Testing Asset Pricing Models under Non-linear Assumptions: Evidence from UK firm level Panel Data

by

YE JIANG

A thesis submitted to

The University of Birmingham

for the degree of

MASTER BY RESEARCH

Department of Accounting and Finance

Birmingham Business School

University of Birmingham

August.2013

ABSTRACT

The thesis reported here offers the first attempt to test the non-linear relation, i.e. the inverted-U curve hypothesis among firm size and stock return, with respect to firms listed on London Stock Exchange, by employing panel data from 1991 to 2011. The methodology of the study contains relevant asset pricing literature review, theoretical and empirical linear and non-linear model specifications, data collection and cleaning, model estimation and testing, and interpretation and analysis of the empirical results. The thesis is organized into 4 chapters, including an introduction, a literature review, an empirical chapter, and a conclusion. The literature review (2nd) chapter critically reviews the theoretical and empirical asset pricing models and some important considerations in testing and estimating asset pricing models. The empirical paper (3rd) chapter overall investigate whether the effect of firm size on stock return is approximated by an inverted-U curve shape relation, rather than linear: at low levels of firm size the effect is positive and at high levels of firm size the effect is negative. Most importantly, the paper propose that the inverse monotonic relation among firm size and stock return may not be the case instead it makes the first attempt at investing an inverted-U curve relationship, such that the presence of firm size leads to lower stock returns only after the levels of firm size become large enough. The fourth chapter offers the main conclusions and contributions on the thesis, and explains some of the implications and limitations of the study; moreover, some promising ideas for extending and improving this study will be discussed.

To my dearest parents

for all of your support and love

ACKNOWLEDGEMENTS

I would like to express my appreciation to many individuals for their successive assistance and support on completing the thesis.

My greatest sincerely acknowledgement is to my supervisor Professor Victor Murinde, for his valuable and patient comments, advice, support and guidance for my thesis. Professor Victor Murinde's view and enthusiasm on research broaden my horizon and his attitudes on work make me rethink my own career development. Although I cannot complete a PhD degree with him, I still want to express my sincere thanks and best wishes to him. My study has also benefited from Dr. Mark Mak and Professor Mathew Tsamenyi, and Professor Robert Cressy, for their suggestions and encouragement.

I wish to thank academic staff and fellow PhD students of the Accounting & Finance Subject Group at the Birmingham Business School who have commented on my work during the PhD student seminars. Thanks to university library staff for their supporting in searching and utilizing databases and resources. My grateful gratitude to Miss Yan CEN, PhD colleague from Department of Economics, who also do me great favor on settling down in Birmingham. Her kindness and assistance to me make me insist on my faith for the study. My thanks also go to Dr. Joanne Locke, from PhD department and Marleen Vanstockem, from research office, who generously provide patient and kind assistance on my study and life in Birmingham.

In addition to those help I get from the University of Birmingham, I also owe credit to my lecturer and friends as well in Xi'an Jiaotong-Liverpool University, Dr Xiaoming DING, Professor Chi-Yi YANG, and Dr. David LIU. Their frank and experienced guidance and suggestion make a significant influence on my professional study and career development.

Finally, and also most importantly, I wish to show my deepest love and caring to my dearest parents, especially my mum. Without her understanding, I cannot go abroad and study at Birmingham. Equally important, wholehearted thanks to my dad, for his love and financial support for my study in UK. I will love them more than before and report back on what they have done for me.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1 INTRODUCTION

1.1 Background and Motivation of the study

In asset pricing literature, William Sharp (1964) and John Lintner (1965) develop the Capital Asset Pricing Model, which also signals the birth of asset pricing models. CAPM[1] was inferred based upon the Markowitz's (1959) theory. The theory pointed out that investors are risk averse and they are trying to find a portfolio, containing risky assets that will maximize the portfolio expected return for a given level of portfolio risk. Capital Asset Pricing Model assumes investors are risk averse and it also states how investors measure the relationship among risk and expected return. Specifically, the model suggests that there is a linear relation among risk premium and expected return, which can be depicted by beta. The basic idea behind the model is that beta is the only risk investors need to bear when holding an asset, and they call it systematic risk or market risk[2]. In 1993, based on CAPM, Fama and French develop a three-factor model, containing two more factors, i.e. book-to-market and size factor. The comprehensive three factors are market factor, which mainly illustrated by CAPM, book-to-market ratio and market capitalization. Fama and French point that book-to-market ratio and market capitalization are the two factors that can affect stock returns and could be used to explain the asset return anomalies, which cannot be

---

[1] CAPM is the acronym for the Capital Asset Pricing Model.

[2] Systematic risk or market risk in finance is defined as the risk cannot be eliminated through diversification commands requires returns in excess of the risk-free rate. A well-diversified portfolio provides returns which correspond with its exposure to systematic risk.

captured by CAPM. There is also a linear relationship among the expected return of the asset and the three factors. Thereafter, Carhart (1997) adds momentum to Fama and French three-factor model, which signals the birth of Carhart Four-Factor model. The forth factor is considered to be related to investor behaviour. Investors will buy those assets perform quite well and sell those perform badly in the past. Investors may make excess profits based on the past performance of the assets. This profit is called momentum profit. The trading intuition behind the model is that it may have some correlation between past and current asset performance, the trading strategies based on the intuition is called momentum trading. Furthermore, Ross, in 1976, had proposed the Arbitrage Pricing Model (APT[3]). The model states that asset return is affected by *k* macroeconomic factors, rather than a single market factor or some factors. There are also some proxies to measure these macroeconomic factors, e.g. exchange rates, inflation rates, GDP, etc. To make a conclusion, all the above asset pricing models are risk and return models.

Bowman (1980) finds a negative relation among risk and return, when concerning different industries. In the standard CAPM, assets with higher risk, will generate higher required rate of return to its investors. Nevertheless, this phenomenon is violated when negative relation is found between risk and return. A possible explanation is that managers are risk seekers, instead of risk averse assumed in the

[3] APT is the acronym for Arbitrage Pricing Model.

standard CAPM. Besides, motivated by the paper suggests that there is a non-linear relationship among the risk and return, more specific, it is a curvilinear shape (Mukherji, Desai and Wright, 2008).

Furthermore and most importantly, motivated by Amel-Zadeh (2011)'s paper by investigating the size anomaly in the German stock market. When applying the three-factor model in the German stock market, the relation between size and return is conditional on the market condition. When in the 'bull' market, firms with smaller size perform better than firms with bigger size; however, the reverse occurs when the market condition become 'bear', smaller firms have worse performance than larger firms. This finding causes concerns on the wide acknowledged unconditional findings of 'a size discount' when applying the three-factor model in the stock market. Although the paper confirms that small capitalization stocks have different performance from big capitalization stocks, it causes research interests on whether the effect of firm size on stock return is not monotonic decreasing, and whether the effect of firm size on stock return is approximated by an inverted-U Curve shape relation, rather than linear.

Hence, the empirical paper in the thesis tests the non-linear relation, i.e. the Inverted-U Curve hypothesis among firm size and stock return within the Fama and French three-factor model, with respect to firms listed on London Stock Exchange, by employing panel data from 1991 to 2011. In the traditional Fama and French three-factor model, smaller firms tend to generate higher returns, and bigger firms will induce lower returns. However, the thesis is to investigate whether the effect of firm size on stock return is approximated by an inverted-U Curve shape relation, rather than linear: at low levels of firm size the effect is positive and at high levels of firm size the effect is negative. The inflexion point for the inverted-U Curve can be treated as the threshold value between firm size and stock return. It represents the critical point at which the bigger firm size will not increase stock return; on the contrary, it will affect stock return negatively. Before the threshold value, stock return is an increasing function of the firm size, and the positive effect exceeds the negative effect at low levels of firm size, nevertheless, after the threshold point, it becomes a decreasing function and the negative effect becomes a dominate position at relatively high levels of firm size.

1.2 Contribution of the study

The main aim of the thesis is to expand the frontier of research in the areas of asset pricing model and its related factors. In specific, to test the non-linear relation, i.e. the Inverted-U Curve hypothesis among firm size and stock return, by employing UK panel data.

My thesis will contribute to the literature through following aspects. Firstly, the study critically reviews the literature, and starts with a comprehensive literature review of asset pricing model issues. Chapter 2 focuses on the theoretical and also the traditional asset pricing models. In addition, recent most important considerations, such as model specification, influential factors etc, in asset pricing models are critically discussed. The literature review in this chapter also concludes some promising research ideas.

Secondly, this study proposes and implements formal research methodology and procedures, like the data generating process and data cleaning process. This is to reach a consistent research methodology as a whole study.

Thirdly, in asset pricing literature, to my best knowledge, related factors tested in asset pricing models are considered to be linear factors as the independent variables. The empirical paper in the thesis offers the first attempt to test the non-linear relation, i.e. the Inverted-U Curve hypothesis among firm size and stock return, by employing UK panel data. Fama and French three-factor model proposes that smaller firms tend to generate higher returns, and bigger firms will induce lower returns. The paper contribute to asset pricing literature by investigating whether the effect of firm size on stock return is approximated by an inverted-U curve shape relation, rather than linear.

The investigation proposes that at low levels of firm size the effect is positive and at high levels of firm size the effect is negative. The inflexion point for the Inverted-U Curve can be treated as the threshold value between firm size and stock return. It represents the critical point at which the bigger firm size will not increase stock return; on the contrary, it will affect stock return negatively. Overall and most importantly, the paper provides the first trial to propose the inverse monotonic relation among firm size and stock return may not be the case by using UK panel data and offers the first attempt to investigate an inverted-U curve relationship, such that the presence of firm size leads to lower stock returns only after the levels of firm size become large enough.

1.3 Structure of the Thesis

The thesis is structured into four chapters, comprised by an introduction, two freestanding research papers (one is literature review chapter, and the other is an empirical paper) and a conclusion.

Chapter one introduces the background and motivation of the study, explains main contributions of the thesis, and puts forward the main empirical procedures of research methodology advocated, and finally the structure of the study.

Chapter two presents the literature review chapter of the whole thesis. The main purpose of this chapter is to discuss some important considerations in testing asset pricing models. Starting from the most typical Capital Asset Pricing Model, then Fama and French three-factor model, Carhart four-factor model, and Arbitrage Pricing Model and the main concerns in these models are also discusssed. In order to illustrate the non-linear risk and return relation, the quadratic relationship among the dependent and independent variables are introduced, such as the relationship between investment and uncertainty (Lensink and Murinde, 2006); the nexus among inflation and finance-growth (Lensink and Murinde, 2006), and the investigation of how the foreign bank entry affect gross domestic investment (Rousseau and Wachtel, 2002). Finally, this chapter makes a summary on the mains issues considered on asset pricing model and asset pricing models in UK stock market; in addition, some promising research ideas based on the above discussion are also proposed.

Chapter three presents the empirical study. The main aim of this chapter is to fill in the gap in risk and return relation by examining the inverted-U Curve hypothesis among firm size and stock return, with respect to firms listed on London Stock Exchange. The panel form dataset will be used. The gap come from the traditional Fama and French three-factor model, it states that smaller firms tend to generate higher returns, and bigger firms have lower returns, accordingly. That is size factor is a linear monotonic factor that affect stock return when applying the three-factor model, the bigger the firm size, the lower the return. However, the study supposed to find a negative and statistically significant coefficient on the quadratic size factor.

Chapter four concludes previous chapters of the thesis. Key findings and important implications of the study are also presented. Moreover, in the end of the study, some limitations, promising ideas for extending and improving this study are discussed.

CHAPTER 2 IMPORTANT CONSIDERATIONS IN TESTING ASSET PRICING MODELS: LITERATURE REVIEW

2.1 Introduction

The Capital Asset Pricing Model of William Sharpe (1963) and John Lintner (1964) marks the birth of asset pricing model. The model claims that the market factor has explanatory power in stock return and takes into account the asset's sensitivity to market risk, as well as the expected return of the market and the expected return of a theoretical risk-free asset. However, Fama and French (1992) put forward a three-factor model, they augment the single market factor model with two more factors, i.e. book-to-market factor and size factor. Book-to-market is defined as the book value over the market value. Size factor is represented by market capitalization, which is the multiplication of number of shares outstanding and share price. Fama and French (1992)[4] also explain that size and book-to-market equity, combined to capture the cross-sectional variation in average stock returns, which is associated with market beta[5], size, leverage[6], book-to-market equity, and earnings-to-price ratio[7]. Despite

---

[4] Ball (1978) and Keim (1988) also find that variables like size, earning-price ratio, leverage, and book-to-market equity can be regarded as different ways of extracting information from stock prices about the cross-section of expected stock returns.

[5] In finance, market beta is a measure of the risk arising from exposure to general market movements as opposed to idiosyncratic factors. For example, the market portfolio of all investable assets has a beta of exactly 1. A beta below I can indicate either an investment with lower volatility than the market, or a volatile investment whose price movements are not highly corrected with the market.

[6] Leverage, in finance, involves buying more of the assets by using borrowed money. The rationale is that the income from the asset will be more than the pay for the cost of borrowing.

[7] The price-to-earnings ratio is defined as market price per share divided by annual earnings per share.

Fama and French three-factor model add more explanatory power than the Capital Asset Pricing Model, the momentum profit cannot be rationally explained. Momentum profit is first recorded in Jegadeesh and Titman (1993). Jegadeesh and Titman (1993) find that buying stocks that perform quite well in the past and sell the stocks that perform poorly could make a huge profit. By studying mutual fund, Carhart (1997) forms a momentum four-factor model to explain the mutual fund returns. The momentum factor is built based on buying past ‘winners’ and selling past ‘losers’ of the portfolios. In order to better explain stock returns, Carhart four-factor model consists of four factors: market, book-to-market, size and momentum. Ross (1976) argued that the market risk of any single asset or portfolio cannot be fully represented by a single market factor or some factors. Ross proposed that stock returns are influenced by *k* risk factors, and he proposed Arbitrage Pricing Model (APT). APT assumes that the return on any asset or portfolio is a linear function of *k* microeconomic factors, such as inflation rates, exchange rates, etc.

In addition, some important considerations in asset pricing models are investigated and various specifications of asset pricing models have been put forward. New risk factors are introduced into the traditional asset pricing models to see whether or how the factors may affect the performance of the models. Literatures also study whether the inclusion of the factors enhance the goodness-of-fit of the model in stock markets. For example, factors such as inflation (Gerard and Mathijs, 2010); investor sentiment

(Chienwei and Chi-Hsiou, 2009); liquidity risk factor (Pastor and Stambaugh, 2003; Acharya and Pedersen, 2005; Elena et al., 2010; and Lam and Tam, 2011); accounting standards homogeneity (Griffin, 2002; Moerman, 2005; Gomez Biscarri and Lopez Espinosa, 2008); foreign exchange risk factor (Apergis et al., 2011); world market risk and country-specific risk (Bali and Cakici, 2010) and heterogeneous tax (Eikseth and Lindset, 2009), etc have been investigated.

The literature review chapter for this study is structured around three building blocks: traditional asset pricing models (i.e. CAPM, FF3[8], etc), non-linear relation in corporate finance literature and important considerations in asset priding models.

2.2 Capital Asset Pricing Model

2.2.1 Introduction

The Sharpe-Lintner CAPM says that the expected value of an asset's excess return (the asset's return minus the risk-free interest rate,$R_i - R_f$) can be explained by its expected risk premium, i.e. beta[9] times expected value of $R_m - R_f$. Capital Asset Pricing Model also claims that beta is the only linear variable that can explain asset expected returns. It is a theoretical model that seeks to explain the relationship

[8] FF3 is the acronym for Fama and French three-factor model.

[9] Beta, by definition, is a measure of volatility of a security or a portfolio in comparison to the market as a whole. Beta represents the asset's sensitivity to non-diversifiable risk (also known as market risk or systematic risk) in the capital asset pricing model.

between risk and expected return in an efficient market[10]. Mathematically, the model postulates that the expected return on a portfolio or an asset equals the return on the risk-free asset plus a risk premium and can be represented by the following equation:

$$E(R_i) - R_f = \beta_i \left[E(R_m) - R_f \right] \tag{2.1}$$

where

$E(R_i)$ = the expected return on the *i*-th risky asset;

$E(R_m)$ = the expected return on the market portfolio[11];

R_f = the return on the risk-free asset; and

β_i = the asset *i*'s market risk.

Theoretically, CAPM is used to determine an asset's expected rate of return. The market risk of CAPM can be represented by beta (β_i). Beta coefficient is the ratio of the covariance between the asset's return and the return of the market portfolio, and the variance of the market, which can be expressed as:

---

[10] In an efficient market, stocks always trade at their fair value on stock exchanges, making it impossible for investors to either purchase undervalued stocks or sell stocks for inflated prices. Hence, it should be impossible to outperform the overall market through expert stock selection or market timing. The only way an investor can possibly obtain higher returns is by purchasing riskier investments.

[11] Market portfolio is a portfolio consisting of a weighted sum of every asset in the market, with weights in the proportions that they exist in the market.

$$\beta_i = \frac{Cov(R_i, R_m)}{Var(R_m)} \qquad (2.2)$$

Where

Cov (R_i, R_m) = the covariance between the asset's return R_i and the return of the market portfolio R_m;

Var (R_m) = the variance of the return of the market portfolio R_m.

2.2.2 Risk vs Return Relationship

In the theoretical and empirical finance literature, asset pricing model proposes that there is a linear relation between risk and expected return. The traditional risk and expected return models also indicate that the higher the risk embedded in the asset, the greater the corresponding return investors will have. The behaviour pattern in the relation is called risk averse[12], which is one of the basic assumptions behind the Capital Asset Pricing Model[13]. However, many empirical studies have found a negative and nonlinear (i.e. curvilinear) relationship among risk and expected return (Bowman, 1980 & 1982; Aaker and Jacobson, 1987; Sinaee and Moradi, 2010;

[12] Risk averse based on the behaviour of humans while exposed to uncertainty to attempt to reduce that uncertainty. Investors, who are risk averse when faced with two investments with a similar expected return but different risks, will prefer the one with the lower risks.

[13] The model assumes that given a certain expected return, active and potential shareholders will prefer lower risk (lower variance) to higher risk and conversely given a certain level of risk will prefer higher returns to lower ones. It does not allow for active and potential shareholders who will accept lower returns for higher risk.

Mukherji et al., 2008; Kliger and Tsur, 2011; Hunjra et al., 2011; etc).

Bowman (1980) finds a negative relation among the risk and expected return within different industries. In the standard CAPM, stocks with higher risk will generate higher required rate of return to investors. Nevertheless, this phenomenon will be violated when the negative relation is found between the risk and stock return. The paper proposed some possible explanations for this ‘paradox’: for the low risk - high return relation, it may be that in each industry, managers manage the firm quite well and always make rational decisions so that they generate higher returns with lower risk. Another potential explanation is that managers are risk seekers, instead of risk averse assumption in the traditional CAPM. Moreover, it is also possible that the market is imperfect, which also violates the efficient market hypothesis in the.

Kliger and Tsur (2011) use the accounting-based risk measure and test the risk-return relation. Using the median of the industry return as a benchmark, the paper finds that when return exceeds the benchmark, the relation among risk and expected return is positive; however, if return below the benchmark point, there will be a negative relation among the risk and expected return.

2.3 Fama and French Three-Factor Model

2.3.1 Introduction

Fama and French (1993) developed a three-factor model. The Fama and French three-factor model improves the explanation power by adding size and book-to-market factors. Thus, the mathematical model is described as:

$$\boldsymbol{R_{p,t} - R_{f,t} = \alpha_p + \beta_p[R_{m,t} - R_{f,t}] + s_p SMB_t + h_p HML_t + \varepsilon_{p,t}} \quad (2.3)$$

Where

$\boldsymbol{R_{p,t}}$ = the weighted stock return on a constructed portfolio *p* in time *t*;

$\boldsymbol{R_{f,t}}$ = the risk-free rate at time *t*;

$\boldsymbol{\alpha_p}$ = intercept, which should not statistically different from 0 if all the unsystematic risks are diversified away;

$\boldsymbol{R_{m,t}}$= return on the market portfolio in time *t*, so $\boldsymbol{R_{m,t} - R_f}$ is the market risk premium;

$\boldsymbol{SMB_t}$ = the proxy for the size factor, which constructed as the difference between the average return of the three portfolios containing the smallest-capitalization stocks and the return of the three portfolios containing the largest-capitalization stocks;

$\boldsymbol{HML_t}$ = the proxy for the BTM[14] factor, which can be constructed as the difference between the average returns of the stock portfolios with a high BTM ratio and the average returns of the stock portfolios with a low BTM ratio;

$\boldsymbol{\beta_p}$ = the coefficient loading for the excess return of the market portfolio over the risk free rate;

$\boldsymbol{s_p}$ = the coefficient loading for the excess average return of portfolios with small market capitalization over the portfolios with big market capitalizations;

$\boldsymbol{h_p}$ = the coefficient loading for the excess average return of portfolios with high book-to-market ratio over portfolios with low book-to-market ratio.

$\boldsymbol{\varepsilon_{p,t}}$ = error term for a constructed portfolio *p* in time *t*, which should be independent of the risk factors included in the model.

Further, the validity and the empirical specification of the model as well as the microeconomic factors that may affect the three-factor model are extensively investigated by researchers.

Ferguson and Shockley (2003) put forward that when constructing the market portfolio in the Capital Asset Pricing Model, only the equity-based claims are taken

[14] BTM is the acronym for book-to-market ratio, which is defined as the ratio of book value over the market value.

into account, hence the appearance of the size and book-to-market anomalies. Therefore, they propose that both the equity and debt claims should been involved into the market portfolio when estimating the expected stock return[15] and the higher expected return are the compensation for the higher risk the stocks bear. Based on Ferguson and Shockley's paper, Chou, Ko and Lin (2010) propose an augmented five-factor model, including market factor (equity only), size, book-to-market, relative distress level and relative leverage level. Moreover, they also investigate simply the Ferguson-Shockly model and Fama-French model, which include the distress and leverage level factor, and the size and book-to-market factor, respectively. The results show that both of the two models cannot explain the normal asset pricing anomalies individually, such as the momentum and book-to-market effect. As a further result reveals that the augmented-five factor model, i.e. the combination of two Fama and French (1993) factors and two Ferguson-Shockly (2003) factors, is able to explain the asset pricing anomalies. In addition, the paper also implies that the Fama and French three-factor model explains more and better for time-series variations of stock returns, whereas the Ferguson-Shockly model captures more in cross-sectional variations in stock returns.

[15]. Ferguson and Shockley (2003) argue that the economy's debt claims, which have been ignored in empirical asset pricing estimation and testing, are associated with firms leverage risk and distress risk. Therefore, it is essential to include both the equity claim and the debt claim when constructing the market portfolio.

Gómez-Biscarri and López-Espinosa (2008) investigate whether the difference in accounting standards across countries has effect on the performance of asset pricing model. Griffin (2002) also presents that the performance of the Fama and French three-factor model is better on the country-specific basis than on the global-specific basis, and the domestic pricing model has less pricing errors than the world-factor model. There is lack of uniform in accounting standards and accounting systems across countries, hence the relatively poor performance of the asset pricing model that uses accounting variables across countries. Based on the fact that the Fama and French three-factor model uses book value, i.e. accounting variable, to construct the value factor, i.e. HML. Gómez-Biscarri and López-Espinosa (2008) analyse whether the accounting variable has effect on the estimation of model parameters and the validity of overall model performance. The paper applied the three-factor model to domestic countries, international samples that across countries and different countries but using same accounting systems (e.g. IASB[16]), and finds that the domestic countries and the IASB cross-countries have similar validity and accuracy of model performance. The empirical analysis indicates that Fama and French three-factor model is accounting specific. Moreover, the more homogeneous the accounting data, the more consistent the parameters estimation will be, hence the better the performance of the model.

[16] IASB stands for International Accounting Standards Board. It is the independent, accounting standard-setting body of the IFRS Foundation, based in London, England. The International Reporting Standards Foundation, or IFRS Foundation, is a non-profit organization. Its main objectives include the development and promotion of the International Financial Reporting Standards (acronym as IFRS) through the International Accounting Standards Board.

In addition to the overall performance of the three factor model, empirical studies also focus on the size effect in the asset pricing tests. Banz (1981) presents the evidence of size effect in the US market and reports that the size effect is non-linear and firms with smallest size have the most influential size effect. However, Brown et al. (1983b) show that there is an approximately linear relation among the firm size and stock return. A huge amount of literature has been developed on the size effect among different countries after Banz (1981)'s paper and numerous papers aimed to investigate the size anomaly in empirical asset pricing tests. Reinganum (1981) finds that the smallest size firms have better performance than the largest size firms by testing NYSE and Amex firms. Lamoureux and Sanger (1989) report a size premium for the Nasdaq stocks. More recently, Amel-Zadeh (2011) documents that the relation among firm size and stock return is conditional on market condition. That is in the up-market, smaller firms outperform larger firms, nevertheless, in the down-market, the inverse occurs.

From above discussion, size factor in the Fama and French three-factor model may not act as a linear influential factor.

2.3.2 The Size Effect

2.3.2.1 Introduction

Recent research on the size effect investigates whether the validity and persistence of size effect in the cross-sectional stock returns rely on the magnitude of firm size. Moreover, how the magnitude of firm size or the portfolios formed according to the magnitude of firm size affect stock return (Amel-Zadeh, 2011; van Dijk, 2011; De Moor and Serch, 2012; etc). Moreover, evidence on the well known size anomaly[17] – the performance of small capitalisation stocks is better than the large capitalisation stocks, is also extensively examined. Since 1990s, two debates are raised on the size effect, one is the size effect arises endogenously as a consequence of systematic risk and the other is the disappearance of size effect (van Dijk, 2011).

2.3.2.2 Evidence on Size Effect

Banz (1981) first documents the size effect in the US stock market and the size effect is nonlinear. However, Merton (1987) finds a positive and linear relationship among firm size and stock return. Reinganum (1981) finds that the smallest size portfolios outperform the largest by using NYSE and Amex firms. Lamoureux and Sanger (1989) report a size premium for Nasdaq stocks and NYSE & Amex stocks. Empirical works

[17] Size anomaly refers to the size effect in the Fama and French three-factor model, which is the small capitalization stocks outperform large capitalization stock.

have shown the existence and persistence of size effect in the corresponding stock market. Nevertheless, the underlying reason behind the size effect or size premium is still uncertain and the research on the size effect is incomplete. Thereafter, Fama and French (1992) examine the size and book-to-market anomaly, which is uncaptured by the Capital Asset Pricing Model of Sharpe (1964) and Lintner (1965). By investigating the US stock market, Fama and French find that smallest firms outperform the largest. Moreover, Fama and Macbeth (1973) regression also indicates that the size and book-to-market factors have significant explanatory power in asset pricing tests.

Additionally, there are many international evidences for the validity and importance of size effect across countries. The result suggests that there is consistent international evidence on the size premium (van Dijk, 2011). Among 19 countries, small size firms outperform large firms in 18 countries and the monthly size premiums are also reported.

2.3.2.3 Possible Explanations for Size Effect

There are many potential explanations for the size effect, but no widely accepted theory, thus it is still regarded as an anomaly. A consensus has not been reached by researchers why smaller firms earn higher returns than larger firms with regard to the predication of traditional asset pricing models. Empirical papers have discussed the question extensively and many published work may have some explanatory power. Banz (1981) proposed that small stocks have relatively less information hence investors may not want to keep them, thus resulting in the high return for small stocks. Roll (1981) argues that smaller firms are riskier than relatively larger firms thus need to have higher returns to compensate for the extra risks investor bear. Barry and Brown (1984) propose that the differences among the returns for small and large firms are partly due to the different ease of information flow among them. Smaller firms may have less information access and less information credibility than larger firms, thus result in investors' perceive of the higher risks smaller firms may have.

There are also many attempts to explain the size effect from the behavioural finance perspective. Dissanaike (2002) uses UK data and find the size effect in the FT500. The paper also indicates that the size effect and momentum effect are related to each other. Besides, the paper puts forward that it is investors' overconfidence and overreaction that cause size anomaly. The paper also provides evidence that the small size firms in UK stock market are also the firms with negative past stock returns performance. Similarly, Hong et al. (2000) conclude that the stocks with smaller size are also the stocks with more prominent momentum effect, which is mainly due to the slow information spread among small firms. Fan and Liu (2005) present that the winner-loser effect contributes to the size effect in explaining stock returns.

2.3.4 The Book-to-Market (BTM) Effect

2.3.4.1 Introduction

Book-to-market (BTM[18]) is a ratio calculated as the book value over the market value. The existence of the size effect and book-to-market effect is also examined by many researchers (Al-Horani et al., 2003; Dissanaike, 2002; Dimson et al., 2003; Fletcher and Kihanda, 2005; Gregory et al., 2001; Hussain et al., 2002; ect). The book-to-market effect exhibits that average stock returns are greater when the book value to market value (BE/ME) is higher and vice versa. It is also denoted as the value premium[19].

2.3.4.2 Empirical Evidence on BTM Effect

Fama and French (1992) find evidence that value stock, that is stocks with higher book-to-market value, have higher returns than growth stocks, which have lower BTM value. Fama and French (1998) document a premium approximately 7.68% per annual to stocks with higher book-to-market values (i.e. value stocks). Groot and Verschoor (2002) find the value premium is alive in emerging stock markets of Korea,

[19] Value premium refers to the greater expected return of value stocks over growth stocks, i.e. the high book-to-market ratio stocks (value stock) have higher expected returns than the low book-to-market ratio stocks (growth stock).

Malaysia and Thailand. Drew and Veeraghavan (2002) test the Fama and French three-factor model in the Malaysia stock market, and find that overall stocks with smaller size and higher book-to-market value generate higher average stock returns than bigger size and lower book-to-market value. Maroney and Protopapadakis (2002) extend previous research and test the three-factor model on the Canada, Germany, Japan, France, the US, the UK, and Australia stock markets. The BTM effect is significant for all countries under consideration. The study also confirms that BTM effect is internationally existed. Barry et al. (2002) investigate 35 emerging markets to test the three-factor model and find the existence of a value premium.

More recently, Malkiel and Jun (2009) confirm the significance of the value premium in Chinese stock market. Cenesizoglu (2011) test how the daily portfolio returns are changing when there are unexpected changes in microeconomic condition, such as GDP, inflation. The paper finds that changes in return on portfolios formed based on high book-to-market and low book-to-market react quite differently to changes in the predefined macroeconomic variables. This also confirms the existence of BTM effect when applying the three-factor model empirically. Fama and French (2012) detail the value premium in average stock returns for 23 development stock markets.

2.4 Carhart Four-Factor Model

2.4.1 Introduction

Momentum is a strong and challenging anomaly in asset pricing tests. Persistent momentum profit is a violation of the efficient market hypothesis[20] and can be partly explained from the behavioural finance perspective. The basic idea behind the momentum effect is that stocks with higher past returns will have higher future returns than stocks with low past returns. Typically it can be defined as a positive relationship among the return of a stock in a certain period with its lagged returns. The momentum investing strategy simply is buying the past winners (higher past returns) and selling past losers (lower or even negative past returns).

Carhart four-factor model says that the asset return can be explained by four factor – that is market, size, book-to-market ratio and momentum factor. The model specification is as follows:

$$\boldsymbol{R_{p,t} - R_{f,t} = \alpha_p + \beta_p\left[R_{m,t} - R_{f,t}\right] + s_p SMB_t + h_p HML_t + w_p WML_t + u_{p,t}}$$

(2.4)

Where

[20] Efficient market hypothesis assets that financial markets are informationally efficient. All the public or private information are publicly available and no one can earn excess returns based on past information. However, the rationale behind momentum profit is that excess profit can be made according to past trading information, that is buying past winners and selling past losers.

$\boldsymbol{R_{p,t}}$ = the weighted stock return on a constructed portfolio *p* in time *t*;

$\boldsymbol{R_{f,t}}$ = the risk-free rate at time *t*;

$\boldsymbol{\alpha_p}$ = intercept, which should not statistically different from 0 if all the unsystematic risks are diversified away;

$\boldsymbol{R_{m,t}}$= return on the market portfolio in time *t*, so $\boldsymbol{R_{m,t} - R_f}$ is the market risk premium;

$\boldsymbol{SMB_t}$ = the proxy for the size factor, which constructed as the difference between the average return of the three portfolios containing the smallest-capitalization stocks and the return of the three portfolios containing the largest-capitalization stocks;

$\boldsymbol{HML_t}$ = the proxy for the BTM factor, which can be constructed as the difference between the average returns of the stock portfolios with a high BTM ratio and the average returns of the stock portfolios with a low BTM ratio;

$\boldsymbol{WML_t}$ **=** the proxy for the momentum factor at time *t*, which is constructed as the difference between the returns on the winners' portfolio and the returns on the losers' portfolio for a given set of stocks;

$\boldsymbol{\beta_p}$ = the coefficient loading for the excess return of the market portfolio over the risk free rate;

$\boldsymbol{s_p}$ = the coefficient loading for the excess average return of portfolios with small market capitalization over the portfolios with big market capitalizations;

$\boldsymbol{h_p}$ = the coefficient loading for the excess average return of portfolios with high book-to-market ratio over portfolios with low book-to-market ratio.

$\boldsymbol{w_p}$ **=** the coefficient loading for the excess average return of winners' portfolios over losers' portfolios;

$\boldsymbol{u_{p,t}}$ = error term for a constructed portfolio *p* in time *t*, which should be independent of the risk factors included in the model.

2.4.2 Momentum Effect

Momentum effect is typically understood as a positive relation among stock returns with time lags in a certain period. Jegadeesh and Titman (1993) find that an excess profit could be made by buying stocks that performed quite well in the past and selling stocks that performed poorly. Momentum trading strategy is built based on buying past 'winners' and selling past 'losers' of the assets.

In addition to momentum on individual stock, momentum on a country level also exists. Bhojraj and Swaminathan (2001) confirm the existence of momentum on a country level. Lewellen (2002) find industry momentum exists. However, individual momentum is also significant in the sample after taken into account the industry momentum. Nijman et al. (2004) test the European stock market to investigate individual, industry and country momentum. They find the existence of individual and industry momentum, but no country momentum.

2.4.3 Possible Explanations for Momentum Effect

Momentum is the empirically observed tendency for rising asset prices to rise further, and falling prices to keep falling. Jegadeesh and Titman (1999) show that stocks with strong past performance continue to outperform stocks with poor past performance in the next period with an average excess return of about 1% per month. Conrad and Kaul (1998) provide empirical results show that momentum profits could be explained by cross-sectional difference in stock returns rather than time-series difference. However, Hong and Stein (1999), Daniel, Hirshleifer and Subrahmanyam (1998) propose that momentum effect is related to investor behaviours. It is the way investors interpret and react to information, and their behaviour may cause overreaction or underreaction.

2.5 Arbitrage Pricing Model

Fama and French (1993) and Carhart (1997) proposed asset pricing models that intend to mitigate problems presented by CAPM. They complete the task by adding additional variables, such as firm size, long-term past return, Earning/Price, book-to-market ratio, etc. According to their conclusion, relation between stock returns and risks are better explained by augmenting the factors to CAPM. In 1976, Ross argued that systematic risk (the risk that cannot be diversified away through portfolio formation) cannot be fully represented or explained by a single common factor (i.e. market factor) or some factors. Instead, Ross (1976) proposed that there are *k* risk factors that may have impact on stock returns, which is the Arbitrage Pricing Model.

Arbitrage Pricing Model also assumes that the expected rate of return on any asset is a linear function of *k* factors.

We can express the model as

$$\boldsymbol{R_i} = \boldsymbol{E}(\boldsymbol{R_i}) + \boldsymbol{U},\ \boldsymbol{U} = \ \mathbf{m_i} + \ \boldsymbol{\varepsilon}_{\mathbf{i}} \qquad (2.5)$$

Where

$E(R_i)$ = Expected return;

U = Unexpected return;

$\mathrm{m_i}$ = Systematic unexpected return;

ε_{i} = Unsystematic unexpected return.

Derived from the definition of APT, this model can be written as the following *k*-factor model:

$$\boldsymbol{R_i = E(R_i) + b_{i1}F_1 + b_{i2}F_2 + b_{i3}F_3 + \cdots + b_{ik}F_k + e_i} \qquad (2.6)$$

Where,

R_i = the random rate of return on the *i*-th asset;

$E(R_i) =$ expected return on the *i*-th asset;

b_{ik} = sensitivity of the *i*-th asset returns to the *k*-th factor;

F_k = *k*-th factor common to the returns of all assets under consideration;

$\varepsilon_i =$ a noise term for the *i*-th asset, with mean zero.

The *F*s in the above equation can be interpreted as the appropriate proxies for the risk factors, or called microeconomic variables, such as the inflation, term structure of interest rates, exchange rates, etc. From mathematical equation (2.6), APT consists of risky asset's expected return and risk premium of a number of macroeconomic factors.

Arbitrage Pricing Model is a straightforward model to explain asset returns. However, for the empirical application of APT in financial circumstances, it is difficult to construct these risk factors involved in the model. Literature in arbitrage pricing model has presented three main methods to test and estimate the model. First, factor analysis, this is a pure statistical procedure for estimating factors and sensitivities of returns to the factors; second, using five macroeconomic variables, Chen, Roll and Ross (1986) use ‘changes in the monthly growth rate of the GDP, changes in default risk premium, changes in interest rate, unexpected changes in price level and changes in expected inflation’ to construct risk factors in the model; thirdly, using firm characteristics to form factor portfolios.

2.6 Modelling of Quadratic Relationship in Corporate Finance

In the asset pricing test, most factors previous research taken into account are linearly related to the stock return in asset pricing models. Mathematically, the independent variables (e.g. liquidity factor, inflation factor, exchange rate risk factor) put forward on the right-hand side of the asset pricing equation are always linearly related to the dependent variable (e.g. stock return) on the left-hand side. However, in corporate finance, the quadratic relationship among the dependent and independent variables are also modelled and testable, such as the relationship between investment and uncertainty (Lensink and Murinde, 2006); the nexus among inflation and finance-growth (Lensink and Murinde, 2006), and the investigation of how the foreign bank entry affect gross domestic investment (Rousseau and Wachtel, 2002).

Lensink and Murinde (2006) investigate the effect of uncertainty on investment by using UK panel data. Their approach suggests that the relationship between uncertainty and investment mimics an inverted-U shape curve, rather than linear. When testing the empirical model, in addition to the linear term for uncertainly, the paper also includes a quadratic term to model non-linearity and expects to have a negative and statistically significant coefficient on the quadratic term. The paper contributes to the literature in the sense that it provides empirical evidence on the threshold effects among uncertainty and investment – at low levels of uncertainty, the effect is positive, nevertheless, after the critical point the level of uncertainty reach, the effect becomes negative.

Furthermore, Lensink and Murinde (2006) show that the degree of foreign bank entry has different effect on gross domestic investment. The quadratic terms are also involved into the threshold equation and expect to have a positive and statistically significant coefficient. The results the paper obtained show a non-linear relation for the effect of foreign bank entry on the gross domestic investment. In specific, the effect can be treated as a U-curve shape relation. Low levels of foreign bank entry have negative effect on gross domestic investment, however, when the level exceeds the critical point, the effect becomes positive, and after the threshold value, more foreign bank entry will stimulate gross domestic investment. Overall, the paper presents that after the foreign bank entry levels reach high enough, the gross domestic investment will be expanded.

In addition to the above papers employing quadratic term to model nonlinearity, Rousseau and Wachtel (2002) investigate how the strength of the relationship among financial sector development and economic growth varies with inflation. The paper propose the effect inflation has on the finance-growth relation is not monotonic and financial sector is at least part of the channel that inflation influences economic growth. By using a panel of 84 countries, the paper finds that inflation indeed has a threshold effect on the finance-growth relationship. Specifically, at low levels of inflation, which is below the threshold point of the inflation rate, say 13% to 25%, financial sector has a positive effect on economic; however, after the threshold point, finance ceases to stimulate economic growth.

2.7 What are important considerations in Asset Pricing?

In asset pricing models literature, numerous researches have been carried out to study the important considerations in asset pricing models and various specifications of the asset pricing models are put forward by researchers. Specifically, they introduce new variables into the traditional asset pricing models to see whether or how the new factors may affect the performance of those models. They also study whether the inclusion of the factors improves the goodness-of-fit of the model in stock markets. For example, researches have included factors such as inflation (Gerard and Mathijs, 2010); investor sentiment (Chienwei and Chi-Hsiou, 2009); liquidity risk factor

(Elena et al., 2010), and illiquidity factor for the Spanish stock market (Jose and Maria, 2006). New model specifications have also been developed to see how those factors will behave in the overall models, examples such as the liquidity-augmented capital asset pricing models (Weimin, 2006); consumption-based asset pricing model (Joachim and Andreas, 2009) and production-based asset pricing model (Frederico, 2010).

In addition, prior researches also use different testing methods to test the reasonable specifications of the asset pricing models in corresponding stock market, such as the Markov Chain Monte Carlo method to estimation and testing in (Manuel and Michael, 2008); Generalized Methods of Moments for estimation and inferences in the emerging stock market (Javed et al., 2010) etc.

When the empirical estimation and application of the Capital Asset Pricing Model have failures to explain the asset anomalies, Fama and French three-factor model improves the explanation power by adding size and book-to-market factors. Further, the validity and the empirical specification of the model as well as the microeconomic factors that may affect the three-factor model are extensively investigated by researchers.

Chou, Ko and Lin (2010) propose an augmented five-factor model, including market factor (equity only), size, book-to-market, relative distress level and relative leverage level. Moreover, they also investigate simply the Ferguson-Shockly (FS) model and Fama-French (FF3) model, which include the distress and leverage level factor, and the size and book-to-market factor, respectively. The results show that both of the two models cannot explain the normal asset pricing anomalies individually, such as the momentum and book-to-market effect. As a further result reveals that the augmented-five factor model, i.e. the combination of two Fama and French (1993) factors and two Ferguson-Shockly (2003) factors, is able to explain these asset pricing anomalies. In addition, the paper also implies that the Fama and French three-factor model explains more and better for time-series variations of stock returns, whereas the Ferguson-Shockly (FS) model captures more in cross-sectional variations in stock returns.

Gómez-Biscarri and López-Espinosa (2008) investigate whether the difference in accounting standards across countries has effect on the performance of asset pricing model. Griffin (2002) also presents that the performance of the Fama and French three-factor model is better on the country-specific basis than on the global-specific basis, and the domestic pricing model has less pricing errors than the world-factor model. There is lack of uniform in accounting standards and accounting systems across countries, hence the relatively poor performance of the asset pricing model that uses accounting variables or across countries. Based on the fact that the Fama and French three-factor model uses book value, i.e. accounting variable, to construct the value factor, i.e. HML, Gómez-Biscarri and López-Espinosa (2008) analyse whether the accounting variable has effect on the estimation of model parameters and the validity of overall model performance. The paper applied the three-factor model to domestic countries, international samples that across countries and different countries but using same accounting systems (e.g. IASB), and finds that the domestic countries and the IASB cross-countries have similar validity of model performance. The empirical analysis indicates that Fama and French three-factor model is accounting specific. Moreover, the more homogeneous and uniform the accounting data, the more consistent the parameters estimation will be, hence the better the performance of the model.

2.8 Conclusion

In addition to the overall performance of the Fama and French three-factor model, empirical studies also focus on the size effect in the asset pricing tests. Banz (1981) presents the evidence of size effect in the US market and reports that the size effect is non-linear and firms with smallest size have the most influential size effect. However, Brown et al. (1983b) show that there is an approximately linear relation among the firm size and stock return by examining 10 size-based portfolios. A huge amount of literature has been developed on the size effect among different countries after Banz (1981)'s paper and numerous papers aimed to investigate the size anomaly in empirical asset pricing tests. Reinganum (1981) finds that the smallest size decile firms have better performance than the largest size decile firms by testing NYSE and Amex firms. Lamoureux and Sanger (1989) report a size premium for the Nasdaq stocks. More recently, Amel-Zadeh (2011) documents a conditional relation among firm size and stock return. That is in the up-market, smaller firms outperform larger firms, nevertheless, in the down-market, the inverse occurs.

In corporate finance literature, asset pricing models, here Fama and French three-factor model works as a cornerstone. In addition, its empirical applications assist firms and investors make rational investment decisions and create great fortune. Hence, research in testing and analyzing asset pricing models, especially the Fama and French three-factor model is significant.

CHAPTER 3 THE INVERTED-U CURVE HYPOTHESIS FOR THE EFFECT OF FIRM SIZE ON STOCK RETURN: EVIDENCE FROM LONDON STOCK EXCHANGE

3.1 Introduction

In finance literature, research on the size effect has been carried out extensively. The more recent research have focused on the robustness of the size effect in asset pricing models and the potential explanation for the size effect, such as the financial distress risk, liquidity risk, dividend yields effect etc, in different capital markets (De Moor and Sercu, 2012; Amel-Zadeh, 2011; van Dijk, 2011; Fama and French, 2012; Lischewski and Voronkova, 2012; etc).

Banz (1981) is the first document the size effect in the US stock market and propose that the size effect is nonlinear. However, in contrary to Banz's nonlinear size effect, Merton (1987) finds a positive and linear relationship among firm size and stock return. The different conclusion triggered research on the shape and robustness of the relation among firm size and stock return.

The empirical paper aims to extend the literature and with the main objective to test the inverted-U curve hypothesis among firm size and stock return under the framework of Fama and French three-factor model, with respect to the nonfinancial listed firms on London Stock Exchange. By using the UK panel data from 1991 to 2011 on an annual basis, the two-stage-least-squares (2SLS) will be used to estimate the coefficients in the regression model. This method is used to consider the endogeneity problem and omitted variable bias, which may arise in regression model.

3.2 Theoretical & Empirical Models

Models involved in the paper are used to test the nonlinear effect of firm size on stock return. In the Fama and French three-factor model, size acts as an explanatory variable that influences stock return. SMB portfolios constructed based on the magnitude of firm size also show that firms with smaller size generate higher stock returns and firms with larger size have lower stock returns respectively. Therefore, size factor is considered as having an inverse monotonic linear effect on stock return. Nevertheless, the effect of firm size on stock return probably depends on the level of the magnitude of firm size. Hence, the paper is proposed to investigate the Inverted-U curve hypothesis: at low levels of firm size, the effect is positive whereas at high levels of firm size, the effect is negative.

3.2.1 Traditional Fama and French three-factor model

In asset pricing and portfolio risk management literature, the three-factor model is developed by Fama and French (1993) to depict the relation among risk factor and stock return. The Capital Asset Pricing Model only uses one factor, i.e. beta, to explain the return of the asset. By contrast, Fama and French use three factors. They notice firms generate higher returns when they have smaller sizes and high book-to-market values. As a consequence, they add two factors to the traditional CAPM, i.e. size and book-to-market. Portfolios are formed based on small capitalization minus large capitalization and high book-to-market minus low book-to-market, which are represented by SMB and HML in the model, respectively. The three-factor model explains more diversified portfolio returns that the CAPM.

Firm size is represented by the market capitalisation of the firm, which is the market share price multiplied by the number of shares outstanding. Book-to-market is the ratio of book value to market value. Empirically, the model specification used in the regression analysis is:

$$\boldsymbol{R_{p,t} - R_{f,t} = \alpha_p + \beta_p\left[R_{m,t} - R_{f,t}\right] + s_p SMB_t + h_p HML_t + \varepsilon_{p,t}} \qquad (3.1)$$

Where

$\boldsymbol{R_{p,t}}$ = the weighted stock return on a constructed portfolio *p* in year *t*;

$\boldsymbol{R_{f,t}}$ = the risk-free rate at time *t*; represented by the 1-Month Treasury bill rate;

$\boldsymbol{\alpha_p}$ = intercept, which should not statistically different from 0 if all the unsystematic risks are diversified away;

$\boldsymbol{R_{m,t}}$= return on the market portfolio in year *t*, represented by the FTSE All Share Index, hence ($\boldsymbol{R_{mt} - R_f}$) is the market risk premium;

$\boldsymbol{SMB_t}$ = the proxy for the size factor at year *t*, which is constructed as the difference between the average return on a portfolio containing small capitalisation stocks and the return on a portfolio containing large capitalisation stocks;

$\boldsymbol{HML_t}$ = the proxy for the book-to-market factor at year *t*, which can be constructed as the difference between the average return on a portfolio with high book-to-market stocks and the average return on a portfolio with low book-to-market stocks;

$\boldsymbol{\beta_p}$ = the coefficient loading for the excess return of the market portfolio over the risk free rate;

$\boldsymbol{s_p}$ = the coefficient loading for the excess average return of portfolios with small market capitalizations over the portfolios with big market capitalizations;

$\boldsymbol{h_p}$ = the coefficient loading for the excess average return of portfolios with high book-to-market ratios over portfolios with low book-to-market ratios;

$\boldsymbol{\varepsilon_{p,t}}$ = error term for a constructed portfolio *p* in year *t*, which should be independent of the risk factors included in the model.

Visually, the Fama and French three-factor model is also like an extension of the CAPM. In fact, the model augments the CAPM with the size effect and the book-to-market effect. The size effect is that on average firms with small market capitalisation exhibit returns that exceed the large firms significantly. The book-to-market effect, also known as the value effect presents that returns are larger the higher the book-to-market ratio and vice versa.

3.2.2 Testing of the Inverted-U Curve hypothesis for the Effect of Firm Size on Stock Return

The factors included in the traditional Fama and French three-factor model are market factor, firm size, and book-to-market. The model states that smaller firms tend to outperform the larger ones, and hence the intuition behind the construction of the SMB portfolios. Nonetheless, it may be the hypothesis that when firm size increases to an extent, its further increase may affect the stock return negatively, that is, after one point (I will call this point threshold point thereafter), the larger the firm size, the lower the stock return will be. In order to investigate the hypothesis, the traditional linear three-factor model will be augmented with the nonlinear term, i.e. the squared size factor. The nonlinear term indicates that there may have an inverted-U curve relationship among the stock return and the firm size. Overall, at low levels of firm size, size has positive effects on stock returns, whereas at high levels of firm size, size has negative effects on stock returns. Moreover, the positive effect exceeds the negative effect when the levels of firm size are not large enough, but the negative effect has a superior position at relatively high levels of firm size. The threshold value for the inverted-U curve is regarded as the inflexion point at which the level of firm size ceases to accelerate stock returns. Before the threshold value, the larger the firm size, the higher the stock return, however, after the threshold, the inverse occurs, i.e. the larger the firm size, the lower the stock return. Therefore, the paper is aiming to investigate a non-linear relation, such that the presence of firm size leads to higher stock returns only before the levels of firm size become large enough.

Based on the above interpretation, the ***risk-factor based pricing model*** need to be estimated is specified as:

$$\boldsymbol{R_{p,t} - R_{f,t} = \alpha_p + \beta_p \left(R_{m,t} - R_{f,t}\right) + s_{1,p}SMB_t + s_{2,p}SMB_t^{\,2} + h_pHML_t + \varepsilon_{p,t}} \qquad (3.2)$$

Where

$\boldsymbol{R_{p,t}}$ = the weighted stock return on a constructed portfolio *p* in year *t*;

$\boldsymbol{R_{f,t}}$ = the risk-free rate at time *t*, represented by the 1-Month Treasury bill rate;

$\boldsymbol{\alpha_p}$ = intercept, which should not statistically different from 0 if all the unsystematic risks are diversified away;

$\boldsymbol{R_{m,t}}$= return on the market portfolio in year *t*, represented by the FTSE All Share Index, hence ($\boldsymbol{R_{mt} - R_f}$) is the market risk premium;

$\boldsymbol{SMB_t}$ = the proxy for the size factor at year *t*, which is constructed as the difference between the average return on a portfolio containing small capitalisation stocks and the return on a portfolio containing large capitalisation stocks;

$\boldsymbol{SMB_t^{\,2}}$ = the squared size risk factor, i.e. the quadratic term to capture the nonlinear effect the firm size may exert on the stock return;

$\boldsymbol{HML_t}$ = the proxy for the BTM factor at year *t*, which can be constructed as the difference between the average return on a portfolio with high book-to-market stocks and the average return on a portfolio with low book-to-market stocks;

$\boldsymbol{\beta_p}$ = the coefficient loading for the excess return of the market portfolio over the risk free rate;

$\boldsymbol{s_{1,p}}$ = the coefficient loading for the excess average return of portfolios with small market capitalizations over the portfolios with big market capitalizations;

$\boldsymbol{s_{2,p}}$ = the coefficient loading for squared size factor;

$\boldsymbol{h_p}$ = the coefficient loading for the excess average return of portfolios with high book-to-market ratios over portfolios with low book-to-market ratios;

$\boldsymbol{\varepsilon_{p,t}}$ = error term for a constructed portfolio *p* in year *t*, which should be independent of the risk factors included in the model.

Besides the risk-factor based asset pricing model, the ***characteristics-based asset pricing model*** will be employed to test the nonlinear hypothesis. Similar to the model used by Amel-Zadeh (2011), the empirical specification is expressed as:

$$\boldsymbol{R_{i,t} - R_{f,t} = \alpha_i + \gamma_1 Beta_{i,t} + \gamma_2 BTM_{i,t} + \gamma_3 Size_{i,t} + \gamma_4 Size_{i,t}^2 +}$$
$$\boldsymbol{\gamma_5 Div_{i,t} + \gamma_6 PTE_{i,t} + \gamma_7 BAS_{i,t} + \gamma_8 TVOL_{i,t} + \varepsilon_{i,t}} \qquad (3.3)$$

Where

$\boldsymbol{Beta_{i,t}}$ = the portfolio beta assigned to each firm;

$\boldsymbol{BTM_{i,t}}$ = book to market ratio, the ratio of the book value of the stock to its market value;

$\boldsymbol{Size_{i,t}}$ = market value of the stock;

$\boldsymbol{Size_{i,t}^2}$ = the squared size factor, which is the interest of the model to test the nonlinear relation among firm size and stock return;

$\boldsymbol{Div_{i,t}}$ = dividend yield;

$\boldsymbol{PTE_{i,t}}$ = price-to-earnings ratios;

$\boldsymbol{BAS_{i,t}}$ = bid-ask-spread of the stock;

$\boldsymbol{TVOL_{i,t}}$ = trading volume;

$\boldsymbol{\varepsilon_{i,t}}$ = the error term.

The factors included in the above equation are all accounting variables and they are also the explanatory variables for stock return. The squared size factor is constructed in the model to examine the nonlinear effect firm size may have on stock return, i.e. the inverted-U curve hypothesis. We expect to have a positive and statistically significant coefficient before the size factor, i.e. $\boldsymbol{\gamma_3} > 0$; and a negative and statistically significant coefficient before the squared size factor, i.e. $\boldsymbol{\gamma_4} < 0$.

3.3 Data & Proposed Methodology

3.3.1 Data Description & Measurement

In corporate finance background, the main empirical methodology is consisted by set of quantitative procedures. The main procedures include literature review, identifying existing models, finalizing measurements of variables, setting the hypothesis; and formal collecting the data. Moreover, modeling and evaluating empirical models by analyzing and running the dataset and interpreting the empirical results and comparing with previous studies are also quite important. Finally we reach a conclusion.

My dataset includes the UK-listed firms from the London Stock Exchange. The data used cover the period from 1991 to 2011. Firms in the financial industry are excluded from the sample[21]. I sorted out the firms that satisfy the basic search criteria employed: Public, Active, Non-ADRs only, Country code: GBR, Currency: GBP (British Pound Sterling), London Stock Exchange (XLON) and only the non-financial industries are included. I select ICB Industry code to specify the industry, including: '1000 Basic Materials', '3000 Consumer Goods', '5000 Consumer Services', '4000 Health Care', '2000 Industrials', '9000 Technology', '6000 Telecommunications', '7000 Utilities' and excluding '8000 Financials'. Finally, I got 1004 UK firms from Thomson ONE Banker with each firm has 21 years period.

The annual frequency data is mainly sourced from Thomson ONE Banker and the London Share Price Database (LSPD) maintained at the London Business School. Specifically, from *Thomson Datastream* I collected the following variables for each firm: stock price, market value (i.e. market capitalization or firm size, hence to construct the SMB factor), market-to-book value (used to calculate the book-to-market ratio, hence to construct the HML factor), dividend yield, price-to-earnings ratio (PER), ask price and bid price (to calculate the bid-ask spread) and trading volume. The stock return, book value per share, book value of total assets,

[21] Firms from financial industries are excluded main due to their rather different book keeping method. This is for consistent purpose.

total sales (i.e. net sales or revenues), leverage ratio (measured by the ratio of long-term debt to market equity), earnings before interest and taxes (to calculate the profitability ratio) are from Thomson Worldscope. The One Month Treasury Bill rate is used to measure the risk-free rate and the market portfolio returns are proxied by the FTSE All Share Index returns. The two market variables are from the LSPD. The definition, measurement and source of the variables used in the study are listed in the ***Appendix Table A1***. The data screening criteria used in previous empirical studies on asset pricing with UK firm are detailed in ***Appendix Table A2.***

The data generating process[22] and the formation of portfolios are critical and meaningful to the study. Similar to Fama and French (1993)[23], Gregory et al. (2009) also present the Fama- French and Momentum factors and some other benchmark portfolio data for UK stock market[24]. These portfolios are employed as the benchmark portfolios in the study.

[22] The detailed data generating process will be presented in ***Appendix Table A3***.

[23] Fama and French data source for US stock market: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[24] In the Factor datasets containing the Monthly and Annual SMB, HML and momentum factors for the UK market 1980-2010. In the portfolio datasets containing the Fama-french and momentum portfolios used to create the SMB, HML and UMD factors and other benchmark portfolios. Data source: http://xfi.exeter.ac.uk/researchandpublications/portfoliosandfactors/files.php

3.3.2 Construction of portfolios

The Fama and French model claims that all stock returns can be explained by three factors roughly: (1) exposure to the market factor (*Rm – Rf*); (2) exposure to the small capitalization stocks (*smb*); and finally (3) exposure to the value (high book-to-market ratio) stocks (*hml*). In the US, the estimation of the SMB and HML factors are standardised and can be easily and freely available from Fama and French's website[25]. However, when applying the Fama and French three-factor model in the UK stock market, there are many papers illustrating the distinct ways of estimating SMB and HML and constructing portfolios. In this section, the most representative and influential factor estimation and portfolio construction methods will be discussed, by taking into account the UK stock markets.

3.3.2.1 Definition of Size & Book-to-Market

Firm size is measured as the number of ordinary shares outstanding multiplied by the stock market price. The existence of the size effect and book-to-market effect in the UK stock market is also examined by many researchers (Al-Horani et al., 2003; Dissanaike, 2002; Dimson et al., 2003; Fletcher and Kihanda, 2005; Gregory et al., 2001; Hussain et al., 2002; ect). In order to measure firm size, all the papers mentioned use the market value at the end of June.

---

[25] The website can be found from footnote 23.

Book-to-market (BTM) is a ratio calculated as the book value over the market value. However, the definition of market value in calculating BTM is ambiguous. Most papers use the market value at the end of December to compute BTM. When concerning the definition of book value of the equity, the differences are quite significant. For instance, Dimson et al. (2003) use the ordinary share capital plus reserves plus deferred and future tax. Three of them measure the book value as equity capital plus reserves minus total intangible assets (Al-Horani et al., 2003, Gregory et al., 2001; Hussain et al., 2002).

3.3.2.2 Portfolio Formation Methods

The most prevail and most popular sorting methods are proposed by Fama and French (1993 and 1996). After assuming the end of June for each year, as a portfolio formation date, and then stocks involved are sorted into two groups, i.e. small (S) and big (B), based on the firm size, i.e. market value. Moreover, stocks are also allocated into three groups according to BTM ratio, i.e. low BTM (L), medium BTM (M) or high BTM (H). Firms with negative BTM are excluded from the formation procedures.

Then six size-BTM portfolios in Table 3.1 (S/L, S/M, S/H; and B/L, B/M, B/H) are created from the combination of two size groups and three BTM groups. Therefore, the size factor (SMB) in the Fama and French three-factor is defined as the difference between the average of returns on the three small portfolios (i.e. S/L, S/M, S/H) and the average of returns on the three big portfolios (i.e. B/L, B/M, B/H). The value factor (HML) is calculated as the difference between the average of the returns on the two high BTM portfolios (i.e. S/H, B/H) and the average of the returns on the two low BTM portfolios (i.e. S/L, B/L). Graphically,

Table 3.1 Six Size & Book-to-Market portfolios (2*3)[26]

	Small SIZE (S)	*Big SIZE(B)*
Low BTM (L)	S/L	B/L
Medium BTM (M)	S/M	B/M
High BTM (H)	S/H	B/H

The above method is also utilized by Al-Horani et al (2003), Dimson et al (2003), Dissanaike (2002), Gregory et al. (2001), and Hussain et al. (2002).

[26] Size is measure as the number of shares outstanding multiplied by the stock market price. Book-to-Market (BTM) is a ratio of the book value over the market value. S/L refers to a portfolio containing small size and low BTM stocks.

Fama and French (2012) investigate international stock returns in North American, Japan, Asia Pacific and Europe, with two main aims: (1) give full illustration about the size, value and momentum effects in the average stock returns for the above developed markets, and the sample covers all size groups of the firms involved; hence presents strong evidence for size groups; (2) investigate how well the three-factor model and four-factor model explain average returns for portfolios formed on size and value, or size and momentum.

One different portfolio formation method is used by Fletcher (2001). Specifically, stocks are sorted into two size groups each year. Then within each size group, stocks are sorted into three BTM groups. The SMB and HML factors are defined the same as mentioned above. The difference among the Fama and French method and the one used by Fletcher (2001) is the number of stocks in each size-BTM portfolio.

3.3.3 My approach

I will follow the procedure described by Fama and French (1993) to form portfolios based on size and book-to-market ratio. Firm size is measured by market capitalization or market value of ordinary shares the firm has. It is defined as the product of stock price and the number of ordinary shares outstanding. The book-to-market ratio is computed as the ratio of a firm’s book value at the fiscal year end in calendar year *t - 1* and its market value at the end of year *t - 1*. Only the BTM with positive values will be involved. Based on firm size, stocks will be assigned as small size or big size, which are two portfolios Small (S) and Big (B). Then according to the value of the Book-to-Market ratios, the same stocks will be sorted into three portfolios, which are Low BTM (L), Medium BTM (M), and High BTM (H). At the intersection of the size factor and BTM factor, six portfolios will be formed (i.e. S/L, S/M, S/H, B/ L, BM, and B/H) to construct the SMB and HML factors.

Basically, SMB (small minus big) is constructed as the difference between the average return of the three portfolios containing the smallest-capitalization stocks and the return of the three portfolios containing the largest-capitalization stocks. Similarly, HML (high minus low) is the simple average of the returns on the high-BM portfolios minus the returns on the low-BM portfolios.

3.3.4 Empirical Estimation & Testing procedure

The structure of the dataset is panel. Totally 1003 firms are concerned and each firm have 21 years period data. In the regression analysis, the empirical procedures include the following main steps:

Step 1 - Descriptive statistics and univariate testing.

This section will present the summary statistics for the panel data of non-financial UK-listed firms. The table should summarise the number of firms, annual stock return, market risk premium, book-to-market ratio, firm size, dividend yield, price-to-earnings ratio, bid-ask-spread, trading volume, total assets, total sales, leverage ratio and profitability ratio. The summary statistics will consist of the number of observations, mean, standard deviation, skewness, and kurtosis of the sample firms.

Step 2 – Correlation matrix.

This section will present the correlations between the explanatory variables. The variables involved include $\mathrm{R_{p,t} - R_{f,t}}$, $\mathrm{(R_m - R_f)}$, SMB, HML, BTMV, Size, DIVY, PTE, BAS, TVOL, $\mathrm{TASSETS}$, TSAL, $\mathrm{LEV}_{\mathrm{i,t}}$, $\mathrm{PRO}_{\mathrm{i,t}}$.

Step 3 – Regression analysis: the regression method includes

Instrumental variable (IV)[27] method is advocated in the thesis, which involves the selection of a variable or a set of variables which are assumed to be exogenous. Then

---

[27] Instrumental variable method allows consistent estimation when the explanatory variables are correlated with the error terms in a regression. Such correlation may occur when the dependent variable causes at least one of the independent variables, when there are relevant explanatory variables which are omitted from the model. In this situation, ordinary linear regression will produce biased and inconsistent estimates. An instrumental variable must be correlated with the endogenous variables, but cannot be correlated with the error term in the model.

the two-stage-least-squares (2SLS) will be used to estimate the coefficients in the regression model. This method is employed to consider the endogeneity problem and omitted variable bias, which may arise in regression model. Amel-Zadeh (2011) proposes that the explanatory power of firm size could be due to several other factors omitted in the CAPM. This unobserved heterogeneity leads to biased and inconsistent estimators when an ordinary least square regression estimation is used to find the linear relationship between the risk and expected return, since the error terms in the regression will be corrected with one or more explanatory variables. Furthermore, Berk (1995) shows that size is usually measured by the market value of the equity, it is not only endogenous, but simultaneously related to expected returns. Therefore, this can induce the so called simultaneity bias. In light of the above discussions, the thesis uses a set of instrumental variables in the two-stage-least-squares estimation, in order to control for the endogenous variable size. In the empirical regression analysis, the panel data of UK firms from 1991 to 2011 will be used on an annual basis. In the first stage of the 2SLS, we regress the size factor on its set of instrumental variables. Then applying the fitted values of size from the first-stage of the 2SLS when running the second-stage regression.

Berk (2000) emphasises the misspecification of size measure by the market value of equity in explaining stock returns since the simultaneous relation with expected returns. He proposes non-market measure of size in asset pricing tests. Hence, the book value of total asset (BA) and the total sales (TS) are used as instrumental variables. Additionally, Fama and French (1996) and Vassalou and Xing (2004) suggest that smaller firms suffer more from financial distress and hence are expected to have higher stock returns. Leverage (LEV), measured as the ratio of long-term debt to market equity and the profitability ratio (PRO), measured as the ratio of earnings before interest and taxes to total sales, are also included as instrumental variables. Finally, since smaller stocks are normally also illiquid stocks and therefore traded less frequently with regard to higher holding period risk (Hamon and Jacquillat, 1999), bid-ask-spread (BAS) and trading volume (TV) are used as instrumental variables for firm size as well. The instrumental variables are the variables that correlated with size but unrelated with the error term in the models (or equations).

The first-step regression should be run as the following equation:

$$\boldsymbol{SIZE_{i,t} = \varphi_0 + \varphi_1 BA_{\,i,t} + \varphi_2 TS_{\,i,t} + \varphi_3 LEV_{\,i,t} + \varphi_4 PRO_{\,i,t} + \varphi_5 BAS_{\,i,t} + \varphi_6 VOL_{\,i,t} + \vartheta_{i,t}} \qquad (3.4)$$

Where

$\boldsymbol{BA_{\,i,t}}$= the book value of the total assets;

$\boldsymbol{TS_{\,i,t}}$= total sales;

$\boldsymbol{LEV_{\,i,t}}$= leverage ratio;

$\boldsymbol{PRO}_{i,t}$= profitability ratio;

$\boldsymbol{BAS}_{i,t}$= bid ask spread;

$\boldsymbol{VOL}_{i,t}$= trading volume.

Then use the fitted value of SIZE from equation (3.4) to run the second-step equation.

$$\boldsymbol{R_{i,t} - R_{f,t} = \gamma_0 + \gamma_1 Beta_{i,t} + \gamma_2 BTM_{i,t} + \gamma_3 Size^*_{i,t} + \gamma_4 Size^{*}{}_{i,t}^{2} + \gamma_5 Div_{i,t} + \gamma_6 PTE_{i,t} + \varepsilon_{i,t}} \qquad (3.5)$$

Where

$\boldsymbol{BTM}_{i,t}$= book to market ratio;

$\boldsymbol{Size^*}_{i,t}$= fitted value of SIZE from first-step regression in equation (3.5);

$\boldsymbol{Size^*}_{i,t}^{2}$= squared of fitted value of SIZE from first-step regression in equation (3.5);

$\boldsymbol{Div}_{i,t}$= dividend yield;

$\boldsymbol{PTE}_{i,t}$= price to earnings ratio.

Step 4 – Robustness check.

In order to do the robustness check, the benchmark portfolios from Gregory et al. (2009) will be employed to test the significance of the regression result.

3.4 Empirical Results & Analysis

Table 3.2 Descriptive statistics for related variables[28] used in this study

This table shows summary statistics of the sample. It presents the number of firms, years, data frequency, asset risk premium, market risk premium, book-to-market, market values, dividend yield, price-to-earnings, bid ask spread, trading volume, book value of total assets, total sales, leverage ratio and the profitability ratio. Whereas, book value of total assets, totals sales, leverage ratio, the profitability ratio, bid-ask-spread and trading volume are instrumental variables for SIZE factor in the Two-stage Least Square regression. The summary consists of mean, standard deviation, skewness, and kurtosis values of the sample. All the values shown in the table are presented to three digital points after the decimal point.

Firms	1004			
Years	21			
Frequency	annual			
Variables	**Mean**	**Std. Dev.**	**Skewness**	**Kurtosis**
R*p*,*t*-R*f*,*t*	2733.244	3319.934	0.837	2.239
MRP	0.049	0.152	-0.947	2.844
BTMV	3405.249	3946.762	0.738	2.062
SIZE	2490.142	2874.931	0.737	2.074
DIVY	158.488	272.980	1.563	4.212

[28] Asset risk premium is the return on the asset minus the risk free rate. MRP, i.e. market risk premium is the return on the market minus the risk free rate. BTMV is the book-to-market ratio. SIZE is firm size. DIVY is dividend yield, which is a financial ratio shows how much a firm pays out in dividends relative to its share price each year. PERATIO is pricing-to-earnings ratio, i.e. it's a valuation ratio of a firm’s current share price compared to earnings per share. BAS is bid ask spread, defined as bid price minus ask price. Whereas bid price is the highest price a prospective buyer is willing to pay for a stock at a particular price and ask price is the lowest value a prospective seller would be willing to accept for a particular stock. TVOL is trading volume, which refers to the number of share s traded in the stock market for a given period of time. TASSETS (BA) is the book value of total assets. TSAL (TS) total sales, i.e. the net sales or revenues. LEV is leverage ratio, which is measured by the long term debt to market equity ratio. PRFT (PRO), i.e. the profitability ratio is measure by the ratio of earnings before interest and taxes (EBIT) over the total sales.

PERATIO	174.929	276.023	1.610	4.342
BAS	748.136	856.109	0.720	2.056
TVOL	1691.940	2359.594	1.189	3.050
TASSETS (BA)	3833.576	4169.824	0.630	1.924
TSAL (TS)	3369.780	3889.354	0.730	2.050
LEV	2079.133	2760.183	1.157	2.900
PRFT (PRO)	3433.239	3966.741	0.733	2.056

Table 3.2 show the descriptive statistics for the main variables used in the analysis and regression. The dataset includes 1003 nonfinancial UK-listed firms from the London Stock Exchange. The data used cover the period from 1991 to 2011, i.e. 21-year annual frequency. Firms in the financial industry, such as banks, equity investment, life insurance, etc are excluded from the sample.

Table 3.3 Correlation coefficient matrix of variables[29] used in chapter 3

	Rpt-Rft	MRP	SMB	HML	BTMV	SIZE	DIVY	PTE	BAS	TVOL	TASSETS	TSAL	LEV	PRFT
Rpt-Rft	1.0000													
MRP	0.0279	1.0000												
SMB	-0.0778	0.2266	1.0000											
HML	-0.0880	0.0201	0.2312	1.0000										
BTMV	0.4836**	-0.0116	-0.0297	-0.0528	1.0000									
SIZE	0.5102**	-0.0218	-0.0571	-0.0816	0.5030	1.0000								
DIVY	0.0549	-0.0348	-0.0554	0.0190	0.0855	0.0605	1.0000							
PTE	0.0868	-0.0164	-0.0201	-0.0448	0.1292	0.0941	0.4230	1.0000						
BAS	0.5097	-0.0453	0.0068	-0.0277	0.5656	0.5438**	0.0462	0.0798	1.0000					
TVOL	0.4419**	0.0004	-0.0681	-0.0952	0.3616	0.4530**	0.0593	0.0929	0.2061	1.0000				
TASSETS	0.5115***	-0.0127	-0.0434	-0.0623	0.5713	0.5297*	0.0597	0.1074	0.5121**	0.4037**	1.0000			
TSAL	0.4995	-0.0035	-0.0318	-0.0453	0.5205	0.4875*	0.0466	0.0878	0.4936**	0.3709**	0.6358**	1.0000		

[29] Variables mentioned in Table 3.3 are defined in footnote 28.

LEV	0.4383	-0.0082	-0.0252	-0.0405	0.4992	0.4344**	0.0400	0.0719	0.4097	-0.3989*	0.4915 **	-0.5398*	1.0000	
PRFT	0.5103	-0.0127	-0.0319	-0.0595	0.4092	0.4908	0.0406	0.0800	0.5185**	-0.4078*	-0.5700**	-0.6212*	0.5250**	1.0000

Notes:

a. The sample contains UK company-year data observed for the period 1991–2011.

b. For each cell, the reported figure is the pair-wise Pearson correlation coefficient between the corresponding variables.

c. * significant at the 10% level; ** significant at the 5% level; and *** significant at the 1% level.

d. All the values shown in Table 3.3 are presented to four digital points after the decimal point.

The pair-wise correlation among the main variables is presented in Table 3.3. It shows that instrumental variables for SIZE are significantly correlated with each other. Specifically, TASSETS is significantly and positively related to BAS and TVOL. TSAL is positively and significantly related to BAS, TVOL and TASSETS. LEV is negatively related to TVOL and TSAL, but positively and significantly related to TASSETS. PRFT is negatively related TVOL, TSAL and TASSETS; and positively and significantly related to BAS and LEV.

3.4.1. Test for the Inverted-U Curve hypothesis

In order to test the inverted-U curve hypothesis, the risk-factor based pricing model and the characteristics-based asset pricing model are employed. We expect to have a positive and statistically significant coefficient before the size factor and a negative and statistically significant coefficient before the squared size factor. Table 3.4 shows the fixed and random regression results from the risk-factor based pricing model.

Table 3.4 Fixed & Random effects regression results for Fama and French three-factor model

This table presents the fixed and random effects regression to test the inverted-U curve hypothesis for 1004 UK-listed firms from London Stock Exchange, which cover the period from 1991 to 2011. The data is in panel form and with annual frequency. mrp is the market risk premium. SMB is size factor in the three-factor model, which is constructed as the difference between the return on a portfolio containing small capitalisation stocks and the return on a portfolio containing large capitalisation stocks. HML is the BTM factor, which defined as the difference between the return on a portfolio with high book-to-market ratio and the return on a portfolio with low book-to-market ratio. SMB2 is the squared size factor. R-squared and adjusted R-squared are employed to test the fitness of the model to the data. Durbin-Watson test is used to detect the presence the autocorrelation from the regression analysis.

Panel A: Fama and French three-factor model: $R_{p,t} - R_{f,t} = \alpha_p + \beta_p[R_{m,t} - R_{f,t}] + s_p SMB_t + h_p HML_t + \varepsilon_{p,t}$		
VARIABLES[30]	**Fixed Effects** $R_{p,t} - R_{f,t}$	**Random Effects** $R_{p,t} - R_{f,t}$
mrp t-statistics	0.9925** (3.1311)	0.9925** (3.1311)
SMB t-statistics	0.3911*** (13.3395)	0.3911*** (13.3395)
HML t-statistics	0.0797*** (13.3307)	0.0797*** (13.3307)
Constant t-statistics	0.0025 (4.9110)	0.0025 (5.2720)
Diagnostic for the traditional FF3 model		
R-squared	Adjusted R-squared	Durbin-Watson
0.492	0.5230	1.9450

Panel B: Fixed & Random effects regression results for three-factor model with non-linear factor: $R_{p,t} - R_{f,t} = \alpha_p + \beta_p\,(R_{m,t} - R_{f,t}) + s_{1,p} SMB_t + s_{2,p} SMB_t^{\,2} + h_p HML_t + \varepsilon_{p,t}$	
VARIABLES	**Random Effects** $R_{p,t} - R_{f,t}$
mrp t-statistics	1.0920** (132.4000)

SMB t-statistics	0.3263*** (3.4310)
SMB2 t-statistics	-0.2483*** (0.0919)
HML t-statistics	0.0893*** (3.3370)
Constant t-statistics	0.0016 (62.4900)

Notes:

a. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

b. All the values shown in Table 3.4 are presented to four digital points after the decimal point.

Table 3.4 presents the fixed and random effects regression to test the inverted-U curve hypothesis. Panel A shows the fixed and random effects regression for the traditional Fama and French three-factor model. The three risk factors, i.e. mrp, SMB and HML are all significant on 5% and 1% significance level, with positive sign. The magnitude of constant is not significant on any significance level. The R-squared is 0.492, which means the regression model fits the date well. The above interpretation about Panel A indicates that generally the data fit the model quite well. The data also shows that small firms outperform large firms, and value stocks with high book-to-market ratio outperform low book-to-market ratio. Then the panel data is employed for subsequent regression analysis.

Moreover, the value of Durbin-Watson is 1.945, which indicates that there is no evidence for the autocorrelation. The random effects assumption is that the individual specific effects are uncorrelated with the independent variables. The fixed effect assumption is that the individual specific effect is correlated with the independent variables. If the random effects assumption holds, the random effects model is more efficient than the fixed effects model. However, if this assumption does not hold (i.e., if the Durbin–Watson test fails), the random effects model is not consistent. Hence, the value of Durbin-Watson indicates that random effects model is more efficient and consistent in the empirical testing.

Panel B shows the random effects regression results (simplified as 'regression results', thereafter) for three-factor model with non-linear factor. The explanatory variables in the regressions are risk factors. The non-linear factor is the quadratic SMB risk factor. The variable is used to test the inverted-U curve hypothesis. The respective t-statistics are given in parentheses. Similar to the regression results in Panel A, mrp, SMB, and HML are all significant on different significance level, all with positive sign as expected. Specifically, for SMB, the coefficient is 0.3263; and the value is significant on 1% significance level. SMB2, the variable the model aims to investigate has a negative sign, with the value -0.2483 in the regression results. However, the coefficient before SMB2 is not significant on any significance level. The sign of SMB and SMB2 show that the inverted-U curve hypothesis among firm size and stock return may be the case under the framework of Fama and French three-factor model, by using the UK panel data. Moreover, the coefficient before SMB2 is significant on the 1% significance level by using the random effects regression.

Table 3.4 shows there is a positive and statistically significant coefficient before the size factor and a negative and statistically significant coefficient before the squared size factor. Therefore, the inverted-U curve hypothesis among firm size and stock return by using UK panel data is supported here. Such that the presence of firm size leads to lower stock returns after the levels of firm size become large enough.

Table 3.5 Regression results for characteristics-based asset pricing model

This table presents the regression results to test the inverted-U curve hypothesis for UK-listed firms from London Stock Exchange, which cover the period from 1991 to 2011. The data is in panel form and with annual frequency. Beta is the portfolio beta assigned to each stock. BTM is the book value of equity to market value of equity. SIZE is the market value of the stock. SIZE2 is the squared SIZE factor. DIVY is the dividend yield, PTE is the price to earnings ratio, BAS is the bid ask spread, and TVOL is the trading volume. BAS and TVOL are used to proxy for the liquidity factor.

Characteristics-based asset pricing model with non-linear factor: $\boldsymbol{R_{i,t} - R_{f,t} = \alpha_i + \gamma_1 Beta_{i,t} + \gamma_2 BTM_{i,t} + \gamma_3 Size_{i,t} + \gamma_4 Size_{i,t}^{2} + \gamma_5 Div_{i,t} + \gamma_6 PTE_{i,t} + \gamma_7 BAS_{i,t} + \gamma_8 TVOL_{i,t} + \varepsilon_{i,t}}$	
VARIABLES	**Random Effects** $\boldsymbol{R_{i,t} - R_{f,t}}$
Beta	0.0012

t-statistics	(0.1280)
BTM t-statistics	0.0104*** (1.6080)
SIZE t-statistics	1.9800*** (3.0102)
SIZE2 t-statistics	-0.1110 (0.0100)
DIVY t-statistics	0.0000 (-0.0076)
PTE t-statistics	0.0001 (-0.0714)
BAS t-statistics	0.0110*** (0.3850)
TVOL t-statistics	0.3530*** (0.9350)
Constant statistics	0.0065 (0.3229)

Notes:

a. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

b. All the values shown in Table 3.5 are presented to four digital points after the decimal point.

Table 3.5 employs the characteristic-based asset pricing model to test the non-linear relation among size and stock return. Among the multiple variables, SIZE2 is the squared SIZE factor, which is the main testing aim of the model and its magnitude and sign are the concentrations of the regression. BAS and TVOL are utilised to proxy for liquidity factor, since stock returns are also determined by liquidity factor. Table 3.5 summarises the regression results. The regression results show that the coefficient on SIZE remains positive and statistically significant (1%-level). As expected, the coefficients on SIZE2 are negative but are statistically insignificant on any significance level.

From Table 3.5, the coefficient before SIZE is 0.9800 and significant on 1% significance level and the coefficient before SIZE2 is -0.1110 but insignificant on any significance level. The sign and significance of SIZE and SIZE2 confirm that there is possible a non-linear relation among firm size and stock return when employing the characteristics-based asset pricing model, however, this nonlinear relation is not significant.

The results so far confirm an insignificant inverted-U curve relationship among the stock return and the firm size, within the Fama and French three-factor model. Overall, low levels of firm size have positive effects on stock returns, whereas high levels of firm size have negative effects on stock returns. Moreover, the positive effect exceeds the negative effect when the levels of firm size are not large enough, but the negative effect has a superior position at relatively high levels of firm size. The threshold value for the inverted-U curve is regarded as the inflexion point at which the level of firm size ceases to accelerate stock returns. Before the threshold value, the larger the firm size, the higher the stock return, however, after the threshold, the inverse occurs, i.e.

the larger the firm size, the lower the stock return. Therefore, the regression testing aiming to investigate a non-linear (inverted-U curve) relationship, such that the presence of firm size leads to higher stock returns only before the levels of firm size become large enough.

3.4.2. Robustness Tests

The findings in previous section show that firm size, as an explanatory factor for stock return, has an inverted-U curve relation with stock return. Fama and French (1995) argue that firm size captures the firm characteristics in stock returns. Furthermore, the explanatory power of firm size could also due to other factors which omitted in the traditional CAPM. Hence, when a simply regression estimation is used to find the relation among firm size and stock return, it possibly leads to biased and inconsistent estimators. Fama and French (1996) also indicate that size is measured by the market value of the equity, so it is endogenous. Moreover, size factor could be simultaneously related to stock returns. Therefore, in order to examine the robustness of previous regression findings, I use the instrumental variables for size factor and employ Two-stage Least Square estimation method.

In the first-step of the Two-stage Least Square estimation, SIZE is regressed on its instrumental variables. According to the literature (Fama and French, 1996; Vassalou and Xing, 2004, etc), book value of the total assets, total sales, leverage ratio, profitability ratio, bid ask spread, trading volume are utilised as instrumental variables. Instrumental variables estimation could give consistent estimators. Then the fitted value of SIZE from the first-step regression, i.e. SIZE* will be used to run the second-step regression. It is expected to have a positive and statistically significant coefficient before SIZE* and a negative and statistically significant coefficient SIZE*2.

Table 3.6 Two-stage Least Square regression results for robustness check

The table presents the first-step and second-step regression results by using Two-stage Least Square method. Panel A shows the regression results of the first-step regression of SIZE on the set of instrumental variables. Standard errors are calculated according to the heteroskedasticity-consistent method suggest by White (1980). Panel B shows the regression results of the second-step regression by using the fitted values of SIZE from the first-step regression. SIZE is the natural logarithm of the market value of equity, BA is the book value of total assets, TS is the total sales, LEV is the leverage ratio, measured as the ratio of total debt to common equity, PRO is the profitability ratio, which measured as the earnings before interest and taxes (EBIT) to total sales. BAS is the bid-ask-spread and VOL is the shares trading volume. All the values shown in Table 3.6 are presented to four digital points after the decimal point.

Panel A: The first-step regression:

$$\boldsymbol{SIZE_{i,t} = \varphi_0 + \varphi_1 BA_{\,i,t} + \varphi_2 TS_{\,i,t} + \varphi_3 LEV_{\,i,t} + \varphi_4 PRO_{\,i,t} + \varphi_5 BAS_{\,i,t} + \varphi_6 VOL_{\,i,t} + \vartheta_{i,t}}$$

VARIABLES	L.SIZE	BA	TS	LEV	PRO	BAS	VOL
COEFFICIENT	0.2390***	0.7393***	0.1772***	-0.0196****	0.0699***	0.1869***	0.0013
T-STATISTICS	(0.0082)	(80.8000)	(19.3700)	(-10.0711)	(4.9250)	(14.0200)	(13.5100)

Panel B: The second-step regression: use the fitted value of SIZE from above equation to run the second-step equation

$$\boldsymbol{R_{i,t} - R_{f,t} = \gamma_0 + \gamma_1 Beta_{\,i,t} + \gamma_2 BTM_{\,i,t} + \gamma_3 Size^{*}_{\,i,t} + \gamma_4 {Size^{*}_{\,i,t}}^{2} + \gamma_5 Div_{\,i,t} + \gamma_6 PTE_{\,i,t} + \varepsilon_{i,t}}$$

VARIABLES	Constant	Beta	BTM	SIZE*	SIZE*2	DIV	PTE
COEFFICIENT	0.0023	-0.0031	-0.0082***	0.0941***	-0.0472	-0.0000	-0.0000

T-STATISTICS	(0.5275)	(-0.6500)	(-0.5080)	(0.0141)	(0.0129)	(-0.4380)	(-0.1030)

Table 3.6 presents the results of Two-stage Least Square regressions. Panel A shows the regression results of the first-step regression of SIZE on the instrumental variables. Panel B shows the second-step regression coefficients. The respective t-statistics are given in parentheses. The explanatory variables in the regressions are firm characteristics. Beta is stock beta, represent the market risk factor. SIZE* is the fitted value of SIZE from the first-step regression, which aiming to run the second-step equation. SIZE*2 is the squared SIZE*. BTM is the book to market ratio. DIV represents the dividend yield calculated as the ratio of dividends per share to price per share and PTE is the price to earnings ratio calculated as the market value of equity divided by earnings before interest and taxes. * denotes statistical significance at the 10% level using a two-tailed test, **denotes statistical significance at the 5% level using a two-tailed test and *** denotes statistical significance at the 1% level using a two-tailed test.

Panel A shows all explanatory variables are statistically significant at the 1% significance level in the first-step regression. The high correlation among firm size and the instrumental variables also indicate that the instrumental variables are an appropriate choice. After the first-step regression, the second step of the Two-stage Least Square regression is run containing the fitted value of size as a regressor, which shown in Panel B. Moreover, the squared fitted value of size is also employed in the model to explore the relationship among firm size and stock return. From Panel B, the coefficient before SIZE* is positive and statistically significant on 1% level and SIZE*2 have a negative coefficient but not statistically significant on any significance level. Similarly, the robustness results also confirm an insignificant inverted-U curve hypothesis: the effect of firm size on stock return is approximated by an inverted-U curve shape relation, rather than linear: at low levels of firm size the effect is positive and at high levels of firm size the effect is negative. But its significance needs to be further examined.

3.5 Concluding remarks

In this chapter, this paper examines the size effect in the UK stock market, which is an intensively hot debated topic investigated by researchers. Many researches in asset pricing model test the validity and persistence of size effect in the cross-sectional stock returns and how firm size or the portfolios formed according to firm size affect stock return (Amel-Zadeh, 2011; van Dijk, 2011; De Moor and Serch, 2012; etc).

Overall, our results suggest that, firstly, the empirical paper confirms the existence of size effect in the UK stock market. By augmenting the Fama and French three-factor model with the squared size factor, i.e. the non-linear risk-factor based pricing model. Moreover, the non-linear characteristics-based asset pricing model is also employed to test the inverted-U curve hypothesis. Generally, we have a positive and statistically significant coefficient before SMB and size factor and a negative but statistically insignificant coefficient before squared SMB and quadratic size factor. However, other characteristics, such as dividend yield, price to earnings ratio and bid ask spread play an insignificant role in explaining the stock return in our model testing. The above interpretation could provide evidence for the non-linear effect of firm size on stock return.

Secondly, the economic rationale of firm size is investigated. In particular, firm size seems capture the variations in book value of total assets, total sales, leverage ratio, profitability ratio, bid ask spread and trading volume. The finding is consistent with Fu (2009) and Fan and Liu (2005) and their paper also propose that firm size captures the variations in firm characteristics components in stock returns. Finally, as Berk (1995) show that size is measured by the market value of the equity, it may simultaneously relate to stock returns. In order to address this simultaneity bias, the paper uses the instrumental variable regression. The regression result generally also presents an insignificant inverted-U curve relationship among the stock return and the firm size.

Overall, low levels of firm size have positive effects on stock returns, whereas high levels of firm size have negative effects on stock returns. Moreover, the positive effect exceeds the negative effect when the levels of firm size are not large enough, but the negative effect has a superior position at relatively high levels of firm size. The threshold value for the inverted-U curve is regarded as the inflexion point at which the level of firm size ceases to accelerate stock returns. Nevertheless, the exact value of the threshold point is out of the investigation scope of the paper. However, the results reported here have important implications for many parties, which will be discussed in the next chapter.

CHAPTER 4 CONCLUSION

4.1 Summary

The thesis empirically investigates how the level of firm size affects the stock return within the framework of linear and nonlinear Fama and French three-factor model, with reference to a large panel of UK-listed firms observed from period 1991 to 2011. Many asset pricing issues have been proposed in the finance literature, such as firm size (Banz, 1981), book-to-market (Fama and French, 1992), momentum (Jegadeesh and Titman, 1993), liquidity (Pastor and Stambaugh, 2003), and asset growth issue (Yao et al., 2011). In addition to the CAPM and the Fama and French three-factor model, many alternative asset pricing models have been developed. Ross (1976) proposed the arbitrage pricing theory (APT). The APT suggests a number of empirical factors and but does not specify the factors. Fama and French (1995) also suggest a five factor model, which includes the market factor, size factor, value factor, default and term spread. Chen et al. (2010) propose a three-factor model, containing the market factor, a return on asset factor and an investment factor. Kim (2006) put forward a two-factor model including the market factor, and the earnings information risk factor; and the paper find that the two-factor model has explanatory power in size effect.

As discussed above, many empirical asset pricing models are tested and investigated in the finance literature. Among them, to my knowledge, most of the models are tested based on the specification that all the factors are linearly related to stock returns in the model. The most classic model is the Fama and French three-factor model, which is also the aim of the thesis.

The purpose of the thesis is to comprehensively evaluate and test the linear and non-linear Fama and French three-factor models in the UK stock market. Basically, we consider the traditional linear Fama and French three-factor asset pricing models. Then in order to test the non-linearity of the size factor, the quadratic size factor is added to the risk-based factor model and the characteristic-based factor model. Therefore, the paper aims to test the inverted-U curve hypothesis among firm size and stock return, with the dataset of UK-listed firms from the London Stock Exchange.

This chapter concludes the thesis by briefly recapping the research questions, concluding the key findings, highlighting the main contributions, discussing the limitations, as well as putting forward some promising ideas for further research.

4.2 Main contributions to the existing literature

The thesis empirically investigates how the level of firm size affects the stock return within the framework of linear and nonlinear Fama and French three-factor model, with reference to a large panel of UK-listed firms observed from period 1991 to 2011, which aims to improve our current knowledge of nonlinearity about asset pricing models. To my knowledge, it also represents one of the first studies to explicitly examine the nonlinear influence of firm size on stock return within the different specifications of Fama and French three-factor model. The nonlinear three-factor model explicitly allows the squared size factor to model nonlinearity. This study reveals new insights into the role played by firm size in the model. Additionally, this thesis offers a measure of nonlinear factor that are currently debated in asset pricing literature and the relation among the nonlinear factor and other economic factors, such as total assets, total sales, leverage ratio, profitability ratio, bid ask spread, and trading volume, etc.

4.3 Key empirical conclusions & implications

4.3.1 Empirical findings and conclusions

From the empirical regression, the finding confirms the existence of size effect in the UK stock market. By augmenting the Fama and French three-factor model with the squared size factor, i.e. the non-linear risk-factor based pricing model and the non-linear characteristics-based asset pricing model, the inverted-U curve hypothesis is investigated specifically. We have a positive and statistically significant coefficient before SMB and size factor and a negative but statistically insignificant coefficient before squared SMB and quadratic size factor. On the other hand, other factors such as dividend yield, price to earnings ratio and bid ask spread play an insignificant role in explaining variations in the stock returns. The above interpretation could provide evidence for the non-linear effect of firm size on stock return. When firms size increase to an extent, its further increase may negatively affect the stock return. After the threshold point, the larger the firm size, the lower the stock return will be.

Furthermore, the economic characteristic of firm size is also tested. Particularly, firm size seems capture the variations in book value of total assets, total sales, leverage ratio, profitability ratio; bid ask spread and trading volume. The finding is consistent with Fu (2009) and Fan and Liu (2005) and their paper also propose that firm size captures the variations in firm characteristics components in stock returns.

Berk (1995) proposes that size is measured by the market value of the equity, it may simultaneously relate to stock returns. Hence, the instrumental variables regression is utilised in the robustness test to address the bias and the regression result also confirms the inverted-U curve relationship among the stock return and the firm size. The threshold value for the inverted-U curve is considered as the inflexion point at which the level of firm size ceases to increase stock returns. However, the exact value of the threshold point is out of the scope of the paper.

4.3.2 Practical implications of the findings

The findings and conclusions presented in this thesis not only contribute to the existing academic literature, but also have broader practical implications, and here the implications for corporate managers will be discussed.

4.3.2.1 Implications for academic literature

In this section, I evaluate the current state of empirical and theoretical research on the size effect. The empirical evidence suggests that a significant size premium exists in both US and non-US stock markets. However, continuous asset pricing tests as well as more robustness checks are required to confirm the existence of a size effect in international stock returns. Here three promising strands of theoretical literature are presented.

First, firm-level investment decision models illustrate an endogenous relation among firm size and stock return. Berk et al. (1999) provide evidence that these models can produce many widely known stock return features. This strand of research is still immature, and it is not sufficiently clear to what extent these models can explain return anomalies uncovered by empirical research on the size effect.

Second, in addition to transaction costs, asset pricing models predict that stock returns also depend on liquidity risk. The empirical evidence indicates that liquidity is an important factor in asset pricing. However, most studies do not explicitly examine whether the size effect or the inverted-U curve hypothesis investigated in the thesis can be explained by liquidity factors. Therefore, the size effect and liquidity interact is an important area for future research.

Third, the size effect can be treated as originate from the behaviour of less rational investors. In Daniel et al. (2001)'s model, investors' overconfidence creates an endogenous relation between size and future returns. Therefore, investors' overreaction or underreaction may be another underlying cause for size effect anomaly, which also worth further research.

4.3.2.2 Implications for corporate managers

In addition to above implications for academic asset pricing literature, we also consider the practical implications for corporate managers.

The empirical findings in this study have distinct implications for companies, especially UK companies and their management. First, in making key corporate investment decisions, managers must be aware of the relation among risk factor and the corresponding return, in order to avoid undesirable loss, which may be caused by an irrational decision. Thus, managers should consider the level of firm size involved in decision making, and meanwhile, consideration must be paid to the influence of other factors.

Second, the empirical evidence on the relationship between firm size and stock return uncovered in this study also has profound implication for corporate decision makers. Our results indicate that, other thing equal, smaller firms will have higher stock returns than bigger firms; however, the level of firm size does matter in explaining this relationship. Managers therefore need to consider strategic interactions among firm size and the corresponding return.

4.4 Limitation of the Study

In general, the paper provides evidences support the inverted-U curve hypothesis among firm size and stock return. On the other hand, it is worth noting that there are several limitations of the research reported in the thesis.

First, the dataset used in the study only includes the UK stock market, and the findings and conclusions of the thesis are largely drawn from the collected dataset. Hence, the first main limitation of this thesis is that the validity of the conclusion may, to some extent, be sensitive to the selection of the sample, the measurement of variables, the specification of the model, the nonlinear representation of the size factor, the choice of the estimation methodology, and the interpretation and understanding of results. Given the fact that a large amount of financial and accounting information is employed to produce the empirical evidence, the findings presented in the thesis may be subject to the managerial manipulation of the reported accounting data. On the other hand, the frequency of the data is annual in order to use the accounting data; therefore, the data frequency may not be precise enough to accurately estimate the variable coefficients.

Secondly, the thesis does not intend to develop new theoretical models on asset pricing models. By using statistical testing methods, we can only find empirical evidence to support or reject the theoretical inverted-U curve hypothesis implied by the existing asset pricing theories. Therefore, theoretical issues in asset pricing modes are not addressed in this thesis.

Thirdly, the study provides some evidences for the non-linear effect of firm size on stock returns, but the thesis does not find exact value for the threshold point for this inverted-U curve shape relation. The exact threshold value for the inverted-U curve is out of the scope of the paper, but its exact value does worth further research and investigation. Besides, the thesis is produced by combining two stand-alone research papers, presented in Chapter 2 and 3. Hence there is some limited amount of overlap among the chapters, in particular the duplication of literature review materials. Nevertheless, it is also necessary to briefly review the relevant previous studies in the empirical chapters and a limited amount of overlap among chapters is allowed. But the amount of overlap has been minimised as much as possible.

Despite the above limitations, this research reveals new insights into the explanation power of the factors proposed by the Fama and French three-factor model, and the way how the factor affects the stock return as the core. Some of the limitations of the research are expected to be addressed by further research.

Reference

Aaker, D.A. and Jacobson, R., 1987. The role of Risk in explaining difference in Profitability. Academy of Management Journal 30 (2), 277 – 296.

Acharya, V. and Pedersen, L.H., 2005. Asset pricing with liquidity risk. Journal of Financial Economics 77, 375–410.

Al-Horani, A., Pope, P. F. and Stark, A. W., 2003. Research and Development Activity and Expected Returns in the United Kingdom. European Finance Review 7, 27–46.

Amel-Zadeh, A., 2011. The Return of the Size Anomaly: Evidence from the German Stock Market. European Financial Management 17 (1), 145–182.

Apergis, N., Artikis, P. and Sorros, J., 2011. Asset pricing and foreign exchange risk. Research in International Business and Finance 25, 308 – 328.

Bali, T. and Cakici, N., 2010. World market risk, country-specific risk and expected returns in international stock markets. Journal of Banking and Finance 34, 1152 – 1165.

Banz, R.W., 1981. The relationship between return and market value of common stocks. Journal of Financial Economics 9, 3–18.

Barry, Ch. and Brown, S., 1984. Differential information and the small firm effect. Journal of Financial Economics 13, 284–294.

Barry, C.B., Goldreyer, E., Lockwood, L. and Rodriguez, M., 2002. Robustness of size and value effects in emerging equity markets, 1985–2000. Emerging Markets Review 3, 1–30.

Berk, J., 1995. A critique of size-related anomalies. Review of Financial Studies 8, 275–86.

Berk, J., 2000. A view on the current status of the size anomaly, in Keim, D. and

Ziemba, W., eds, Security Market Imperfections in World Wide Equity Markets. Cambridge: University Press, 90–115.

Berk, J.B., Green, R.C. and Naik, V., 1999. Optimal investment, growth options, and security returns. Journal of Finance 54, 1553–1607.

Bhojraj, S. and Swaminathan, B., 2001. Macro-momentum: Evidence of predictability in international equity markets. Johnson Graduate School Working Paper SSRN (273569).

Bowman, H. E., 1980. The risk/return paradox for strategic management. Sloan Management Review 21 (3), 17-31.

Bowman, H. E., 1982. Risk seeking by troubled firms. Sloan Management Review 23 (4), 33-42.

Brown, P., Kleidon, A.W. and Marsh, T.A., 1983b. New evidence on the nature of size related anomalies in stock prices. Journal of Financial Economics 12, 33–56.

Carhart, M. 1997. On persistence in mutual fund performance. Journal of Finance 52, 57–82.

Cenesizoglu, T., 2011. Size, book-to-market ratio and macroeconomic news. Journal of Empirical Finance 18, 248-270.

Chan, K. and Chen, N., 1991. Structural and return characteristics of small and large firms. Journal of Finance 46, 1467–1484.

Chen, L., Novy-Marx, R. and Zhang, L., 2010. An Alternative Three-factor Model. Working paper.

Chienwei, H. and Chi-Hsiou, H., 2009. Investor sentiment as conditioning information in asset pricing. Journal of Banking & Finance 33, 892–903.

Chou, P.-H., Ko, K.-C. and Lin, S.-J., 2010. Do relative leverage and relative distress really explain size and book-to-market anomalies? Journal of Financial Markets 13, 77–100.

Conrad, J., and Kaul,G., 1998. An anatomy of trading strategies. Review of Financial Studeies 11, 489-519.

Daniel, K., Hirshleifer, D., and Subrahmanyam, A., 1998. Investor psychology and security market under-and over reactions. Journal of Finance 53, 1839-1886.

Daniel, K., Hirshleifer, D., and Subrahmanyam, A., 2001. Overconfidence, arbitrage, and equilibrium asset pricing. Journal of Finance 56, 921–965.

De Moor, L. and Sercu, P., 2012. The smallest firm effect: An international study. Journal of International Money and Finance, doi:10.1016/j.jimonfin.2012.04.002.

De Groot, C.G. and Verschoor, W.F., 2002. Further evidence on Asian stock return behavior. Emerging Markets Review 3, 179–193.

Dimson, E., Nagel, S. and Quigley, G., 2003. Capturing the Value Premium in the United Kingdom. Financial Analysts Journal 59, 35–45.

Dissanaike, G., 2002. Does the size effect explain the UK winner-loser effect? Journal of Business Finance and Accounting 29, 139–54.

Drew, M.E. and Veeraghavan, M., 2002. A closer look at the size and value premium in emerging markets: Evidence from the Kuala Lumpur stock exchange. Asian Economic Journal 17, 337-351.

Eikseth, H. and Lindset, S., 2009. A note on capital asset pricing and heterogeneous taxes. Journal of Banking and Finance 33, 573 – 577.

Elena, A., Hendrik, B. and Ivalina, K., 2010. Liquidity biases in asset pricing tests. Journal of Financial Economics 96, 215–237.

Fama, E.F. and French, K.R., 1992. The cross-section of expected stock-returns. Journal of Finance 47, 427–465.

Fama, E.F. and French, K.R., 1993. Common risk factors on the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Fama, E. and French, K., 1995. Size and book-to-market factors in earnings and returns. Journal of Finance 50, 131–155.

Fama, E.F. and French, K.R., 1996. Multifactor explanations of asset pricing anomalies. Journal of Finance 51, 55–84.

Fama, E.F. and French, K.R., 1998. Value versus growth: the international evidence. Journal of Finance 53, 1975–1999.

Fama, E.F. and French, K.R., 2012. Size, value, and momentum in international stock returns. Journal of Financial Economics 105, 457–472.

Fama, E.F. and MacBeth, J., 1973. Risk and return: some empirical tests. Journal of Political Economy 81, 607–636.

Fan, X. and Liu, M., 2005. Understanding size and book-to-market ratio: an empirical exploration of Berk's critique. Journal of Financial Research 28, 503–518.

Ferguson, M.F. and Shockley, R.L., 2003. Equilibrium ‘‘anomalies’’. Journal of Finance 58, 2549–2580.

Fletcher, J., 2001. An Examination of Predictable Risk and Return in UK Stock Returns. Journal of Economics and Business 53, 527–546.

Fletcher, J. and Kihanda, J., 2005. An examination of alternative CAPM-based models in UK stock returns. Journal of Banking & Finance 29, 2995–3014.

Frederico, B., 2010. Production-based measures of risk for asset pricing. Journal of Monetary Economics 57, 146–163.

Fu, F., 2009. Idiosyncratic risk and the cross-section of stock returns. Journal of Financial Economics 91, 24–37.

Gerard, A. M. and Mathijs, A. van D., 2010. Inflation risk and international asset returns. Journal of Banking & Finance 34, 840–855.

Gomez Biscarri, J. and Lopez Espinosa, G., 2008. The influence of differences in accounting standards on empirical pricing models: An application to the Fama–French model. Journal of Multinational Financial Management 18, 369–388.

Gregory, A., Harris, R.D.F. and Michou, M., 2001. An Analysis of Contrarian Investment Strategies in the UK. Journal of Business Finance and Accounting 28 (9&10), 1931-1228.

Gregory, A., Tharyan, R. and Huang, A., 2009. The Fama-French and Momentum Portfolios and Factors in the UK. Xfi Centre for Finance and Investment, University of Exeter Business School Working paper series, http://ssrn.com/abstract=1506960.

Griffin, J.M., 2002. Are the Fama and French factors global or country specific? Review of Financial Studies 15, 783–803.

Hamon J. and Jacquillat, B., 1999. Is there value-added information in liquidity and risk premiums? European Financial Management 5, 369–393.

Hong, H., Lim, T. and Stein, J.C., 2000. Bad news travels slowly: size, analysts, and the profitability of momentum strategies. Journal of Finance 55, 265–295.

Hong, H., and Stein, J. C., 1999. A unified theory of under-reaction, momentum trading and overreaction in asset markets. Journal of Finance 54, 2143-2184.

Hunjra, A.I., et al., 2011. Risk and Return Relationship in Stock Market and Commodity Prices: A Comprehensive Study of Pakistani Markets. World Applied Sciences Journal 13 (3), 470-481.

Hussain, S.I., Toms, J.S. and Diacon, S., 2002. Financial distress, market anomalies and single and multifactor asset pricing models: New evidence. Available at SSRN: http://ssrn.com/abstract=313001

Javed, I., Robert, B. and Don, U.A. G., 2010. Testing conditional asset pricing models: An emerging market perspective. Journal of International Money and Finance 29, 897–918.

Jegadeesh, N. and Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. Journal of Finance 48, 65–91.

Jegadeesh, N. and Titman, S., 1999. Profitability of Momentum Strategies: An Evaluation of Alternative Explanations. NBER Working paper 7159.

Joachim, G., and Andreas, S., 2009. Asset pricing with a reference level of consumption: New evidence from the cross-section of stock returns. Review of Financial Economics 18,113–123.

Jos´e, L. M. M. and Mar´ıa del, M. M. Q., 2006. The role of an illiquidity risk factor in asset pricing: Empirical evidence from the Spanish stock market. The Quarterly Review of Economics and Finance 46, 254–267.

Kim, D., 2006. On the information uncertainty risk and the January effect. Journal of Business 79, 2127–2162.

Kliger, G. and Tsur, I., 2011. Prospect Theory and Risk-Seeking Behavior by Troubled Firms. The Journal of Behavioural Finance 12, 29–40.

Lam, K.S.K. and Tam, L.H.K., 2011. Liquidity and asset pricing: Evidence from the Hong Kong stock market. Journal of Banking and Finance, doi:10.1016/j.jbankfin.2011.01.015.

Lamoureux, C.G. and Sanger, G.C., 1989. Firm size and turn-of-the-year effects in the OTC/Nasdaq market. Journal of Finance 44, 1219–1245.

Lensink, R. and Murinde, V., 2006. Does foreign bank entry really stimulate gross domestic investment? Applied Financial Economics 16, 569–582.

Lensink, R. and Murinde, V., 2006. The Inverted-U Hypothesis for the Effect of Uncertainty on Investment: Evidence from UK Firms. The European Journal of Finance 12 (2), 95–105.

Lewellen, J., 2002. Momentum and autocorrelation in stock returns. Review of Financial Studies 15, 533-573.

Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. The Review of Economics and Statistics 47 (1), 13–37.

Lischewski, J., and Voronkova, S., 2012. Size, value and liquidity. Do They Really Matter on an Emerging Stock Market? Emerging Markets Review 13, 8–25.

Malkiel, B. and Jun, D., 2009. The “value” effect and the market for Chinese stocks. Emerging Markets Review 10, 227–241.

Manuel, A. and Michael, M., 2008. Testing Conditional Asset Pricing Models Using a Markov Chain Monte Carlo Approach. European Financial Management 14(3), 391–418.

Markowitz, H., 1952. Portfolio selection, Journal of Finance 7, 77–91.

Maroney, N and Protopapadakis, A., 2002. The book-to-market and size effects in a general asset pricing model: Evidence from seven national markets. European Financial Review 6, 189-221.

Merton, R.C., 1987. A simple model of capital market equilibrium with incomplete information. Journal of Finance 42, 483–510.

Moerman, G.A., 2005. How domestic is the Fama–French three factor model? An application to the Euro Area. Working paper, Erasmus University, Rotterdam.

Mukherji, A., Desai, B.A. and Wright, P., 2008. A Contingent Relationship between Risk and Return: Toward A Behavioral Model of Decision Making. Institute of Behavioural and Applied Management.

Nijman, T.E., Swinkels, L.A.P. and Verbeek, M., 2004. Do countries or industries explain momentum in European? Journal of Empirical Finance 11, 461-481.

Pastor, L. and Stambaugh, R., 2003. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642–685.

Reingaum, M.R., 1981. Misspecification of capital asset pricing: empirical anomalies based on earning yield andmarket values. Journal of Financial Economics 9, 19–46.

Roll, R., 1981. A possible explanation of the small firm effect. Journal of Finance 36, 879–888.

Ross, S.A. 1976. Arbitrage theory of capital asset pricing. Journal of Economic Theory 13, 341–360.

Rousseau, P.L. and Wachtel, P., 2002. Inflation thresholds and the finance–growth nexus. Journal of International Money and Finance 21, 777–793.

Sharpe, W.F., 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. Journal of Finance 19 (3), 425–442.

Sinaee, H. and Moradi, H., 2010. Risk-Return Relationship in Iran Stock Market. International Research Journal of Finance and Economics, ISSN 1450-2887 Issue 41.

van Dijk, M.A., 2011. Is size dead? A review of the size effect in equity returns. Journal of Banking & Finance 35, 3263–3274.

Vassalou, M. and Xing, Y., 2004. Default risk in equity returns. Journal of Finance 59, 831–69.

Weimin, L., 2006. A liquidity-augmented capital asset pricing model. Journal of Financial Economics 82, 631–671.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.

Yao, T., Yu, T., Zhang, T., Chen, S., 2011. Asset growth and stock returns. Pacific-Basin Finance Journal 19, 115–139.

Appendix Table A1: Definition, Measurement and Source of Data for Variables used in Estimation and Testing

All the variables are collected on an annual basis for 1004 UK-listed nonfinancial firms from London Stock Exchange and the data period covered is from 1991 to 2011. The main databases used are Thomson ONE Banker (i.e. Thomson Datastream and Thomson Worldscope) & LSPD.

Variables	Definition	Measurement /Construction	Data Source	Item Name from Database
P_t	The closing share price at time t	Share price at time t	Thomson Datastream	DS.PriceClose
$R_{i,t}$	The return on asset i for time t	Worldscope Item Name: TOTAL INVESTMENT RETURN	Thomson Worldscope	WS.TotalInvestmentReturn
Market Value (MV)	The value of a firm as determined by the market price of its issued shares and common shares outstanding.	Market capitalization. Share price multiplied by number of shares.	Thomson Datastream	DS.MarketValue
Market-to-Book ratio	Market value of a firm as a percentage of its total equity.	Market price compared to the price of the firm on its books, calculated as (MV/BV) x100%	Thomson Datastream	DS.MarketToBookValue
Book Value per share	the book value of the stock at the firm's fiscal year end	Worldscope Item Name: BOOK VALUE PER SHARE	Thomson Worldscope	WS.BookValuePerShare
Book-to-Market ratio	The ratio of book value to market value Book Value is defined as the equity capital and minus total intangibles assets.	Book value/market value	By calculation	BookToMarket(BV/MV)

	Market value is defined as the market capitalisation of the firm; it is calculated as: share market price multiplied by number of shares outstanding.			
$Size_{i,t}$	Firm size for firm i at time t	Market capitalization = Market share price x No of shares outstanding	Thomson Datastream	DS.MarketValue
$Size_{i,t}^2$	Squared size factor	$Size_{i,t}$ x $Size_{i,t}$	By calculation	MarketValuexMarketValue
$R_{f,t}$	Risk-free rate at time t	UK Treasury Bill rate (One Month)	LSPD	TBR (One Month)
$R_{m,t}$	Return on the market at time t	FTSE All Share Index returns[31]	LSPD	FTSEAllSharePer(%)
$R_{m,t} - R_{f,t}$	Market risk premium is defined as the difference between the expected return on a market portfolio and the risk-free rate	$R_{m,t} - R_{f,t}$	By calculation	MarketRiskPremium(MRP)
SMB_t	Size risk factor for period t	Constructed as the difference between the average return on a portfolio containing small stocks and the return on a portfolio	---	SMB(SizeFactor)

[31] The FTSE All-Share is a market-capitalisation weighted index representing the performance of all eligible firms listed on the London Stock Exchange's main market, which pass screening for size and liquidity. Today the FTSE All-Share Index covers approximately 98% of the UK's market capitalisation.

		containing the large stocks		
HML_t	Book to market risk factor for period t	Constructed as the difference between the average return on a portfolio with high book-to-market stocks and the average (or weighted) return on a portfolio with low book-to-market stocks	---	HML(ValueFactor)
Dividend Yield ($Div_{i,t}$)	A financial ratio that shows how much a company pays out in dividends each year relative to its share price.	The annualized dividend rate expressed as a percentage of the share price, at the end of the year.	Thomson Datastream	DS.DividendYield
Price-to-earnings ratio ($PTE_{i,t}$)	A valuation ratio of a firm's current share price compared to earnings per share (EPS).	The price divided by the earnings rate per share at the end of the year.	Thomson Datastream	DS.PERatio
Bid Ask Spread ($BAS_{i,t}$)	PriceAsk: Asking price at close of market. The lowest value a prospective seller would be willing to accept for a particular security.	Bid ask spread is used to proxy for the liquidity factor. Bid-Ask Spread = PriceBid – PriceAsk	Thomson Datastream	DS.PriceAsk
	PriceBid: Bid price at close of market. Highest price a prospective buyer is willing to pay for a security at a particular price.			DS.PriceBid

			---	BidAskSpread
Trading Volume (TVOL $_{i,t}$)	The number of shares or contracts traded in a security or in an entire market during a given period of time, here is the end of the year period.	This shows the number of shares traded for a stock on a particular period. Trading volume is used to proxy for the liquidity factor.	Thomson Datastream	DS.Volume
Possible Instrumental Variables for Size:				
Book Value of Total Assets (BA $_{i,t}$)	Total Assets represent the sum of total current assets, long term receivables, investment in unconsolidated subsidiaries, other investments, net property plant and equipment and other assets. It represents the non-market measures of size.	Worldscope Item Name: TOTAL ASSETS	Thomson Worldscope	WS.TotalAssets
Total Sales (TS $_{i,t}$)	Represent gross sales and other operating revenue less discounts, returns and allowances. It represents the non-market measures of size.	Worldscope Item Name: NET SALES OR REVENUES	Thomson Worldscope	WS.Sales
Leverage (LEV $_{i,t}$)	The leverage ratio is measure by the long term debt-to-market equity ratio. Leverage is to model default risk.	Worldscope Item Name: LONG TERM DEBT % COMMON EQUITY	Thomson Worldscope	WS.LTDebtPctCommonEquity

Profitability Ratio (PRO $_{i,t}$)	Earnings Before Interest and Tax (EBIT): represents the earnings of a company before interest expense and income taxes.	Worldscope Item Name: EARNINGS BEFORE INTEREST AND TAXES (EBIT)	Thomson Worldscope	WS.EarningsBeforeInterestAndTaxes
	Total Sales: represent gross sales and other operating revenue less discounts, returns and allowances.	Worldscope Item Name: NET SALES OR REVENUES	Thomson Worldscope	WS.Sales
	Profitability ratio measures the company's use of its assets and control of its expenses to generate an acceptable rate of return, which is to model default risk. Calculated as: Profitability Ratio (PRO $_{i,t}$) = $\frac{\text{Earnings Before Interest and Tax (EBIT)}}{\text{Total Sales}}$			

Appendix Table A2: Data screening criteria used in previous empirical studies on asset pricing with UK firm level data

Author (Year)	Aim of the Paper	Data Sources & Screening Criteria
Kostakis, Muhammad, Siganos (2011)	To investigate higher co-moments and asset pricing on London Stock Exchange. Specifically, the paper aims to study whether coskewness and cokurtosis premia are priced in the UK stock market and examine asset pricing implications of investors' preference with respect to return's skewness and kurtosis for shares listed on LSE.	a. All UK listed and de-listed common shares available in Thomson Datastream from January 1986 to December 2008, including both listed and dead firms, the dataset is free of any potential survivorship bias; The paper ends up with a final sample of 3,501 shares. b. Screening criteria: (1) exclude firms with market value less than £5 million and firms that we cannot obtain return data for at least 60 consecutive months that are necessary to estimate their coskewness and cokurtosis values; (2) further exclude unit trusts, investment trusts and ADRs; c. The market portfolio returns are proxied by the FTSE All Share Index returns; and the risk-free rate by the 10 UK interbank rate;
Kassimatis (2011)	The paper aims to investigate whether the different type of risk preference creates a premium in UK stock market.	a. the dataset is from the UK stock market and the period covers August 1998 to January 2009. The data is extracted from Datastream; b. Four filters for the data: (1) equities listed on exchanges outside the UK are deleted, (2) non-common equities such as ADRs, warrants, etc, are deleted; (3) zero returns resulting from 'dead' stocks are deleted; (4) high returns reversed in the next period are checked and corrected if they are incorrect.

Morelli (2011)	To joint conditionality test the beta-return relationship based on the UK stock market. The role of beta in explaining stock returns is tested conditionally upon the sign of the excess market return.	a. The sample consists of 300 securities listed continuously on the UK stock exchange over the period 1st January 1980 to 31st December 2006; b. Data is obtained from the London Share Price Database. All monthly security returns are adjusted for dividends and stock splits. c. The three-month T-Bill rate is used to proxy for the risk-free rate. A value-weighted average of all 300 securities is used to proxy for the market portfolio.
Michou, Mouselli, and Stark (2010)	To test the effectiveness of Fama and French three-factor model in capturing risk in UK.	a. The sample period covered for firm returns is from July 1980 to April 2003; The empirical analysis uses annual accounting data from Datastream; b. Screening criteria: (1) include firms that have been de-listed from the exchange due to merger or bankruptcy etc; (2) exclude firms with more than one class of share, firms with negative book-to-market ratios, and firms that belong to the financial sector (banks, insurance companies, investment funds, unit trusts and property companies.
Fletcher and Kihanda (2005)	Examine the alternative CAPM-based models (e.g. the conditional and unconditional versions of the model) in UK stock returns	a. The excess returns of nine UK industry portfolios and the return on a one month UK Treasury Bill are used as the primitive assets.. b. The industry portfolios are Thomson Financial Datastream Level 3 industry sectors and are value-weighted. The Treasury Bill return is collected from the London Share Price Database. c. The return on the one-month Treasury Bill and the lagged dividend yield on the Financial Times All Share (FTA)[32] index are used as the information variables.

[32] The FTA index is a value-weighted index of the largest companies on the London Stock Exchange. Dividend yield is collected from LSPD.

Gregory and Michou (2009)	To explore the industry cost of equity capital for the UK industries, obtained from a conditional CAPM, the Carhart (1997) four-factor model, and the Al-Horani, Pope and Stark (2003) R&D model.	a. Data is sourced from Datastream and the London Share Price Database, covers period from 1975 to 2005 and the SEDOL number is used to match the companies between the two databases. b. Incorporate data for all companies, including de-listed companies either due to bankruptcy, acquisition or various other reasons in order to minimise survivorship bias.
Gregory (2007)	To investigate how low is the UK equity risk premium	a. The equity data, returns on Treasury Bills, and Yield to Redemption data on UK Government Gilts are required. b. The data source of earnings, price and dividend information for the early years (pre the formation of the FT All Share Index) is the Global Financial Data (GFD)[33] database.
Gregory, Harris and Michou (2003)	To carry out a comprehensive investigation on the relationship among the returns to value investment strategies and various macroeconomic variables in a multi-factor asset pricing model.	a. UK stocks, for which market capitalization and returns, and book-to-market ratios are available on the LSPD database and Datastream respectively. The study uses December year t-1 market capitalization and accounting data. b. the book value of assets, operating cash flows, earnings per share and three years' past sales are available on Datastream.
Gregory, Harris and Michou (2001)	To analyze the performance of contrarian investment strategies in the UK and documents that value strategies generate both statistically and economically significant	a. The sample period covered is from January 1975 to December 1998. b. The empirical analysis uses annual accounting data[34] from Datastream, and return data from the London Share Price Database. The firms are matched between the two databases by their SEDOL number.

[33] GFD is also used as the source of data for early gilt yields, prior to long run bond yields being available on Datastream. The early gilt yields are based on Consol yields. GFD is also the source of inflation data prior to the UK retail price index being available on either Datastream or from the Office of National Statistics.

	excess returns in the UK.	c. Screening criteria: (1) exclude firms with more than one class of ordinary share, firms with negative BM, EP, CP, and firms that belong to the financial sector (banks, insurance companies, investment funds and property companies); (2) include firms that have been de-listed from the exchange due to merger or bankruptcy etc.
Hon and Tonks (2003)	To test momentum in the UK stock market; in specific, to investigate the presence of abnormal returns by advocating trading strategies that exploit the predictability of stock price movements	a. The data is taken from the London Share Price Database tape of returns of UK companies from January 1955 to December 1996. b. There are a total of 1571 securities in the sample starting in January 1955, and as securities enter and leave the Exchange over the next 40 years, there are over 6600 securities in total over the entire sample period.
Vivian (2007)	Employing Fama and French (1992), the paper aims to examine the UK equity premium over more than 100 years (1901 – 2004)	a. The main data sample covering the period 1900-2002 is taken from the Barclays Equity-Gilt Study. It covers firms listed on the London Stock Exchange. b. The Barclays equity index[35] for the period 1900-1962 comprises the 30 largest shares by market capitalisation in each year and is rebalanced annually. From 1963-2002 the data is derived from the FTSE All-Share Index, it is then augmented for 2003-4 from Datastream, who also quote the index.
Lee, Liu and Strong (2007)	To evaluate the Fama-French three-factor model in the UK to determine whether the	a. The study covers the 25-year period from July 1977 to June 2004. The full sample comprises over 4,700 firms, including dead and surviving firms, based on the intersection of

[34] The accounting variables used in the analysis are book-to-market value of equity (BM), earnings yield (EP), cash flow yield (CP) and average sales growth over the previous three years (SG). Returns for each company, including dividends, are compounded annually, and are adjusted for changes in capital structure.

[35] The Barclays equity-price index is value-weighted with the weights of constituent companies being proportional to their market capitalisation.

	characteristics or covariance risk better explain the size and value premiums and show that small and value stocks earn higher stocks returns in spite of their return covariance	the London Share Price Database and Datastream. b. Factor/portfolio construction: construct the three factors for our UK sample following the method of Fama and French (1993)[36].
Dimson, Nagel, and Quigley (2003)	To investigate how is the value premium captured in the United Kingdom	a. The source of share price and listing information was from the London Share Price Database. It also includes all non-surviving firms and is, therefore, free of survivor bias. Stocks selected are officially listed on the LSE and foreign firms are excluded. Investment trusts were also excluded. Listing information, monthly returns, and monthly market values are obtained from the LSPD. b. the study linked the LSPD with accounting information from the database described in Nagel (2001). It combines data from three sources. The first source is Datastream, the second one is from the Cambridge/DTI database, which covers U.K. manufacturing firms (see Meeks and Wheeler 1999). For firms not in Cambridge/DTI or Datastream, balance sheets are collected from the official stock exchange yearbooks. In total, about 100,000 firm-years of accounting data are gathered.
Bauer, Cosemans and	To investigate the conditional asset pricing	a. The MSCI data set consists of the monthly return and book and market value for a sample

[36] (1) The market premium (MKT) is the difference between the market return and the risk free return (given by the 3-month UK Treasury bill yield). (2) SMB (Small Minus Big) is the difference between the returns on low and high market capitalization portfolios neutral with respect to book-to-market. Size is measured at the end of June of year t. (3) HML (High Minus Low) is the difference between the returns on high and low book-to-market portfolios neutral with respect to size. Book-to-market is measured at the end of December of year $t - 1$. (4) To obtain these factors, the study constructs six value-weighted portfolios based on the intersections of stocks sorted independently on size and book-to-market. To account for the dominance of small-cap stocks in the UK market and their negative correlation with book-to-market, the study follows the portfolio breakpoints of Dimson et al. (2003). The breakpoint for size is the 70th percentile, giving small (S) and big (B) groups. The breakpoints for book-to-market are the 40th and 60th percentiles, giving low (L), medium (M), and high (H) groups.

Schotman (2010)	and stock market anomalies in Europe.	of common stocks from 16 European countries that covers approximately 80% of European stock market capitalization. All variables are denominated in euros. b. The stocks are listed on the exchanges of Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, The Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the UK. The number of stocks per country ranges from 37 for Ireland to 519 for the UK. The MSCI data is free from survivorship bias as it includes historical data for firms that are delisted over time.
Fletcher (1997)	The paper examines the conditional relationship among beta and return in UK stock returns, and the significance of the relation between beta and return depends upon whether the excess market return is negative or positive.	a. Returns were collected for securities included in the London Business School Share Price Database between January 1975 and December 1994.. b. The monthly return on a 30-day UK Treasury Bill was used as the risk-free return obtained from LSPD.

Appendix Table A3: Data Generating Process

Standard formal research procedures are crucial for the whole thesis, especially the Data generating process. I started working on accessing and extracting data from the key databases: Datastream, Thomson ONE Banker, and EXCEL Add-In, which are the main databases and tools for my data. I need the *t* (time series) and the *n* (cross-section) in order to obtain a comprehensive panel of many UK firms observed for a long period of time. The data period covers 1991 to 2011, on an annual basis.

Data generating process		
	Database	**Main work undertaken**
Step 1	Datastream	From Datastream, the primary criteria employed are: Category: Equities; Market: United Kingdom; Currency: UK Pound; Exchange: London Stock Exchange; and Activity: Active. The data listed on Datastream are according to country and industry, there are totally 33 industries in the UK stock market.[37]

[37] The 33 industries including Aerospace & Defence, Alternative Energy, Automobiles & Parts, Beverages, Chemicals, Construction & Materials, Electricity, Electronic & Electrical Equipment, Fixed Line Telecommunications, Food & Drug Retailers, Food Producers, Forestry & Paper, Gas, Water & Multiutilities, General Industries, General Retailers, Health Care Equipment & Service, Household Goods & Home Construction, Industrial Engineering, Industrial Metals & Mining, Industrial Transportation, Leisure Goods, Media, Mining, Mobile Telecommunications, Oil & Gas Producers, Oil Equipment & Service, Personal Goods, Pharmaceuticals & Biotechnology, Software & Computer Services, Support Services, Technology Hardware & Equipment, Tobacco, and Travel & Leisure. Financial related industries are excluded mainly due to their

Step 2	Datastream	The variables identified and intended to generate are: market value, total assets, total sales, leverage, earnings before interests and taxes, ask price, bid price, trading volume, market to book ratio, dividend yield, price-to-earnings ratio, total investment return & price. I have generated 33 spreadsheets from Datastream for the 33 non-financial industries. However, the data I collected from Datastream have many N.A.s, and not all the variables have their corresponding values. Moreover, the data generated from Datastream are not easy to transfer into the panel form, which is necessary for STATA regression.
Step 3	Thomosn ONE Banker	I collected the firm level data at the annual frequency for the variables in my study. Firstly, I sorted out the firms that satisfy the basic search criteria employed: Public, Active, Non-ADRs only, Country code: GBR, Currency: GBP (British Pound Sterling), London Stock Exchange (XLON) and only the non-financial industries are included. I select ICB Industry code to specify the industry, including: '1000 Basic Materials', '3000 Consumer Goods', '5000 Consumer Services', '4000 Health Care', '2000 Industrials', '9000 Technology', '6000 Telecommunications', '7000 Utilities' and excluding '8000 Financials'. After executing the above criteria, I got 1004 UK firms from Thomosn ONE Banker, with the Firm KEY & Name (both full name and

different book-keeping method, including Banks, Equity Investment Instruments, Equity Warrants, Financial Service (Sector), Life Insurance, Non-equity Investment Instrument, Non-life Insurance, Real Estate Investment & Services, Real Estate Investment Trusts, 9 in total.

		abbreviation) spreadsheet.
Step 4	Thomosn ONE Banker	Secondly, I specify the variables for my empirical model, which need to be generated from the database (with the name from database in bracket); i.e. Price (DS.PriceClose), DividendPerShare (WS.DividendsPerShare), Market Value (DS.MarketValue), Market-to-Book ratio (DS.MarketToBookValue), Book Value Per Share (WS.BookValuePerShare), Dividend Yield (DS.DividendYield), Price-to-Earnings Ratio (DS.PERatio), Price Ask (DS.PriceAsk), Price Bid (DS.PriceBid), Trading Volume (DS.Volume), Book Value of Total Assets (WS.TotalAssets), Total sales (WS.Sales), Leverage Ratio (WS.LTDebtPctCommonEquity), and Earnings Before Interest and Taxes (WS.EarningsBeforeInterestAndTaxes). DS refers to Thomson Datastream, and WS refers to Thomson Worldscope. For the market return, R*m*, I use the return on the FTSE All share Index, and for R*f*, the risk-free rate, I use the one month return on Treasury Bills. I used EXCEL-Add In function to log in the database and select all the above variables from Thomson Datastream and Thomson Worldscope.
Step 5	Thomosn ONE Banker	Finally, in total, I got 21084 rows for 1004 firms, with 21 years (1991-2011) for each firm (1004 x21=21084).

Step 6	Thomosn ONE Banker	Next, I carried out necessary data computations to obtain the empirical counterparts of the variables that are required for estimation and testing of the model in the merged excel spreadsheet. Both the database-generated and computation-generated variables are included in one spreadsheet.
Step 7	---	Furthermore, I also get the Fama-French factors, portfolios and other benchmark portfolio data, constructed by Gregory et al. (2009). The dataset contains annual SMB and HML factors for the UK market 1980-2010 & the Fama-French portfolios used to create the SMB and HML factors and other benchmark portfolios. The dataset is similar to Fama and French (1993) for the US, whereas Gregory et al. (2009) is for the UK.

Note: I have generated a comprehensive master file of observable data for my empirical research. I have also screened and cleaned the data to insure continuity and consistency.