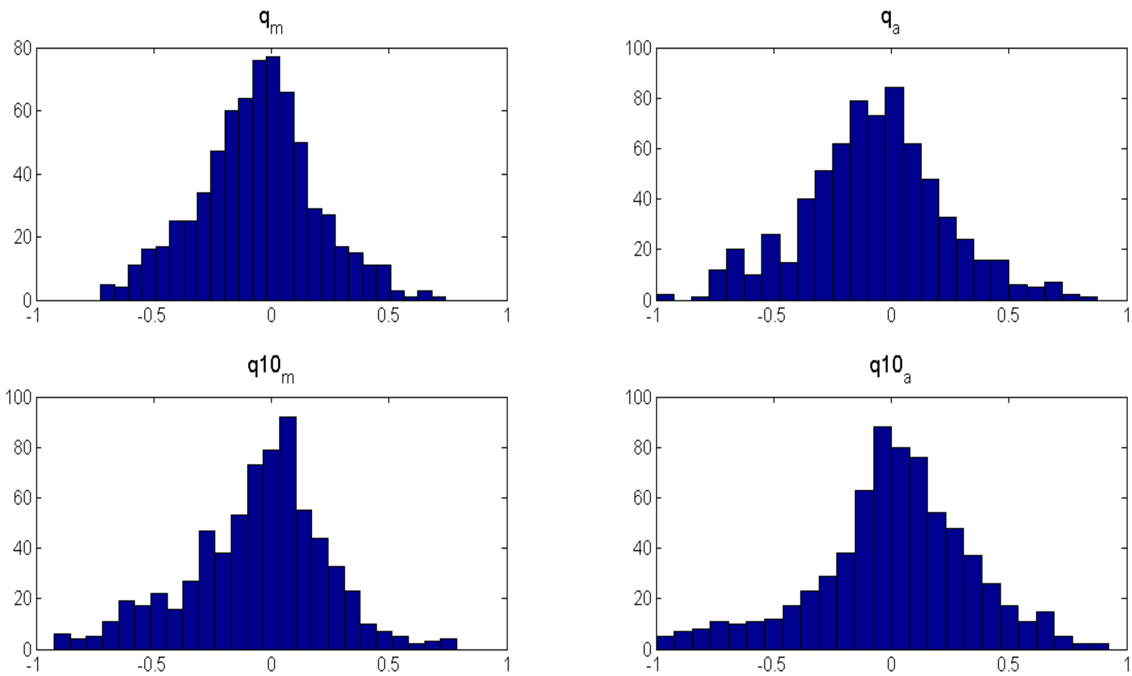

Figure 5.2 The Frequency Distribution of Investor Behaviours Developed from ETF500

6. Thesis Conclusion

6.1. Legacy of the Thesis

This thesis extends ETF research to the topics of asset pricing, mutual fund performance, and market prediction. Although plenty of studies have focused on these topics, no study approaches them through the ETFs. Due to the special characteristics of the ETF market, it provides more useful information that needs to be investigated. First, the ETF market dramatically increases the investment universe and securitizes many non-tradable and illiquid assets. This property makes the ETF market superior to the stock market as the proxy for the theoretical portfolio, according to Roll's (1977) critique. Second, a passively managed index ETF normally tries to match the performance of the underlying index as closely as possible, which makes the underlying index tradable like other stocks listed on the stock exchange. For example, an ETF can provide the order book information of the underlying index.

The main question of chapter 3 is whether the information contained in ETFs contributes to new asset pricing models. Based on the first characteristics of the ETF market, a parsimonious asset pricing model was developed through identifying sources of risk from the ETF market. The performance of this model can compete with the five-factor model of Fama and French (2015) and the q-factor model of Hou *et al.* (2015). More importantly, this research provides an alternative approach for describing anomalies at the firm level. For instance, Dittmar and Lundblad (2017) summarise 55 anomaly portfolios. In the period of January 1963 to December 2007 many anomalies were uncovered by previous studies: size, book-to-market, momentum, financial distress, liquidity, and so on. Particularly after the 2007-08 financial crisis, investment and profitability anomalies were documented in some empirical research. Current studies continue to search for sources of risk to construct risk factors to describe most anomalies.[30] However, this research shows that the ETF market contains more information about the significant alphas.

[30] KUEHN, Simutin and Wang (2017) develops a labour capital asset pricing model by investigating the impact of labour search on the cross-section of equity returns. He, Kelly and Manela (2017) tries to use shocks to the equity capital ratio of financial intermediaries to explain the cross-sectional variation in expected returns.

The main question of chapter 4 is whether the ETF-factor models are valid benchmarks for measuring the mutual fund performance. Conventional models (e.g. the three-factor model of Fama and French (1993), the four-factor model of Carhart (1997) and the q-factor model of Hou, Xue and Zhang (2015)) and ETF-factor models are used to measure the US equity fund performance. The plausible results provide support for the validity of the ETF-factor model proposed in chapter 3. In addition, chapter 4 develops passive relative benchmarks for measuring the fixed-income fund performance. This attempt is important for two reasons. First, it extends the work of Gebhardt *et al.* (2005) and other studies. Second, it is the first time to construct a risk factor from the bond ETFs. Moreover, this research contributes to the work of Fama and French (2010) on mutual fund performance by providing more empirical results.

The main question of chapter 5 is whether a new measure of investor behaviour can predict the Chinese stock market. Given the trading variables provided by the index ETFs, new measures of Chinese investor behaviour are developed using the ask/bid volumes of the ETF500, ETF300, and ETF50. This new measure of investor behaviour successfully extends ETF research to the field of behavioral finance and enriches the existing literature of behavioral finance. The 2015-16 turbulence in the Chinese stock market shows that investors do not make decisions following traditional financial theory. When the Chinese stock market slumps, the real economy still performs satisfactorily. There are numerous reasons why this turbulence occurs, but one of them is the irrationality of investors. This research uses the information provided by index ETFs as tools to investigate how investors make irrational financial decisions. The findings of this chapter enriches the literature of investor sentiment and asset pricing (see e.g. Brown and Cliff (2005), Baker and Wurgler (2006)).

6.2. Summary of Contributions

This thesis extends ETF research to the fields of asset pricing, mutual fund performance, and market prediction. First, chapter 3 proposes a two-factor model through identifying risk factors from the ETF market. This approach is in line with modern portfolio theory. In other words, chapter 3 attempts to describe the anomalies under the framework of traditional finance. Second, chapter 4 applies the ETF-factor model to measuring the US equity mutual fund performance and develops pricing models for measuring the US fixed-income mutual fund performance. Finally, chapter 5 develops a new measure of investor behaviour using the ask/bid volumes provided by the index ETFs. This approach is in line with behavioral finance. Chapter 5 shows that the measure of investor behaviour has explanatory power on the contemporaneous market return. Moreover, the market can be predicted by the analysis of investor behaviour. More detailed findings are presented as follows:

Chapter 3 identifies the sources of risk from the ETF market because it increases the investment universe and securitizes more non-tradable and illiquid assets. Based on the types of ETFs, five value-weight portfolios were constructed: the equity, bond, commodity, gold, and alternative ETF portfolios. Based on the GRS statistics, a long position in the commodity ETF portfolio and a short position in the bond ETF portfolio help describe the returns of seven sets of FF 25 double-sort portfolios. Thus, a risk factor is constructed that is the difference between the return on the commodity ETF portfolio and the return on the bond ETF portfolio. This factor describes the anomalies related to profitability, size, and liquidity, at least in the sample period used. Moreover, the excess return on the ETF market is used to replace the market factor in CAPM. This alternative market factor captures the size and value premiums. By including these two factors, the ETF-factor model with factors comprised of ETFs is developed. Based on the GRS tests, this ETF-factor model is superior to the FF five-factor and q-factor models in describing the returns on FF portfolios.

The findings of chapter 3 provide more evidence to the existing literature of empirical assert pricing model. The ETF-factor model is an alternative option of the five-factor model (Fama and French, 2015), the q-factor model (Hou, Xue and Zhang, 2015) and the four-factor model (Carhart, 1997). The cross-sectional regression results confirm the superior performance of the five-factor model (Fama and French, 2015), while the time-series regression results show that the ETF-factor models are at least as good as the five-factor model and the q-factor model in the sample period. In summary, this chapter provides some innovative ideas to the asset pricing research.

Chapter 4 applies the ETF-factor models, along with other conventional models, to measuring the US equity mutual fund performance. The performance of aggregate portfolios and individual funds are investigated. In aggregate, US equity funds perform very close to the market and exhibits no abnormal return in the period April 2009 to December 2016. However, the simulation results conclude the existence of inferior managers during the same period. Compared to the FF five-factor and q-factor models, the ETF-factor results are plausible. This confirms the validity of the ETF-factor model. In addition, this chapter identifies the systematic risk from the bond ETFs and constructs the bond market factor. Combined with bond risk factors discussed in previous studies, several relevant bond pricing models to measure the US fixed-income fund performance are proposed. The simulation results show that the six-factor model is the best for examining the performance of individual corporate and government fixed-income funds.

The findings of chapter 4 provide more empirical evidence to the existing literature of mutual fund performance. The ETF-factors represent the investment opportunities that can be accessed by investors, particularly those retail investors. This is the biggest advantage of using ETF-factors because the long-short portfolios in Fama and French factor models are not accessible in practice. Threfore, the abnormal returns measured by the Fama and French factor models do not provide any useful information to the investors. In addition, the passive benchmarks for fixed-income funds enrich the literature of pricing models for bonds. Finally, the results of measured manager skills confirm the conclusion of Fama and French (2010) that inferior managers do exist.

Chapter 5 attempts to explain the stock anomalies and predict the market based on analysis of investor behaviour. A new measure of the investor behaviour is developed using the bid/ask volumes provided by the index ETFs. Combined with the previous day's S&P 500 return, it is found that the measure of investor behaviour has strong explanatory power on the contemporaneous half-daily index return. In addition, the measure of investor behaviour has in-sample and out-of-sample forecasting ability. Another interesting phenomenon is that the measure of investor behaviour tends to continue within a day. Thus, a two-step prediction process is proposed based on this continuity. The out-of-sample R squared coefficient is employed to measure the prediction performance. The measured results show the robust forecasting ability of the measure of investor behaviour in the two-step prediction process. The investor behaviour reflected by the index ETFs is the key to explaining or predicting the market.

The findings of chapter 5 enrich the literature of behavioural finance and applies the analysis of investor behaviour to actual trading. The measure of investor behaviour quantifies the herding behaviour in the Chinese stock market (Yao *et al.*, 2014). The predictive model proposed in this chapter enriches the literature of market prediction (e.g. Huang, Jiang and Zhou (2015)). Apart from the contribution to the academic research, the predictive model can lead to trading strategy in practice. But more efforts will be needed to develop a profitable strategy.

6.3. Possible Future Studies

This thesis shows the roles that ETFs can play in asset pricing, mutual fund performance, and market prediction. In the 1990s, the ETF market was trivial. ETFs were seen merely as new types of investment vehicles. With the dramatic growth of ETFs, more studies focused on the impact of ETFs on the stock market via liquidity, volatility, transaction costs, and so on. Particularly in the last 2 years, the days of the traditional fund industry are going away, and the age of ETFs is coming. However, ETF research is still limited, particularly when considering the abundant information provided by ETFs.

This research shows that it is possible to study other topics via the ETF market. First, this thesis shows that the ETF factors capture the documented market anomalies. Thus, it will be worthy to investigate the validity of conventional risk factors. For example, many studies show that source of the profitability premium is not systematic risk. However, the profitability of public companies is significantly related to the commodity price movement in the long run. Thus, what is the source of the profitability premium for the volatile commodity market? Further, it is shown that the size and value premiums are captured by the excess return on the ETF market after the latest financial crisis. Thus, will the size and value factors be necessary for the pricing models if there is a better market factor?

Second, this thesis shows that the value-weight portfolio of US equity funds is a replication of the stock market portfolio, which indicates that the performance of superior managers is balanced by that of inferior managers. However, the simulation results only confirm the existence of inferior managers. Thus, it will be interesting to investigate this conflict. Are some managers are lucky enough to balance the performance of inferior managers? Or are most inferior managers operating small funds? Furthermore, this thesis proposes several models for measuring the US fixed-income fund performance. Thus, is it possible to identify the sources of risk and achieve a consensus about the pricing model for bonds?

Third, this thesis shows that it is possible to develop a new measure of investor behaviour by using the information provided by the index ETFs. And this investor behaviour has explanatory power and forecasting ability on the stock market. Thus, it will be worthwhile to explore what other useful information is provided by ETFs. For example, is it possible to analyze other statistics provided by ETFs to forecast the future movement of the underlying indexes? In addition, the put-call ratio is a popular tool for reflecting the overall sentiment of the market. Thus, it will be interesting to investigate the relationship between this thesis' measure of investor behaviour and the put-call ratio. Finally, as the investor behaviour has strong explanatory power on the stock market, it will be interesting to explore the role of the investor behaviour in asset pricing.

7. Reference

Acharya, V.V., Amihud, Y., Bharath, S.T., 2013. Liquidity risk of corporate bond returns: conditional approach. Journal of Financial Economics 110, 358-386

Acharya, V.V., Pedersen, L.H., 2005. Asset pricing with liquidity risk. Journal of financial Economics 77, 375-410

Ackert, L.F., Tian, Y.S., 2008. Arbitrage, liquidity, and the valuation of exchange traded funds. Financial markets, institutions & instruments 17, 331-362

Agapova, A., 2011. Conventional mutual index funds versus exchange-traded funds. Journal of Financial Markets 14, 323-343

Agarwal, V., Hanouna, P., Moussawi, R., Stahel, C., 2016. Do ETFs increase the commonality in liquidity of underlying stocks? Work. Pap., Villanova Univ

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of financial markets 5, 31-56

Amihud, Y., Mendelson, H., 1986. Asset pricing and the bid-ask spread. Journal of financial Economics 17, 223-249

Amihud, Y., Mendelson, H., 1989. The Effects of Beta, Bid‐Ask Spread, Residual Risk, and Size on Stock Returns. The Journal of Finance 44, 479-486

Ang, A., Hodrick, R.J., Xing, Y., Zhang, X., 2006. The cross‐section of volatility and expected returns. The Journal of Finance 61, 259-299

Baker, M., Wurgler, J., 2006. Investor sentiment and the cross‐section of stock returns. The Journal of Finance 61, 1645-1680

Balakrishnan, K., Bartov, E., Faurel, L., 2010. Post loss/profit announcement drift. Journal of Accounting and Economics 50, 20-41

Bandopadhyaya, A., Jones, A.L., 2006. Measuring investor sentiment in equity markets. Journal of Asset Management 7, 208-215

Bandopadhyaya, A., Jones, A.L., 2008. Measures of investor sentiment: A comparative analysis put-call ratio vs. volatility index. Journal of Business & Economics Research 6, 27-34

Banerjee, A., Pitarakis, J.-Y., 2014. Functional cointegration: definition and nonparametric estimation. Studies in Nonlinear Dynamics & Econometrics 18, 507-520

Banz, R.W., 1981. The relationship between return and market value of common stocks. Journal of financial economics 9, 3-18

Barber, B.M., Huang, X., Odean, T., 2016. Which factors matter to investors? Evidence from mutual fund flows. The Review of Financial Studies 29, 2600-2642

Barberis, N., Shleifer, A. and Vishny, R., 1998. A model of investor sentiment1. Journal of financial economics, 49(3), pp.307-343.

Barberis, N. and Thaler, R., 2003. A survey of behavioral finance. Handbook of the Economics of Finance, 1, pp.1053-1128.

Barras, L., Scaillet, O., Wermers, R., 2010. False discoveries in mutual fund performance: Measuring luck in estimated alphas. The journal of finance 65, 179-216

Basu, S., 1977. Investment performance of common stocks in relation to their price‐earnings ratios: A test of the efficient market hypothesis. The journal of Finance 32, 663-682

Ben-David, I., Franzoni, F., Moussawi, R., 2014. Do ETFs increase volatility? National Bureau of Economic Research

Ben-David, I., Franzoni, F., Moussawi, R., 2016. Exchange Traded Funds (ETFs). National Bureau of Economic Research

Berk, J.B. and Van Binsbergen, J.H., 2015. Measuring skill in the mutual fund industry. Journal of Financial Economics, 118(1), pp.1-20.

Berk, J.B., Van Binsbergen, J.H., 2016. Assessing asset pricing models using revealed preference. Journal of Financial Economics 119, 1-23
Berndt, A., Douglas, R., Duffie, D., Ferguson, M., Schranz, D., 2005. Measuring default risk premia from default swap rates and EDFs.
Bhattacharya, A., O'Hara, M., 2016. Can etfs increase market fragility? effect of information linkages in etf markets.
Blake, C.R., Elton, E.J., Gruber, M.J., 1993. The performance of bond mutual funds. Journal of business, 371-403
Bradley, H., Litan, R.E., 2010. Choking the Recovery: Why New Growth Companies Aren't Going Public and Unrecognized Risks of Future Market Disruptions.
Brennan, M.J., Subrahmanyam, A., 1996. Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. Journal of financial economics 41, 441-464
Brooks, C., Fernandez-Perez, A., Miffre, J., Nneji, O., 2016. Commodity risks and the cross-section of equity returns. The British Accounting Review 48, 134-150
Brown, G.W., Cliff, M.T., 2005. Investor sentiment and asset valuation. The Journal of Business 78, 405-440
Burdekin, R.C., Redfern, L., 2009. Sentiment effects on Chinese share prices and savings deposits: The post-2003 experience. China Economic Review 20, 246-261
Campbell, J.Y., Thompson, S.B., 2007. Predicting excess stock returns out of sample: Can anything beat the historical average? The Review of Financial Studies 21, 1509-1531
Carhart, M.M., 1997. On persistence in mutual fund performance. The Journal of finance 52, 57-82
Cederburg, S., O'Doherty, M.S., 2015. Asset-pricing anomalies at the firm level. Journal of Econometrics 186, 113-128
Charteris, A., 2013. The price efficiency of South African exchange traded funds. Investment Analysts Journal 2013, 1-11
Chen, J., 2002. Intertemporal CAPM and the cross-section of stock returns.
Chava, S., Purnanandam, A., 2010. Is default risk negatively related to stock returns? The Review of Financial Studies 23, 2523-2559
Chen, G.-M., Kim, K.A., Nofsinger, J.R., Rui, O.M., 2003. Does investor sophistication influence investing behavior and trading performance? Evidence from China. Unpublished Working Paper
Chen, G., Strother, T.S., 2008. On the contribution of index exchange traded funds to price discovery in the presence of price limits without short selling.
Chen, H., Chong, T.T.L., She, Y., 2014. A principal component approach to measuring investor sentiment in China. Quantitative Finance 14, 573-579
Chi, L., Zhuang, X., Song, D., 2012. Investor sentiment in the Chinese stock market: an empirical analysis. Applied Economics Letters 19, 345-348
Chow, G.C., Liu, C., Niu, L., 2011. Co-movements of Shanghai and New York stock prices by time-varying regressions. Journal of Comparative Economics 39, 577-583
Chu, Q.C., Hsieh, W.L.G., 2002. Pricing efficiency of the S&P 500 index market: Evidence from the Standard & Poor's Depositary Receipts. Journal of Futures Markets 22, 877-900
Cici, G., Gibson, S., 2012. The performance of corporate bond mutual funds: Evidence based on security-level holdings. Journal of Financial and Quantitative Analysis 47, 159-178
Cochrane, J.H., Piazzesi, M., 2002. Bond risk premia. National Bureau of Economic Research
Cuthbertson, K., Nitzsche, D., O'Sullivan, N., 2008. UK mutual fund performance: Skill or luck? Journal of Empirical Finance 15, 613-634
Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic‐based benchmarks. The Journal of finance 52, 1035-1058

De Bondt, W.F., Thaler, R.H., 1987. Further evidence on investor overreaction and stock market seasonality. Journal of finance, 557-581
De Jong, F., Driessen, J., 2012. Liquidity risk premia in corporate bond markets. The Quarterly Journal of Finance 2, 1250006
Dittmar, R.F. and Lundblad, C.T., 2017. Firm characteristics, consumption risk, and firm-level risk exposures. Journal of Financial Economics, 125(2), pp.326-343.
Driessen, J., 2004. Is default event risk priced in corporate bonds? The Review of Financial Studies 18, 165-195
Elton, E.J., Gruber, M.J., Blake, C.R., 1995. Fundamental economic variables, expected returns, and bond fund performance. The Journal of Finance 50, 1229-1256
Engle, R.F., Sarkar, D., 2006. Premiums-discounts and exchange-traded funds. ETFs and Indexing 2006, 35-53
Fama, E.F., French, K.R., 1992. The cross‐section of expected stock returns. the Journal of Finance 47, 427-465
Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of financial economics 33, 3-56
Fama, E.F., French, K.R., 1995. Size and book‐to‐market factors in earnings and returns. The Journal of Finance 50, 131-155
Fama, E.F., French, K.R., 1998. Value versus growth: The international evidence. The journal of finance 53, 1975-1999
Fama, E.F., French, K.R., 2008. Dissecting anomalies. The Journal of Finance 63, 1653-1678
Fama, E.F., French, K.R., 2010. Luck versus skill in the cross‐section of mutual fund returns. The journal of finance 65, 1915-1947
Fama, E.F., French, K.R., 2016. Dissecting anomalies with a five-factor model. The Review of Financial Studies 29, 69-103
Fan, J., Gijbels, I., 1995. Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. Journal of the Royal Statistical Society. Series B (Methodological), 371-394
Fan, J., Zhang, W., 1999. Statistical estimation in varying coefficient models. Annals of Statistics, 1491-1518
Fang, V., Hung, C.-H.D., 2014. Corporate bond prices and idiosyncratic risk: Evidence from Australia. Journal of International Financial Markets, Institutions and Money 33, 99-114
Foye, J., 2017. Testing Alternative Versions of the Fama-French Five-Factor Model in the UK.
Gao, L., Kling, G., 2008. Corporate governance and tunneling: Empirical evidence from China. Pacific-Basin Finance Journal 16, 591-605
Gastineau, G.L., 2004. The benchmark index ETF performance problem: a simple solution. ETFs and Indexing 2004, 62-69
Gebhardt, W.R., Hvidkjaer, S., Swaminathan, B., 2005. The cross-section of expected corporate bond returns: Betas or characteristics? Journal of Financial Economics 75, 85-114
Gibbons, M.R., 1982. Multivariate tests of financial models: A new approach. Journal of financial economics 10, 3-27
Glosten, L.R., Nallareddy, S., Zou, Y., 2016. ETF trading and informational efficiency of underlying securities.
Gokmenoglu, K.K., Fazlollahi, N., 2015. The interactions among gold, oil, and stock market: evidence from S&P500. Procedia Economics and Finance 25, 478-488
Griffin, J.M., Ji, X., Martin, J.S., 2005. Global momentum strategies. The Journal of Portfolio Management 31, 23-39
Groenewold, N., Tang, S.H.K., Wu, Y., 2003. The efficiency of the Chinese stock market and the role of the banks. Journal of Asian Economics 14, 593-609

Gruber, M.J., 1996. Another puzzle: The growth in actively managed mutual funds. The journal of finance 51, 783-810

Guedj, I., Huang, J., 2009. Are ETFs replacing index mutual funds. In: 19th Annual Conference on Financial Economics and Accounting [Working Paper]. American Finance Association annual meeting

Guo, K., Sun, Y., Qian, X., 2017. Can investor sentiment be used to predict the stock price? Dynamic analysis based on China stock market. Physica A: Statistical Mechanics and its Applications 469, 390-396

Gutierrez, R.C., Maxwell, W.F., Xu, D., 2009. On economies of scale and persistent performance in corporate-bond mutual funds.

Hahn, J., Lee, H., 2006. Yield spreads as alternative risk factors for size and book-to-market. Journal of Financial and Quantitative Analysis 41, 245-269

Hamm, S.J., 2014. The effect of ETFs on stock liquidity.

Harper, J.T., Madura, J., Schnusenberg, O., 2006. Performance comparison between exchange-traded funds and closed-end country funds. Journal of International Financial Markets, Institutions and Money 16, 104-122

Haugen, R.A., Baker, N.L., 1996. Commonality in the determinants of expected stock returns. Journal of Financial Economics 41, 401-439

He, Z., Kelly, B. and Manela, A., 2017. Intermediary asset pricing: New evidence from many asset classes. Journal of Financial Economics, 126(1), pp.1-35.

Henriksen, E., Kværner, J.S., 2017. Commodity Prices and Industry Profitability.

Ho, C., Hung, C.-H., 2009. Investor sentiment as conditioning information in asset pricing. Journal of Banking & Finance 33, 892-903

Horowitz, J.L., Loughran, T., Savin, N.E., 2000. The disappearing size effect. Research in Economics 54, 83-100

Hou, K., Szymanowska, M., 2013. Commodity-based consumption tracking portfolio and the cross-section of average stock returns. Unpublished working paper, Ohio State University, Erasmus University Rotterdam

Hou, K., Van Dijk, M.A., 2014. Resurrecting the size effect: Firm size, profitability shocks, and expected stock returns.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach. The Review of Financial Studies 28, 650-705

Houweling, P., Mentink, A., Vorst, T., 2005. Comparing possible proxies of corporate bond liquidity. Journal of Banking & Finance 29, 1331-1358

Hsu, P.-H., Lee, H.-H., Liu, A.Z., Zhang, Z., 2015. Corporate innovation, default risk, and bond pricing. Journal of Corporate Finance 35, 329-344

Huang, D., Jiang, F., Tu, J., Zhou, G., 2014. Investor sentiment aligned: A powerful predictor of stock returns. The Review of Financial Studies 28, 791-837

Huang, D., Jiang, F., Tu, J., Zhou, G., 2015. Investor sentiment aligned: A powerful predictor of stock returns. The Review of Financial Studies 28, 791-837

Huberman, G., Kandel, S., 1987. Mean‐variance spanning. The Journal of Finance 42, 873-888

Israeli, D., Lee, C.M., Sridharan, S.A., 2017. Is there a dark side to exchange traded funds? An information perspective. Review of Accounting Studies 22, 1048-1083

Ivanov, S.I., Jones, F.J., Zaima, J.K., 2013. Analysis of DJIA, S&P 500, S&P 400, NASDAQ 100 and Russell 2000 ETFs and their influence on price discovery. Global Finance Journal 24, 171-187

Jaffe, J., Keim, D.B., Westerfield, R., 1989. Earnings yields, market values, and stock returns. The Journal of Finance 44, 135-148

Jagannathan, R., Wang, Z., 1996. The conditional CAPM and the cross‐section of expected returns. The Journal of finance 51, 3-53
Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. The Journal of finance 48, 65-91
Jegadeesh, N., Titman, S., 2001. Profitability of momentum strategies: An evaluation of alternative explanations. The Journal of finance 56, 699-720
Jensen, M.C., 1968. The performance of mutual funds in the period 1945–1964. The Journal of finance 23, 389-416
Jiang, Y., Guo, F., Lan, T., 2010. Pricing Efficiency of China's Exchange-Traded Fund Market. Chinese Economy 43, 32-49
Jostova, G., Nikolova, S., Philipov, A., Stahel, C.W., 2013. Momentum in corporate bond returns. The Review of Financial Studies 26, 1649-1693
Kaplanski, G., Levy, H., Veld, C., Veld-Merkoulova, Y., 2015. Do happy people make optimistic investors? Journal of Financial and Quantitative Analysis 50, 145-168
Keim, D.B., Stambaugh, R.F., 1986. Predicting returns in the stock and bond markets. Journal of financial Economics 17, 357-390
Koijen, R.S., Lustig, H., Van Nieuwerburgh, S., 2017. The cross-section and time series of stock and bond returns. Journal of Monetary Economics
Kosowski, R., Timmermann, A., Wermers, R., White, H., 2006. Can mutual fund “stars” really pick stocks? New evidence from a bootstrap analysis. The Journal of finance 61, 2551-2595
Krause, T., Ehsani, S., Lien, D., 2014. Exchange-traded funds, liquidity and volatility. Applied Financial Economics 24, 1617-1630
KUEHN, L.A., Simutin, M. and Wang, J.J., 2017. A labor capital asset pricing model. The Journal of Finance, 72(5), pp.2131-2178.
Laurence, M.M., Cai, F., Qian, S., 1997. Weak-form efficiency and causality tests in Chinese stock markets.
Lemmon, M., Portniaguina, E., 2006. Consumer confidence and asset prices: Some empirical evidence. The Review of Financial Studies 19, 1499-1529
Lewellen, J., Nagel, S. and Shanken, J., 2010. A skeptical appraisal of asset pricing tests. Journal of Financial economics, 96(2), pp.175-194.
Li, E.X., Livdan, D., Zhang, L., 2009. Anomalies. The Review of Financial Studies 22, 4301-4334
Li, H., 2007. International linkages of the Chinese stock exchanges: A multivariate GARCH analysis. Applied Financial Economics 17, 285-297
Li, W., Rhee, G., Wang, S.S., 2016. Differences in herding: Individual vs. institutional investors. Pacific-Basin Finance Journal
Lin, C.-C., Chan, S.-J., Hsu, H., 2006. Pricing efficiency of exchange traded funds in Taiwan. Journal of Asset Management 7, 60-68
Lin, H., Wang, J., Wu, C., 2011. Liquidity risk and expected corporate bond returns. Journal of Financial Economics 99, 628-650
Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. The review of economics and statistics, 13-37
Litterman, R.B., Scheinkman, J., 1991. Common factors affecting bond returns. The Journal of Fixed Income 1, 54-61
Liu, W., 2006. A liquidity-augmented capital asset pricing model. Journal of financial Economics 82, 631-671
Longstaff, F.A., Mithal, S., Neis, E., 2005. Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market. The Journal of Finance 60, 2213-2253
Lustig, H., Van Nieuwerburgh, S., Verdelhan, A., 2013. The wealth-consumption ratio. The Review of Asset Pricing Studies 3, 38-94

Madhavan, A., Sobczyk, A., 2015. Price dynamics and liquidity of exchange-traded funds.
Markowitz, H., 1952. Portfolio selection. The journal of finance 7, 77-91
Marshall, B.R., Nguyen, N.H., Visaltanachoti, N., 2013. ETF arbitrage: Intraday evidence. Journal of Banking & Finance 37, 3486-3498
Martikainen, T., Puttonen, V., 1996. Call-put signal predicts Finnish stock returns. Applied Economics Letters 3, 645-648
Massa, M., Yadav, V., 2015. Investor sentiment and mutual fund strategies. Journal of Financial and Quantitative Analysis 50, 699-727
Mensi, W., Beljid, M., Boubaker, A., Managi, S., 2013. Correlations and volatility spillovers across commodity and stock markets: Linking energies, food, and gold. Economic Modelling 32, 15-22
Meziani, A.S., 2016. Exchange-Traded Funds. Springer.
Moskowitz, T.J., Grinblatt, M., 1999. Do industries explain momentum? The Journal of Finance 54, 1249-1290
Nguyen, C., Bhatti, M., 2015. Investor sentiment and idiosyncratic volatility puzzle: evidence from the Chinese stock market.
Ni, Z.-X., Wang, D.-Z., Xue, W.-J., 2015. Investor sentiment and its nonlinear effect on stock returns—New evidence from the Chinese stock market based on panel quantile regression model. Economic Modelling 50, 266-274
Pástor, Ľ., Stambaugh, R.F., 2003. Liquidity risk and expected stock returns. Journal of Political economy 111, 642-685
Petajisto, A., 2017. Inefficiencies in the pricing of exchange-traded funds. Financial Analysts Journal 73, 24-54
Dittmar, R.F. and Lundblad, C.T., 2017. Firm characteristics, consumption risk, and firm-level risk exposures. Journal of Financial Economics, 125(2), pp.326-343.
Roll, R., 1977. A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory. Journal of financial economics 4, 129-176
Rosenberg, B., Reid, K., Lanstein, R., 1985. Persuasive evidence of market inefficiency. The Journal of Portfolio Management 11, 9-16
Ross, S.A., 1976. The arbitrage theory of capital asset pricing. Journal of economic theory 13, 341-360
Schwert, G.W., 2003. Anomalies and market efficiency. Handbook of the Economics of Finance 1, 939-974
Sharpe, W.F., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. The journal of finance 19, 425-442
Stambaugh, R.F., 1982. On the exclusion of assets from tests of the two-parameter model: A sensitivity analysis. Journal of financial economics 10, 237-268
Stambaugh, R.F., Yu, J., Yuan, Y., 2012. The short of it: Investor sentiment and anomalies. Journal of Financial Economics 104, 288-302
Stambaugh, R.F., Yu, J., Yuan, Y., 2014. The long of it: Odds that investor sentiment spuriously predicts anomaly returns. Journal of Financial Economics 114, 613-619
Thomas, J.K., Zhang, H., 2002. Inventory changes and future returns. Review of Accounting Studies 7, 163-187
Titman, S., Wei, K.J., Xie, F., 2004. Capital investments and stock returns. Journal of financial and Quantitative Analysis 39, 677-700
Van Dijk, M.A., 2011. Is size dead? A review of the size effect in equity returns. Journal of Banking & Finance 35, 3263-3274
Wang, X.L., Shi, K., Fan, H.X., 2006. Psychological mechanisms of investors in Chinese Stock Markets. Journal of Economic Psychology 27, 762-780

Yao, J., Ma, C., He, W.P., 2014. Investor herding behaviour of Chinese stock market. International Review of Economics & Finance 29, 12-29