Long-term Abnormal Stock Performance:
UK evidence

Yan Huang

Submitted by Yan Huang to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Finance in May 2012.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature:………………Yan Huang………………………

Acknowledgement

Upon the accomplishment of this thesis, which marks the end of my journey in obtaining my PhD, I have mixed feelings. The most overwhelming sense among them is my gratitude towards the people who have kept encouraging and supporting me over the last few years.

I am sincerely thankful to Professor Alan Gregory, who not only nurtured and guided me as a student, but also provided me with invaluable help and advice in my research work. Without his support and help, this thesis would not have been possible. My thanks also go out to Professor Richard Harris and Dr Rajesh Tharyan for their valuable suggestions, constructive criticisms and extensive discussions on my work.

I feel heavily indebted to my parents. They have invested so much in me. Although they are physically on the opposite side of the globe, they have accompanied and encouraged me with their love and care to walk through every step of this journey.

I am also thankful to Celeste and Vincent Handford, who look after me like parents. I would not enjoy my life in the UK as I do now without their love and support.

My appreciation also goes to Dr Jane Shen. Studying the PhD can be a lonely experience, her friendship with a great deal of support and encouragement has been very important to me.

Finally, I would like to share my accomplishment with Xuefeng Sun, who has been incredibly supportive and helpful since the beginning of this journey.

Abstract

One of the most controversial issues for long-term stock performance is whether the presence of anomalies is against the efficient market hypothesis. The methodologies to measure abnormal returns applied in the long-run event studies are questioned for their reliability and specification. This thesis compares three major methodologies via a simulation process based on the UK stock market over a period of 1982 to 2008 with investment horizons of one, three and five years. Specifically, the methodologies that are compared are the event-time methods based on models (Chapter 3), the event-time methods based on reference portfolios (Chapter 4), and the calendar-time methods (Chapter 5).

Chapter 3 covers the event-time approach based on the following models which are used to estimate normal stock returns: the market-adjusted model, the market model, the capital asset pricing model, the Fama-French three-factor model and the Carhart four-factor model. The measurement of CARs yields misspecification with higher rejection rates of the null hypothesis of zero abnormal returns. Although the application of standard errors estimated from the test period improves the misspecification, CARs still yield misspecified test statistics. When using BHARs, well-specified results are achieved when applying the market-adjusted model, capital asset pricing model and Fama-French three-factor model over all investment horizons. It is important to note that the market model is severely misspecified with the highest rejection rates under both measurements.

The empirical results from simulations of event-time methods based on reference portfolios in Chapter 4 indicate that the application of BHARs in conjunction with p-value from pseudoportfolios is appropriate for application in the context of long-run event studies. Furthermore, the control firm approach together with student t-test statistics is proved to yield well-specified test statistics in both random and non-random samples. Firms in reference portfolios and control firms are selected on the basis of size, BTM or both. However, in terms of power of test, these two approaches have the least power whereas the skewness-adjusted

test and bootstrapped skewness-adjusted test have the highest power. It is worth noting that when the non-random samples are examined, the benchmark portfolio or control firm needs to share at least one characteristic with the event firm.

The calendar-time approach is suggested in the literature to overcome potential issues with event-time approaches like overlapping returns and calendar month clustering. Chapter 5 suggests that both three-factor and four-factor models present significant overrejections of the null hypothesis of zero abnormal returns under an equally-weighted scheme. Even for stocks under a value-weighted scheme, the rejection rate for small firms shows overrejection. This indicates the small size effect is more prevalent in the UK stock market than in the US and the calendar-time approach cannot resolve this issue. Compared with the three-factor model, the four-factor model, despite its higher explanatory power, improves the results under a value-weighted scheme. The ordinary least squares technique in the regression produces the smallest rejection rates compared with weighted least squares, sandwich variance estimators and generalized weighted least squares. The mean monthly calendar time returns, combining the reference portfolios and calendar time, show similar results to the event-time approach based on reference portfolios. The weighting scheme plays an insignificant role in this approach.

The empirical results suggest the following methods are appropriately applied to detect the long-term abnormal stock performance. When the event-time approach is applied based on models, although the measurement of BHARs together with the market-adjusted model, capital asset pricing model and Fama-French three-factor model generate well-specified results, the test statistics are not reliable because BHARs show severe positively skewed and leptokurtic distribution. Moreover, the reference portfolios in conjunction with p-value from pseudoportfolios and the control firm approach with student t test in the event-time approach are advocated although with lower power of test. When it comes to the calendar-time approach, the three-factor model under OLS together with sandwich variance estimators using the value-weighted scheme and the mean monthly calendar-time abnormal returns under equal weights are proved to be the most appropriate methods.

Table of Content

List of Figures and Tables

Chapter 3

Chapter 4

Chapter 5

Chapter 1: Introduction

1.1 Background on abnormal performance

An event study is an empirical methodology to analyze stock performance following an event. This includes three key factors: the event, abnormal returns during the event period and the significance of abnormal returns. This methodology assumes an efficient market, an unexpected event and the independence of events for a firm. The efficient market hypothesis advocates that stock prices respond immediately to all the information including past, public and insider information (Fama, 1970). Investors cannot gain abnormal returns in an efficient market, especially through technical analysis based on past stock price patterns, since the stock price reflects everything. The event studies can be traced back to the research by Dolley (1933) who discusses the change of stock prices responding to stock splits. Ball and Brown (1968) document a positive relationship between stock performance and accounting income based on stock returns on the announcement date of a firm's annual earning. Fama, Fisher, Jensen and Roll (1969) support market efficiency with no abnormal returns by investigating stock performance after the announcement of stock splits. They are considered to be the pioneers who established the process of event studies by studying the relationship between new information and stock returns. An event has two categories: firm-specific and Macro-events. Firm-specific events normally reflect company policies such as initial public offering (Ritter, 1991; Loughran and Ritter, 1995), takeover (Agrawal and Mandelker, 1990; Gregory, 1997; Bruner, 1999), seasoned equity offering (Loughran and Ritter, 1995,Brav, Geczy and Gompers, 2000), stock splits (Dharan and Ikenberry, 1995; Byun and Rozeff, 2003), share repurchases (Lakonishok and Vermaelen, 1995), acquisitions (Agrawal, Jaffe and Mandelker, 1992), dividend announcements (Michaely, Thaler and Womack, 1995), CEO or an auditor appointment (Defond et al., 2005; Weber et al., 2008), and compensation plan policy (Yeo et al., 1999). When there is a Macro-level event, this has a different influence on a series of firms in the same industry or market. Examples are change of legislation (Schumann, 1988), and financial crisis (Baek et al., 2004). Apart from stock returns, event studies extend to stock volatility (DeFusco et al., 1990; Jayaraman and Shastri, 1993), trading volume of a stock (Karafiath, 2009). Securities rather than stock, such as bonds, are examined for the effect of an event (Asquith and Wizman, 1990; Gugler et al., 2004). Other research areas besides financial markets, such as accounting, management, marketing, economics, widely apply event studies to measure the effect of an event on certain variables. To sum up, this methodology is widely discussed and used in both literature and practice. Stocks or portfolios are investigated within the framework of event studies to test the pattern of post-event performance in order for investors to make more effective investment decisions.

Over the past few decades, an increasing number of empirical studies based on the methodology of event studies with the application of different models and statistical inferences have been conducted. The evidence of anomalies documented both in short- and long-term time horizons following corporate events casts doubt on the market efficiency hypothesis (Dodd, 1980; Mitchell and Lehn, 1990; Higson and Elliot, 1998; Ritter, 1991; Gregory, 1997; Loughran and Vijh, 1997). However, initiated by Fama (1998), a rapidly growing research argues that the anomalies originate from the misspecification of models and statistical tests applied. The study of long-term stock performance following an event is considered to be a joint test. It not only tests market efficiency but also seeks for a well-specified equilibrium asset pricing model. This controversy is traced back to the basic concept of abnormal return which is defined as the difference between actual stock return and expected return. Various pricing models are proposed to predict expected returns. Among these are the capital asset pricing model, the market model, and the Fama-French three-factor model. But these models cannot fully capture the real returns even with carefully constructed controlling factors such as size and book-to-market ratio (Fama, 1998). This results in unreliable benchmarks to measure abnormal returns. In order to resolve the 'bad model' issue, some studies establish a reference portfolio or a single control firm with similar characteristics such as size and book-to-market ratio as the event firm (Lyon and Barber, 1997). Furthermore, Fama (1998) suggests the application of the calendar-time approach. But this approach is subject to criticism for ignoring investors' experiences and lower power of test in the long run (Ang and Zhang, 2004).

Apart from an appropriate pricing model, the application of statistical inferences is still an on-going debate. The conventional parametric test is questioned on its assumption of normal distribution with zero mean and constant variance. The violation of this strict assumption leads to inefficient and flawed estimates (Connolly and McMillan, 1989; Abad and Rubia, 1999). Parametric tests are therefore modified in consideration of heteroskedasticity and skewness to test the abnormal stock performance (Brown and Warner, 1980, 1985; Poulsen, 1991). However, the adjustment still gives biased results if a severe asymmetry exists. It has been proposed that nonparametric tests should be applied in event studies because no requirement of a normal distribution of abnormal returns. The sign test advocated by Brown and Warner (1980), is widely applied in event studies analysis (Kim and Schatzberg, 1987; Agrawal and Mandelker, 1990; McWilliams, 1990). It tests the median of abnormal returns with null hypothesis of equal numbers of positive and negative abnormal returns. Moreover, the rank test introduced by Corrado (1989) ranks the order of abnormal returns including estimation and event periods (Campbell and Wasley, 1993; Corrado and Zivney, 1992; Cowan and Sergeant, 1996). The bootstrapped process is incorporated in statistical inferences in event studies (Lyon, Barber and Tsai, 1999), particularly when the sample is small. Here, the underlying concept is to construct an empirical distribution which is closely matched with the real distribution by randomly selecting subsamples to compute a large number of test statistics. A transformed version involving the bootstrapping procedure is the empirical p value derived from pseudoportfolios (Lyon, Barber and Tsai, 1999). The null hypothesis of this test is the mean of portfolio returns is equal to the mean returns for the pseudoportfolios. Still, the nonparametric tests are inappropriate to be applied in event studies due to its lower power and imprecision (Edgington, 1995; Freidlin and Gastwirth, 2000).

So far, no universal agreement has been reached regarding the best model to predict future stock performance. It is impossible to completely eliminate the" bad model" issue but it is feasible to minimize it. As stated by Fama (1998) "all models for expected returns are incomplete descriptions of the systematic patterns in average returns", hence, an open question still exists regarding the bad model. Additionally, the statistical inferences in conjunction with abnormal returns from models are subject to debate. For instance, although the bootstrapped skewness-adjusted t-test introduced by Lyon, Barber and Tsai (1999) is a popular approach to apply with buy-and-hold returns, Mitchell and Stafford (1999) criticize it for its lack of cross-sectional independence. It is worth noting that long-term studies tend to have more critics and limitations than short-term studies, especially in terms of the statistical inferences. Fama (1991) claims short-term event studies are “the cleanest evidence I have on efficiency” while Kothari and Warner (1997) indicate “long-horizon studies require extreme caution”; this is consistent with the findings of studies related to misspecified tests for long-time horizons undertaken by Lyon and Barber (1996). The power of test is also subject to criticism, especially in the long run.

1.2 Two popular approaches in event studies

Two major approaches in the literature exist regarding how to conduct event studies: the event-time approach and the calendar-time approach. The event-time approach has two variants: models-based approach and reference portfolios approach including control firms as benchmarks. Similarly, the model-based approach, primarily consisted of the Fama-French three-factor model and Carhart four-factor model, and the mean monthly abnormal returns are two main methods in the calendar-time studies. The literature covering event studies still differs in opinion of which methodology in conjunction with statistical references yields well-specified test statistics and has higher power. It is important to note that the measurement of stock abnormal returns over a period of time is different when applying the even-time approach and the calendar-time approach. The most commonly applied measurements of abnormal returns in the event-time approach are cumulative abnormal returns (CARs) and buy-and-hold returns (BHARs), which accumulate abnormal returns of individual stocks over the investment horizon and then sum these up to achieve portfolio returns. However, the calendar-time approach employs average abnormal returns which sum up individual stock returns in a calendar month based on an equally- or value-weighted scheme to achieve a series of monthly portfolio returns for regression. The summary of key methods for event studies on abnormal performance is illustrated in **Figure 1.1**.

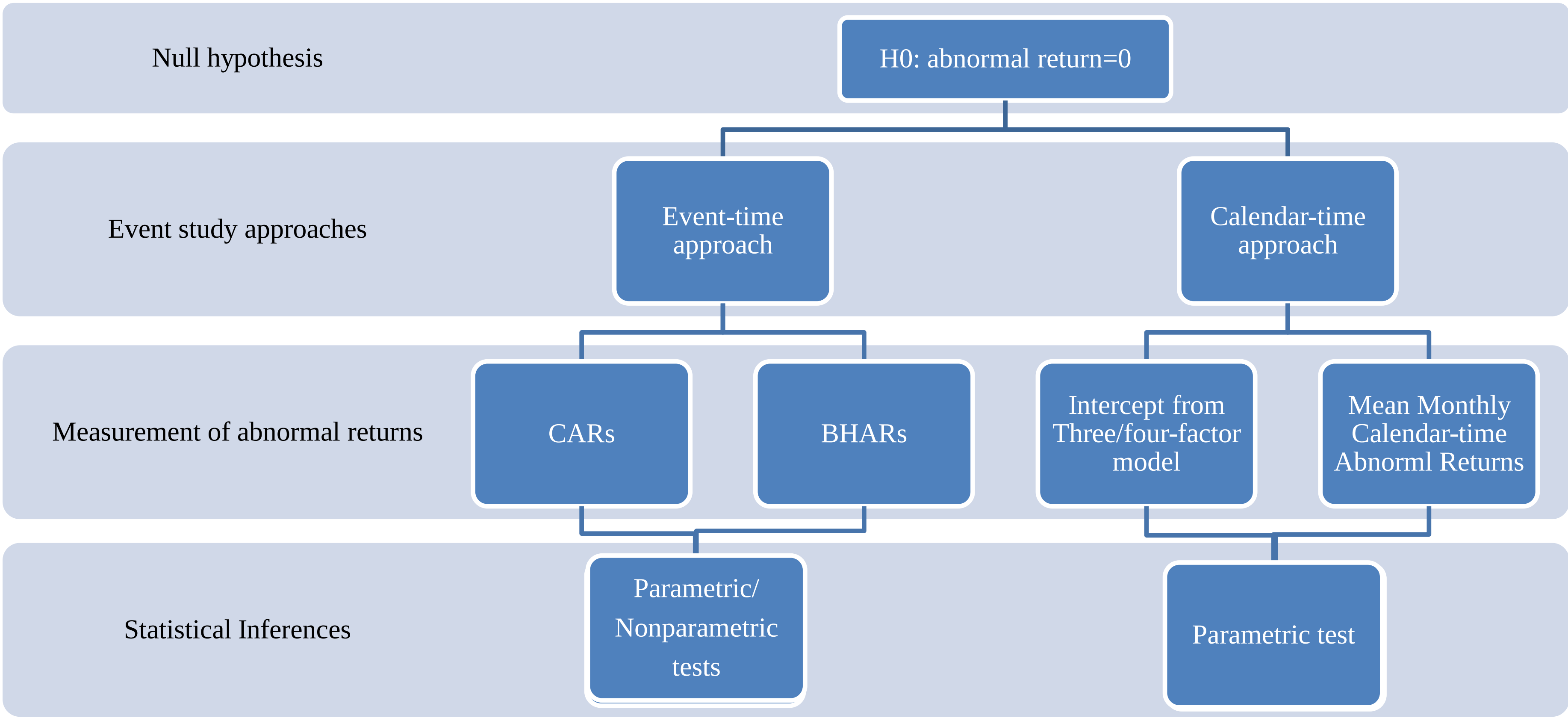
1.2.1 Event-time approach

The event-time approach measures the stock performance following a major corporate event based on event time, particularly for a portfolio of event firms. Therefore, an initial setup of an event date exists which is based on the occurrence of the event. If there is a portfolio of stocks, the event dates for all stocks are all set at the same starting point noted as time 0 for accumulation of stock returns on certain dates in order to test the null hypothesis of zero abnormal returns on event dates for the portfolio. Suppose a portfolio consists of firms A and B. The event months for A and B are January, 2000 and February, 2000. Those two event months are both denoted as t=0. The portfolio return in the event month 0 is $\mathrm{R}_{\mathrm{t=0}}^{\mathrm{P}} = \mathrm{R}_{\mathrm{t=0}}^{\mathrm{A}} + \mathrm{R}_{\mathrm{t=0}}^{\mathrm{B}}$. The exact event months are ignored. Since the abnormal return is defined as the difference between stock returns and benchmark returns, the benchmarks are categorized into three types: asset pricing models, reference portfolios and a control firm. Kothari and Warner (1997) examine long-term stock performance with the application of four asset pricing models: market-adjusted model, market model, capital asset pricing model and Fama-French three-factor model. Meanwhile, Lyon and Barber (1997) establish reference portfolios and a control firm as benchmarks matching by size, book-to-market ratio or both. The statistical inferences include a parametric test which bears an assumption of normal distribution and a nonparametric test which has no requirement of distribution. In an event-time approach, both cumulative return and buy-and-hold return can be utilized as a measurement of long-run abnormal returns over multiple periods. Lyon, Barber and Tsai (1999) advocate the application of buy-and-hold returns together with the bootstrapped skewness-adjusted test based on reference portfolios for its ability to control three biases: listing, rebalancing and skewness bias. However, Mitchell and Stafford (2000) argue this combination of measure of returns and test statistics still violates the assumption of cross-sectional independence of returns. Moreover, the bootstrapped distribution seems to be too ‘slim’ for statistical inferences. Mitchell and Stafford also show evidence of causes, namely industry and calendar clustering, for the violation. Lyon, Barber and Tsai (1999) further highlight another cause, average return clustering. When a company has more than one event over the study period, its own returns overlaps to create bias when applying the event studies.

1.2.2 Calendar-time approach

As far as the viewpoint regarding the weakness of the event-time approach is concerned, Jaffe (1974) and Mandelker (1974) propose the construction of a portfolio of stocks according to calendar months so as to eliminate the issue of cross-sectional dependence of returns. Lyon, Barber and Tsai (1999) advocate the calendar-time approach due to its advantage of well-specified tests in non-random samples. The conventional formula applied is a transformed version of the Fama-French three-factor model (1993). Ang and Zhang (2004) maintain that the Carhart four-factor model (1997) incorporating the momentum factor is not superior to the three-factor model. The calendar-time approach assumes managers rebalance their portfolio monthly with returns already taking cross-correlation of returns into account. Take five years as the study period and IPO as the event. Firms which experience events in the previous 60 months of the calendar month are included in the portfolio. The number of firms varies every month which could potentially cause heteroskedasticity (Lyon, Barber and Tsai, 1999). This issue can be overcome by using the weighted least squares technique in regression with consideration of the number of firms in each calendar month as weights (Ang and Zhang, 2004) or with White's (1980) correction or generalized least squares (Gregory, Guermat and Al-Shawawreh, 2008). Regarding regression, the equally- or value-weighted portfolio monthly returns are taken as one of the inputs in the formula. The factor models such as the Fama-French three-factor model are employed to obtain the intercept; this is an indicator of abnormal returns. Conventional test statistics are applied on the intercept due to approximated normal distribution of this estimator (Mitchell and Stafford, 2000). Despite the opinion expressed by Fama (1998), this approach is strongly supported by Brav and Grompers (1997), Mitchell and Stafford (2000). Furthermore, Loughran and Ritter (2000) pinpoint the fact that the model may underestimate abnormal returns when underperformance clusters together with the low power of test. A variant of the calendar-time approach, named mean monthly calendar-time abnormal return, is a combination of event-time reference portfolios and calendar time (Lyon, Barber and Tsai, 1999). In each calendar month, individual stock abnormal return is computed as the difference of stock returns and reference portfolio returns. The grand mean of portfolio returns is used to measure and test long-term abnormal stock performance.

Figure 1.1 Two approaches to test long-run abnormal returns



1.3 Motivation and contribution

Long-run event studies are considered to be more problematic than event studies on short windows, especially in statistical inferences (Fama, 1991; Kothari and Warner, 1997). Tuch and O'Sullivan (2007) list three major issues in long-run event studies, namely, unreliable test statistics due to joint tests, thin trading for some stocks, and overlapping events. Long-term stock performance is topical in empirical finance. Two different views exist: the presence of anomalies and the absence of anomalies with an underlying assumption of an efficient market. Researches focus mainly on known events and event dates, or simulated event studies with unknown event dates. Most of the research is based on the US stock market due to the comprehensively large number of US firms with a long history. The simulation process with random event dates based on historical return data is advocated to support Fama' argument (1998) concerning the misspecification of models and biased statistical inferences in long-run horizons (Kothari and Warner, 1997; Lyon and Barber, 1997; Lyon, Barber and Tsai, 1999; Mitchell and Stafford, 2000; Loughran and Ritter, 2000; Ang and Zhang, 2004). These key papers are all based on the US market. This gap gives me the first motive to initiate the research on the UK stock market over the period of 1982 to 2008 which has a relatively large number of small companies. The London Stock Exchange has two major markets: the Main Market with mostly large-cap firms, and the Alternative Investment Market (AIM) filled mainly with small-cap firms due to its more flexible regulatory requirements. Note that the Unlisted Securities Market (USM) only runs from 1980 to 1996. For instance, the Main Market totals 1080 UK listed firms with around 55.9% of firms with a market value over £50m in 2008. The AIM is mostly filled with small firms; approximately 73% of firms have a market value of less than £25m in 2008. With this special characteristic, it is interesting to discover whether the UK stock market, with a significant number of small firms, is still efficient, compared to prior research based on other markets when applying different methodologies within a framework of simulation process. Since small firms prevail in the UK stock market, I examine small size effect in detail via robustness checks using different approaches; I will then attempt to reduce this effect via the winsorization process in order to take outliers of returns into account or use the benchmark matched by size.

A number of studies are conducted on long-term stock performance following specific corporate events in the UK. Most of them reveal the presence of abnormal returns. However, Fama (1998) argues the asset pricing model does not fully capture the characteristics of a firm. This model misspecification may potentially result in imprecise abnormal returns. The arguments regarding misspecified models of expected return in event-time studies initiate a discussion of the calendar-time portfolio approach within a calendar-time framework. Recent studies by Kothari and Warner (1997) focus on expected returns measured by different asset pricing models; however, Lyon, Barber and Tsai (1999) and Ang and Zhang (2004) only investigate abnormal stock performance in long-run horizons based on reference portfolio and control firm benchmark in event-time studies and calendar-time approach in calendar-time studies. No study yet exists which comprehensively examines all the methodologies including models and statistical inferences to test long-term abnormal stock returns covering both event-time and calendar-time studies. Therefore, I attempt to conduct a more comprehensive simulation process in both event-time and calendar-time studies based on methodologies applied by Kothari and Warner (1997), Lyon and Barber (1997), Lyon, Barber and Tsai (1999) and Ang and Zhang (2004). The event-time studies include expected returns measured as a proxy from different asset pricing models or based on reference portfolios and a control firm with similar characteristics as sample firms. The calendar-time studies consist of the conventional calendar-time portfolio approach and mean monthly calendar-time abnormal returns, which has benchmarks based on different criteria. Cumulative abnormal returns and buy-and-hold abnormal returns are both applied to measure the stock performance over one-, three- and five- years. Furthermore, the statistical inferences cover parametric tests (conventional student-t test and skewness-adjusted test) and nonparametric tests (bootstrapped skewness-adjusted test statistics, empirical p value from pseudoportfolios, sign test, and rank test). The power of test is also investigated in three investment horizons. Both random samples and non-random samples such as small-/large-cap firms, firms with low-/high-book-to-market ratios are examined. Therefore, this thesis aims to provide a comprehensive framework of long-term abnormal stock performance study based on the historical simulation process.

Last but not least, regarding the calendar-time approach, only ordinary and weighted least squares techniques are adopted by Ang and Zhang (2004) in the simulation procedure so as to deal with the issue of heteroskedasticity. The generalized least squares technique introduced by Gregory, Guermat and Al-Shawawreh (2010) and the White (1980) correction test have not yet been utilized via the simulation process in prior literature. I, therefore, apply all four techniques with correction to heteroskedasticity in order to identify which technique outperforms the others.

1.4 Structure of the thesis

This thesis is structured with six chapters: introduction, literature review, event-time approach: simulation based on models, event-time approach: simulation based on reference portfolios, calendar-time approach and conclusion.

Chapter 2 starts with evidence of long-term abnormal performance documented by prior literature in the last few decades to show the overwhelming and extensive empirical studies. Due to the broad range of events investigated in the literature, I only discuss the most popular three topics, namely, initial public offering, seasoned equity offering, and mergers and acquisitions, in the context of long-run abnormal stock performance. Then, I move on to discuss the widely accepted and controversial methodology: event studies. The procedures of event studies are summarized in seven steps according to Campbell (1997). The first question in event studies is: what is a normal return? It is an expected return which can be measured by different models such as capital asset pricing model, factor models or even a proxy from reference portfolios. The models suffer strong criticism in the literature due to their inability to capture all the characteristics of a stock. The second question is how to measure abnormal returns. Regardless of a single stock or portfolio, cumulative abnormal returns and buy-and-hold abnormal returns are employed to measure abnormal returns across time or firms. Cumulative abnormal returns are subject to debate on three biases: new listing bias, measurement bias and skewness bias, while buy-and-hold returns are criticized for their new listing bias, rebalancing bias and skewness bias as summarized by Lyon and Barber (1997).

The third question is whether the statistical inferences successfully test the significance of abnormal returns in event studies. Several kinds of test statistics including parametric and nonparametric tests are proposed to fit in the context of event studies. I further discuss the issues of event studies such as cross-sectional abnormal returns, skewness and so on. In the last two sections, I review prior research on the event-time and the calendar-time approaches in detail by examining studies, especially those utilizing simulation processes, such as Kothari and Warner (1997), Lyon and Barber (1997) Mitchell and Stafford (1999), Lyon, Barber and Tsai (2000).

Chapter 3 covers an empirical study of the event-time approach based on models. The first section explains the source and process of data with an illustration of preliminary descriptive statistics of the raw data. Five models: market-adjusted model, market model, capital asset pricing model, Fama-French three-factor model and Carhart four-factor model, are examined in this chapter. The statistical inferences include conventional student t test with an assumption of a normal distribution of abnormal returns and Wilcoxon signed-rank test, a non-parametric test that is distribution-free. After performing programs in STATA, a well-functioned and popular software applied in a simulation process, I discuss the empirical results in two sections: random samples and non-random samples. In random samples, I report results of the rejection frequency at significance levels of 1% and 5% over one-, three- and five-year investment horizons, respectively. Both cumulative abnormal returns and buy-and-hold returns are discussed for comparison. Furthermore, two issues including distribution of abnormal returns and sample selection bias which potentially bring in biases in the results are examined. Non-random samples such small firms, large firms are examined for robustness check.

Chapter 4 details event-time studies based on reference portfolios, including control firms as benchmarks. To begin with, the raw data is the same methods but with different screening criteria. Secondly, I closely follow the research methodology applied by Lyon, Barber and Tsai (1999). The first section discusses the setup of reference portfolios. It is important to justify the benchmark portfolios which have different criteria. I eventually have reference portfolios established by size, by book-to-market ratio, by size and book-to-market ratio, and equally weighted market portfolios. Moreover, I have a single control firm, matched by size, book-to-market ratio or both as the event firm. The second part of the methodology covers statistical inferences which have conventional test statistics together with adjustment of skewness and nonparametric tests such as bootstrapped skewness-adjusted test statistics, empirical p value from pseudoportfolios and Wilcoxon signed-rank test. The simulation process is conducted in 250 samples of 200 firms, a small number of samples compared with studies by Lyon, Barber and Tsai (1999). This is due to the smaller number of firms in the UK stock market compared with the US. The empirical results are discussed in two sections: random samples and non-random samples. In random samples, both rejection rates of cumulative and buy-and-hold abnormal returns are discussed based on different tests. Moreover, the power of test with reference portfolios by size and BTM when applying BHARs over three investment horizons is examined. The non-random sample includes firms with small/large size, low/high book-to-market ratio, and industry clustering. Furthermore, I explore two issues in this method. One is industry clustering with sample firms clustering in the same industry. The other is calendar-time clustering with sample firms sharing the same event month.

Chapter 5 demonstrates the effectiveness of the calendar-time approach to deal with cross-sectional abnormal returns. Initially, the data, which is similar as the one applied in chapter 4, is briefly discussed. Additionally, the construction of reference portfolios is discussed in details. Secondly, I discuss research methodology according to Lyon, Barber and Tsai (1999) and Ang and Zhang (2004) in two parts. The first part illustrates models applied such as the Fama-French three-factor model, Carhart four-factor model and mean monthly calendar-time abnormal returns. It also includes the simple test statistics from regressions and the simulation process. Note that the portfolio returns in calendar months, rather than individual stock returns, are applied in the regression. The empirical results are discussed in two sections: conventional calendar-time approach and mean monthly calendar-time abnormal returns. Both sections have two subsections - random samples and non-random samples. As for the conventional calendar-time approach, in random samples, I focus on the rejection frequency based on equally weighted and value weighted portfolios. Both coefficients and power of test are discussed. Furthermore, I analyze the effect of winsorization which is applied to deal with outliers in stock returns. All tests including ordinary least squares, weighted least squares and generalized least squares in both weighting schemes are examined. In non-random samples, firms with small/large size, low/high book-to-market ratio, industry clustering, calendar-time clustering, and overlapping returns are discussed. The second empirical results are based on the mean monthly calendar-time abnormal returns. Both random sample and non-random samples are examined under equal weights and value weights with the application of benchmarks matched by size, BTM and both.

Chapter 6 summarizes empirical findings and discusses limitations and further potential research topics. Moreover, the implication for the finance industry is discussed.

Chapter 2 Literature Review

2.1 Introduction

"In general terms, the theory of an efficient market is concerned with whether prices at any point in time 'fully reflect' all available information. "

----Fama (1970)

Financial market anomalies, which are contrary to market efficiency hypothesis, are extensively documented and exploited in the literature. In the context of event studies, much controversy exists regarding whether there are abnormal returns over a long-time horizon, defined as one to five years. The methodology of event studies that examines stock performance following an event is summarized in detail by MacKinlay (1997), Campbell et al. (1997) and Kothari and Warner (2006). Initiated from event studies carried out by Fama, Fisher, Jensen and Roll (1969) based on stock split, attention is attracted to investigate stock performance with the application of event studies following major corporate events such as earning announcement, initial public offering, seasoned equity offering, takeovers. Extensive evidence regarding stock market anomalies with consistent abnormal stock returns is documented over the past few decades.[1] These anomalies are considered to indicate strong evidence of market inefficiency. However, as suggested by Fama (1970), the market equilibrium model is tested jointly with market efficiency. Therefore, the anomalies with overrejection of the null hypothesis of absence of abnormal returns could be potentially caused by misspecification of models and biased statistical inferences. Fama (1998) reviews prior research on event studies and concludes the market efficiency hypothesis holds due to the apparent frequency of underreaction as well as overreaction, and of post-event persistency of abnormal returns as post-event reversals. Moreover, he emphasizes the fact that long-run abnormal returns are sensitive to methodologies applied. This argument arouses the interest of academia and suggests a shift of attention to search for an improved model which can

[1] Earning announcement (Jones and Litzenberger, 1970), IPOs(Ritter, 1991; Loughran and Ritter, 1995); SEOs(Loughran and Ritter, 1997; Jegadeesh, 2000); mergers (Jensen&Ruback,1983; Magenheim&Mueller,1987;); Acquisitions(Mitchell and Stafford,2000)

completely capture the characteristics of stocks in conjunction with unbiased statistical inferences.[2] However, until now, no consensus exists regarding which model is superior. It is worth noting that the short-term event studies are less problematic than long-run event studies because longer time horizons compound errors and produce unreliable statistical inferences.

The literature regarding event studies is categorized into two groups: even-time approach and calendar-time approach. The event-time approach can be further divided into two variants. The first, advocated by Kothari and Warner (1997), is based on normal return defined as a proxy from different pricing models. The authors demonstrate severe misspecification of models with overrejections of the null hypothesis of zero abnormal returns across all models. The second event-time approach based on reference portfolios is proposed by Lyon and Barber (1997, 1999). They suggest reference portfolio returns be taken as a benchmark return to compare with observed stock returns in order to detect abnormal returns. Reference portfolios are constructed based on size, or book-to-market ratio, or both, or equally weighted market portfolios, or a single control firm. To deal with the assumption of normal distribution when applying conventional test statistics, they introduce non-parametric tests such as bootstrapped skewness-adjusted test statistics and empirical p-value from pseudoportfolios which do not require distribution of abnormal returns. However, this methodology is subject to criticism of cross-sectional abnormal returns and 'poor model'. Moreover, it does not work as well in non-random samples as in random samples. Therefore, they examine long-term stock performance based on the calendar-time approach, first introduced by Jaffe (1974) and Madelker (1974). Although this methodology identifies dependence of abnormal returns and works well in non-random samples, it is under criticism for three major issues. Firstly, it fails to take into consideration investors' experience regarding buy-and-hold returns. It also has lower power of test when introducing abnormal returns over periods. Lastly, the models based on the Fama-French three-factor model and Carhart four-factor model are still misspecified; this "can be attributed to the inability of firm size and book-to-market ratio to capture all of the misspecifications of the Capital Asset Pricing Model" (Lyon, Barber and Tsai, 1999).

[2] Kothari&Warner (1997), Lyon and Barber (1997, Barber, Lyon and Tsai (1999), Mitchell and Stafford (2000), Loughran and Ritter(2000), Ang and Zhang (2004)

2.2 Evidence of abnormal returns in the long-run event studies

The contrary views of the market efficiency hypothesis suggests not only that the price of stock does not fully and immediately reflect new information in a short run, but also indicates that abnormal performance exists in the long-term (one to five years). Since the studies undertaken by Ritter (1990) reveal findings of underperformance of IPOs in three years, a growing body of literature has shifted focus from short-term to long-term stock performance studies. The evidence of long-run abnormal returns with mixed results following corporate events such as equity issuance, dividends, stock splits, is widely documented in prior research. I subsequently discuss three of the most popular corporate events: initial public offering, seasoned equity issuing and mergers and acquisitions.

2.2.1 Initial Public Offerings (IPOs)

Most previous studies focus to a greater extent on initial stock performance based on the event date rather than aftermarket performance over a longer time horizon. However, earlier studies by Stoll and Curley (1970) document underperformance of only small IPO firms, whereas Ibbotson (1975) reveals consistency of market efficiency in the aftermarket but positive returns in the first year, negative returns in the next three years and positive in the fifth year. Furthermore, Buser and Chan (1987) show a positive performance of 6.2% in the first year and 11.2% in two years by applying IPOs during the period covering 1981 to 1985. Aggarwal and Rivoli (1990) test a sample of 1598 IPOs during the period of 1977 to 1987 and document the average returns of -13.73% from a period covering the close of the first trading day until the 250 trading days later. A developing literature is emerging focusing on long-run abnormal performance since the first paper published by Ritter (1991) who extends the post-IPO period to three years and finds consistent results of underperformance compared to previous studies. In his studies, 1526 IPOs covering the period from 1975 to 1984 are examined. The average return for the sample is 34.47%, while the average return for the benchmark of firms with similar characteristics is 61.86% over the time horizon of three years. The following study undertaken by Loughran (1993) examines 3,556 firms going public in the US market from 1967 to 1987. Underperformance of IPOs is evidenced with 17.29% of IPOs' returns and 76.23% of benchmark return over a six-year investment horizon.

Furthermore, Loughran and Ritter (1995) extend the study period to 1990. Sample firms going public are matched with a benchmark firm by size. Approximately -7% of abnormal returns is documented over five years. Rajan and Servaes (1997) compare post-IPO stock returns with market returns measured by NYSE and AMEX indices and show underperformance of firms going public with returns of 17%, whereas the benchmark return is 47.1% in a five-year investment horizon. Incorporating stocks from NASDAQ as a benchmark together with NYSE and AMEX, Carter, Frederick and Singn (1998) demonstrate findings of stock returns of IPO firms which are 19.92% lower when compared with the benchmark. A rapidly growing body of research applied similar methodology on different markets other than the US stock market and show consistent empirical results as Ritter (1991). Levis (1993) uses the IPO data during the period of 1980-1988 from the London Stock Exchange to test the robustness of previous studies. A sample of 712 IPOs shows underperformance over three years even with three different benchmarks-the Financial Times Actuaries All Share Index, the Hoare Govett Small Companies Index and the All Share Equally Weighted Index. The magnitude of abnormal performance ranges from -8.31% to -22.96% with the application of cumulative abnormal returns. Furthermore, the performance tends to be poorer when extending the post-event period to four and five years. When it comes to BHAR measurement, negative performance is not as excessive as the US evidence. Levis extends the empirical analysis in different stock markets such as the USA and Brazil. Similarly, the Finnish market documented substantial underperformance for IPOs with -22.4% of total average sample returns and -1.6% of benchmark returns over a three-year time horizon (Keloharju, 1993). Other markets are also explored for anomalies of IPOs, such as New Zealand (Firth, 1997), Hong Kong (McGuinness, 1993), Australia (Lee et al., 1996). Following the methodology outlined by Loughran and Ritter (1995), Brav and Grompers (1997) document underperformance of venture-backed and nonventure-backed IPOs based on a wide range of benchmarks over the period of 1972-1992. Applying the conventional buy-and-hold approach, empirical results indicate venture-backed IPOs tend to outperform nonventure-backed IPOs because of more apparent information, higher shareholding by institutional investors and capitalists' reputation. The negative abnormal returns are not consistent over the whole study period under investigation. Taking equally-weighted venture-backed IPO returns as an example, the underperformance dates from 1978 onwards. From 1976 to 1978, the trend of positive returns is evident. Although underperformance of IPOs is found when applying the buy-and-hold approach, this is not statistically significant when applying the calendar-time approach. Furthermore, portfolios of similar size/BTM to IPO firms demonstrate poor performance, indicating that underperformance is not caused by an IPO event. It is commonly believed that stocks which go public tend to have high initial returns and lower returns in the long-run compared with non-IPO firms.

Nevertheless, some studies argue that the universal proposal of underperformance of IPOs does not stand in certain circumstances. Brav and Gompers (2000) find no evidence of underperformance for firms going public over the period of 1975 to 1992 in five-year investment horizons. Furthermore, Schultz (2003) illustrates the underperformance of IPOs is subject to "pseudo market timing'; this means firms tend to go public in A similar timeline due to "high market returns and high returns for comparable firms". He suggests applying calendar-time returns which distribute equal weights to calendar months rather than event-time returns which give equal weights to IPOs. His findings imply no underperformance of IPOs is found when measuring returns in calendar-time. This conclusion suggests the underperformance of IPOs according to prior studies does not offer sufficient evidence concerning the argument against market efficiency. However, Ang, Gu and Hochberg (2007) argue that the underperformance of IPOs stands the test of the robustness checks in both calendar-time and event-time approaches. Furthermore, they suggest the underperformance of IPOs measured in calendar time "reappears" in longer investment horizons. Grompers and Lerner (2003) continue to question the justification of underperformance of IPOs if applying alternatives. They investigate IPOs from 1935 to 1973 before the NASDAQ was launched. The research subject in their study covers a small sample of firms, which is also taken as an attribute to go public by Loughran and Ritter (1995). Their findings reveal evidence of underperformance of IPOs with the application of value-weighted event-time buy-and-hold abnormal returns and no evidence when an equal weighting scheme is applied. However, Ang, Gu and Hochberg (2007) utilize a Markov switching model to oppose the idea of measurement issues.

2.2.2 Seasoned Equity Offerings (SEOs)

Empirical studies reach universal findings of negative stock returns on the event date and underperformance in the long-run following an equity issuance. Smith (1977) initially documents underpricing of firms issuing equities with a sample of 328 firms. Subsequent research in the 1980s completes models to study price reactions post-SEOs (Parsons and Raviv, 1985; Bhagat and Frost, 1986; and Giammarino and Lewis, 1989). Loughran and Ritter (1995) examine a sample of 3,702 SEOs over the period of 1970 to 1990. Firms undertaking SEOs have -8% of abnormal returns. Even the average annual returns of sample firms do not exceed the returns of T-bills. Five different benchmarks are applied to test the stock performance. It is shown that abnormal performance is more noticeable when using the S&P00 Index. Apart from the market size, book-to-market ratio is also used as criteria for matched firms. Nevertheless, limited explanation can be attributed to the underperformance after the control. It is worth noting that underperformance is more obvious when an active market exists for issuance. Spiess and Affleck (1995) find consistent evidence of underperformance for both IPOs and SEOs by examining firms over the period of 1975 to 1989. Particularly for SEOs, the difference of average returns between sample firms and benchmark firms with similar characteristics is -32.3% over five years. Furthermore, persistent underperformance is documented when controlling different factors such as size and book-to-market ratio. On the basis of previous studies on SEOs, Lee (1997) initiated a question of whether there is any relation between long-run returns of firms with SEOs and insider trading over three years. He classifies the sample into two categories: primary offerings (>=50% IPOs) and secondary offerings (<50%). If the issuer sells shares before issuing and the stock underperforms when compared with benchmark controlling for size, industry and book-to-market ratio, then this indicates that the issuer knows the stock is overvalued. Underperformance is found for both the primary offerings and secondary offerings by the author. But this only works for top executives who sell shares before issuing. When it comes to those who purchase shares, no evidence suggests underperformance for secondary offering with executives' purchases. Jegadeesh (2000) matches SEOs in the US stock market over the period of 1970 to 1994 with benchmark firms by size, book-to-mark ratio, earnings-to-price ratio, lagged six-month returns and lagged 36-month returns.

Benchmark returns are measured based on both value-weighted and equal-weighted schemes. Strong evidence of underperformance of SEOs over five years is shown by the author in different benchmarks. For instance, SEOs underperform 39.6% with a benchmark matched by both size and book-to-market ratio and 47.5% with a size-matched benchmark.

Similarly, SEOs also suffer criticism on the issue of misspecified benchmarks. It is common to use the reference portfolios approach to match the firm experiencing SEO with similar characteristics. Among those firms in a reference portfolio, it also includes firms which go public. In this case, since IPOs underperform in the long run, as suggested by Ritter (1991), the return of a benchmark portfolio tends to have lower return which gives a misleading conclusion regarding the performance of SEOs (Loughran and Ritter, 2000). Jegadeesh (2000) shows abnormal return of -25.7% for SEOs when excluding IPOs in the benchmark and of -36.6% when including IPOs in the benchmark under the value-weighted scheme. Moreover, the author argues misspecification of factor models, taken as proxies of benchmark returns, result in the misleading conclusions of underperformance of SEOs. Furthermore, Eckbo, Masulis and Norli (2000) demonstrate statistically insignificant underperformance of SEOs as a result of benchmarks which do not adjust risks. Multiple equity offerings, which occur more commonly in banking industries due to capital ratio requirements, show no evidence of abnormal performance in the first sales (Slovin, Shushka, and Polonchek, 1991). The authors also present evidence of strong underperformance when a bank meets capital requirements in the first and follow-up SEOs. However, D'Mello, Tawatnuntacha, and Yaman (2002) include other non-financial firms apart from banks in the sample and reach the opposite conclusion. They state that firms display statistically significant negative abnormal returns in the first sales and insignificant results in the subsequent sales.

2.2.3 Mergers and Acquisitions

The empirical findings regarding stock performance following mergers and acquisitions for acquirer and target are mixed since Roll initially indicates potential non-zero abnormal returns in 1986. Underperformance of stock prices in the long run for acquiring firms is evidenced with extensive research. Based on the US merges, Langetieg (1978), Asquith (1983) and

Magenheim and Mueller (1988) present significantly negative abnormal returns for acquirers after acquisition in one- to three-year investment horizons. Significant negative post-merger performance with -0.1026% CAAR over five years after adjustment of size and beta is found by Agrawal, Jaffe and Mandelker (1992). With control of size and book-to-market ratio, Anderson and Mandelker (1993) indicate statistically significant CAARs with approximately -0.0956% and -0.0931% with adjustment of size and book-to-market ratio over a five-year time horizon, respectively. Two-year CAARs are studied by applying six models based on the UK stock market over the period of 1984 to 1992, (Gregory, 1997). The two-year CAARs range from -11.82% to -18% with statistical significance that suggests strong evidence of underperformance for acquirers. Moreover, the magnitude of negative abnormal returns in the UK stock market is found to be larger than that in the US stock market. Loughran and Vijh (1997) firstly introduce buy-and-hold return based on a benchmark of control firms to test abnormal performance of the target and the acquirer. Their results covering the period of 1970 to 1989 show that the five-year abnormal returns are statistically insignificant for acquirers. The abnormal returns for acquiring firms using stock financing and cash after completion of acquisitions are -24.4% and 18.5% in five years, respectively. To incorporate adjustment of systematic risk, size and book-to-market ratio, Rau and Vermaelen (1998) investigate 2,823 merger and 316 tenders in the US market from 1980 to 1991 and show mergers bear a negative abnormal return of -4.0%. Moreover, Mitchell and Stafford (2000) examine 2,068 transactions between 1961 and 1993 with the application of the Fama-French three-factor model in three years. Mergers reveal a significant negative abnormal return of -0.04% based on the equally weighted scheme and -0.03% based on *the* value weighted scheme. Moeller et al. (2003) apply the buy-and-hold return measurement in the event-time approach and calendar-time approach based on 12,023 US acquisitions during the period of 1980 to 2001 and document significant negative abnormal returns in three years.

Although it seems the idea of underperformance for acquirers in the long run is well proven, other evidence indicates different views. Franks, Harris and Titman (1991) demonstrate significant positive abnormal returns for small deals based on 399 acquisitions in the US market over the period of 1975 to 1984. Note that they adjust systematic risk and size.

Furthermore, Rau and Vermaelen (1998) provide similar findings in a larger sample size based on merger and tender offers. Besides significant negative and positive abnormal returns, acquiring firms are claimed to exhibit insignificant abnormal returns after takeovers. This suggests there is no impact from takeovers on long-term stock performance of acquiring firms. In early studies, Mandelker (1974) applies the factor model introduced by Fama and MacBeth (1973) to examine 241 mergers in the New York Stock Exchange from 1941 to 1962. Insignificant abnormal returns over the first 10 and 20 months after merge are exhibited. Dodd and Ruback (1977) record insignificant test statistics for abnormal returns for CAAR over five years with the application of the market model. Frank, Harris and Titman (1991) apply a multifactor model and find insignificant underperformance over three years. Similar results were found by Loderer and Martin (1992) with longer investment horizon than five years but significantly negative three-year abnormal returns. With regard to the consideration of cross-sectional abnormal returns, Mitchell and Stafford (2000) still document no significant abnormal returns for acquirers over three years with a sample of 2,068 US deals. This view is echoed by Ikenberry et al. (2000), Andre, Kooli and L'her (2004), Dutta and Jog (2009) who investigate the Canadian acquisitions from the 1980s to 2000s and find no significant abnormal returns for acquiring firms in the long run.

2.3 Event studies

2.3.1 Background of event studies

"The cleanest evidence on market-efficiency comes from event studies……event studies can give a clear picture of the speed of adjustment of prices to information."

-Fama (1991)

The efficient market hypothesis, summarized by Fama (1970), illustrates that the price of a stock reflects all the relevant information including public and private information fully and simultaneously. This means when a macroeconomic event such as interest rate policy or a corporate-level event such as an announcement of earning takes place, the price of the stock should be incorporated with the news flow immediately. However, anomalies following events are widely documented for both short- and long-term stock performance with the applications of event studies which examine the impact from an event on the firm. Most empirical studies focus on short-run event studies which examine stock performance around the event date, with daily stock returns. Since the event window is short, the daily expected return is close to zero; this produces more reliably statistical inferences with limited negative impact from models applied to achieve expected returns. Moreover, the underlying assumption of short-run event studies implies even if there is a delayed reaction of prices, it disappears quickly since the event window is short. However, a vast body of literature disagrees with this assumption, arguing that stock performance over longer periods can test market efficiency more accurately if the market takes a longer time to digest the new information. Long-term stock performance bears with a joint test of a well-specified equilibrium model and market efficiency (Fama, 1970). Therefore, although long-term stock performance following major corporate events is extensively studied, researchers cast doubt on the reliability of statistical inferences in long-run test statistics due to risk adjustments in expected returns, independence of abnormal returns and the power of test.[3] A fruitful body of empirical studies based on the methodology of event studies not only focuses on accounting and finance but also other areas such as marketing, economics, law, management, and history.

Event studies provide a satisfactory solution to investigate how the prices react to the announcement of information including managerial decisions such as dividend policies, equity issues. The methodology of event studies can be traced back to a seminal article published by Dolley in 1933. A sample of 95 firms experiencing stock splits from 1921 to 1931 is examined for the price change following the event. Approximately 60% of firms have price increases, while 30% of firms experience negative stock returns. The remaining 12 firms have no significant change in prices. Inspired by Dolley's study, event studies attract more and more academics' attention in the 1940s and 1950s (Myers and Bakay, 1948, Austin Barker, 1956, Ashley, 1962). In 1960s, the empirical work by Ball and Brown (1968) examine the relationship of income to stock prices with the application of event studies. Three models are adopted to predict the movement of forecasted income and chi-square is used to examine the relation between income and abnormal performance index. If the actual earnings changes

[3] Kothari and Warner (2008)

are larger than those which are estimated from models, this is an indication that the stock return is expected to be positive. The authors conclude the price reflects approximately 80% to 90% of news before the earnings announcement, whereas income explains 50% of stock price changes. Moreover, the market adjusts stock price to its original level in 3 to 6 months after the announcement, suggesting market efficiency still holds in this circumstance. Following this, Fama, Fisher, Jensen and Roll (1969) offer a standard framework of event studies which is cited 2092 times by 2010. This is an event-time analysis which should be differentiated from a calendar time analysis. By taking an event date as an announcement date of a stock split, they study 622 firms with 940 splits excluding dividends in the US market during the period of January 1927 to December 1959. The market model is applied by regressing logarithm stock returns on logarithm market returns which use CRSP NYSE market portfolio returns as a proxy. The residual from the regression is taken as a proxy of abnormal return. The test period is from the previous 29 months before announcement date to 30 months thereafter. The test statistics are based on cumulative average abnormal returns across firms and months. They arrive at a similar conclusion to Ball and Brown (1980), that the stock price reflects the information event before the formal announcement.

Brown and Warner (1980, 1985) initially shed light on the statistical properties of daily and monthly stock returns when conducting event studies. Over the last two decades, a growing body of research follows. This questions the reliability of long-run event studies so as to support the market efficiency hypothesis with the argument that anomalies come from misspecification of models and statistical inferences. Kothari and Warner (1997) study long-term stock performance with unknown event dates via simulation. The market-adjusted model, market model, Fama-French three-factor model and capital asset pricing model are evidenced with misspecification tests with overrejection of the null hypothesis of zero abnormal returns in both cumulative abnormal returns and buy-and-hold returns. Meanwhile, Lyon and Barber (1997) construct reference portfolios as a benchmark to detect abnormal returns. The three types of benchmarks are “a reference portfolio, an appropriately matched control firm, and an application of the Fama-French three-factor model”. The control firm approach matched by size and book-to-market ratio together with conventional test statistics is claimed by the authors to produce a well-specified test and reduce biases such as the new listing bias, rebalancing bias and skewness bias. Fama (1998) summarize literature regarding event studies to test the market efficiency with the argument that "long-term return anomalies are sensitive to different models". Based on prior research, Lyon, Barber and Tsai (1999) propose nonparametric tests in addition to parametric tests so as to take in account the issue of skewness which violates the assumption of normal distribution. They conclude that the buy-and-hold abnormal returns in conjunction with bootstrapped skewness adjusted test statistics outperform. Furthermore, they advocate the application of the calendar-time approach, initially introduced by Jaffe (1974) and Mandelker (1974), to deal with the cross-sectional dependence abnormal returns. However, they argue that this approach does not reflect investors' experience as does the buy-and-hold return approach. Although the reference portfolio approach, control firm approach and calendar-time approach produce specified results in random samples, when it comes to non-random samples, they lose their power. Mitchell and Stafford (2000) question the approach proposed by Lyon et al. (1999), with an argument of violation of independence of abnormal returns. On examination of specific corporate events including mergers, seasoned equity offerings and share repurchases, they reach a conclusion even with the consideration of "positive cross-correlations of event-firm abnormal returns", no evidence of abnormal returns is found. They again strongly advocate the application of the calendar-time approach. Loughran and Ritter (2000) imply a "positive model", such as the control firm approach and reference portfolio approach, is inappropriate in its application to test market efficiency, while a "normative model" such as the asset pricing model is more suitable. However, the multi-factor model, such as the Fama-French three factor model, is claimed to be unsuitable because it underestimates abnormal return based on the value-weighted scheme when the corporate event is managerial decisions. Ang and Zhang (2004) conduct a simulation process to test the null hypothesis of zero long-term abnormal return, based on the calendar-time approach and buy-and-hold approach. They demonstrate the loss of power when the investment horizon lengthens. The control firm approach together with sign test is claimed to present the best performance. Although a heated discussion arose between defenders of market efficiency and opponents of market efficiency with fruitful published and upcoming articles, no consensus has been reached.

Long-run event studies, which normally have event windows lasting for more than 12 months, are classified into two types of analysis. One is event-time analysis which includes both cumulative abnormal returns and buy-and-hold abnormal returns using benchmark returns based on factor models or characteristics matched firms. For a single firm A, suppose it has an IPO in January, 2000, with a monthly return of 5%. The event month of January 2000 is denoted as month 0. If the study period is one year, then after the event month, months 1 to 12 are allocated to match with February, March and so on. This system is more straightforward and useful when there is a portfolio with multiple firms. Suppose firm B has an IPO in February, 2000, with a monthly return of 10%. As before, the event month of firm B is February, 2000, which is denoted as month 0. March, 2000 for firm B is then denoted as month 1. When assessing the cumulative returns of two firms in the event-time analysis, the portfolio return is computed based on the timeline of denoted months instead of the actual month. That is to say, the portfolio return in month 0 (event month) is 5% + 10% = 15%. The alternative analysis is based on calendar time, which calculates calendar-time returns for a portfolio of event firms based on an equally-weighted or value-weighted scheme by the application of multiple-factor models. The timeline is based on calendar time rather than allocated event time. Firms are assumed to experience an event in the previous n months prior to the calendar month. Applying the previous example of firm A and firm B. In the calendar month of February, 2000, since firm A has an event in January 2000, this is included in the portfolio. However, if firm B has an event in February, it has to be included in the portfolio when it comes to the calendar month of March. Therefore, in March, 2000, the portfolio comprises both firm A and B. After collecting the portfolio returns across calendar months, the regression based on the multifactor model is run to achieve the intercept which is a proxy to test if the long-term abnormal return is zero. As previously discussed, both approaches have advantages and disadvantages, making it difficult to determine which is superior.

Campbell, Lo and MacKinlay (1997) outline seven basic procedures to conduct event studies. Firstly, an event and event window need to be identified. An event includes both Macro-level and company-specific events. For instance, if the management of a firm decides to have cash dividends issued, this can be taken as an event. The date, when the announcement of the dividend policy is made, is considered as the event date. Researchers are interested in the stock price movement either before or after the event date, normally over a period. If it is a short-lived effect study, stock return on the event date or one or two days following the announcement is investigated. If it is a long-term stock performance study, stock returns can be extended to between one and seven years. Short-horizon studies focus on how rapidly the news is reflected into the price, whereas long-horizon studies focus on tests of market efficiency and specification of an equilibrium asset pricing model. The timeline of event studies includes an estimation window, event window and post-event window. Most long-run event studies have a short event window, for example, one month or one day (Lyon and Barber, 1997). When applying economic models, coefficients in the models are estimated based on stock returns in the estimation window; these can be defined as pre-event or post-event returns. The second step concerns sample selection. After identifying an event, a sample of firms needs to be carefully chosen according to the purpose of studies. For instance, when estimating the coefficients in a model, such as CAPM, availability of returns over the estimation window is required. The firms with missing returns over the estimation period need to be excluded from the sample (Kothari and Warner, 1997). Furthermore, firms with similar characteristics may need to be excluded from the sample due to potential bias from inter-correlation. For instance, when a firm experiences multiple events within the event period, its returns overlap the study period. This indicates cross-sectional dependence of returns which brings up the issue of unreliable statistics. Alternatively, if firms in the same industry cluster in one sample, returns of firms have an impact on each other which violates the assumption of conventional statistical inferences. Kothari and Warner (1997) argue potential biases come from the data screen. Thirdly, the measurement of normal and abnormal returns needs to be determined. The abnormal returns are simply the difference between actual stock returns and expected returns. As for actual returns, two approaches exist to achieve this through prices. One is simple return, which is sum of capital gain and dividend gain relative to previous prices. The other is logarithm returns; these apply natural logarithms of simple returns. Log return is widely applied in time-series, while simple return works well in panel data. There are many benchmark settings for expected returns. These can be market return under different weighting schemes such as FTSE ALL SHARE, or non-event reference portfolio returns based on similar characteristics such as size; alternatively a control firm matched with similar characteristics or a proxy from asset pricing models such as the market model. Moreover, the accumulation of abnormal returns can be categorized as cumulative abnormal returns and buy-and-hold returns. Fourthly, the estimation process is conducted. When the asset pricing model is applied to achieve normal returns, the parameter estimates can be achieved from regression of single stock returns on control factors. The parameter estimates are achieved from regression of returns on multiple factors. Returns over the pre-event and post-event periods can be applied in the regression. However, the pre-event window is commonly applied in the estimation process due to its lesser impact from the event compared with the post-event window. Fifthly, test statistics are computed after obtaining the abnormal returns. The null hypothesis is assumed to be zero mean abnormal return, which suggests an absence of abnormal performance. The two most popular streams of statistical inferences are parametric tests which bear an assumption of normal distribution and nonparametric tests that do not have requirement of distribution. It is worth comparing which test has the higher power to test the abnormal returns. Lastly it comes to presentation of results and conclusion.

2.3.2 What is a normal return?

The normal return is defined as an expected return when a firm does not experience an event. Particularly for long-run event studies, normal returns need to be carefully calculated since small errors accumulate. The measurement of a normal return can be classified into three groups: models which capture the major risks and characteristics of event firms, reference portfolios matched with event firms with similar characteristics and lastly, the control firm approach. The five popular models are constant-mean-return model, market model, market-adjusted model, capital asset pricing model (CAPM) and multi-factor model such as the Fama-French three-factor model or Carhart four-factor model. The normal returns of an event firm over an event window are estimated, based on coefficients from regression of returns of the firm over an estimation period. Therefore, it is necessary to identify a model which is well specified in order to capture as completely as possible the price impact from the event. The widely applied criteria to construct reference portfolios include size, book-to-market ratio, size and book-to-market ratio, market index. The reference portfolios normally contain non-event stocks with similar characteristics. However, those benchmarks may only reflect some commons shared among stocks.

2.3.2.1 Constant-mean-return model

Constant-mean returns are the addition of the mean return of a security and a residual which follows normal distribution with expected value to be zero and constant variance. Although it is a simple model, it yields results which are well-specified compared to other sophisticated models (Brown and Warner, 1985). The formula is shown as follows:

$$E(R_{i,t}) = \mu + \epsilon_{i,t},$$
$$\epsilon_{i,t} \sim N(0, \sigma_{\varepsilon,i}^2) \quad (2.1)$$

where $R_{i,t}$ is the individual stock return at time t and μ is the mean return of the stock over the period. The error term is assumed to follow normal distribution with zero mean and constant variance. Since there is no other coefficient apart from the mean return, there is no need to estimate over the estimation period. The application of this model is straightforward but it ignores additional factors such as the market which also have significant impact on stock returns.

2.3.2.2 Market model

The market model improves the constant mean return model by controlling market risk. It reflects a linear relationship between individual stock return and market portfolio return as follows:

$$E(R_{i,t}) = \alpha_i + \beta_i R_{m,t} + \epsilon_{i,t},$$
$$\epsilon_{i,t} \sim N(0, \sigma_{\varepsilon,i}^2) \quad (2.2)$$

where $R_{m,t}$ is the market return measured as different indices returns, based on either an equally-weighted or value-weighted scheme, while $\epsilon_{i,t}$ captures the unsystematic risk that is endogenous to individual stocks. The stock return is assumed to be correlated with the systematic risk which can be reduced through diversification. The parameters of the market model are estimated by running regression over the estimation period excluding event date in order to avoid the influence of events in the estimation. The application of the market model reduces the variance of abnormal returns. The R square subtracted from regression illustrates the significance power of the independent variable when explaining the dependent variable. The market model is described as a single factor model which assumes asset returns are explained by the return on the market portfolio.

When the coefficient of market returns β_i is one and α_i is zero, the market model becomes the market-adjusted model.

$$E(R_{i,t}) = R_{m,t} + \epsilon_{i,t} \tag{2.3}$$

This simple model assumes the expected return of an individual stock can be forecasted by the market return over a period of time. That is, the underlying assumption of this model is that both the individual stock and the whole market share the same systematic risk.

2.3.2.3 Capital asset pricing model (CAPM) and Arbitrage pricing model (APT)

The capital asset pricing model (CAPM) proposed by Sharpe (1964) and Lintner (1965) based on Markowitz's portfolio theory (1959), is an additional stream to measure normal returns. CAPM takes expected excess return of an individual stock as a function of the excess return of market portfolios,

$$E(R_{i,t}) = R_{f,t} + \beta_i (R_{m,t} - R_{f,t}) \tag{2.4}$$

where $R_{f,t}$ is the risk-free rate which is generally measured as a 3-month Treasury bill while β_i is the systematic risk of the firm that measures the stock's reactions to movements of the entire market. It initially describes the linear relationship between systematic risk and return for individual assets. The CAPM is an example of an equilibrium model which suggests if a stock diverges from its equilibrium price, it will eventually get back to the equilibrium price via the market mechanism. However, this method is controversial for its rigid assumptions, such as homogenous expectation of all investors, no frictions to trading and unlimited short-selling. Once these assumptions are violated, the results could deviate. Moreover, beta is found to be statistically significant with small value, especially for small firms (Lintner 1965). The main reasons stems not only from measurement errors but also skewness. The correlation of beta and unsystematic risk is also under discussion (Fama and Macbeth, 1973). Furthermore, the linear relationship between stock return and beta is questioned (Black et al., 1972). It is worth mentioning that Roll (1977) argues that CAPM is designed to test the efficiency of the market portfolio. Empirical studies of anomalies address the imperfect explanation of expected return implied in CAPM (Basu, 1977; Banz, 1981; and Fama and French, 1992). Other factors which could affect the expected return of a security include size, book-to-market ratio, earning-to-price ratio, cash flow-to-price ratio, sales growth, (Banz 1981, Stattman 1980, Lakonishock et al., 1994). Therefore, the arbitrage pricing model (APT) introduced by Ross (1976) with the assumption of no arbitrage opportunities deals with the previous issues. It bears an underlying assumption that undertaking infinitely large positions can exploit any perceived mispricing, causing asset prices to adjust immediately to their equilibrium values. It illustrates the relationship of equilibrium between expected returns for well-diversified portfolios and their multiple sources of systematic risk,

$$E(R_{i,t}) = R_{f,t} + \beta_{1,t}\vartheta_{1,t} + \beta_{2,t}\vartheta_{1,t} + \cdots + \beta_{n,t}\vartheta_{n,t} \tag{2.5}$$

where ϑ stands for the expected risk premium associated with each risk factor. $\vartheta_{n,t}$ equals the risk premium for a portfolio with factor sensitivity equal to 1 to factor n and factor sensitivity equal to zero for the remaining factors. CAPM is considered as a special restrictive case of APT in which has only one risk factor, and that one factor is restricted as the only market risk factor. The APT approach confirms more than one factor is attributed to stock return but it is difficult to identify which specific factors are appropriate and the number of factors involved (Friend and Gultekin, 1984).

2.3.2.4 Multifactor models: The Fama-French three-factor model and Carhart four-factor model

A multifactor model which incorporates more than one factor affecting stock returns is a popular application in both event-time and calendar-time studies. The factors are portfolios that explain the variance in asset returns. Typical examples are the Fama-French three-factor model and Carhart four-factor model.

The Fama-French three-factor model

Fama and French (1993) assume market risk premium, size premium and book-to-market premium capture the key drivers of a firm's return,

$$E(R_{i,t}) = R_{f,t} + \beta_{1,t}(R_{m,t} - R_{f,t}) + \beta_{2,t}SML_t + \beta_{3,t}HML_t \quad (2.6)$$

where SML is the size factor which represents the difference in returns between small size portfolios and large size portfolios, while HML is the difference between returns of portfolios with high book-to-market ratio and portfolios with low book-to-market ratio. $R_{m,t}$ is the value-weighted market portfolio return. The coefficients are estimated by regression over the estimation period. Fama and French (1993) claim the three-factor model to be an equilibrium model if investors take size and book-to-market ratios into consideration when making investment decisions. Furthermore, Fama (1998) documents the consistency of the three-factor model when comparing it with the APT model. However, Loughran and Ritter (2000) argue the three-factor model is not an equilibrium model since it only detects anomalies in financial markets and fails to test market efficiency. In particular, small firms with lowest book-to-market ratios represent significant anomalies which cannot be fully explained by the three factors. This argument is further tested and advocated by Mitchell and Stafford (2000). Therefore, an increasing number of papers start to focus on improvement of the model. Although additional factors including macroeconomic factors are introduced, small size effect is still an open question (Eckbo et al., 2000).

The Carhart four-factor model

To resolve Jegadeesh and Titman's (1993) proposal regarding the issue of returns momentum proposed, Carhart (1997) expands the three-factor model to a four-factor model by introducing a fourth factor which captures the momentum anomalies,

$$E(R_{i,t}) = R_{f,t} + \beta_{1,t}(R_{m,t} - R_{f,t}) + \beta_{2,t}SML_t + \beta_{3,t}HML_t + \beta_{4,t}UMD_t \quad (2.7)$$

where UMD_t is the difference between returns of winners and losers. However, this fails to explain the anomalies of firms with small market values. The factor models remain in the spot light over the past twenty years even with the argument of the limited range of reduction in variance of stock returns when introducing more factors (MacKinlay, 1997).

2.3.2.5 Reference portfolios

Apart from using models which predict normal returns, reference portfolios with non-event firms that have similar characteristics to event firms, for example, industry, size, book-to-market ratio, market index, is an alternative. Researchers believe that matching can control the cross-sectional differences in normal returns based on the assumption of similar riskiness for both an event firm and a control firm. That is to say, in the case of an economic shock, firms with similar characteristics are generally affected in a similar way. For instance, if there is an announcement of stricter carbon emission requirements for airplanes, the stock price airline companies in the same industry will potentially have negative performance due to increased operating costs. However, it could be the case that when a shock arises from an event, the returns of two firms will be affected in different ways. In that case, it is biased to apply this approach to measure abnormal performance since the fundamentals of the two firms are different. Matching techniques for reference portfolios can be extended from size to market index. Lyon and Barber (1997) construct reference portfolios based on size, book-to-market ratio, size/book-to-market ratio and equally-weighted market index. The majority of their empirical results based on simulation show positive skewness of abnormal returns. When compared with the control firm approach, reference portfolios are not superior. Moreover, the cross-sectional abnormal returns still exist when this approach is applied. Ang and Zhang (2004) further test long-term abnormal stock returns in a simulation process based on reference portfolios with the incorporation of nonparametric tests. They provide "strong evidence that reference portfolio benchmarks overestimate event firms' long-term return".

2.3.2.6 Control-firm approach

The control firm approach applies similar rules to reference portfolios. It matches event firms with one or multiple control firms according to similar characteristics. Lyon el al. (1999) find the control firm approach together with bootstrapped test statistic is better specified than other methods, whereas Ang and Zhang (2004) advocate that a single control firm approach with sign test produces higher power. However, the control firm approach is subject to debate since the repeatedly used portfolio of control firms violates the assumption of independent returns. If each control firm is restricted to be selected only once, this could result in a drop-out of sample firms due to an absence of matching firms (Spiess and Affleck-Graves, 1995, Mitchell and Stafford, 2000). Moreover, the application of the control firm approach is subject to criticism based on small sample and strict matching criteria. In some cases, it may be difficult to identify a control firm for the event firm. In empirical studies, the reference portfolio approach has been under attack for its matching criteria. Banz (1981) casts doubt on matching a technique based on size only by documenting small firm effect with higher returns than big firms. Furthermore, Fama and French (1992) add one more matching criteria, book-to-market ratio, as one of the significant characteristics to capture the stock return. Different matching criteria generate different results. Empirical studies undertaken by Brav and Gromper (1997) show no abnormal performance in post-IPO period, even when the benchmark is related to the basis size and book-to-market ratio. This viewpoint is contradicted by Loughran and Ritter (1995), when examining the existence of abnormal performance based only on size.

In conclusion, asset pricing models, reference portfolios or the control firm approach to predict normal return offer both advantages and disadvantages. The discussion regarding which is superior has been continuing since the last century. The main argument for the models approach is summarized by Fama (1970). The studies on stock performance pre- or post-event not only test the market efficiency but also test the specification of an equilibrium pricing model. This is defined as a bad-model issue. Fama (1998) recommends solutions for two distinct types of bad-model problems: the imperfect model and choice of estimation period. The market model (Fama et al., 1969) and comparison period approach (Masulis, 1980), with the estimation period out-of-sample in order to obtain reasonable expected returns, are suggested because no constraints are placed on the cross-section of expected returns. However, it is difficult to define ‘normal estimation period’, as some firms have unusual anomalies before the events. If avoiding the period selection, the cross-section of expected returns will affect the justification of three-factor models by imperfectly capturing normal returns. Even when applying the control firm approach, the bad-model issue cannot be solved entirely due to additional factors such as small pool of matching firms (Spiess and Affleck-Graves, 1995). Especially when conducting long-run event studies, the bad model problem compounds to generate significant errors. There is still no agreement regarding which model perfectly describes stock returns. Fama (1998) summarizes as follows: “bad-model problems are unavoidable, and they are more serious in tests on long-run returns”.

2.3.3 Measurement of abnormal returns

The abnormal return is simply the difference between actual stock returns and normal returns. Due to the different definitions of normal return as previously discussed, results could vary.

$$\mathrm{AR}_{i,t} = \mathrm{R}_{i,t} - \mathrm{E}(\mathrm{R}_{i,t}) \qquad (2.8)$$

Once abnormal returns for asset i in specific dates or months is identified, the accumulation method needs to be carefully chosen since researchers are more interested in the performance of the stock or a portfolio of stocks over a period. Abnormal returns can be accumulated either by stocks or time.

2.3.3.1 Event-time approach: Cumulative abnormal returns vs. Buy-and-hold abnormal returns

Two major return metrics commonly applied in research are cumulative abnormal return (CARs), and buy-and-hold return (BHARs). CARs cumulate abnormal returns over a period by simply summing up abnormal returns. BHARs firstly calculate compounding returns of stock returns and benchmark returns individually and then take the difference,

$$\mathrm{CAR}_i(t_1, t_2) = \sum_{t=t_1}^{t_2} \mathrm{AR}_{i,t_n} \qquad (2.10)$$

$$\mathrm{BHAR}_i(t_1, t_2) = \prod_{t_1}^{t_2}(1 + \mathrm{R}_{i,t_n}) - \prod_{t_1}^{t_2}(1 + \mathrm{R}_{b,t_n}) \qquad (2.11)$$

where R_{b,t_n} is the benchmark return which is the normal returns predicted from models or reference portfolios. CARs is under criticism as it ignores the fact that investors normally buy a stock and hold it for a certain time rather than instantly sell it. Buy-and-hold returns correct CARs’ weakness by assuming investors hold the same portfolio over the investment period. Therefore, BHARs is mostly employed in long-run event studies by compounding short-term abnormal returns over the holding period. However, Fama (1998) questions BHARs by arguing that asset pricing models do not explain which interval is more appropriate to estimate expected returns. He summarizes three arguments which advocate shorter intervals in event studies. The first is that the distribution of returns tends to be more normally distributed in short-run compared with long run. Secondly, monthly returns are generally applied to test market efficiency. Lastly, BHARs maintain a compounding effect which disguises the actual speed of price adjustment (Mitchell and Stafford, 1997).

It is still controversial regarding which return metric is superior in empirical studies due to potentially biased statistical inferences. CARs is considered to have fewer statistical problems than BHARs, thus magnifying the "bad model" issue suggested by Fama (1998). Ritter (1991) initially proposes the application of both BHARs and CARs in event studies. Kothari and Warner (1997) compare tests for abnormal returns over a period of time through simulation with the application of CAR and BHAR methods. Both methods generate misspecified test statistics but BHARs shows higher rejection rates. Although their results indicate fewer problems when applying CARs, Lyon and Barber (1997) prefer BHARs to CARs due to biased estimators obtained by regressing BHARs on CARs in one year. They underline three biases for both BHARs and CARs, respectively. Regarding CARs, the first is measurement bias, based on the fact that CARs are a biased predictor of BHARs when running the regression of BHARs on CARs over one year. The second bias for CARs is new listing bias which suggests incorporation of new firms in reference portfolios. Since Ritter (1991) identifies the underperformance of new listed stocks, the abnormal returns are overestimated if new stocks are included in the benchmark. This results in positively biased statistical inference. However, Brav et al. (2000) argue that there is a marginal change in the intercepts and R square of the model when new firms are excluded. The last basis for CARs is skewness bias derived from skewed returns of individual stocks and non-skewed returns of reference portfolios. As Mitchell and Stafford document in a seminal article in 2000, BHARs tend to be more positively skewed. However, when determining if the null hypothesis of zero abnormal return is rejected, the abnormal returns are assumed to follow normal distribution. This leads to negative bias in test statistics. Although they point out that the bias reduces when fewer firms are involved, the assumption of normal distribution works better with large samples. Lyon, Barber and Tsai (1999) propose three tests to overcome the issue of skewness. One is the skewness-adjusted test proposed by Johnson (1978). The other two aiming at constructing empirical distribution of abnormal returns are the bootstrapped skewness-adjusted test which randomly selects subsamples with replacement to achieve critical values (Lyon, Barber and Tsai, 1999) and empirical p value from pseudoportfolios built up with firms having similar characteristics to event firms (Brock et al., 1992, Ikenberry et al., 1995). The major difference between the bootstrapped procedure and empirical distribution is the null hypothesis. The bootstrapped procedure tests if the mean abnormal return is zero, whereas empirical distribution tests if the mean abnormal return is the same as the mean abnormal return for n pseudoportfolios. With respect to BHARs, apart from new listing bias and skewness bias, it is subject to rebalancing bias because investors are assumed to hold the event securities over a period of time; however, the composition of reference portfolio varies with new-listing and delisting stocks. Lyon and Barber (1997) suggest the control firm approach, which matches event firms with control firms by similar characteristics such as size, book-to-market ratio, is appropriate to eliminate new listing bias, rebalancing bias and skewness bias. However, the authors shed light on the lower power of test in the control firm approach as the variance of difference between returns of the event firm and control firm is higher than that of the difference between returns of the event firm and reference portfolios. This could be resolved when the sample size is large. Fama (1998) also outlines the compounding affect of BHARs that suggests stock returns increase with the length of time horizon. Brav (1997) further tests the inference based on BHARs and argues the cross-sectional returns accumulate errors through compounding especially over a long time horizon. Although he attempts to establish corrected BHARs incorporating the correlation of returns, there is no clear solution since the number of covariance is larger than the number of returns. Mitchell and Stafford (2000) add one more weakness regarding BHARs. They argue that there is overrejection of the null hypothesis of zero BHARs when firms have cross-sectional dependent abnormal returns. The CAR method is also subject to criticism regarding spurious bias..

2.3.3.2 Calendar-time approach: Average abnormal returns

An average abnormal return (AAR), which gives equal weight to individual stocks, aggregates abnormal returns across firms at time t and takes the average. It is used to test the null hypothesis of zero mean cross-sectional abnormal returns. The calendar-time portfolio approach is one variation of AARs. For N stocks, the AAR is calculated as follows:

$$\mathrm{AAR_t} = \frac{\sum_{i=1}^{N} \mathrm{AR_{i,t}}}{\mathrm{N}} \tag{2.9}$$

This is normally used to test the null hypothesis of zero single or multiple stock abnormal returns at time t. It is commonly applied in short-run event studies for individual stocks. Aggregation by time is to test if the mean abnormal return over a period of time is zero.

The calendar-time portfolio approach (Jensen-alpha approach), another measurement of long-run abnormal returns, is proposed to deal with issues of cross-sectional dependence of returns (Loughran and Ritter, 1995, Kothari and Warner, 2008). It is a variation of AARs introduced by Jaffe (1974) and Mandelker (1974). It considers the cross-sectional returns across securities, whereas BHARs assumes independence of stock returns. Take an event window of five years as an example. For each calendar month, a portfolio of firms which experience events in the last five years is constructed on an equal- or value-weighted base. Then the Fama-French three-factor model or Carhart four-factor model are applied to derive the intercept which indicates the average abnormal return of the event portfolio. The null hypothesis is the intercept; this is a proxy of abnormal returns, which is from the regression is zero. The heteroskedasticity problem could be resolved by applying the weighted least squares technique instead of ordinary least squares (Fama, 1998; Ang and Zhang, 2004). Loughran and Ritter (2000) criticize the fact that the calendar-time approach underestimates abnormal returns when events are clustering in a month. They are against applying this approach because of its underestimation of the timing of managerial decisions. Moreover, the lower power of test to detect long-run abnormal returns under this method is concluded. But Mitchell and Stafford (2000) debate that abnormal performance hardly changes when the number of issuances increases.

2.3.4 Test statistics

The two commonly applied tests in event studies are parametric tests which support the assumption of normal distribution and nonparametric tests being distribution free. No agreement is reached regarding which test is superior. Whereas some research indicates parametric tests are well-specified when testing abnormal performance of stocks and have higher power compared with nonparametric tests (Brown and Warner, 1985; Jain, 1986), other studies highlight the advantages of nonparametric tests such as loose assumption of distribution of abnormal returns.

2.3.4.1 Parametric tests

Regarding parametric tests, the most widely applied is the Patell t-test, introduced by Patell (1976) and advocated by many subsequent studies such as Dodd and Warner (1983). This is a standardized test which aims at constant variance of abnormal returns. To summarize the major issues concerning properties of abnormal returns when applying the Patell test, Binder (1998) concludes “the abnormal return estimators 1) are cross-sectionally (in event time) correlated, 2) have different variance across firms, 3) are not independent across time for a given firm or 4) have greater variance during the event period than in the surrounding periods”. Brown and Warner (1980, 1985) document cross-sectional abnormal returns and propose two versions of test statistics. One is the time-series standard deviation test which is a crude adjustment of standard deviation based on a single variance estimate of the time series in the estimation period rather than the event period. The other is the cross-sectional standard deviation test; as indicated by its name, this employs the portfolio time-series standard deviation in test statistics. The violation of constant variances of abnormal returns is documented in literature extensively (Beaver, 1968; Fama, 1976; Collins and Dent, 1984; Bernard, 1987). To deal with this issue, Boehmer, Musumeci and Poulsen (1991) suggest applying cross-sectional prediction errors on the event date with an assumption of proportionate of variance on the event date to variance over the estimation period. A further solution which also corrects the autocorrelation of abnormal returns is the calendar-time approach introduced by Jaffe (1974) and Mandelker (1974). This calculates mean portfolio returns based on calendar month. The autocorrelation of abnormal returns of stocks due to the same estimates has limited impact on statistical inference when the event period is short (Karafiath and Spencer, 1991; Cowan, 1993). Mikkelson and Partch (1988) develop corrected test statistics by using autocorrelation and credit this to Craig Ansley.

As mentioned above, the Patell test is assumed to have student-t distribution with zero mean and constant variance in N degrees of freedom. Although this approach is popular in short-term and long-run event studies (Brown and Warner, 1985; Corrado, 1989; Haw, Pastena and Lilien, 1990; Kothari and Warner, 1997; Lyon and Barber, 1997), it also suffers criticism for its unrealistic assumption of normal distribution. Most studies show evidence of fat tail and right skewness of abnormal returns (Lyon and Barber, 1997). This results in higher rejection in the upper-tail test and lower rejection in lower-tail test (Cowan and Sergeant, 1996). The test is misspecified together with low power in this circumstance (Maynes and Rumsey, 1993). Hall (1992) and Sutton (1993) develop a transformed skewness-adjusted test based on the original version introduced by Johnson (1978). Although it does not totally resolve the issue, it inspires the following research to focus on non-parametric tests which release the assumption of normal distribution.

2.3.4.2 Non-parametric tests

The popular non-parametric tests include sign test, rank test, signed-rank test, bootstrapped skewness-adjusted test and pseudoportfolios approach.

Sign test

Fisher advocates sign test, a simple binomial test, in a seminal paper in 1921. Brown and Warner (1985) utilize the test with null hypothesis of median of abnormal returns equals to zero in short-term event studies and document misspecification. However, this version of the sign test approximates a normal distribution when the sample is larger than 25. It violates the idea of relaxing the assumption of normal distribution. Therefore, Cowan (1992) formulates a generalized version of the sign test on the basis of models developed by McConnell and Muscarella (1985). Although this test recognizes the real distribution of abnormal returns, it requires abnormal returns in both estimation period and event period. The null hypothesis is that the number of stocks with positive CARs in the estimation period is equal to the number of stocks with positive CARs in the event period. This test is proved to be well-specified and powerful by Cowan (1992) with the application of simulation on NYSE-AMEX stocks over the period of 1972 to 1990. This result is confirmed by the following study based on the

Asia-Pacific stock market by Corrado and Truong (2008). When the distribution of abnormal returns is not normally distributed, the sign test produces biased inferences.

Rank test

The rank test, introduced by Corrado (1989), is claimed to have higher power to detect abnormal returns than parametric tests (Corrado and Zivney, 1992, Cowan, 1992, Campbell and Wasley, 1993). Corrado (1989) employs simulation of the rank test to test abnormal returns with return data of stocks listed on the NYSE and AMEX in a one-day event window in order to resolve the issue of non-normal distribution of abnormal returns. He firstly ranks abnormal returns of individual firms over both the estimation and event periods and allocates ranks to each abnormal return. Rank one means the smallest abnormal return. Then, he calculates the mean rank over both periods. The rank test statistics over the event window is computed to test the null hypothesis of zero abnormal returns. Cowan (1992) and Campbell and Wasley (1993) extend the rank test to test CARs by summing up with an assumption of independent stock returns ranks. This test is subject to criticism over the lower power of test when large abnormal returns scatter over the study period (Cowan, 1992), since the rank test does not consider the magnitude of abnormal returns.

Bootstrap approach

Efron and Tibshirani (1993) detail the theory of bootstrap to achieve accurate statistics applied in prior studies. The bootstrapped process is applied initially in event studies by Marais (1984). The author conducts a bootstrapping process with four steps on residuals from market models based on returns of a single firm to resolve the problem of non-normalities of residuals. Note that a Monte Carlo simulation is run in Marais' (1984) study rather than historical simulation developed by Butler and Frost (1992). A covariance matrix of the residuals is incorporated into the bootstrapping process based on historical simulation to test the null hypothesis of zero abnormal returns by Chou (2004). Kramer (2001) develops a bootstrapping process to deal with small samples with test statistics defined as the difference between a student t-test and the mean t-test of N firms. Moreover, Hein and Westfall (2004) extend the bootstrapping by taking into account autocorrelation. To construct an empirical distribution together with adjustment of right-skewness, Lyon, Barber and Tsai (1999) transform the original version into the bootstrapped skewness-adjusted test. It bootstraps 1000 times with subsamples of a quarter of the original sample size. The skewness-adjusted tests are computed based on these subsamples. It is worth noting that the conventional test statistics involved in the calculation have the null hypothesis of the equality of mean abnormal return of subsample and mean abnormal return of total sample. With 1000 bootstrapped skewness-adjusted t tests, an empirical distribution is established to decide critical values at different significance levels. The critical values are used to compare with the skewness-adjusted tests of the original sample. As summarized by Serra (2002), the bootstrapping procedure is generalized into three steps.

1. Determine N/division factor. This is denoted as the size of the subsample. The division factors mostly applied in empirical studies are 2 or 4 which means the number of observations in a subsample is half or a quarter of that of the original sample.
2. N/division observations are randomly selected from the original sample with replacement n times to form a series of subsamples. A test statistic is computed for each subsample. An empirical distribution with expected value and standard deviation is constructed from a series of n test statistics from the simulation procedure. This empirical distribution is approximated more closely to the real distribution of the sample I study, compared with the assumption of normal distribution.
3. Critical values and confidence intervals for different levels are achieved by ranking n test statistics from smallest to largest and taking percentiles according to different significance levels. The rejection of the null hypothesis at a significance level of α or with a confidence level of $1-\alpha$ requires test statistics from the original sample to compare whether it is either less than $\alpha/2$ or greater than $(1-\alpha/2)$ of the percentile; these are the critical values of simulated test statistics.

Empirical p value from pseudoportfolios

Empirical p value from pseudoportfolios, firstly applied by Brock et al. (1992), follows the bootstrapping process proposed by Efron (1982). This approach aims at establishing an empirical distribution of observed variables in order to deal with the issue of non-normal distribution. It is widely employed in empirical studies by Ikenberry et al. (1995), Ikenberry, Rankine and Stice (1996), Rau and Vermaelen (1996), Lee (1997) and Lyon, Barber and Tsai (1999). The procedures applied in event studies to detect abnormal returns can be summarized as follows (Lyon, Barber and Tsai, 1999):

1. Identify the null hypothesis. The null hypothesis is the mean return of sample firms is equal to the mean return from n pseudoportfolios.
2. For each sample, each event firm is matched with a pseudo-firm randomly selected with a replacement from reference portfolios which contains stock with similar characteristics to event firms. This is actually a bootstrapping process. Eventually, the sample will be filled with pseudo-firms for further study.
3. The mean abnormal return based on buy-and-hold returns or cumulative returns is computed.
4. Steps 2 and 3 are repeated 1000 times to obtain 1000 mean abnormal returns. These returns are used to construct an empirical distribution of mean abnormal returns in order to identify critical values for determination of whether the null hypothesis is rejected. The rejection of the null hypothesis at a significance level of α or with a confidence level of $1-\alpha$, requires the mean abnormal returns from the original sample to compare a t either less than $\alpha/2$ or greater than $(1-\alpha/2)$ of the percentile, which are the critical values of simulated results.

2.4 Literature on long-run event studies based on event-time approach

Most event studies focus on the traditional event-time approach which calculates the short-run and long-run stock performance based on event time. For individual firms, event studies can be applied with specific event time or simulated uncertain event time. By constructing a portfolio with a number of firms with different event dates, the event-time studies generally assumes all the event dates of firms as time 0, putting all the firms on the same starting point rather than considering individual event time separately. Asset pricing models, reference portfolios and control firm approaches are three popular applications for long-run stock performance. Long-run event-time studies with time horizon spanning from one to five years have attracted academics' attention since the 1990s with the finding of underperformance of IPOs as documented by Ritter (1991). With the conclusion of misspecified models and statistics drawn by Kothari & Warner (1997), asset pricing models applied to estimate normal returns seem to fade out gradually in the framework of long-run event studies. However, the reference portfolios and control firm approach have evoked more interest in recent years (Lyon, Barber and Tsai, 1999, Ang and Zhang, 2004). There is still no consensus regarding which approach is more superior to detect long-term stock performance.

Ritter (1991) firstly uses measurements of CARs and BHARs to investigate performance of 1,526 IPOs in the US market during from 1975 to 1984 over three years which exclude the event month. Firms delisted over the investment horizons are only tracked until the last day they list in the market. Benchmark returns which compare with actual IPOs' returns to obtain abnormal returns are the CRSP value-weighted NASDAQ index, the CRSP value-weighted Amex-NYSE index, listed firms matched by industry and size and an index of the smallest size decile of the New York Stock Exchange. The average raw returns start at positive returns of 14.3% and maintain an upward trend over three years. CARs cumulating average matching firm-adjusted returns in the first four months are positive with an upward trend until the third month, but start decreasing to negative figures from the fifth month until the 36th month. Regarding BHARs, mean returns of IPOs and matching firms based on CRSP NASDAQ and Amex-NYSE indices are 34.47% and 61.86% respectively over a three-year holding period. Moreover, the distribution of returns of IPOs seems more right-skewed than matching firms. Regarding matching firms by industry, most industries show long-term underperformance of IPOs except for financial institutions, drugs firms and airlines. On average, there is an underperformance of IPOs over three years when the matching criteria is industry wise. The author also looks at aftermarket performance by year of issuance from 1975 to 1984. On average, the wealth relative is 0.831 indicating an underperformance of IPOs. However, only half of the years exhibit underperformance while the remaining years show outperformance. Furthermore, when grouping IPOs in term of ages, that is the difference between the year of offer and the year of founding, underperformance is documented for sample of all firms with an age of less than 20 years. Outperformance is shown for all firms with an age of more than

20 years, excluding oil and gas firms and financial institutions. The author raises three unresolved issues: the persistence of underperformance in longer investment horizons, the generality of underperformance of IPOs and the relationship between short-term underpricing and long-term underperformance. The following studies by Levis (1993), Loughran and Ritter (1995), Rajan and Servaes (1997), Brav and Gomper (1997) and Espenlaub et al. (1998) present findings similar to Ritter's (1991) regarding IPO performance in the long run. Ritter's initiation of long-term stock performance evokes interests for academics to explore other events apart from IPOs. [4]

Instead of examining specific events, Kothari and Warner (1997) randomly select 250 samples of 200 firms with different event months between 1980 and 1989. The company included in the sample has to meet the requirements of at least 24 consistent returns before the event date. The test periods studied are one, two and three years. In order to control the survival bias, long-term performance is tracked until the data is not available within the test period. Four models, the market-adjusted model, market model, capital asset pricing model, and Fama-French three-factor model are applied to test abnormal performance. The cross-sectional average abnormal returns are tested for statistical significance with the assumption of zero abnormal returns. Both CARs and BHARs are applied in this study. If the model is well specified, the rejection rate of 250 samples for the null hypothesis of zero abnormal performance is expected to be 5% at the 5% significance level and 1% at the 1% significance level. Additionally, abnormal returns are expected to be normally distributed with zero mean and constant variance. When CARs is employed, all four models tend to be misspecified in different time horizons with an uptrend rejection rate when the time horizon is increasing at both 1% and 5% significance levels. The market model seems to be the worst misspecified with the highest rejection rate of 35.2%; the market-adjusted model is the less misspecified with the lowest rejection rate of 18.4% over three years at 5% significance level among four models. Moreover, all these models show higher rejection rates when the

[4] Seasoned equity offering: Loughran and Ritter (1995),Lee (1997),Jegadeesh (2000),Loughran and Ritter (2000), D'Mello, Tawatnuntacha, and Yaman (2002)
Mergers and Acquisitions: Agrawal, Jaffe and Mandelker (1992), Anderson and Mandelker(1993),Gregory (1997),Loughran and Vijh (1997),Moeller et al. (2003),Kooli and L'her (2004)

alternative hypothesis is positive abnormal performance, which indicates positive skewness of abnormal returns. The magnitude of abnormal returns measured by four models increases with the length of investment periods and violates the null hypothesis of zero CARs. The FF model captures the highest positive CAR of 3.91% while CAPM has the lowest CAR of 3.32% in three years. A similar trend is documented for cross-sectional standard deviation of test statistics for four models, suggesting that the overrejection could be caused by a small standard error. Apart from CARs, the authors also apply buy-and-hold abnormal returns in the process of simulation. Similar results to CARs are found with the exception of the market model with outstanding estimates of a 76.8% rejection rate over three years. With respect to the magnitude of abnormal returns over 36 months, BHARs tend to have higher positive means together with more significant right-skewness than CARs.

Kothari and Warner (1997) also attempt to discuss the possible reasons for the misspecification. Pre-event consistent returns requirement which includes newly listed firms with poorer returns and drop-out firms with low standard deviation over the estimation period are studied. After considering the biases, the adjusted test indicates that negative performance is rejected too infrequently with the rejection rate of 0.4% for CAPM. Furthermore, the authors undertake robustness checks with non-random samples consisting of firms classified on the basis of book-to-market ratio and size. The study period is traced further back to 1970 rather than 1980 and the low book-to-market firms are assumed to be those with a ratio of 0.8 or less. Low book-to-market firms show negative abnormal returns and less misspecified test statistics than random samples except for the market model over 36 months. The severe misspecification of the market model shows a rejection rate of 96.8% for low book-to-market samples. It is worth mentioning that a more frequent negative abnormal performance is documented for all four models. However, for high book-to-market samples, the positive mean abnormal performance with the highest return of 25.6% in the market model, much greater than the random samples with abnormal return of 3.66%, is consistent with the previous findings of the small firm effect (Brav and Gompers, 1997). In this case, the null of negative CARs is rejected with a much higher rejection rate; furthermore, all four models are more severely misspecified when compared with random samples. Other tests using different characteristics to match firms, including controlling firms for their size and book-to-market ratio, reach similar conclusions of misspecification with previous approaches in the article. The authors propose two approaches to deal with the misspecification. One is to apply nonparametric procedures, for example, bootstrap procedures outlined by Ikenberry, Lakonishok and Vermaelen (1995). This means drawing random samples to estimate the actual sample returns without the assumption of distribution of returns. The problem regarding this approach is that the firms used to construct the distribution need to be matched with sample firms not only by characteristics but also bias. The other approach is calendar time portfolios which could resolve the issue of event time clustering and cross-sectional problems. But this method is still contentious when considering behavioural timing.

Apart from the application of the asset pricing model for abnormal returns, Lyon and Barber (1997) further extend the long-run stock performance studies on the basis of the application of the reference portfolio approach and control firm approach in terms of different benchmarks in the simulation process. They firstly discuss four possible causes for the misspecifications in test statistics of CARs and BHARs in event-time studies. To begin with, they investigate the measurement of abnormal returns, including BHARs and CARs. A sample of 10,000 monthly returns from 1963 to 1993 is randomly selected from NASDAQ and NYSE/AMEX. The observations are compared with a benchmark of an equally weighted market index. Then 100 portfolios of 100 firms are constructed according to the ranking of annual BHARs. For each portfolio, the mean difference between CARs and BHARs is calculated. It is apparent that the difference decreases when the annual BHAR gets closer to 28% but becomes negative and decreases once the annual BHAR exceed 28%. This indicates a case of larger annual BHAR when comparing this with CAR. It could be the result of a compounding problem. In order to find out which approach is better, the authors regress the annual BHAR on CAR by randomly drawing 20,000 returns using an equally weighted market index as a benchmark. Statistically significance from zero and one for intercept and slope coefficients imply that CARs are a biased predictor of BHARs. This is identified as ‘measurement bias’ for CARs. Secondly, both CARs and BHARs suffer ‘new listing bias’ which incorporates the effect of newly listed firms after the event month. Ritter (1991) shows the long-run underperformance of IPOs after listing. This suggests that the reference portfolio as a benchmark has a lower return than a sample firm due to inclusion of new listed firms in the benchmark but not in the sample when random selection occurs. This eventually leads to positively biased abnormal returns. Thirdly, 'skewness bias' is documented for both measurements, especially BHARs. The mean and median of BHARs are -0.48% and -7.23%, respectively. Regarding CARs, these figures are 0.82% and -0.99%. BHARs tend to have more severe positive skewness issues than CARs due to the compounding measurement. This bias has a negative impact on test statistics when the sample mean is positive, and vice versa. Additionally, rebalancing bias exists in BHARs. If the equally-weighted benchmark is applied, it needs to balance the portfolio every month by selling securities outperforming the market and buying those underperforming the market. Eventually this will lead to negative bias in BHARs. In conclusion, both CARs and BHARs suffer new listing bias and skewness bias. But CARs is also subject to measurement bias while BHARs is subject to rebalancing bias.

To detect the long-term abnormal returns, Lyon, Barber and Tsai (1999) screen the data from CRSP NYSE/AMEX/NASDAQ by excluding firms which do not have ordinary common stock and book-to-market ratio in year t-1 together with those which do not have size in year t and book-to-market ratio in year t-1. They then establish reference portfolios of benchmarks for abnormal returns based on a range of benchmarks including size, book-to-market, both criteria and an equally weighted market index. However, reference portfolios are still subject to different biases previously discussed when choosing BHARs or CARs. So the authors propose an approach of control firms which match the benchmark portfolios with sample firms by specific firm characteristics of size or book-to-market ratio. 200 event dates are randomly selected without replacement firstly, then 200 firms are randomly selected to match with event dates.[5] The process continues 1000 times to achieve 1000 samples of 200 firms. Both CARs and BHARs are examined over one, three and five holding periods. Regarding CARs, the four reference portfolio approaches all show positive bias, while the equally weighted market index has the most pronounced bias compared with others. The rejection rate

[5] The simulation process without replacement of n event months is also conducted in studies by Barber, Lyon and Tsai (1999). Lyon and Barber (1997) claim that "results are robust to sampling with replacement".

increases with the increase in time horizons. Well-specified test statistics are found in the control approach with approximate zero mean in all the time horizons. It is worth noting that the power of test is undertaken by introducing abnormal returns from -20% to 20% at an interval of 5% to each sample firm; this reveals a more powerful test in the reference portfolio methods. In contrast, BHARs shows negative bias in all test statistics with higher rejection rate on the null hypothesis of zero BHARs and in favour of the alternative hypothesis of significantly negative BHARs. This could be attributed to skewness bias. The control firm approach still shows well-specified t-statistics. Moreover, the introduction of additional returns to individual firms for empirical power of test suggests the control firm approach does not work better than the reference portfolio approach but it does have well-specified t-statistics. Considering the sampling biases, the authors use non-random samples including smallest/largest size deciles and lowest/highest book-to-market deciles. The control firm approach still presents well-specified test statistics in most cases. In summary, test statistics of abnormal returns using reference portfolios are misspecified but the control firm approach outperform all others.

Taking a more comprehensive look at biases which are discussed by Lyon and Barber (1997), the authors attempt to find out how to improve tests. They apply two ways to control the biases by investigating the NYSE/AMEX/Nasdaq monthly returns over the period of 1973 to 1994. One is to combine buy-and-hold measurement and event studies. The first approach of buy-and-hold measurement is to calculate the average return for each portfolio, then compound it over a period. This cannot eliminate the rebalancing bias since the portfolio is assumed to be rebalanced monthly. Neither can the new listing bias. Because the long-run returns with newly listed firms are downwardly biased (Ritter, 1991). Therefore, the authors construct another buy-and-hold measurement to deal with the issues. They firstly compound the returns of each security in the portfolio and then sum them up. In this way, the rebalancing bias and new listing bias can be eliminated as the same securities in the portfolio are assumed to be held over the period. To solve the issue of delisted securities during the period, the missing returns are replaced with the average monthly return of reference portfolios. With regard to skewness bias, the authors propose one parametric test and two nonparametric tests for random samples. These are skewness-adjusted test statistics, bootstrapped skewness-adjusted t-statistic which constructs a distribution of t-statistics and critical values on the basis of the bootstrapped samples (Sutton, 1993), and the empirical p value from pseudoportfolios that generates an empirical distribution of mean long-run abnormal returns (Ikenberry et al, 1995). The measurement of abnormal returns is defined as the difference of BHARs of individual stock and expected return for the security. Two benchmarks are employed based on size and book-to-market ratio. This includes the reference portfolios and control firm approach. To test the specification of test statistics, 1000 samples of 200 firms are randomly selected. The event date is selected first without replacement and then the firm. If the test is well-specified, the null hypothesis of zero mean return will be rejected 50 times among 1000 samples at 5% significance level.

As for random samples, all tests are negatively biased, particularly when rebalanced size and book-to-market portfolios are applied. Among all tests, the bootstrapped skewness-adjusted t-statistics and empirical p value using buy-and-hold portfolios, and conventional t-statistics using the control firm approach are well-specified. The empirical power of these three methods is tested by adding an incremental 5% abnormal return from -20% to 20%. The bootstrapped skewness-adjusted t-statistic and empirical p value yield improved power when compared with the control firm approach. As for non-random samples, the simulation process is carried out in the largest/smallest size deciles, highest/lowest book-to-market, high/low prevent returns and industry-clustering firms. The previous three well-specified t-statistics still outperform when considering the size effect. However, only the control firm approach yields well-specified test statistics among firms with lowest book-to-market ratios, whereas all three methods yield with highest book-to-market ratios. Moreover, they test the persistence of abnormal returns over a long horizon by ranking 6-month returns and forming deciles for simulation. It is found that test statistics of firms with high 6-month pre-event returns are positively biased in one year. However, this trend reverses over three and five years. When taking into account industry clustering, overrejection of the null hypothesis is found in all test statistics. In a word, although not perfect, the bootstrapped skewness adjusted t-statistics and empirical p value appear reasonably well specified. Furthermore, both appear to be more powerful for abnormal returns detection. However, industry clustering and firms with overlapping returns seem to cause severe misspecification problems.

Mitchell and Stafford (2000) question the application of bootstrapped skewness adjusted t-statistics for BHARs; this is also advocated by Lyon et al., (1999). The assumption of independence of abnormal event-firm returns is debatable as in practice the corporate events normally cluster over a specific period by industry. This dependence problem could result in overstated statistical inferences. The authors investigate samples of firms having corporate events such as SEOs, merger and share repurchases from 1958-1993 in the US stock market. To control the cross-sectional dependence, firms with multiple events over three years will be kept when the first event occurs. The abnormal return is defined as the difference between individual event firm buy-and-hold returns and the benchmark buy-and-hold returns excluding the event firm, as in previous studies. Initially the assumption of independence of abnormal returns is applied. Due to positive skewness of BHARs which indicates the null hypothesis is disconfirmed, the authors suggest testing the null hypothesis of zero mean BHARs over three years by using the pseudo-sample approach. The firm with the same characteristics as the event firm, for instance, the same size and book-to-market ratio distribution, is generated for one new mean BHAR. The process is repeated 1000 times in order to establish an empirical distribution to test the null hypothesis. Both value-weighted and equally-weighted schemes are studied.

As for acquirers, the equally-weighted post-event mean BHAR is -0.1 with a p-value of 0.164 which does not reject the null hypothesis, while the value-weighted post-event mean BHAR with a return of -0.038 with a p-value of 0.027 reject the null hypothesis at 5% significance level. Moreover, the growth and value firms expose non-rejection of null zero mean BHAR even at 10% significance level under both weighting schemes. When it comes to a merger financed with or without stock under the equally-weighted scheme, the zero mean BHAR is rejected at 5% significance level with a p-value of 0 and 0.047 respectively. However, under the value-weighted scheme, the merger financed without stock does not reject the null hypothesis. It is interesting to note that both mergers financed with stock show lower average sample returns than the merger financed without stock; however, the mean BHAR reaches the opposite conclusion. For the equally-weighted seasoned equity issuers, the mean BHARs apart from growth firms all show a p-value of zero; this rejects the null hypothesis. But the value-weighted scheme for SEOs only have value firms showing significant non-zero mean BHAR. Value-weighted mean BHARs tend to not reject the null hypothesis for equity repurchase. The results for equally-weighted BHARs are mixed for share repurchases.

Secondly, the authors turn to access the reliability of statistical inference of BHARs. By stimulating firms to match event firm with similar characteristics, the empirical distribution is constructed. The distribution of equally-weighted mean BHAR for SEOs seems symmetric and normally distributed. But both empirical distributions are rejected in the JB test. The bootstrapped distribution and empirical distribution do not exhibit much difference, indicating insignificant influence from residual variance of firms. High probabilities exist that the events for different firms occur closely, particularly according to industry. For instance, when the economy is in good condition, a high percentage of private firms may choose to go public due to the potential higher value of the firm. This creates problems of clustering of abnormal returns, thereby violating the assumption of independence of abnormal returns.

Additionally, this could result in overstated statistical inferences that are mainly caused by the correlation of abnormal returns. All three corporate events show increasing average positive correlations with longer time horizons. For example, the annual correlation of BHARs for mergers is 0.0175 when the monthly figure is 0.002. The authors propose another version of t-statistics with adjustment of cross-sectional dependence. The statistical inference declines significantly from -6.05 to -1.49 for SEOs and from 4.86 to 1.91 for repurchases after the adjustment. When considering the correlation of BHARs, no evidence appears to exist regarding abnormal returns for all three corporate events. Facing the problem of independence, the authors propose two methods to deal with this situation. One is to utilize the covariance structure by considering the cross-sectional dependence; the other is to apply the calendar time approach.

Loughran and Ritter (2000) further investigate the impact of weighting schemes on the measurement of abnormal returns. They argue that if evidence of abnormal returns exists, some models should have superior power to capture this compared to other models. The definition of mean annual return differs from previous studies. They separate three-year abnormal returns into three individual years, then compound each one-year buy-and-hold return to arrive at mean abnormal returns. Attempting to locate the approach which has better specified returns, the authors study the misvalued stocks by intentionally introducing abnormal returns and conducting a simulation procedure. For instance, for undervaluation simulation, a random firm is selected each month and 5% is deducted from its market value. 14% is added in the stock monthly return over three years following the event month; this suggests correction of undervaluation. An additional procedure of adjusting returns is employed when the buy-and-hold market return in the previous one year is less than -10%. Five additional firms are randomly chosen from the population without large firms in the highest quintile of market value. -50% is subtracted from the market value of these five firms, with a correction of 1.94% each month over a three year time horizon. As for overvaluation simulation, 5% is added to the event firm's market value and 14% monthly correction is subtracted from returns. In this case, five firms are selected when the market return is larger than 30% in the previous one year; subsequently, 100% is added to their market value.

The difference of raw returns before and after adjustment of valuation is 16.1% and -14.6% for undervalued firms and overvalued firms respectively when observations are equally weighted. If the methodology is well specified, the mean abnormal returns should be close to the differences of 16.1% and -14.6%. Value-weighted market-adjusted returns show an annual mean return of 22%. On the other hand, the equally-weighted market-adjusted returns have an annual mean return of 6.6% for undervalued firms, highlighting the small company effect which generates higher returns than large firms. Furthermore, size matched reference portfolio mean annual returns are 14.8% and 14.2%, which are close to 16%, before and after accounting for misvaluation. The decreased figures indicate that the size effect does not improve the misspecification. Lastly, the reference portfolio matched with both size and BV/MV ratio has 14% and 11.8% annual returns before and after induced abnormal returns.

The misspecification of declining mean returns could be caused by incorporating the benchmark firms with higher returns. Similar results are found for overvalued firms. In conclusion, the size-matched reference portfolio approach is better specified than the size- and BV/MV matched portfolio approach.

Cowan and Sergeant (2001) argue that the conventional t-test allows for pair wise dependence. But since the research is carried out through simulation, the event firms could cluster around a certain period of time. Apart from this, the practice could result in positive cross-sectional dependence. If the conventional t-test is applied, the inference could be higher with low variance due to ignorance of positive cross-sectional dependence. Therefore, they propose the two-group test with null hypothesis of equality of means of stock returns and benchmark returns to control the dependence on the assumption of independence for two populations. Both portfolio returns and benchmark returns do not correct the pairwise and cross-sectional dependence, hence the test will be more powerful. For the conventional student t-test, similar results are documented based on the equally-weighted scheme as outlined in Lyon et al. (1999). The rejection rate of the null hypothesis of zero abnormal returns decreases with the increase in length of time. Furthermore, the test shows negatively biased figures with higher rejection rates on lower-tail tests. But the rejection rates of lower-tail tests decrease when the size of sample increases, indicating that the sample size effect brings down skewness bias. When the value-weighted scheme is examined, positive skewness is found with higher rejection rates on the upper-tail, except when the sample size is as small as 50 firms. For instance, the rejection rate in favour of alternative hypothesis of positive abnormal returns is 32.7%, while the lower-tail test has a 0.2% rejection rate over five years. The misspecified t-test is more severe to a level of 20.2% rejection rate over five years when the time horizon lengthens. This indicates that the value-weighted scheme generates much higher overrejection rates which result in more misspecified tests compared with other benchmarks.

The control firm approach does not outperform equal-weighted benchmark and it shows positive skewness. It also represents decreasing rejection rates when the time horizon extends as opposed to the results found by Lyon et al. (1999). But when sample size increases from 50 to 1000 firms, it worsens the misspecification with more than 5% rejection rates. When it comes to the two-group test, a similar skewness issue is found to the paired difference test, but the magnitude of rejection rates is smaller. Still, the value-weighted benchmark is the most unfavourable approach, especially when there are 1000 firms in one sample. The control firm approach tends to have a rejection rate close to 5%. It can be concluded that the two-group test corrects misspecification compared with conventional tests. Additionally, the authors firstly suggest winsorization of abnormal returns; these allocate less weight to extreme returns rather than drop them out on purpose to improve the specification of tests. Three standard deviations from mean seem to produce better specified tests than other choices. The means of winsorized abnormal returns on the basis of size and BV/MV matching benchmark show negative results for the equally-weighted scheme and positive results for the value-weighted scheme. But a significant improvement exists regarding right-skewness and leptokurtosis. For example, over a three-year holding period, the equal-weighted mean before winsorization is 0.016 with skewness and kurtosis of 6.772 and 97.39 respectively; after winsorization it is -0.015 with skewness and kurtosis of 3.755 and 41.259. However, the control firm benchmark documents positive mean both before- and after- winsorization with improvement in skewness and kurtosis. Although winsorization contributes close to zero abnormal returns, the distribution of abnormal returns is still asymmetric, especially in small sizes.

Most previous studies only focus on sample sizes of 50 and 200 firms. The authors argue that if the sample size increases, this could possibly result in better specified tests. Therefore, a sample size of 1000 firms is investigated along with a sample size of 25, 50 and 100 firms. The right-skewness and leptokurtosis are still significant for both raw returns and abnormal returns before- and after- winsorization but improve when sample size increases. For instance, the abnormal returns with a reference portfolio using size and BV/MV as benchmark document that skewness and kurtosis are 6.772 and 97.39 when the sample size is 50 firms and 0.517 and 0.911 when the sample size is 1000 firms under an equally-weighted scheme. Instead of monthly data as in the previous research, data employed stems from daily stock returns. The winsorized paired difference tests demonstrate much higher rejection rates in upper-, lower-, and two-tailed tests. However, the rejection rate of winsorized two-group two-sided tests ranges from 3.1% to 7.4% for the value-weighted benchmark and control firm approach. In this case, the equal-weighted benchmark does not perform better. It is worth mentioning that the tests are still conservative with less than 5% rejection rates in most cases.

The population consists of firms from NYSE, AMEX and NASDAQ over the period of 1965 to 1995. Firms without data of market value and book value are excluded. For example, the closed-end funds firms. In contrast to Baber et al. (1999), the simulation procedure is conducted by selecting the firm firstly, then the event date. Moreover, the event firm is randomly chosen in proportion to the number of trading days. 1000 samples with 50, 200 and 1000 firms follow the random selection procedure; they test for one, three and five years. The benchmark construction differs from Lyon et al. (1999). Rather than compounding cross-sectional average returns, they simply average the returns under both weighting schemes over the period in order to eliminate the new-listing and rebalancing bias. With the introduction of abnormal returns over three years, the winsorized two-group test under the value-weighted scheme shows the highest power to capture abnormal returns. To conclude, Cowan and Sergeant (2000) advocate a combination of value-weighted benchmark, winsorized abnormal returns and two-group test to improve misspecification although this is not applicable in some cases. Moreover, the size effect can generate a positive or negative effect on test statistics and cannot be ignored.

Ang and Zhang (2004) update the simulation study on long-run event studies by applying the four-factor model and nonparametric tests. The data applied is ordinary stocks from CRSP from 1980 to 1997. 250 samples of 200 firms are randomly selected with replacement as Kothari and Warner (1997) to test the null hypothesis of zero one-, three- and five-year abnormal returns. In order to examine the effect of different sample sizes, 250 samples of 200 firms follow the same simulation process (Lyon et al., 1999). Initially, the abnormal return is defined as in previous studies by comparing the difference of buy-and-hold returns of individual security and buy-and-hold returns of benchmark buy-and-hold returns. The reference portfolio benchmark is employed in five different ways. The first is the reference portfolio approach, as in Lyon et al. (1999), by matching firms with similar size and BE/ME. An alternative is the control firm approach which only has one single firm in the benchmark portfolio. Another approach incorporates market beta as an additional matching criteria. The market beta is calculated by regressing the control firm 24-month returns on the value-weighted market index. Then the market beta is ranked to categorize three same size sub-portfolios. Fourthly, ten firms, closely correlated to the event firm before the event date, are chosen as the benchmark portfolio. The correlation coefficients of the control firm portfolio returns and event firm returns are investigated. The ten firms with the highest figures are kept to construct the benchmark portfolio. Lastly, the most-correlated firm with the event firm is selected as the benchmark.

By examining the descriptive statistics of five benchmarks, the mean and median returns are close to zero when applying the single firm approach; the magnitude of these returns increases with the other three alternatives with the increasing time horizon. For example, when applying the reference portfolios of firms with similar characteristics, the mean return is 0.034, while the median is -0.25. This generates the overestimation bias. Moreover, the leptokurtic abnormal returns are found with a kurtosis of more than three for all approaches. It is worth noting that no relevant research exists regarding kurtosis adjustment for long-term stock performance. The skewness for the single firm benchmarks is close to zero; in contrast, others are positive and increase when the time horizon increases. Additionally, the standard deviation indicates that volatility of returns is lower for the three reference portfolio benchmarks and increases over the time horizon.

In summary, the three reference portfolio approaches show positive skewness with negative median returns; this indicates that mean abnormal returns are close to zero, while the two single firm approaches present symmetric distribution with leptokurtosis. In terms of test statistics, both parametric and non-parametric tests are conducted. Normal student t-statistics, skewness-adjusted t-test, bootstrapped t-test and sign test are applied. When it comes to two-tailed test, for most benchmarks, the student t-statistics yield other tests. The bootstrapped t-test shows little improvement in the specification, contrary to Lyon et al.'s results (1999). It is worth mentioning that the single firm approach with the highest correlation coefficient is well specified, while the skewness-adjusted test and sign test are misspecified for most approaches. However, the single firm approach displays well specified sign tests with overrejection rates in the range of 4.8 to 5.2.

In addition, the power test is conducted, to compare the ability of a test to reject the null hypothesis when it is false. -20% to 20% abnormal returns at an interval of 5% are introduced to the holding period return of an event firm. The power of test for all different tests and reference portfolio approaches tends to decline when the length of study period extends. For instance, the student t-test of the control firm approach decreases from 44% rejection rate in one year to 8.8% rejection rate in five years when introducing 10% abnormal return to the event firm. The authors complete the power of test for different horizons as most of previous studies only investigate one time period. Lyon et al. (1999) only document one-year rejection rates, such as 43% for student t-test of a control firm when 10% abnormal return is introduced. Furthermore, several additional findings are documented. The general size and BE/ME portfolios and the benchmark portfolios with size, BE/ME and market beta display little difference in the power of test. This indicates that the extra criteria of market beta cannot better capture abnormal returns compared to the general approach with only size and BE/ME criteria. Moreover, the sign test of the single firm approach with pre-event highest correlation with event firms shows no significant improvement when compared with that of the single firm approach matching with size and BE/ME. Overall, the student t-test and sign test in combination with reference portfolios matching with size and BE/ME yield the highest power to capture abnormal returns. Additionally, the authors highlight the power of test for small firms that are withdrawn from the smallest quintile sorted by size. It is interesting to discover that the sign test for the most-correlated single firm approach has a similar result in small firm samples to random firm samples. This suggests that this test is robust for small firms. All the tests relating to small firms have lower power to test the abnormal returns than tests of random firms.

For the event-time approach in long-run event studies, although anomalies are documented in prior research, a strong argument exists regarding the "joint-test" issue which considers long-run event studies not only test the market efficiency but also the specification of the equilibrium model which captures normal returns of stock returns. Fama and French (1998) conclude that the previous literature regarding long-term stock anomalies does not have solid reasons to challenge market efficiency hypothesis because anomalies tend to be sensitive to methodologies including models for returns and statistical inferences. They focus on two issues, namely bad-model problems in the long-term stock performance, and return metric. As for bad-model problems, they argue that the asset pricing models could not perfectly capture the expected returns with characteristics similar with the event firms. For instance, small size firms show higher returns when compared to expected returns estimated from CAPM. Even though the model is true, average returns could represent sample-specific patterns which cause biases. The authors suggest resolving the bad model issue by estimating expected returns with the application of models matched by firm-specific characteristics. This not only employs returns outside the event period to conduct the regression in order to obtain expected returns, but also matches event firms with non-event firms which share similar characteristics. Size or size/book-to-market is commonly applied as the criteria to identify matched firms. Although matching non-event firms by size/book-to-market benchmarks is widely used in literature, it does not completely capture stock returns (Carhart, 1997).

Regarding the return metric, although BHARs capture investors' experience by assuming investors purchase the stock and hold it for a period of time before realizing the returns compared with CARs, this approach compounds returns to generate abnormal returns even when there is no abnormal performance. BHARs in combination with bootstrapped test statistics are advocated by Lyon et al. (1999); they also highlight its stronger reliability to produce inferences for CARs. BHARs also bear the assumption of independent returns when returns are actually cross-sectional especially when there is clustering in the same industry; therefore, the calendar-time approach is proposed to resolve the problem. Moreover, the choice of weighting scheme affects the results of abnormal performance. Value-weighted returns seem to reduce the abnormal performance or even show no abnormal performance compared with equally-weighted returns as the small firm effect is taken into consideration. However, the authors attribute the small firm effect to the causes of bad-model problem.

A further argument they advocate is that market efficiency indicates an even distribution of overreaction and underreaction anomalies, extensively documented in the literature. The investor overreaction is initially discovered by DeBondt and Thaler (1985). They construct a portfolio with firms which are identified as losers or winners, then rebalance the portfolio by buying winners and selling losers to achieve high returns. Long-run returns indicate mean-reversing from high to low and low to high due to investor overreaction. This is classified as behavioural finance theory and is an alternative to the market efficiency hypothesis. Empirical findings exist regarding underperformance of post-event long-run returns. Take IPO as an example; most previous studies show underperformance of a stock following IPO (Ritter, 1991). However, those firms listed public tend to have a strong performance before IPO, thus giving managers the opportunity to attract investors with high prices. Once listed, the market efficiency starts to correct the price since the market realizes the issue prices are too high. This could be attributed to investor overreaction on the event date, and mean reversion to reflect market efficiency. If the market efficiency works, underreaction should be as frequently documented as overreaction.

Ball and Brown (1968) initially show stock prices react to the announcement one year later following the announcement date. Other following studies also document positive long-term post-event returns (Lakonishock and Vermaelen, 1990, Ikenberry et al., 1996, Desai and Jain, 1997). Since some other anomalies are difficult to locate to categories of overreaction and underreaction, it can be concluded that behavioural finance theory does not provide answers to why overreaction and underreaction are discovered in different circumstances while market efficiency hypothesis attributes the anomalies to the anomalies as investors' luck. Furthermore, they compare individual previous event studies such as IPOs, SEOs, mergers, stock splits etc. and summarize that fact the abnormal returns could disappear when making changes to the methodology applied. To conclude, long-term anomalies cannot reject market efficiency since overreaction is documented as frequently as underreaction. Moreover, anomalies are sensitive to methodology. When there is a change in methodology, for instance, when weighting schemes shift from equally-weighted returns to value-weighted returns, abnormal returns could disappear.

2.5 Literature on long-run event studies based on calendar-time approach

Along with the popularity of event-time studies, the issue of cross-sectional dependence of returns exists (Mitchell and Stafford, 2000). A growing literature body promotes the application of the calendar-time portfolio approach, based on the three-factor model proposed by Fama and French (1993). This model assumes that expected return is captured by characteristics of size, book-to-market ratio and market risk premium respectively. If the model is specified, the intercept should be zero. However, if the intercept is statistically significant from zero, this indicates market inefficiency and model misspecification implying that three factors are not perfectly attributed to the expected return. The basic concept of the calendar-time approach is to regress portfolio returns on different factors such as size, book-to-market and market risk. The portfolios of firms in each calendar month are those which experience an event in the last n years or months. N is denoted as the study period. Fama (1970) summarizes the application of the three-factor model as a joint-test of both market efficiency and model specification. Kothari and Warner (1997) suggest the use of the calendar-time approach to deal with the problem of event time clustering. The application of this approach is strongly advocated by Fama (1998) to deal with the "bad model problem", cross-sectional dependence of abnormal returns and biased statistical inferences. To address the issue of cross-sectional dependent returns, Baber, Lyon and Tsai (1999) illustrate misspecification of test statistics for two extreme types of cross-sectional dependence which are calendar clustering and overlapping return calculation. Consequently, the calendar-time approach, initiated by Jaffe (1974) and Mandelker (1974), is introduced to deal with the issue (Fama, 1997). One flaw in this approach is that investor experience cannot be captured by the model. There are other two variants of the calendar-time approach. One is mean monthly calendar-time abnormal returns proposed by Barber, Lyon and Tsai (1999). This combines reference portfolios and calendar time instead of event time when calculating cumulative abnormal returns. The alternative is the Carhart four-factor model, initially tested in long-term abnormal performance by Ang and Zhang (2004).

Loughran and Ritter (1995) apply the calendar-time approach based on the Fama-French three-factor model over five years for equity issuers and non-issuers to take the cross-sectional dependence issue into consideration. They expect the intercept of the three-factor model to be zero since the abnormal performance is captured by other factors such as size, book-to-market ratio and market risk. The deviation from zero of intercept seems higher in most categories of firms under equally-weighted schemes, ranging from 0.02 to -0.47. By taking the intercept difference between issuers and non-issuers, issuers show underperformance. Moreover, large size firms, either issuers or non-issuers, have higher abnormal returns than small firms. Also, the adjusted R square for issuers and non-issuers is all above 90%, indicating a strong fit of the model. It is worth mentioning that small non-issuers under both value-weighted and equally-weighted schemes have test statistics which cannot reject the null hypothesis of zero intercept at 5% significance level.

Brav and Gromper (1997) compare the annual return of IPO firms and indices benchmarks from January to December each year. With the assumption of independent IPOs, monthly returns are defined as portfolio returns of firms experiencing IPOs in the past five years in a calendar month. Compounding monthly returns of IPO portfolios relative to benchmarks are listed. The results are mixed with underperformance and outperformance spreading in different years. Therefore, the authors propose application of the Fama-French three-factor model to identify the violation of independence assumption. The regression presents no abnormal performance for venture-backed IPOs on both weighting schemes. However, the nonventure-backed IPOs exhibit an intercept of -0.0052 with test statistics of -2.8 and -0.0029, with a test statistic of -1.84 based on equal weight and value weight, respectively. This implies that the choice of weighting schemes produces different results, thus meeting the expectation from the authors that nonventure-backed IPOs underperform in five years compared to venture-backed IPOs. The reasons for this could be better corporate governance of the event firm and reputation of venture capitalists. It is also worth mentioning that the value-weighted scheme tends to reduce abnormal performance, especially for small firms. To answer the question of whether close-ended fund discount is a source of underperformance, a combination of the three-factor model and an additional variable-change in discount which is a proxy of investor sentiment, is adopted. Statistically significant negative non-zero return is documented only for nonventure capital IPOs on equal weight with a significant negative coefficient of discount factor. Investor sentiment then becomes a possible cause for the underperformance of IPOs. Furthermore, severe underperformance exhibited by small and low BTM firms may be due to unexpected shocks, investor sentiment and asymmetric information.

Barber, Lyon and Tsai (1999) advocate the calendar-time approach for two major reasons. One is that it deals with the cross-sectional dependence of sample observations. The other is that it is robust to non-random samples. In the Fama-French three-factor model, they randomly select 1000 samples of 200 firms to test the specification of the model. The equally-weighted and value-weighted portfolio returns are calculated by assuming the sample stock is held over a period of one, three and five years from the event month. But the portfolio returns are based on calendar time. That is to say, if the study period is 12 months, then for each calendar month, the portfolio consists of firms which experience events in the past one year. Therefore, the number of firms is expected to vary. This would cause the potential biased inference recommended by Ang and Zhang (2004) to correct by applying weighted least squares instead of ordinary least squares. The results based on different weighting schemes are mixed. The rejection rates of the null hypothesis of zero intercept extracted from the model on the basis of an equally-weighted scheme for random samples and small/large samples are less than five percent at 5% significance level, indicating the test is conservative. Both equally-weighted and value-weighted abnormal returns for firms with low book-to-market ratios generate misspecified test statistics with rejection rates ranging from 9.8 to 22.8 at 5% significance level. Under the category of poor pre-event returns, the value-weighted scheme underperforms the equally-weighted scheme with over-rejection rates. When sample firms are industry clustering or calendar clustering, this approach reduces the misspecification but could not eliminate it. It is also worth noting that the time horizon in the calendar-time approach does not affect the inferences as much as the buy-and-hold approach. Apart from the Fama-French three-factor model, the authors also employ the average monthly calendar time approach. In this case, they obtain abnormal returns for each sample firm in each calendar month by taking the difference between actual stock return and reference portfolio return, which is matched with similar size and book-to-market ratios. Then for each month, the average monthly portfolio return is calculated based on different weighting schemes. The results indicate a more conservative than traditional calendar-time approach with fewer rejection rates in most cases. Although an improvement to detect long-run abnormal returns in the calendar-time approach is evident, the authors argue that this approach does not incorporate investors' experience as do BHARs in event-time studies.

Loughran and Ritter (2000) defend the market efficiency hypothesis by proposing that abnormal performance is documented due to sensitivity of different methodologies. That is, if there is mispricing, different methodologies are expected to have the same power to capture abnormal returns. Specifically, they challenge the Fama-French three-factor model for three major reasons. Firstly, since the calendar-time approach weights event months equally regardless of whether a month has heavy or light event activities, tests based on this should have lower power than weighting firms equally. The second issue relates to weighting schemes. Take small firms as an example. If small firms show abnormal performance, the portfolio return measured, based on an equally-weighted scheme, gives more weight to small firms, resulting in more significant abnormal returns than value-weighted returns. Moreover, abnormal performance is more common in small firms due to high bid-ask spread; this causes higher transaction costs and stronger market impact when compared with large firms. The final argument is benchmark contamination which indicates that the factors applied in the Fama-French three-factor model do not exclude the sample firm.

The research design is quite different from previous studies. A simulation procedure is conducted by randomly drawing 1000 samples from one firm with a replacement for each month over the period of January 1973 to December 1996. Each sample has 288 calendar months. The number of firms in each sample varies due to trigger returns. Before random selection, the population is carefully screened by keeping only firms with a market price of no less than $3.00 when the portfolio is constructed. Moreover, the sample firm is expected to have a book value at t-1 and t-2. Since the authors are keen to illustrate the low power of the calendar-time approach, they introduce +-5% change in the market value of sample firms over three years. For undervaluation, 5% is subtracted from the sample firm's market value. The sample firm is expected to recover from the loss over 36-months with a monthly return of 0.14% which should be added to the sample firm. Market returns are computed annually by compounding the equally-weighted returns in 12 months. If the market return is below -10%, five more firms will be randomly selected from the population, excluding firms in the largest size category. Then the total market value of these five firms is assumed to be undervalued by 50%. This is expected to correct with 1.94% each month over three years. For overvaluation, a similar procedure is conducted except that the trigger market return is above 30%. Once five extra firms are selected, their total market value increases 100%. A monthly return of 1.91% is deducted over 36-months for these five firms. The magnitude of annual percentage returns is compared between equally-weighted returns and an equally-weighted time period.

The mean raw returns of undervalued- and overvalued firms are 37.1% and 6.2%, respectively. After incorporating trigger returns, the mean returns for undervalued- and overvalued firms are 21% and 20.8% with differences from original raw returns of 16.1% and -14.6%. The mean returns based on different approaches are expected to be close to the differences. Mean size-adjusted return after mispricing considered for undervalued firms is 14.2%, whereas true abnormal returns are 16.1%, indicating that size could not perfectly capture the abnormal return. Moreover, the mean size/BTM-adjusted return for undervalued firms is 11.8% which deviates more from the original abnormal return. This indicates that benchmark returns are higher than sample returns. The low mean returns when weighting the time-period equally under both value-weighted and equally-weighted schemes suggest that the correlation of mispricing and event activities are ignored by the Fama-French three-factor model. Additionally, when a calendar month has only one or two event firms, large standard errors drive down test statistics not to reject the null hypothesis.

Apart from the simulation procedure, they also conduct the calendar-time approach by using samples of IPOs and SEOs that are categorized as new issues. They argue that the model is biased since the factors contain the event firms. Therefore, event firms are withdrawn from factors along with adjustment of heteroskedasticity according to White's method. For new issues, the magnitude of negative abnormal returns increases when withdrawing new issues from size and BTM factors. For instance, the statistically significant monthly abnormal return is -0.4% on an equal weight base for the conventional calendar-time approach and -0.56% when adjusting both SMB and HML. The finding confirms the contamination effect which reduces the level of abnormal returns from the factors in the regression. It is worth noting that the magnitude of equally-weighted returns tends to be higher than value-weighted returns.

When it comes to event activities, the authors test high-volume and low-volume markets with regular factors and purged factors. The definition of low-volume is below the median of percentage of new issues relative to listed firms. The supply response hypothesis, which suggests more evident abnormal performance in an event-clustering time period, is proven to be correct. It is supported by evidence of a larger magnitude of negative abnormal performance of new issuers based on the three-factor model with regular factors and purged factors in the hot market compared to the cold market. When examining IPO portfolios, statistically insignificant returns based on regular factors in the cold market are found, while statistically significant negative abnormal returns are documented for size/BTM-factor-corrected regression in the hot market. This, in a sense, explains the underperformance of IPOs in previous studies. However, seasoned equity issuers present statistically significant underperformance in all cases. The authors question whether the cause of abnormal performance may arise from misspecification of the Fama-French three-factor model and examine the performance of growth stock and value stock in different size categories. The results indicate that small growth stock underperforms with significant negative abnormal returns, while small value stock outperforms with significant positive abnormal returns. In a word, Loughran and Ritter (2000) disapprove of the application of the Fama-French three-factor model which weighs event time equally. They use a simulation procedure to highlight the low power of the model to capture abnormal performance.

Moreover, by examining IPOs and SEOs, they conclude that underperformance becomes more evident in the hot market when most firms make managerial decisions regarding a certain corporate event in a certain period of time in order to take advantage of a window of opportunity.

Mitchell and Stafford (2000) cast doubt on the buy-and-hold approach with bootstrapped test statistics proposed by Lyon et al. (1999) and argue that the assumption of independence of observations is violated. Therefore, they suggest the calendar-time approach to deal with the issue of cross-sectional returns. Moreover, they advocate the application of the calendar-time approach which is inconsistent with Loughran and Ritter (2000). Samples of firms experiencing mergers, SEOs and share repurchases between July 1961 and December 1993 are examined on the basis of equally-weighted and value-weighted schemes with the requirement of at least 10 firms in one calendar month. The firm with events occurring more than once in three years from the first event month is kept only for observation during the first event. The regression based on 25 equally-weighted and value-weighted benchmark portfolios shows significant mispricing with a rejection of the null hypothesis of zero intercept in most cases, especially small firms with low book-to-market ratios. The highest level of negative abnormal return using equally-weighted and value-weighted schemes is -37% and -49%, whereas the lowest level of positive intercept is 26% and 15%, respectively.

In order to discover the impact from the event itself, expected abnormal performances which are not captured by the three factors in the model and abnormal performance from other sources are separated to constitute the intercept. The authors average the intercept of 1000 samples of 200 non-event firms which have similar characteristics of size and book-to-market ratio, as event firms and take it as the part which is attributable to the event. Then this average figure is used as the null to construct an "adjusted intercept" test with original intercept and standard error from regression. If the source of abnormal performance only comes from a corporate event, the adjusted intercept test statistics cannot be rejected. In summary, two intercepts are investigated. One is the original intercept coming from regression of the three-factor model. The other is the difference between the original intercept and intercept from regression based on non-event firms with similar characteristics, comparable with the buy-and-hold approach. Regarding acquirers, the intercept of a full sample post-merge based on equal weights is -0.20%, which is statistically significant from zero with a test statistic of -3.7; this improved to be -0.14% when matched with non-event firms.

Underperformance of post-merge acquirers over three years is evidenced. However, value weighted portfolio regression reveals insignificant results of original and adjusted intercept for acquirers post-merge. The pre-event acquirers present significant positive abnormal returns. Moreover, firms financed with stock under both weighting schemes underperform, whereas those financed in other ways such as cash are fairly priced. The intercept of growth firms with low book-to-market ratio on an equal weighting base is -0.37% with a test statistic of -3.64. However, the adjusted intercept with a test statistic of -1.76 indicates that the abnormal performance is not due completely to the effect of the event. The opposite case occurs on the value weighted base. Value firms with high book-to-market ratios show statistically insignificant zero average monthly portfolio returns on both weighting schemes. The regressions show a relatively high power of the model with large adjusted R square.

With respect to firms experiencing SEOs, value-weighted portfolio regression suggests no abnormal performance over three years. But on an equal weight base, statistically significant negative returns exist post-event, even with the exclusion of utilities. As for equity repurchases, value-weighted portfolios also show no abnormal performance, while open market and value firms on an equally-weighted base have significant positive abnormal returns which are not explained fully by the event. The authors list four major problems regarding the calendar-time approach. The first is industry clustering which could result in biased estimates. Another flaw is heteroskedasticity, caused by the varying number of firms in each calendar month. This will lead to unbiased but inefficient estimates. One further concern is that the calendar months are weighted equally which ignores the month when the market has heavy activities with many firms experiencing one specific event such as IPOs. The last argument relates to the loss of power when applying the calendar-time approach. To deal with heteroskedasticity, the authors propose the minimum requirement of 10 firms in each event month as well as weighted least squares technique. However, the assumption of independent residuals for each firm in this approach conflicts with the assumption of the calendar-time approach that is supposed to deal with cross-sectional returns. Therefore, bootstrapped non-parametric test statistics proposed by Horowitz (1996) are applied to achieve new critical values to determine the statistical significance of coefficients for different cases.

For event activities clustering, the authors add two dummy variables into the three-factor model as proxies of hot and cold markets. Firms categorized into the hot market with heavy event activities belong to 70th percentile of all listed monthly activities; these are defined as the percentage of the number of firms experiencing events relative to the total number of firms listed. Firms which lie below 30th percentile of monthly activities are classified as being in the cold market. The results show no abnormal performance for hot and cold markets under both weighting bases in all cases except seasoned equity issuers with value-weighted portfolio in the cold markets. The exception has an ordinary intercept of 0.32% with a test statistic of 1.79 and adjusted intercept of 0.42% with a test statistic of 2.34.

Apart from the traditional calendar-time approach, the authors apply the average calendar-time abnormal returns, defined as the difference between portfolio return and benchmark return. This approach is similar to the approach by Lyon et al. (1999). But the proxy of benchmark returns takes the portfolio return from regression on three factors except for 25 portfolios with similar characteristics. Similar results are found such as the traditional calendar-time approach when the benchmark is portfolio returns from the three-factor model. However, the magnitude of abnormal returns is smaller in the average calendar-time abnormal returns. When the benchmark is reference portfolios, there is no abnormal performance for acquirers post-merge under both weighting schemes as well as growth firms. The value-weighted portfolios, regardless of the three-factor benchmark or size/book-to-market reference portfolio benchmark, are fairly priced for seasoned equity issuers in all cases. Equity repurchasers on an equally-weighted scheme for both benchmarks have statistically significant positive returns post-event and in the open market. The value-weighted scheme tends to eliminate abnormal performance for SEOs and share repurchases.

The authors examine the power of the calendar-time approach and the calendar-time abnormal return approach by introducing extra abnormal returns ranging from -20% to 20% over three years. Equally-weighted portfolios tend to have higher power than value-weighted portfolios in both approaches. On an equal weighted base, the calendar-time approach is preferred. On a value weighted base, the calendar-time abnormal return approach is given preference. The authors challenge Loughran and Ritter's (2000) findings and argue that the calendar-time approach has higher power to detect abnormal performance than the buy-and-hold approach. In conclusion, Mitchell and Stafford (2000) advocate the application of the calendar-time approach to deal with cross-sectional returns. However, they do admit that the bad model issue indicates the three-factor model could not perfectly capture the abnormal performance with intercept by decomposing the intercept into event-driven and other sources.

Ang and Zhang (2004) compare the Fama-French three-factor model with the four-factor model including the momentum factor proposed by Carhart (1997). The weighted least squares technique is applied along with ordinary least squares to see if there is any improvement in order to deal with the varying number of firms in the regression. 250 samples of 200 firms are randomly drawn from the population to achieve the rejection rates of the null hypothesis of zero intercept as no abnormal performance. Both OLS and WLS in the four-factor model generate a much higher rejection frequency, ranging from 6.4 to 32.8 percent, over the period of all time horizons. Moreover, the three-factor model using the OLS technique indicates improvement of power of procedures over the WLS technique, especially when the time horizon lengthens. It can be concluded that the three-factor model is the preferred method to the four-factor model. However, the three-factor model using both the OLS and WLS technique shows conservative test statistics since the rejections rate of the null hypothesis of zero abnormal returns are fewer than 5 percent.

In order to compare with the power of buy-and-hold approach based on reference portfolios, they introduce an effective holding period return to each sample firm; this is the difference in a firm's holding return between before and after monthly added abnormal return, instead of nominal return as in other studies. That is, the level of effective induced return for a one-year sample is 24.2%, while the nominal level of induced return is 20%. The power of both models using different techniques declines when the time horizon lengthens. For example, when incremental abnormal return is 24.2%, the three-factor model using OLS captures 99.6% of the abnormal performance in one year period but 49.2% in a five-year period. Still, WLS presents a higher power of test with higher rejection frequency than OLS, particularly in a longer holding period. For instance, the rejection frequency is 74.4% for OLS and 90.8% for WLS when the nominal introduced abnormal return is 20% over three years. When it comes to negative introduced returns, all tests indicate lower power compared to the case of positive introduced returns. Take a three-year sample for example; the rejection frequency of the null hypothesis is 59.6% when the induced return is -20% and 74.4% when the induced return is 20%. The authors argue that this asymmetry could result from lack of knowledge regarding rebalancing costs. It is worth commenting that a conclusion is reached that the most-correlated firm approach with the sign test is almost equal in power to the three-factor approach with WLS over one year but has greater power over three and five years.

To test the robustness of these approaches, samples of small firms are examined. Similar results are documented to random samples. The three-factor model with WLS is still preferred since it yields better specified tests and shows higher power. But small firms cause more significant loss on power of test over three- and five-years than random firms. The authors also discuss the reason why the four-factor model is discarded even with an extra momentum factor. They investigate the magnitude of the average intercept from regression and compare the adjusted R square. The magnitude of mean intercept for the three-factor model is close to zero, whereas the intercept for the four-factor model deviates from zero. The intercept tends to increase with a longer time horizon especially for the four-factor model. However, the four-factor model shows improvement in the goodness of fit of the model with higher adjusted R square although it shows larger magnitude of abnormal return in small firms. In summary, the three-factor model with WLS dominates in the calendar-time approach but shows asymmetry when compared with the buy-and-hold reference portfolio approach.

Still, some research does not support the application of the calendar time approach. Barber and Lyon (1999) debate that the three-factor model assumes stable coefficients which could be easily violated; furthermore, survivor bias exists since coefficients require data availability during the estimation period. Moreover, although misspecification for samples from the same industry can be reduced, this cannot be eliminated due to bad-model problems. From the perspective of weighting schemes, Loughran and Ritter (2000) indicate that the calendar time approach has low power to capture abnormal returns. Managers are considered to be able to consider the behavioural timing when making decisions. Furthermore, the calendar-time approach requires a varying number of firms in different calendar times. This raises the problem regarding heteroskedasticity which influences the inferences. The variance of the regression residuals could constantly change, since when the number of firms in a portfolio increases, the portfolio variance declines. In this case, the estimate would be inefficient. Although Ang and Zhang (2004) propose a solution of weighted least squares estimation which allocate weights to the square root of the number of firms in the portfolios, Mitchell and Stafford (2000) argue that the weighting scheme, that weights each event equally, shares the same purpose as the calendar-time approach when considering cross-sectional returns. Therefore, Mitchell and Stafford (2000) suggest there should be at least 10 firms in a portfolio in any given month to control the effects of heteroskedasticity. Hou et al. (2000) highlight an additional proposal to deal with this issue. The generalized autoregressive conditional heteroskedasticity model (GARCH) is applied to model the time-varying variance of the regression residual.

Chapter 3 Event-time approach: Simulation based on models

3.1 Research questions and hypotheses

Null hypothesis: the cross-sectional cumulative abnormal return or buy-and-hold abnormal return is zero over investment horizons of one-, three- and five years.

This chapter aims to detect long-term abnormal stock performance in the UK stock market over a period from 1982 to 2008 with the application of the event-time approach based on pricing models, namely, the market-adjusted model, market model, capital asset pricing model, Fama-French-three-factor model and Carhart-four-factor model. Following the methodology applied by Kothari and Warner (1997), the simulation process with 250 samples of 200 firms is conducted. In contrast to traditional event studies which use specific events such as IPOs, mergers and acquisitions, the event study applied in this chapter simulates 200 event months over the period of October, 1982 to October, 2007 when the investment horizon is one year; to October, 2005, when the investment horizon is three years; and to October, 2003 when the investment horizon is five years. Since the event months are randomly selected, a possibility exists that there will be more than two identical event months. Moreover, events are not specified. Therefore, sample firms could experience a wide range of events, such as SEOs. This is similar to what occurs in the real world where different firms experience a wide range of events on different dates. The event study not only tests the market efficiency hypothesis, but also the equilibrium pricing model. If the model is well-specified and the market is efficient, the null hypothesis should not be rejected at different significance levels. However, the conclusion of an efficient market cannot be drawn when there is overrejection of the null hypothesis because the model applied may be misspecified. To identify whether the test statistics are well-specified, the percentage of rejection rates is computed. Firstly, the test statistic of average cross-sectional CARs or BHARs for each sample is compared with the critical value. Then if the percentage of rejection rates in 250 samples is 5% at a significance level of 5%, the model is well-specified.

3.2 Data

The starting point of this study is a list of UK non-financials with book value, market value and industry code over the period of 1982 to 2007 from studies undertaken by Gregory, Tharyan and Huang (2009). The dataset excludes financials such as banks, insurance, investment trusts, and financial services. This thesis investigates non-financials in order to reduce potential noises from financials due to different accounting treatments. Moreover, the data includes stocks traded not only in the Main Market in the London Stock Exchange (LSE), but also in the Alternative Investment Market (AIM) launched in 1995 and Unlisted Securities Market (USM) run from 1980 to 1996. The Main Market totals 1,080 UK listed firms with approximately 55.9% of firms with a market value of over £50m in 2008. The AIM witnesses an increase of listings including both the UK and international firms with a market value of £2382.4m in 1995 to 1,546 with market value of £37,698.9m in 2008. It is filled with mostly small firms, approximately 73% of firms having a market value of less than £25m in 2008. There is a trend of shift from the Main Market to the AIM after its launch due to its more flexible regulatory requirements. For instance, 40 firms completed this shift in 2005. The USM is established by LSE to accommodate very small firms or firms that do not meet listing requirements between 1980 and 1996. The number of listings peaks at 448 with a market value of £8,975m in 1989. Gregory, Tharyan and Huang (2009) obtain monthly market value from the London Stock Price Database (LSPD) established by the London Business School. This database covers all listed stocks including dead companies in the UK market from 1955. However, it does not provide accounting data such as book value. To complete the dataset, the book value of a firm is supplemented from the Datastream with a code of 305 which demonstrates the total share capital and reserves excluding preference capital, together with Hemscott. The remaining missing book values are hand collected from the Stock Exchange Year Books issued by the London Stock Exchange by Gregory and Huang (2009). The missing returns of the original data remain the same without any replacements.

To determine the timeline for the collection of book values, I take the research carried out by Agarwal and Taffler (2008) for reference. They investigate UK firms from 1979 to 2002 and conclude that 30th September of each year is a better choice as the monthly date to establish portfolios because 22% of firms' accounting year ends in March, whereas 37% of firms release their report in December. Consequently, I form the portfolios in September of each year based on available market values and book values of firms. To ensure that the required accounting information is available at the time of the portfolio formation, a six-month lag is assumed as Lyon, Barber and Tsai (1999). This means the book value of a firm is from its latest balance sheet having its accounting year in or before March of year t whereas the market value is based on the figure at the end of September. Therefore, the first step to screen the data is to keep only a firm with both market value and book value for the purpose of grouping and robustness checks. But firms with negative book value are excluded. The market value and book value with zero value, which indicates relatively small firms with market values less than £1million, are replaced with 0.5million. Since there are only a few of firms with market value or book value as zero, the replacement value may not have too much impact on the results. The last procedure to clean the data is to include only firms with ordinary shares labelled as zero value of U2 in the LspdU file from the LSPD database.

Table 3.1 presents a summary of the whole sample in each year between 1982 and 2007, which I quote as all firms in the following studies. N1 represents the number of firms from Gregory, Tharyan and Huang (2009). N2 sums up the number of firms without book value or market value or both. It peaks at 545 with 423 firms without book value in 2007. N3 lists firms with both book value and market value, while N5 reveals the final screened number of firms each year. N4 lists the number of firms without ordinary shares. **Figure 3.1** shows a graph of the number of firms in three scenarios: the number of firms excluding financials from Gregory, Tharyan and Huang (2009), the number of firms with both market value and book value, and the number of finalized sample firms. It can be seen that there is an uptrend for all datasets from 1982 to 2007. The big gap between original data and screened data in the 1990s and 2000s indicates firms with incomplete data for book value and market value cluster. The average number of firms over the last 26 years after the adjustment is approximately 1500 annually. Eventually I arrive at a population sample of 4977 firms over 26 years. **Table 3.2** provides the descriptive statistics of raw stock returns of the screened sample. The total number of returns observations over the period of 1980 to 2008 is 549,588 with 19,572 missing returns accounting for approximately 3.6% of total observations. The reason for including 1980 and 1981 is because the following research undertaken by Kothari and Warner (1997), 24-month prior-event returns are required to compute the standard errors for test statistics. The mean return approximates 1.25% with skewness of 34.93 and kurtosis of 6580. **Table 3.3** demonstrates the frequencies of returns of the final sample. The bin values represent the range of values in the dataset. The first bin is -100% which is defined as the minimum return in the dataset. The frequency of 1 in the first bin suggests there is one observation falling into this category. The second bin has values greater than -100% but smaller than or equal to -93%. There are 22 observations meeting this criterion. Same rules are applied to the rest. Most stock returns concentrate in the range of -49% to 52%, particularly 1% to 8% (around 51% of the total sample). It is important to note that 27 observations have returns of more than 500%. The maximum return of 4500% is from a firm with the g1 code such as the gl code of 9705 in January 2000. Without exclusion of these large stock raw returns, the cumulative stock returns, especially with the application of the buy-and-hold approach, are expected to achieve extremely large results. **Figure 3.2** graphs the histogram of returns of all firms to illustrate the distribution of return observations with right-skewness and high kurtosis. From the distribution, I expect the results of abnormal returns to be severely right-skewed if firms with extreme returns are included in the sample.

In order to make comparisons of different approaches, the test period is from October, 1982 to September 2008. The post-event periods of one-, three- and five years are examined. September of each year is the time for dividing firms into categories matching with ranked size and book-to-market ratio. But returns are tracked from October onwards. Therefore, regardless of any approaches I employ, returns data starts from October, 1982. The ending month depends on the event window. If the post-event window lasts for one year, the last event month is October, 2007. If it is three years, October, 2005 is the last month in the simulation process. If it is five years, October, 2003 is the last event month. With the application of the event-time approach as Kothari and Warner (1997) and Lyon, Barber and Tsai (1999), the event month is included in the test period. However, the calendar-time approach tests abnormal returns after the event-month. The details are listed in **Table 3.4.**

For event-time studies based on models in this thesis, I exclude firms which do not have 24 month prior-event returns as Kothari and Warner (1997), in order to have sufficient observations in the estimation period to conduct regressions and achieve stable standard errors based on the estimation periods for test statistics. Moreover, firms are required to have at least one return observation in the test period. In a word, a firm is selected only if it has consecutive 25-month returns. The final sample in this chapter has 3,978, 3,500 and 3,243 firms over one-, three- and five-year investment periods, respectively. The number of firms decreases with the length of investment periods because I need to make sure that the last month of the stock return ends in September, 2008. Following the research undertaken by Kothari and Warner (1997), if a firm is delisted within the test period, its returns are tracked according to the availability of data in order to minimize the survival bias. For instance, suppose the test period is 36 months. A firm is delisted in the 15th month. Then the estimation of abnormal returns of this firm ends in month 15. Moreover, the missing values are kept as they are without any replacement. However, in the subsequent chapter which applies the event-time approach based on reference portfolios and a single control firm, the missing values are replaced with the average returns of reference portfolios. This further suggests if a firm is delisted, its returns over the delisted period are filled with the average returns of reference portfolios. Therefore, an event firm has a complete return dataset over the event window. Taking the previous example, the rest 21-month returns are replaced with the average returns of reference portfolios which the firm belongs to rather than simply taken as missing values. It is important to note that since there is no requirement of returns in the estimation period, the final sample of firms in the following two chapters has 4,977 firms.

3.3 Research methodology

3.3.1 Models

I choose simple returns to start with because these are portfolio additive, whereas log returns are time additive. In this thesis, I examine both buy-and-hold returns and cumulative returns. Simple returns are applied since log returns are not appropriate for cross-sectional studies. LSPD provides log stock returns as follows:

$$\log R_{i,t} = \ln\left(\frac{P_{i,t}+D_{i,t}}{P_{i,t-1}}\right) \tag{3.1}$$

$P_{i,t}$: The last traded price in month t

$D_{i,t}$: The dividend paid during month t and adjusted to a month-end basis

$P_{i,t-1}$: The last trade price in month t-1

Therefore, I convert log returns to simple returns by using the exponential formula:

$$R_{i,t} = \exp\left(\log R_{i,t}\right) - 1 \tag{3.2}$$

The general formula for abnormal return is defined as follows:

$$AR_{i,t} = R_{i,t} - E(R_{i,t)} \tag{3.3}$$

If the market is efficient, abnormal returns should be equal to zero for all investments. That is, based on the given risk level, securities should be priced on average so the actual returns are equal to expectation from investors. The market-adjusted model, market model, capital asset pricing model, Fama-French three-factor model and Carhart four-factor model are applied in this chapter to detect the long-run abnormal returns.

Market-adjusted model:

$$AR_{i,t} = R_{i,t} - R_{m.t} \tag{3.4}$$

$R_{m.t}$ is denoted as market return, measured as equally-weighted market portfolio return.

Market model:

$$AR_{i,t} = R_{i,t} - (\alpha_i + \beta_i R_{m,t}) \tag{3.5}$$

α_i and β_i are achieved by regressing monthly returns for security i on $R_{m,t}$ over the estimation period.

Capital asset pricing mode

$$AR_{i,t} = R_{i,t} - \left\{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\right\} \tag{3.6}$$

$R_{f,t}$ is the risk-free rate which uses three-month market rates of Treasury bills as a proxy. β_i is the coefficient resulting from regression of stock returns on the difference of market returns and risk-free investment returns with intercept of risk-free rate over the estimation period. This model contains the single market portfolio-based risk factor.

Fama-French three-factor model

$$AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_{i,1}RPM_t + \beta_{i,2}HML_t + \beta_{i,3}SMB_t\} \quad (3.7)$$

With the background of CAPM with a single risk factor, Fama and French (1993) incorporate two additional risks which influence the stock prices. HML_t is the difference between returns on portfolios of high and low book-to-market-ratio securities, whilst SMB_t is the difference between returns on portfolios of small and large market value securities. RPM_t is the difference between returns on a market index and risk-free investment. β_s are beta factors corresponding to different control factors and are achieved by regression over the estimation period. The factors are obtained from Exeter University's website provided by Gregory, Tharyan and Huang (2009).

Carhart four-factor model

The argument of the similarity between the three-factor model and CAPM evokes academics' interest to explore a model to incorporate more controlling factors. Therefore, a 4^{th} factor is introduced by Carhart (1997) representing the difference between returns on portfolios of securities that perform well and poorly in the short term. It is a momentum factor to capture the short-lived, (defined as one year period), momentum anomaly.

$$AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_{i,1}RPM_t + \beta_{i,2}HML_t + \beta_{i,3}SMB_t + \beta_{i,4}PR12_t\} \quad (3.8)$$

$PR12_t$ is the difference between average returns on the two high prior-return portfolios and the two low prior-return portfolios. The momentum factor is also available on the Website of Exeter University.

Cumulative abnormal returns

Once individual stock returns are computed, the returns across time and stocks need to be accumulated since I am interested with the mean abnormal return of a portfolio. Therefore, I have two measurements of accumulation of stock returns in a portfolio. One is known as cumulative abnormal returns (CARs), which simply average all stocks returns in a portfolio in month t, and then add up all returns over the study period as follows:

$$AR_{p,t} = \frac{\sum_{i=1}^{N} AR_{i,t}}{N} \quad (3.9)$$

$$\mathrm{CAR}_{p,T} = \sum_{t=0}^{T-1} \mathrm{AR}_{p,t} \tag{3.10}$$

N is the number of firms in the portfolio in month t. The number of firms varies in the test period since stocks are only tracked for available returns. If a firm is delisted during the test period, it needs to be excluded from the portfolio when it dies.

Buy-and-hold returns

Another measurement of abnormal returns is buy-and-hold returns (BHARs) which compounds individual stock returns over the holding periods. The mean BHARs is calculated by taking the average BHARs of stocks in the portfolio

$$. \ \mathrm{BHAR}_{i,T} = \prod_{t=0}^{T-1} (1 + \mathrm{BHAR}_{i,t}) - 1 \tag{3.11}$$

$$\mathrm{BHAR}_{p,T} = \frac{\sum_{i=1}^{N} \mathrm{BHAR}_{i,T}}{N} \tag{3.12}$$

When conducting BHARs, the number of firms is 200 over the test period because of the initial screened requirement of at least 25-month returns including the event month. As can be seen from the formula, CARs computes monthly portfolio returns in a month based on equal weights firstly, and then sums up over the investment period, whereas BHARs calculates individual BHARs over the holding period firstly and then take the average of BHARs of all stocks in the portfolio.

3.3.2 Statistical inferences

Parametric test

The null hypothesis is that the mean monthly CARs or BHARs is zero over the holding period. The conventional test statistics assume normal distribution of CARs and BHARs with zero mean and constant variance. Regarding CARs, test statistics is given as:

$$t = \frac{\mathrm{CAR}_{P,T}}{\sigma(\mathrm{CAR}_{p,t}) * \sqrt{T}} \tag{3.13}$$

When the standard deviation of CARs is based on estimation period, it is shown as:

$$\sigma(\mathrm{CAR}_{p,t}) = \sqrt{\sum_{t=-24}^{-1} \left(\mathrm{CAR}_{p,t} - \overline{\mathrm{CAR}_{p,t}}\right)^2 / 23} \tag{3.14}$$

where

$$\overline{\mathrm{CAR}_{p,t}} = \frac{1}{24} \sum_{t=-24}^{-1} \mathrm{CAR}_{p,t} \tag{3.15}$$

However, when the standard deviation is defined over the test period, it is given as:

$$\sigma\left(\mathrm{CAR_{p,t}}\right) = \sqrt{\sum_{t=0}^{T-1}\left(\mathrm{CAR_{p,t}} - \overline{\mathrm{CAR_{p,t}}}\right)^2/(T-1)} \tag{3.16}$$

where

$$\overline{\mathrm{CAR_{p,t}}} = \frac{1}{T-1}\sum_{t=0}^{T-1}\mathrm{CAR_{p,t}} \tag{3.17}$$

The test statistics for determination of whether null hypothesis of mean BHARs uses standard deviation over the test period is as follows:

$$t = \frac{\mathrm{BHAR_{P,T}}}{\sigma\left(\mathrm{BHAR_{p,t}}\right)}, \tag{3.18}$$

Where

$$\sigma\left(\mathrm{BHAR_{p,t}}\right) = \frac{1}{199}\sqrt{\sum_{t=0}^{T-1}\left(\mathrm{BHAR_{i,T}} - \mathrm{BHAR_{p,T}}\right)^2/(T-1)} \tag{3.19}$$

Wilcoxon Signed-rank test

Apart from parametric tests, I also examine a popular nonparametric test that is a combination of sign test and rank test. It is known as the Wilcoxon Signed-rank test, introduced by Wilcoxon (1945). It bears null hypothesis of an equal number of positive and negative returns under a framework of binominal test. Firstly, the absolute values of 200 CARs or BHARs are ranked in each sample. Rank 1 relates to firms with the lowest CARs or BHARs. Then signs are allocated to observations. Afterwards, the additions of ranks with a positive sign and with a negative signs are computed. Then the sum of these two figures is calculated for the purpose of determining whether or not the null hypothesis of zero median is rejected. The process is expressed as follows:

$$\mathrm{SR_N} = \sum_i \mathrm{CAR_i^+} \tag{3.20}$$

where

$$\mathrm{SR_N} \sim \mathrm{N}(\mathrm{E}(\mathrm{SR_N}), \sigma^2(\mathrm{SR_N})) \tag{3.21}$$

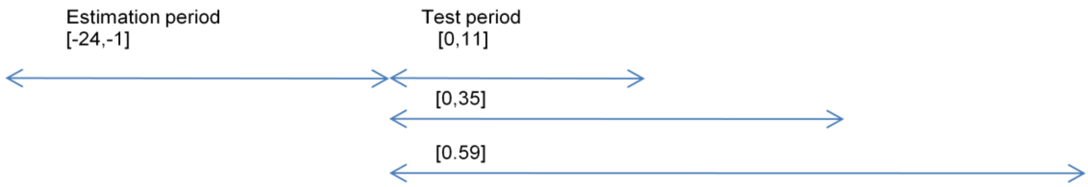
$$\mathrm{E}(\mathrm{SR_N}) = \mathrm{N}(\mathrm{N}+1)/4 \tag{3.22}$$

$$\sigma^2(\mathrm{SR_N}) = \mathrm{N}(\mathrm{N}+1)(2\mathrm{N}+1)/24 \tag{3.23}$$

$\mathrm{CAR_i^+}$ is the positive rank of the absolute value of CARs. When N is large, test statistics follow a normal distribution. N is denoted as the total number of firms. The test statistic Z is calculated based on mean and variance in order to compare this with critical values at a significance level of 5% both on one-sided and two-sided tests. For each sample, I achieve one Z. The number of Z totals 250 over one-, three- and five-year holding periods. The rejection frequency among 250 samples is computed. A similar process is conducted for BHARs.

3.3.3 Simulation process

Following the simulation process by Kothari and Warner (1997), firms in the population need to have at least 25-month returns including returns in the event month. For example, since the test period is from October 1982 to September 2008, firms without 25-month consecutive returns over the period of October 1980 to October 2007 are excluded when examining the stock performance over one year. The simulation process selects 250 samples of 200 firms randomly. I randomly select 200 firms with replacement from the population. For each firm, I randomly select an event month between October 1982 and October 2007 for one year, October 1982 to October 2005 for three years and October 1982 to October 2003 for five years. One firm may appear more than once in the sample with multiple event months. This potentially causes the issue of overlapping returns. The whole period is divided as an estimation period and test period which includes event month. The timeline is illustrated as follows:



As can be seen from the above graph, the same estimation period, defined as the previous 24 months prior the event month, is applied in all investment periods. The test periods are denoted as from month 0, the event month, to month 11 when the study period is one year. The returns of firms are tracked for one year, three years and five years over the whole study period. To control the survivorship bias, the performance of a security is only estimated according to the data availability even when calculating the buy-and-hold returns. If a firm does not have 12-month returns following the event month, I only examine as many monthly returns as it has. For each sample of 200 firms, I test the null hypothesis of zero mean CARs or BHARs at 1% and 5% significance level. Then the rejection frequency is computed. The rejection frequency is 5% at 5% significance level when the test is well-specified. If it is less than 5%, the test is relatively conservative. If it is more than 5%, the test is anti-conservative. Both two-sided and one-side test statistics are investigated.

3.4 Simulation on random samples

3.4.1 Cumulative abnormal returns (CARs)

The rejection frequency of the null hypothesis of zero mean CARs is reported in **Table 3.5** at a significance level of both 1% and 5%. The rejection rates in two-tailed tests increase with the length of investment horizons in all five models. For instance, the rejection rate for the market-adjusted model at 5% significance level increases from 12.4% to 24.8% from one year to five years. The mean CARs and test statistics of CARs show a similar pattern to the rejection rates. All models show severe misspecification with overrejection on the null hypothesis. The market model underperforms other models with the highest rejection rates of 26%, 40% and 50% in one-, three- and five years at a significance level of 5%. Furthermore, it shows the highest mean CAR of 250 samples increasing from 2.6% to 12.2% from one year to five years, together with the highest standard deviation of CARs in the range of 0.06 to 0.209. The mean test statistics of CARs in the market model also are the highest among other models. The Carhart four-factor model is the second underperformer among all the models, with a rejection rate of 28% in one year, increasing to 44.4% in five years at a significance level of 5%. The lower rejection rates in the Fama-French three-factor model suggest that the incorporation of momentum factor in the four-factor model generates more severe misspecification. The less misspecified results generated by the market-adjusted model indicates a simple model that is better applied to detect long-term stock performance, rather than factor models which control different factors. Regarding one-tailed test, CARs in all models demonstrate positive biased test statistics with higher rejection rates on the right tail at both 1% and 5% significance levels. Furthermore, the rejection rates of alternative hypothesis of negative CARs are close to 5% except for the two factor models in one year and apart from the market model in three and five years. This evidence suggests the absence of abnormal returns. The findings are in line with studies undertaken by Kothari and Warner (1997).

However, the magnitude of rejection rates is smaller in one year and larger in three years in the UK market from 1982 to 2008, compared with the US market covering the period of 1980 to 1989. Kothari and Warner (1997) document that the rejection rates in two-tailed test are 12.4% and 20.8% in one year and three years at a significance level of 5% when applying the capital asset pricing model on the US data. The corresponding rejection rates in the studies are 13.6% and 18%. Apart from the conventional test statistics, I also conduct the Wilcoxon signed-rank test under CARs in **Table 3.7** to avoid the violation of the assumption of normal distribution. Over the investment horizon of one year, all models show specified test statistics in both one-tailed and two-tailed tests apart from the market model with a rejection rate of 92%. However, misspecification is indicated for all models in three and five years. The rejection rates increase with the length of investment horizons. Moreover, the overrejections on the right tail suggest that the median of CARs is positive; this is consistent with the distribution of CARs in a sample of 50,000 firms discussed in the subsequent section.

The application of standard deviation in the computation of test statistics over the estimation period results in an upwardly biased test statistic when compared with the application of standard deviation over the test period. The number of firms in the estimation period is constant due to the requirement of at least 25-month returns. However, the number of firms over the test period is changing over the time, as returns are tracked according to the availability of data. Therefore, the standard deviation based on the test period increases when the sample size is smaller. **Table 3.6** exhibits the specification of test statistics for five models over three investment horizons. An improvement of misspecification with lower rejection rates in all models is indicated. The market model still generates the most severely misspecified results with the highest rejection rate; the four-factor model follows this. The rejection rate in two-tailed test in one year when applying the market model based on standard error over the test period is 19.2% which is smaller than the case when the standard error is computed over the estimation period (26%). The market-adjusted model and capital asset pricing model outperform other models over three investment horizons. The market-adjusted model has rejection rates of 8.8%, 6% and 9.2%, while the capital asset pricing model has rejection rates of 8.8%, 6% and 8.8% in one-, three- and five years respectively. It is interesting to note that the Fama-French three-factor model indicates the lowest rejection rate of 7.6% among all models in five years but has a rejection rate of 12.8% in one year. Moreover, the misspecification is more severe in one year except for the market model. Regarding one-tailed test, positively biased returns are still presented in most cases apart from the capital asset pricing model and the two factor models in one year, which tend to be symmetric. The smaller magnitude of mean statistics of CARs compared with the case of taking standard deviation over the estimation period indicates larger standard deviation over the test period. This confirms the findings by Kothari and Warner (1997) who apply the market-adjusted model and find a rejection rate of 18.4% in three years when the standard error is computed over the estimation period and 9.6% when the standard error is calculated over the test period.

3.4.2 Buy-and-hold returns (BHARs)

The empirical results of BHARs in **Table 3.8** indicate inconsistent findings as with the prior studies. The most distinguishing feature is that except for the market model and the four-factor model, the rejection rates of the null hypothesis of zero mean BHARs for all other models at significance levels of 1% and 5% present well-specified test statistics. At a significance level of 5% in two-tailed test, the market-adjusted model shows rejection rates of 6.8%, 4.4% and 8.8%, whereas the capital asset pricing model presents rejection rates of 4.8%, 3.6% and 4.4% in one-, three- and five years. Moreover, the Fama-French three-factor model yields a rejection rate of 6% in one and three years and 3.6% in five years. Although the four-factor model has a lower rejection rate of 4.8% in one year when compared with the three-factor model, it yields rejection rates of 12.8% and 14.8% in three and five years. Although the market model generates a lower rejection rate in one year (19.2%) for BHARs compared with CARs (26%), it reveals the most severe misspecification with the highest rejection rates of 62.4% and 60.8% over three and five years. Compared with CARs, it is interesting to note that the rejection rates are mostly lower in BHARs except for the market model which shows rejection rates of 19.2% in BHARs and 26% in CARs when the investment horizon is one year, and 62.4% in BHARs and 40% in CARs in three years, and 60.8% in BHARs and 50% in CARs in five years. Kothari and Warner (1997) also document similar patterns. However, the misspecification is still indicated for all models over the investment horizon of three years in their findings, which is in contrast to the evidence. In consequence, I advocate the application of BHARs rather than CARs in the event-time approach based on models of the market-adjusted model, capital asset pricing model and three-factor model to detect long-term stock performance.

Regarding symmetry of test statistics under BHARs, the market-adjusted model shows negatively biased test statistics over the holding periods of three and five years; in contrast, the market model and four-factor model display positively biased test statistics. However, a symmetric test statistics under BHARs with rejection rates of 5.6% on both sides in one year is documented in the Fama-French three-factor model, but develops into an asymmetric pattern with positive biased test statistics in three and five years. Moreover, the capital asset pricing model indicates negatively biased test statistics in one year. With respect to the mean BHARs, this increases according to the length of holding periods, as in CARs in most cases except for the market-adjusted model and the two factor models. The magnitude of mean BHARs is mostly larger than that of mean CARs apart from the market-adjusted model. Unlike positive mean CARs in all models, BHARs has three exceptions with negative mean BHARs. The market-adjusted model has mean BHARs of -0.4% and -0.5% in one year and five years, respectively. It is important to note that the market model still presents the highest mean BHARs as CARs. Moreover, the mean BHARs of the market model reaches 413.4% in five years. This may be due to the random sampling. The magnitude of the mean BHARs documented by Kothari and Warner (1997) in the US stock market over the investment horizon of three years is larger in all models except for the market model when compared with the UK data. For instance, the mean BHARs is 1.7% when applying the market-adjusted model in the UK stock market and 4.38% in the US stock market. However, the corresponding figures when applying the market model are 63.9% in the UK stock market and 27.8% in the US stock market.

The Wilcoxon signed-rank test reported in **Table 3.9** demonstrates severe misspecification for BHARs in two-tailed tests at a significance level of 5% in all models except the market model in three and five years, with rejection rates of 3.6% and 5.6%, respectively. The null hypothesis of zero median with equal frequencies of positive and negative BHARs is overrejected in most cases. The market-adjusted model, instead of the market model, has the highest rejection rates. Moreover, the higher rejection rates on the left tail indicate that the median of BHARs in 250 samples is significantly smaller than zero. This is consistent with the distributional properties discussed in the subsequent section.

3.4.2.1 Power of test in BHARs

With a conclusion that the market-adjusted model, capital asset pricing model and Fama-French three-factor model in conjunction of BHARs are appropriate to apply in long-run event studies, it is interesting to find the outperformer in terms of power of test, which is defined as the ability of a test to detect an effect, if the effect exists. In other words, power is the probability that the null hypothesis should be rejected when it is false. Therefore, I induce abnormal returns in the range of -20% to 20% with an interval of 5% to individual stocks over the holding period. If the test is powerful, it should reject the null hypothesis of zero BHARs since there is abnormal return. Therefore, the percentages of rejection rates over all investment horizons are expected to be 100% if the test is superior. **Table 3.10** reports the empirical results. First of all, the rejection rates for all models decline when the holding period lengthens. For instance, the percentage of rejection rates when applying the market-adjusted model is 98.8%, 95.2% and 86% when the holding periods are one-, three- and five years. Second, the test shows higher power with higher rejection frequency for all models when abnormal returns induced are negative. The rejection rate is 85.2% when -20% is induced and 61.6% when 20% is induced with the application of the capital asset pricing model over five years. Third, the rejection rates increase when higher absolute values of abnormal returns are induced. The rejection rates are 38.8% when 15% of abnormal return is induced and 54.8% when 20% of abnormal return is induced with the application of the Fama-French three-factor model over three years. Last, the market-adjusted model shows the highest power of test with the highest rejection rates compared with the capital asset pricing model and Fama-French three-factor model over all holding periods. The rejection rate is 86% when applying the market-adjusted model over five years whereas the corresponding figures for the capital asset pricing model and Fama-French three-factor model are 85.2% and 80.8%, respectively.

3.5 Simulation on non-random samples

3.5.1 Large/Small size

Small firms are in the size decile 1 with the smallest market value in September of year t, whereas large firms are in the size decile 10 with the largest market value. **Tables 3.11-3.12** report the rejection frequency for small firms under the measurement of CARs and BHARs over three investment horizons in all five models. Regarding CARs, all models generate close to 100% rejection rates, which indicates more severe misspecification compared with random samples. For example, the rejection rates are 99.6% when applying the market-adjusted model and 99.2% when applying the capital asset pricing model in five years. The market model is still the most misspecified model with the highest rejection rates of 100% in all three investment horizons. Moreover, there are positively biased test statistics with rejection rates close to 100% on the right tail for all models. Additionally, the misspecification in two factor models is the most severe in three years but the least severe in five years. The mean of CARs increases with the length of investment horizons. For instance, the market-adjusted model has mean CARs of 27.5% in one year, 54.7% in three years and 68.1% in five years. The magnitude of mean CARs is significantly higher compared with random samples. Furthermore, the means of test statistics under the null hypothesis of zero mean CAR in all models are larger than 3 with high variability. Regarding BHARs, the misspecification is improved significantly with lower rejection rate but still exists. The market-adjusted model generates a rejection rate of 98.8% in CARs and 50.4% in BHARs in one year. The market model is still the worst performer with the highest rejection rates. The test statistics are positively biased but with relatively lower rejection rates on the right tail. It is worth mentioning that the means of BHARs when applying the market model and four-factor model are highly abnormal with large value in five years. This could be the issue of random sampling which selects stocks with high raw returns compounded to achieve much higher returns.

The percentages of rejection rates with the application of five models for large firms are reported in **Tables 3.13-3.14**. The rejection rates under CARs are much smaller than small firms. Surprisingly, the severe misspecification is still documented in large firms over all investment horizons. The market-adjusted model yields a rejection rate of 42.2% in one year. The market model still generates the most severe misspecification. Although BHARs improves the misspecification from CARs, the misspecification is indicated. Some findings distinguish the large firms from small firms. The first is that test statistics under both measurements are negatively biased except the case when applying the market model in five years. Secondly, the means of CARs and BHARs are negative or zero for large firms whereas the corresponding figures are positive for small firms. Thirdly, when applying BHARs, the market model, unexpectedly demonstrates a significant fall in rejection rates from 32.8% in one year to 4% in five years, while the market-adjusted model shows an uptrend of rejection from 16.4% in one year to 90% in five years. Moreover, the three-factor model, as well as four-factor model, show severe misspecification in three years and five years but are well-specified in one year with a rejection rate of 4.8%. When compared with the three-factor model, the four-factor model improves misspecification in longer investment horizons. Fourthly, apart from the market model, the severity of misspecification increases with the length of investment horizons for large firms. However, the severest misspecification is found for small firms when the investment horizon is three years. Last but not least, the magnitude of mean BHARs is smaller than CARs with the application of the market-adjusted model, CAPM and three-factor model. The converse situation exists when applying the market model and four-factor model.

Since the models in the studies do not generate specified test statistics for large firms as prior research, it is worth examining firms in other size deciles apart from size 1 and 10. The empirical results with the application of the market-adjusted model and asset pricing model in BHARs at a significance level of 5% on both one-tailed and two-tailed tests are presented in **Table 3.15**. Regarding the market-adjusted model, firms in the size decile 4 and 5 show specified results over three holding periods whereas other deciles indicate misspecification, particularly size 1, 2, 9 and 10. For size decile 4, the rejection rate is 8% in one year, 4.4% in three years and 7.6% in five years. Similar results are documented for size decile 5. However, size decile 1 shows a rejection rate of 30.4% in five years. Interestingly, the corresponding figure for size decile 2 is 60%. Furthermore, size deciles with large firms exhibit large rejection rates. For instance, size 9 presents a rejection rate of 56% while size 10 shows a rejection rate 97.2% in five years. The mean BHARs decrease gradually from positive to negative figures from size 1 to 10. In the holding period of five years, the mean BHARs is 1.1% for size 5 but turns into red from size 6 (-1.3%) to size 10 (-15.9%). It is interesting to note that the asymmetry of BHARs shift from positive bias to negative bias with size. BHARs presents a positive bias with higher rejection rates on the right tail for size 1, 2 and 3 but display a negative bias for the rest sizes when the holding period is five years. With respect to the capital asset pricing model, similar results are found as the market-adjusted model. However, the rejection rates are slightly smaller in most cases compared with the findings when applying the market-adjusted model. The misspecification documented in both small and large firms when applying the UK data can be attributed to the resampling of firms. The small and large firms continue business longer than the medium firms over the period of 1982 to 2008. Based on the simulation procedure established by Kothari and Warner (1997), firms are selected with replacement, which indicates that a firm may appear more than once with multiple event months. Therefore, the repetition of small or large firms is more frequent than the medium firms. Due to the size effect, the misspecification with high rejection rates for small firms is explainable, especially when firms with high returns are selected several times. However, the misspecification of large firms reveals the fact that large returns are resampled several times in some circumstances.

3.5.2 High/Low book-to-market ratio (BTM)

Firms with low book-to-market ratios are those in BTM decile 1, while firms with high book-to-market ratios are those in BTM decile 10 based on market value in September of each year t and a 6-month lagged book value. **Tables 3.16-3.17** reveal the rejection rates for firms with the lowest BTM. I document the following findings. First, same as the random samples, the market model generates the most misspecified test statistics which yield rejection rates close to 100% over three investment horizons in CARs. However, the measurement of BHARs significantly eliminates the misspecification in the market model with the length of holding periods. The rejection rate is 88% in one year and 6.4% in five years at a significance level of 5%. Second, the market-adjusted model, surprisingly, yields more misspecified test statistics with the measurement of BHARs when compared with CARs. However, BHARs improve the misspecification for other models as the random samples. The rejection rate is 85.2% in five years under CARs and is 94.4% under BHARs. Third, all means of CARs and BHARs are negative and close to zero except the case when the four-factor model is applied in five years. The abnormally high mean BHARs can be attributed to the random sampling with firms with high raw returns. Last but not least, abnormal returns, regardless of the measurements, are negatively biased with higher rejection rates on the left tail. This is in line with the findings documented by Kothari and Warner (1997) but with much higher rejection rates. The results of severe misspecification with much higher rejection rates when applying the pricing models based on the UK data, compared with the results found by Kothari and Warner (1997) are consistent with the analysis of the large firms in the previous section. Firms with the lowest BTM contain firms with high market values. Therefore, the high raw returns of large firms and the repetition of these firms contribute the high rejection rates. It is worth mentioning that Kothari and Warner (1997) define firms with low BTM as firms with BTM below 0.8 whereas I decile firms into ten groups according to their BTMs and take the group with the smallest BTM as firms with low BTM. **Tables 3.18-3.19** exhibit the rejection frequencies under the measurement of CARs and BHARs for a sample of firms with the highest BTM. Similar findings are documented as studies undertaken by Kothari and Warner (1997). First of all, the rejection rates are higher for firms with the highest BTM when compared with firms with the lowest BTM. Secondly, both CARs and BHARs are positively biased with higher rejection rates on the right tail. Last but most importantly, the means of CARs and BHARs are positive. The mean BHARs of 1500% with the application of the market model and 100% with the application of the four-factor model may be caused by the process of random sampling. It is important to note the rejection rates are higher based on the UK data compared with studies based on the US data. These findings are plausible since firms with the highest BTM contain small firms which suffer the size effect. To further examine the anomalies in the sample of firms with the highest BTM, I conduct the simulation process on the remaining BTM deciles with the application of the market-adjusted model and capital asset pricing model under BHARs. The results are displayed in **Table 3.20**. BTM 6 yields well specified test statistics over three holding periods when the market-adjusted model is applied. The absolute values of means of BHARs are the highest in both BTM 1 and 10. For instance, the means of BHARs are -29.1% and 30.1% in BTM 1 and BTM 10, respectively. The means of BHARs increase from negative values in BTM 1 to positive values in BTM10. Moreover, test statistics shifts from negative biased to positive biased with an increase in value of BTMs. Additionally, the large value of the means of test statistics in BTM 1 and BTM 10 confirms high rejection rates in the results. Similarly, the capital asset pricing model demonstrates a pattern of the highest rejection rates in BTM 1 and BTM 10. However, BTM 5 yields well-specified test statistics with rejection rates close to 5% at a significance level of 5% whereas BTM 6 bears with overrejection in three and five years. For both models, BTM 5, 6, 7 generate well-specified test statistics in one year. This indicates that the evidence of absence of abnormal returns for firms which are in the boundary which classify a value firm with high BTM and a growth firm with low BTM when the holding period is one year is apparent. Even when the holding period lengthens to three or five years, these firms have lower rejection rates than firms with the highest or lowest BTM. Since firms with the lowest BTM contain large firms with large market value, the anomaly of high rejection rates can be explained in an analogous fashion as large firms. Apart from the repetition of large firms in the simulation and large value of returns of large firms, large firms may also be allocated to other BTM deciles. This allocation is taken as another reason to explain the high rejection rates in other BTM deciles such as BTM 3.

3.6 Causes of misspecification

3.6.1 Distribution of abnormal returns

I examine the distributions of abnormal returns over three investment horizons by applying the method suggested by Kothari and Warner (1997). Instead of randomly selecting 250 samples of 200 firms with replacement, I draw 50,000 firms with event months randomly without replacement from 1982 to 2008. Five models are applied over the investment periods of one, three and five years. I obtain 50,000 CARs and BHARs. I then construct the distributions of long-run abnormal returns in order to compare the effect of skewness on the misspecification in CARs and BHARs. **Table 3.21** presents the empirical results. The means of CARs and BHARs increase with the length of investment horizons except for the case when applying the market-adjusted model in five years. The magnitude of mean BHARs is mostly higher than the magnitude of mean CARs. This is attributed to the compounding effect when applying BHARs. All models, regardless of the measurement of abnormal returns, show positive skewness of abnormal returns. Moreover, BHARs indicate severe right-skewed distribution with extremely large value of abnormal returns over the holding period. The compounding effect accumulates over a longer time horizon. The abnormally large mean CARs and BHARs, which differ from the previous studies, are attributed to random sampling. The market model generates the highest mean BHARs compared with other models. This explains the overrejection of the null hypothesis in random samples. The four-factor model factor yields with the second highest mean BHARs among all models. The negative value of medians for BHARs suggests that the nonparametric tests, such as sign test with an assumption of zero median, are not appropriate in their application to test the abnormal returns. This confirms the findings in random samples with the application of the Wilcoxon signed-rank test. BHARs generate more leptokurtic distributions than CARs. Although BHARs significantly violate the assumption of normal distribution, it still yields less misspecified results with lower rejection rates compared with CARs when conducting the simulation process on random samples, even when CARs uses standard errors based on the test period. This finding is inconsistent with prior research. Therefore, although BHARs is superior to adjust misspecification, test statistics are not reliable because BHARs show severe positively skewed and leptokurtic distribution.

3.6.2 Sample selection bias

The requirement of 25-month stock returns excludes firms which go public or are delisted during the estimation period or are listed over the test period. As suggested by Ritter (1991), the IPOs exhibit long-term underperformance after listing. The exclusion of new listed firms, therefore, yields higher abnormal returns, leading to an upwardly biased test statistics. To examine the level of sample selection bias, I follow the procedures introduced by Kothari and Warner (1997) to compare the mean returns over the test period with different requirements of the availability of return data over the estimation period. I start with no requirement of return data. In September 1982, I include only firms with nonmissing return observation in this month. Then the cumulative returns of individual stocks are computed in 1 month, one, two and three years. If a firm does not survive over the whole test period, the return data is tracked until it is delisted. Then the average return over the study period is computed. The same process is repeated in the next calendar month until October, 2007 when the study period is one year. Then the mean of the average cumulative returns is reported in **Table 3.22**. The number of observations of returns declines from 303,970 to 263,505 when there is no requirement on pre-event data and when 48-month pre-event returns are required. The grand means of cumulative returns over the test period increase with the length of the test period. When there is a much stricter requirement of survival period, the returns also increase with an implication of higher returns for survivors, regardless of the measurement of returns. This is in line with the findings by Kothari and Warner (1997). The survivor bias could potentially result in higher stock return, leading to an upward biased test statistics. Since the event-time approach based on models applies asset pricing models to achieve expected returns which mostly require 'clean' returns in the estimation period in order to estimate the coefficients in the regression, it is inevitable that restrictions exist regarding returns over the estimation period. Lyon and Barber (1997) propose another event-time approach with reference portfolios or a control firm as benchmarks to avoid the sample selection bias. However, because reference portfolios and control firms are matched with event firms with similar characteristics such as book-to-market ratio and size, this still requires firms with returns during the period when conducting the criteria. Therefore, the sample selection bias cannot be eliminated.

3.7 Summary

With the application of the UK data, this chapter duplicates the methodology, which estimate expected stock returns via pricing models, introduced by Kothari and Warner (1997) so as to detect the long-term abnormal stock returns. The original data from studies undertaken by Gregory, Tharyan and Huang (2009) is screened with criteria of firms with both positive market value and book value listed in the Main Market, the Alternative Investment Market and the Unlisted Securities Market. Since the robustness check requires book-to-market ratio, market value of a firm in September of each year and book value of a firm with six-month lag from September of each year are identified. The monthly stock returns are downloaded from London Stock Price Database. Particularly, the monthly risk-free rate from the same database is a proxy of three-month market rate of Treasury bills. Following the data requirement as suggested by Kothari and Warner (1997), firms in the population need to have consecutive 24-month prior-event returns. The stock return on the event date is also a must. The study period is from October, 1982 till September 2008. The simulation process is conducted by randomly selecting 250 samples of 200 firms. The post-event returns of a firm are tracked until it is delisted in order to detect the long-term abnormal returns over the investment horizons of one-, three- and five years. The expected returns are estimated based on different pricing models, namely, the market-adjusted model, market model, capital asset pricing model, Fama-French three-factor model and four-factor model. The null hypothesis is zero cross-sectional cumulative abnormal return (CAR) or buy-and-hold abnormal returns (BHAR) based on equal weights. If a test is well-specified, the rejection rate is 5% at a significance level of 5%. However, if a test is conservative if fewer than 12 null hypotheses are rejected whereas a test is anticonservative if more than 12 null hypotheses are rejected. The comparison of statistics based on standard errors estimated by stock returns over estimation window and event window is shown. Furthermore, the power of test with introduction of shocks is conducted to test the reliability of the test statistics. Apart from random samples, non-random samples such as firms with small/large size or low/high book-to-market ratio are examined. Potential causes of misspecification such as distributional properties of abnormal returns and sample selection bias which questions the data screening process of 25-month consecutive returns are documented.

Cumulative Abnormal Returns (CARs)

For random samples, CARs using standard errors estimated over the estimation period or test period both present severe misspecified results. However, when using the standard errors based on the test period, the misspecification is improved. The number of firms in the estimation period is constant due to the requirement of at least 25-month returns. However, the number of firms over the test period is changing over the time, as returns are tracked according to the availability of the data. Therefore, the improvement of misspecification when applying the standard error based on the test period is because the increase of standard deviation when the sample size is smaller. The market model under CARs yields the most misspecified test statistics with highest rejection rates. These findings are consistent with the studies carried out by Kothari and Warner (1997). When it comes to non-random samples of firms with high/low book-to-market ratio, all models yield misspecified test statistics. The rejection rates, especially for firms with low BTM, are much higher than the findings documented by Kothari and Warner (1997). The anomaly of firms with low BTM can be explained by large returns of firms with large market value and the resampling of these firms since theses large firms continue business longer than small firms. This conclusion is further supported by the evidence of rejection rates of ten BTM deciles which indicate BTM decile 5, 6 and 7 generate well-specified test statistics in one year. When the investment horizons lengthen to three and five years, the rejection rates for firms with medium BTM are smaller than firms in the two extreme ends. It is worth mentioning that Kothari and Warner (1997) define firms with low BTM as firms with BTM below 0.8 whereas I decile firms into ten groups according to their BTMs and take the BTM decile one with the lowest BTM as firms with low BTM. Small and large firms also show similar pattern as firms with low or high BTM. Regarding the distributional properties of CARs, the means increase with the length of investment horizons except for the case when applying the market-adjusted model in five years. All models show positive skewness of abnormal returns. This is potential cause of misspecification with higher rejection rates of null hypothesis of zero CARs. Additionally, the grand means of cumulative returns over the test period increase with the length of the test period. When there is a much stricter requirement of survival period, the returns also increase. This sample selection bias could potentially lead to an upward biased test statistics.

Buy-and-Hold Returns (BHARs)

The rejection rates in BHARs are significantly reduced compared with CARs in most models except for the market model. This is inconsistent with previous findings of severe misspecification in BHARs (Kothari and Warner, 1997). The results suggest BHARs can be applied effectively to reduce misspecification. However, the server asymmetric pattern of test statistics still exist regardless of models applied. This implies although BHARs using standard deviation based on the test period generate well-specified test statistics, it is not reliable. For non-random samples, the market model in most circumstances generates the most severe misspecification. Although the small size effect is significant with the highest rejection rates, large firms also overreject the null hypothesis of zero mean BHARs. This can be attributed to the issue of resampling which indicates repetition of the same firm which have high returns with different event months in one sample. Another finding is that small firms and firms with high BTM show positive skewness of abnormal returns. The similar level of misspecification for small firms and firms with high BTM may be due to the fact that some small firms fall into the category of high BTM. The anomalies of high rejection rates for firms with the lowest BTM are explained in an analogous fashion as large firms. I further examine the potential causes for misspecification. Non-random samples exhibit mostly positive skewness of abnormal returns. To construct the distribution of BHARs over three investment horizons with a random sample of 50,000 firms, BHARs show more severe positively skewed and fat-tailed distribution than CARs. This is proved by the Wilcoxon Signed-rank test which examines the medians of BHARs. Another cause of misspecification is the sample selection bias which suggests the requirements of pre-event returns lifts up the stock returns with survivors in the test period. The buy-and-hold returns show larger magnitude than the cumulative returns due to compounding effect.

To conclude, the market-adjusted model, capital asset pricing model and Fama-French three-factor model are advocated to be applied together with the measurement of BHARs to detect the long-term abnormal stock performance, especially in random samples. However, test statistics derived from these models are not reliable because BHARs show severe positively skewed and leptokurtic distribution.

Figures and Tables of Chapter 3

Table 3.1 Description of sample firms

This table reports the number of firms in the screening process. The sample firms applied in this study cover a period of 1982 till 2008. The original data (N1) with firms trading in the London Stock Exchange including alternative investment market and unlisted securities market is obtained from studies undertaken by Gregory, Tharyan and Huang (2009). To mitigate impact from financials due to different accounting treatment, financials are excluded. I keep only firms with both market value and book value (N3). For firms with zero market value or book value, I keep them in the sample and replace the value of 0 with 0.5. The last screened process is to keep firms with ordinary shares by cross-matching firms with U2 equals 0 in LspdU file from the London Stock Price Data (LSPD).

year	N1	mcap=0	mcap<0 &missing	bv=0	bv<0	missing bv	N2	N3	N4	N5 (all firms)
1982	1,081	58	-	-	3	-	3	1,078	3	1,075
1983	1,091	30	-	-	4	-	4	1,087	3	1,084
1984	1,118	43	-	-	4	-	4	1,114	3	1,111
1985	1,139	21	-	-	4	-	4	1,135	3	1,132
1986	1,361	49	-	-	9	-	9	1,352	3	1,349
1987	1,490	34	-	-	13	-	13	1,477	3	1,474
1988	1,582	31	-	-	18	-	18	1,564	3	1,561
1989	1,612	22	-	-	24	-	24	1,588	5	1,583
1990	1,595	34	-	-	41	-	41	1,554	6	1,548
1991	1,522	10	-	-	46	-	46	1,476	4	1,472
1992	1,454	18	-	-	58	-	58	1,396	4	1,392
1993	1,436	6	-	-	73	-	73	1,363	4	1,359
1994	1,579	2	-	-	90	-	90	1,489	4	1,485
1995	1,695	9	-	-	83	-	83	1,612	7	1,605
1996	1,797	3	-	-	90	-	90	1,707	8	1,699
1997	1,880	3	-	-	101	-	101	1,779	7	1,772
1998	1,908	6	-	-	105	-	105	1,803	7	1,796
1999	1,813	10	-	-	99	1	100	1,713	8	1,705
2000	1,746	2	-	1	118	1	119	1,627	10	1,617
2001	1,751	-	-	1	76	-	76	1,675	10	1,665
2002	1,761	-	-	-	93	-	93	1,668	9	1,659
2003	1,703	-	-	-	135	-	135	1,568	7	1,561
2004	1,709	-	-	-	170	-	170	1,539	7	1,532
2005	1,908	-	-	-	192	-	192	1,716	7	1,709
2006	2,076	-	-	-	191	-	191	1,885	7	1,878
2007	2,288	1	-	-	122	423	545	1,743	4	1,739

* N1: number of firms excluding financials from Gregory, Tharyan and Huang (2009); N2: number of firms without book value or market value or both; N3: number of firms with both book value and market value; N4: number of firms without ordinary shares; N5: number of finalized firms (excluding financials, with BV and MV, with ordinary shares)

Figure 3.1 Number of firms before and after screening

This graph shows the number of firms in three scenarios: firms after exclusion of financials, firms with both market value and book value after exclusion of financials, and firms in the final sample of the studies after the screening process. The study period ranges from 1982 to 2007 with returns tracking from 1980 to 2008. The starting point is the data excluding financials from the data applied by Gregory, Tharyan and Huang (2009). The data contains most of the non-financials with market value, book value and industry code trading in the main market, alternative investment market (AIM) and unlisted securities market (USM). Note that the USM only lasts from 1980 till 1996, whereas AIM starts from 1995. I initially keep only firms with both book value and market value in order to carry out a robustness check for small/large firms and firms with low/high book-to-market ratio. I then exclude firms without ordinary shares by cross-matching firms with U2 equals 0 in LspdU file from London Stock Price Data (LSPD).

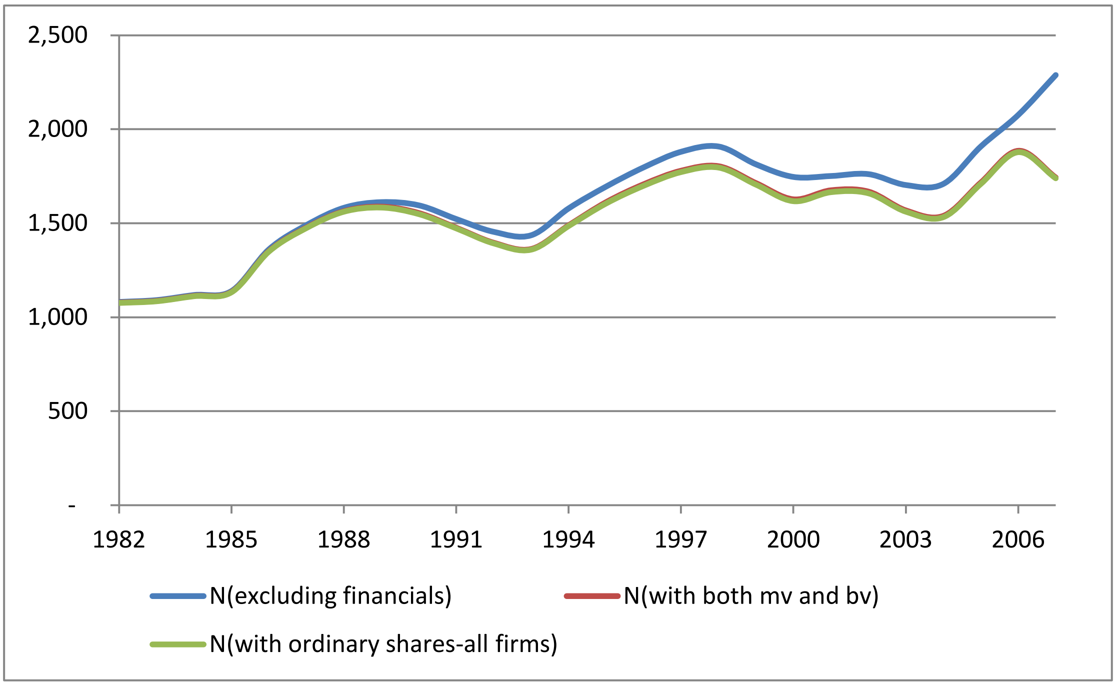


Table 3.2 Descriptive statistics of returns of screened sample

This table reports the descriptive statistics for the final sample of firms applied in the studies. The final sample has 4,977 firms from 1980 to 2008 with returns of 549,588. There are 19,572 missing returns which accounts for 3.6% of the total sample. There is one extreme positive outlier with a return observation of 4500% for the firm with g1 code as 9705 on January, 2000.

	Descriptive statistics
Mean	0.0125
Standard Error	0.0003
Standard Deviation	0.1916
Kurtosis	6580
Skewness	34.93
Range	46.71
Minimum	-0.9963
Maximum	45.72
Count	549,588

Table 3.3 Frequencies of returns

This table shows the distribution of raw stock returns of all firms. The bin values represent the range of values in the dataset of returns of the final sample. The first bin is -100% which is defined as the minimum stock return. The frequency of 1 in the first bin suggests there is one observation falling into this category. The second bin has values greater than -100% but smaller than or equal to -93%. There are 22 observations meeting this criterion. Same rules are applied to the rest. Approximately 50% of observations are in the range of 1% to 8%. There is one extreme positive outlier with a return observation of 4500% for the firm with g1 code as 9705 on January, 2000.

Bin1	Frequency1	Bin2	Frequency2	Bin3	Frequency3	Bin4	Frequency4
-1.00	1	0.71	737	2.41	11	4.36	3
-0.93	22	0.77	451	2.47	6	4.55	2
-0.87	45	0.83	319	2.53	11	4.61	1
-0.81	67	0.89	265	2.60	5	4.80	2
-0.74	138	0.96	155	2.66	2	4.87	1
-0.68	128	1.02	1,135	2.72	6	5.06	5
-0.62	384	1.08	94	2.79	4	5.37	1
-0.55	464	1.15	110	2.85	3	5.69	1
-0.49	2,049	1.21	77	2.91	4	6.00	2
-0.43	995	1.27	78	2.98	3	6.25	1
-0.37	2,448	1.34	95	3.04	25	6.51	1
-0.30	5,227	1.40	27	3.10	1	6.88	1
-0.24	9,030	1.46	35	3.16	6	7.01	2
-0.18	15,821	1.53	74	3.23	1	7.70	1
-0.11	34,238	1.59	20	3.29	1	7.83	1
-0.05	73,830	1.65	24	3.35	3	8.33	1
0.01	175,074	1.71	37	3.42	1	8.52	1
0.08	105,791	1.78	11	3.48	1	8.96	1
0.14	57,044	1.84	16	3.54	5	9.47	1
0.20	28,793	1.90	13	3.61	3	11.23	1
0.26	13,729	1.97	9	3.67	2	12.24	1
0.33	7,279	2.03	98	3.73	2	12.68	1
0.39	5,343	2.09	11	3.79	2	12.94	1
0.45	2,821	2.16	11	3.86	2	15.96	1
0.52	2,801	2.22	12	3.92	3	23.02	1
0.58	1,089	2.28	3	4.05	10	45.72	1
0.64	847	2.34	12	4.24	2		

Figure 3.2 Distribution of returns of the screened sample

This graph shows the distribution returns of the final sample of firms. The final sample contains 4,977 firms between 1980 and 2008 with returns observations of 549,588. There are 19,572 missing returns which accounts for 3.6% of the total sample. Approximately 50% of observations are in the range of 1% to 8%. There is one extreme positive outlier with a return observation of 4500% for the firm with g1 code as 9705 on January, 2000. The histogram illustrates the frequencies of returns observations, which suggests the distribution of raw returns.

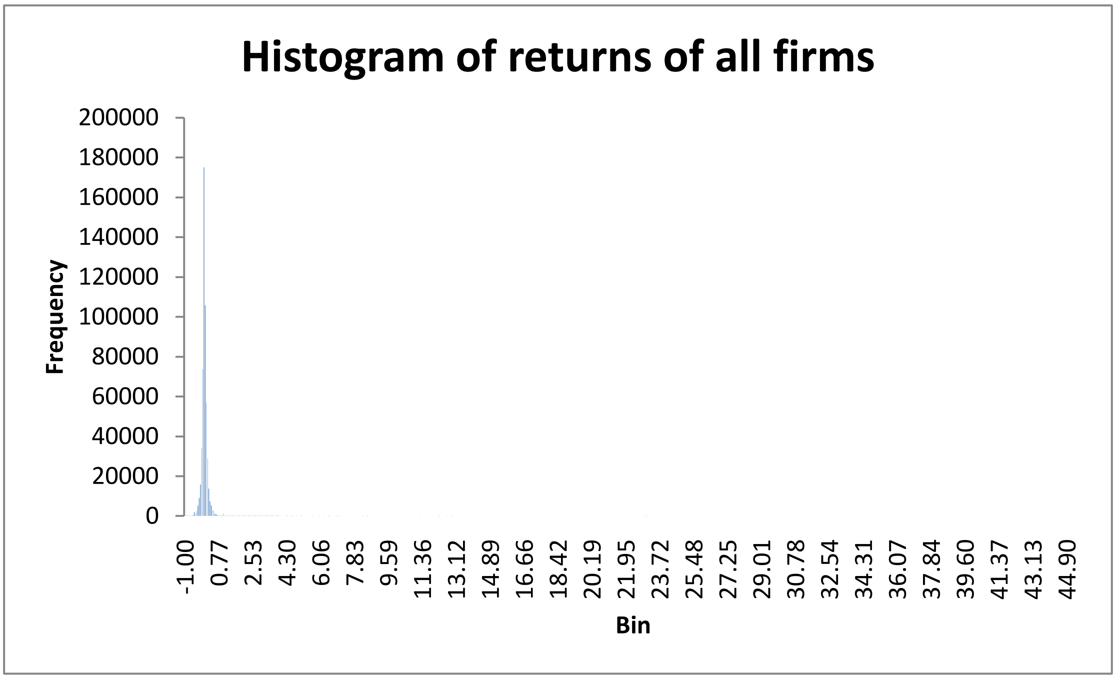


Table 3.4 Summary of event months and test periods for event-time approach and calendar-time approach

This table shows the starting and ending months over three investment horizons based on different approaches. I conduct a simulation process on the event-time approach and calendar-time approach following studies undertaken by Kothari and Warner (1997), Lyon, Barber and Tsai (1999) and Ang and Zhang (2004). All methods share the same test period from October, 1982 to September, 2008 based on the UK data. The event-time approach includes event-month return in the test period while the calendar-time approach tracks returns over the test period after the event month.

Methodology			**First event month**	**Last event month**	**Test period**	**Notes**
Event-time approach (1982.10-2008.10)	Kothari&Warner (1997) Period:1980.1-1989.12	One year	Oct, 1982	Oct, 2007	Oct, 1982-Sep, 2008	Prior 24-month returns plus event month return are required.
		Three years	Oct, 1982	Oct, 2005	Oct, 1982-Sep, 2008	
		Five years	Oct, 1982	Oct, 2003	Oct, 1982-Sep, 2008	
	Lyon, Barber&Tsai (1999) Period: 1973.7-1994.12	One year	Oct, 1982	Oct, 2007	Oct, 1982-Sep, 2008	Reference portfolios are formed based on market value and book value in September each year.
		Three years	Oct, 1982	Oct, 2005	Oct, 1982-Sep, 2008	
		Five years	Oct, 1982	Oct, 2003	Oct, 1982-Sep, 2008	
Calendar-time approach (1982.10-2008.10)	Ang&Zhang(2004) Period: 1980.1-1992.12	One year	Sep, 1982	Sep, 2007	Oct, 1982-Sep, 2008	Event month return is not included in the simulation.
		Three years	Sep, 1982	Sep, 2005	Oct, 1982-Sep, 2008	
		Five years	Sep, 1982	Sep, 2003	Oct, 1982-Sep, 2008	

Table 3.5 Rejection frequency of the null hypothesis of zero mean cumulative abnormal returns (CARs) - Standard error over the estimation period

This table reports the rejection frequency of the null hypothesis of zero mean cumulative abnormal returns in 250 samples over the period of 1982 to 2008. Both one-tailed and two-tailed tests are examined at significance levels of 1% and 5%. To begin with, a firm is excluded from the population if it does not have 25 monthly returns. Then 250 firms are selected randomly. For each firm, I randomly select a random month between October 1982 and October 2007 for one year, October 1982 to October 2005 for three years and October 1982 to October 2003 for five years. The estimation period is defined as the previous 24 months before the event month. The test period starts from the event month. The investment horizons are one, three and five years. If a firm does not survive the whole test period, I only track the available returns. Therefore, the missing returns remain the same. The returns over the estimation period are used to estimate the coefficients in the models through regression. I employ four models including the market-adjusted model, market model, Fama-French three-factor model and Carhart four-factor model. The abnormal returns are measured by cumulative abnormal returns (CARs). The standard error in test statistics is estimated based on the abnormal returns over the 24-month estimation period. CARs are calculated as $\mathrm{CAR}_{p,T} = \sum_{t=0}^{T-1} \mathrm{AR}_{p,t}$, where $\mathrm{AR}_{p,t} = \frac{\sum_{i=1}^{N} \mathrm{AR}_{i,t}}{N}$. The test statistics to test the null hypothesis of zero mean CARs is: $t = \frac{\mathrm{CAR}_{P,T}}{\sigma(\mathrm{CAR}_{p,t}) * \sqrt{T}}$., When the standard deviation of CARs is based on estimation period, it is shown as: $\sigma(\mathrm{CAR}_{p,t}) = \sqrt{\sum_{t=-24}^{-1}\left(\mathrm{CAR}_{p,t} - \overline{\mathrm{CAR}_{p,t}}\right)^2/23}$

Where $\overline{\mathrm{CAR}_{p,t}} = \frac{1}{24}\sum_{t=-24}^{-1} \mathrm{CAR}_{p,t}$.

Model	1%			5%			Mean CARs	SD of CARs	Mean Tests	SD of Tests
	α=0	α<0	α>0	α=0	α<0	α>0				
	Panel A: one year									
Market-adjusted model	4.0	2.4	3.2	12.4	7.6	9.6	0.003	0.050	0.068	1.287
Market model	14.8	2.8	14.0	26.0	5.2	26.4	0.026	0.060	0.682	1.577
CAPM	5.2	3.2	4.4	13.6	8.8	13.6	0.005	0.054	0.115	1.424
FF three-factor model	12.0	6.4	8.8	20.0	10.0	15.2	0.003	0.058	0.073	1.581
Carhart four-factor model	15.2	8.4	12.4	28.0	16.4	19.6	0.006	0.066	0.161	1.841

Table 3.5 continued

Model	1%			5%			Mean CAR	SD of CAR	Mean Tests	SD of Tests
	α=0	α<0	α>0	α=0	α<0	α>0				
	Panel B: three years									
Market-adjusted model	8.0	0.0	12.0	16.8	0.8	23.6	0.062	0.081	0.907	1.199
Market model	26.4	4.0	29.2	40.0	10.0	41.2	0.078	0.133	1.278	2.120
CAPM	9.2	0.8	11.2	18.0	3.2	24.8	0.044	0.088	0.689	1.381
FF three-factor model	16.0	1.6	18.8	26.0	5.6	31.2	0.054	0.098	0.879	1.569
Carhart four-factor model	22.4	3.6	22.8	35.2	7.6	35.2	0.050	0.115	0.861	1.915
	Panel C: five years									
Market-adjusted model	14.8	0.4	18.8	24.8	1.2	32.8	0.095	0.109	1.128	1.347
Market model	39.6	6.0	37.2	50.0	10.0	46.0	0.122	0.209	1.614	2.662
CAPM	14.4	1.2	15.6	24.0	3.6	27.2	0.065	0.121	0.824	1.542
FF three-factor model	19.2	1.6	22.8	30.0	4.4	34.4	0.085	0.127	1.138	1.715
Carhart four-factor model	28.8	3.6	31.2	44.4	6.8	43.6	0.102	0.155	1.394	2.194

Note that the models applied are listed below:

Market-adjusted model: $AR_{i,t} = R_{i,t} - R_{m.t}$;

Market model: $AR_{i,t} = R_{i,t} - (\alpha_i + \beta_i R_{m,t})$;

Capital asset pricing model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i (R_{m,t} - R_{f,t})\}$

Fama-French three-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i (R_{m,t} - R_{f,t})\}$;

Carhart four-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_{i,1} RPM_t + \beta_{i,2} HML_t + \beta_{i,3} SMB_t\}$

$R_{m.t}$ is denoted as market return which is measured as an equally-weighted market portfolio return, α_i and β_i are achieved by regressing monthly returns for security i on $R_{m,t}$ over the estimation period. $R_{f,t}$ is the risk-free rate which uses three-month market rates of Treasury bills as a proxy. HML_t is the difference between returns on portfolios of high and low book-to-market-ratio securities, whilst SMB_t represents the difference between returns on portfolios of small and large market value securities. RPM_t is the difference between returns on a market index and risk-free investment. β_s are beta factors corresponding to different control factors and are achieved by regression over the estimation period.

Table 3.6 Rejection frequency of the null hypothesis of zero mean cumulative abnormal returns (CARs) - Standard error over the test period

This table reports the rejection frequency of the null hypothesis of zero mean cumulative abnormal returns in 250 samples over a period of 1982 to 2008. Both one-tailed and two-tailed tests are examined at significance levels of 1% and 5%. To begin with, a firm is excluded from the population if it does not have 25 month returns. Then 200 firms are selected randomly. For each firm, I randomly select a month over the period of October 1982 to October 2007 for one year, October 1982 to October 2005 for three years and October 1982 to October 2003 for five years. The process is repeated 250 times. The estimation period is the previous 24 months before the event month. The test period starts from the event month. The investment horizons are one, three and five years. If a firm does not survive the whole test period, I only track the available returns. The returns over the estimation period are used to estimate the coefficients in the models through regression. I employ four models including the market-adjusted model, market model, Fama-French three-factor model and Carhart four-factor model. The standard error in test statistics is estimated based on the abnormal returns over the estimation period.

CARs are calculated as $\mathrm{CAR}_{p,T} = \sum_{t=0}^{T-1} \mathrm{AR}_{p,t}$ where $\mathrm{AR}_{p,t} = \frac{\sum_{i=1}^{N} \mathrm{AR}_{i,t}}{N}$.

The test statistics to test the null hypothesis of zero mean CARs is: $t = \frac{\mathrm{CAR}_{P,T}}{\sigma(\mathrm{CAR}_{p,t}) * \sqrt{T}}$,

When the standard deviation of CARs is based on the test period, it is shown as: $\sigma(\mathrm{CAR}_{p,t}) = \sqrt{\sum_{t=-24}^{-1}\left(\mathrm{CAR}_{p,t} - \overline{\mathrm{CAR}_{p,t}}\right)^2 / 23}$

Where $\overline{\mathrm{CAR}_{p,t}} = \frac{1}{24}\sum_{t=-24}^{-1} \mathrm{CAR}_{p,t}$.

Model	**1%**			**5%**			**Mean CARs**	**SD of CARs**	**Mean Tests**	**SD of Tests**
	α=0	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**				
						Panel A: one year				
Market-adjusted model	2.4	3.6	2.0	8.8	8.0	5.6	0.003	0.050	-0.015	1.187
Market model	7.6	2.8	8.4	19.2	6.0	20.8	0.026	0.060	0.529	1.463
CAPM	4.0	2.8	2.8	8.8	8.4	8.8	0.005	0.054	0.032	1.253
FF three-factor model	4.4	3.2	2.0	12.8	10.0	10.8	0.003	0.058	0.015	1.275
Carhart four-factor model	4.8	4.4	4.8	15.6	10.8	11.2	0.006	0.066	0.075	1.383

Table 3.6 continued

Model	1% α=0	1% α<0	1% α>0	5% α=0	5% α<0	5% α>0	Mean CAR	SD of CAR	Mean Tests	SD of Tests
					Panel B: three years					
Market-adjusted model	1.6	0.0	3.6	6.0	0.0	13.6	0.062	0.081	0.659	0.896
Market model	12.8	2.0	14.4	26.0	5.2	27.2	0.078	0.133	0.821	1.512
CAPM	2.0	0.0	2.8	6.0	1.2	10.0	0.044	0.088	0.449	0.980
FF three-factor model	2.8	0.0	4.8	8.4	1.2	14.0	0.054	0.098	0.527	1.011
Carhart four-factor model	3.2	1.2	6.8	11.2	2.4	15.2	0.050	0.115	0.486	1.162
					Panel C: five years					
Market-adjusted model	1.6	0.0	3.6	9.2	0.0	18.8	0.095	0.109	0.773	0.896
Market model	22.4	4.4	22.4	33.6	6.8	34.4	0.122	0.209	0.963	1.765
CAPM	2.0	0.0	3.2	8.8	2.0	13.6	0.065	0.121	0.505	0.983
FF three-factor model	2.0	0.0	3.6	7.6	0.8	12.4	0.085	0.127	0.620	0.943
Carhart four-factor model	3.2	0.4	6.0	13.2	2.4	17.6	0.102	0.155	0.717	1.113

Note that the models applied are listed below:

Market-adjusted model: $AR_{i,t} = R_{i,t} - R_{m.t}$;

Market model: $AR_{i,t} = R_{i,t} - (\alpha_i + \beta_i R_{m,t})$;

Capital asset pricing model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$

Fama-French three-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$;

Carhart four-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_{i,1}RPM_t + \beta_{i,2}HML_t + \beta_{i,3}SMB_t\}$

$R_{m.t}$ is denoted as market return which is measured as an equally-weighted market portfolio return, α_i and β_i are achieved by regressing monthly returns for security i on $R_{m,t}$ over the estimation period. $R_{f,t}$ is the risk-free rate which uses three-month market rates of Treasury bills as a proxy. HML_t is the difference between returns on portfolios of high and low book-to-market-ratio securities, whilst SMB_t represents the difference between returns on portfolios of small and large market value securities. RPM_t is the difference between returns on a market index and risk-free investment. β_s are beta factors corresponding to different control factors and are achieved by regression over the estimation period.

Table 3.7 Specification of Wilcoxon signed-rank test in random samples

This table reports the rejection frequency in 250 samples of 200 firms when applying the Wilcoxon signed-rank test which is a combination of sign test and rank test. It bears the null hypothesis of an equal number of positive and negative returns under a framework of a binomial test. Firstly, the absolute values of 200 CARs are ranked in each sample. Rank 1 is with firms with the lowest CARs. Then signs are allocated to observations. Afterwards, the additions of ranks with a positive sign and with a negative sign are computed. Following this, the sum of these two figures is calculated for the purpose of determining whether the null hypothesis of zero median is rejected or not. The process is expressed as follows:

$$\mathrm{SR_N} = \sum_{\mathrm{i}} \mathrm{CAR_i^+}$$

Where $$\mathrm{SR_N} \sim \mathrm{N}(\mathrm{E}(\mathrm{SR_N}), \sigma^2(\mathrm{SR_N}))$$

$\mathrm{E}(\mathrm{SR_N}) = \mathrm{N}(\mathrm{N}+1)/4)$

$$\sigma^2(\mathrm{SR_N}) = \mathrm{N}(\mathrm{N}+1)(2\mathrm{N}+1)/24$$

$\mathrm{CAR_i^+}$ is the positive rank of the absolute value of CARs. When N is large, test statistics follows a normal distribution. N is denoted as the total number of firms. The test statistics Z are calculated based on mean and variance in order to compare with critical values at a significance level of 5% on both one-sided and two-sided tests. For each sample, I achieve one Z. The number of Z totals 250 over one-, three- and five-year holding periods. The rejection frequency among 250 samples is computed.

5%	**CARs**		
	α=0	**α<0**	**α>0**
	Panel A: One year		
Market-adjusted model	6.0	5.6	4.4
Market model	92.0	0.0	92.0
CAPM	4.4	3.2	6.8
FF three-factor model	6.0	4.4	8.0
Carhart four-factor model	5.2	5.2	5.6
	Panel B: Three years		
Market-adjusted model	10.8	0.0	17.6
Market model	94.0	0.0	94.0
CAPM	9.2	0.4	13.6
FF three-factor model	10.0	0.0	17.2
Carhart four-factor model	10.0	0.4	18.0
	Panel C: Five years		
Market-adjusted model	20.0	0.0	28.0
Market model	94.4	0.0	94.4
CAPM	19.6	0.0	24.0
FF three-factor model	16.8	0.0	26.8
Carhart four-factor model	14.4	0.8	20.8

Table 3.8 Rejection frequency of the null hypothesis of zero mean buy-and-hold abnormal returns (BHARs) - Standard error over the test period

This table reports the rejection frequency of the null hypothesis of zero mean buy-and-hold abnormal returns in 250 samples over a period of 1982 to 2008. Both one-tailed and two-tailed tests are examined at significance levels of 1% and 5%. To begin with, a firm is excluded from the population if it fails to have 25 month returns. Then 200 firms are selected randomly. For each firm, I randomly select a random month over the period of October 1982 to October 2007 for one year, October 1982 to October 2005 for three years and October 1982 to October 2003 for five years. The process is repeated 250 times. The estimation period is the previous 24 months before the event month. The test period starts from the event month. The investment horizons are one, three and five years. If a firm does not survive the whole test period, I only track the available returns. The returns over the estimation period are used to estimate the coefficients in the models through regression. I employ four models including market-adjusted model, market model, Fama-French three-factor model and Carhart four-factor model. The abnormal returns are measured by buy-and-hold returns (BHARs). The standard error in test statistics are estimated based on the abnormal returns over the estimation period. BHARs are calculated as: $\mathrm{BHAR}_{p,T} = \frac{\sum_{i=1}^{N} \mathrm{BHAR}_{i,T}}{N}$, where $\mathrm{BHAR}_{i,T} = \prod_{t=0}^{T-1}(1 + \mathrm{BHAR}_{i,t}) - 1$.

The test statistics to test the null hypothesis of zero mean BHARs is: $t = \frac{\mathrm{BHAR}_{P,T}}{\sigma(\mathrm{BHAR}_{p,t})}$.

The standard deviation of BHARs is based on test period as: as: $\sigma(\mathrm{BHAR}_{p,t}) = \frac{1}{199}\sqrt{\sum_{t=0}^{T-1}(\mathrm{BHAR}_{i,T} - \mathrm{BHAR}_{p,T})^2/(T-1)}$.

	1%			**5%**						
Model	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**	**Mean BHARs**	**SD of BHARs**	**Mean Tests**	**SD of Tests**
						Panel A: one year				
Market-adjusted model	2.0	2.8	0.0	6.8	12.0	1.6	-0.004	0.053	-0.347	1.079
Market model	2.8	0.0	7.2	19.2	0.4	32.8	0.087	0.083	1.225	0.852
CAPM	1.6	1.6	1.2	4.8	5.6	2.8	0.014	0.072	0.021	1.034
FF three-factor model	1.2	1.2	0.8	6.0	5.6	5.6	0.018	0.069	0.092	1.035
Carhart four-factor model	0.0	0.0	1.6	4.8	0.8	12.0	0.130	0.329	0.740	0.885

Table 3.8 continued

Model	1%			5%			Mean BHARs	SD of BHARs	Mean Tests	SD of Tests
	α=0	α<0	α>0	α=0	α<0	α>0				
							Panel B: three years			
Market-adjusted model	1.2	2.0	0.4	4.4	7.6	2.4	0.017	0.083	-0.023	1.024
Market model	23.2	0.0	38.8	62.4	0.0	78.4	0.639	3.211	2.159	0.628
CAPM	0.4	1.2	0.4	3.6	2.8	4.4	0.044	0.128	0.230	0.972
FF three-factor model	0.4	1.2	0.0	6.0	4.0	8.0	0.039	0.258	0.353	1.021
Carhart four-factor model	2.0	0.4	4.4	12.8	0.4	23.2	0.185	0.243	1.100	0.840
							Panel C: five years			
Market-adjusted model	3.6	5.2	0.0	8.8	11.2	1.2	-0.005	0.076	-0.299	1.128
Market model	22.8	0.0	35.2	60.8	0.0	73.2	4.134	6.640	2.140	0.662
CAPM	1.2	1.2	0.0	4.4	4.0	6.4	0.057	0.162	0.193	1.007
FF three-factor model	1.2	0.8	0.4	3.6	5.6	6.0	-0.005	0.774	0.209	1.002
Carhart four-factor model	2.4	0.8	3.2	14.8	0.8	24.4	0.835	8.059	1.098	0.859

Note that the models applied are listed below:

Market-adjusted model: $AR_{i,t} = R_{i,t} - R_{m.t}$;

Market model: $AR_{i,t} = R_{i,t} - (\alpha_i + \beta_i R_{m,t})$;

Capital asset pricing model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$

Fama-French three-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$;

Carhart four-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_{i,1} RPM_t + \beta_{i,2} HML_t + \beta_{i,3} SMB_t\}$

$R_{m.t}$ is denoted as market return which is measured as an equally-weighted market portfolio return, α_i and β_i are achieved by regressing monthly returns for security i on $R_{m,t}$ over the estimation period. $R_{f,t}$ is the risk-free rate which uses three-month market rates of Treasury bills as a proxy. HML_t is the difference between returns on portfolios of high and low book-to-market-ratio securities, whilst SMB_t represents the difference between returns on portfolios of small and large market value securities. RPM_t is the difference between returns on a market index and risk-free investment. β_s are beta factors corresponding to different control factors and are achieved by regression over the estimation period.

Table 3.9 Specification of Wilcoxon signed-rank test in random samples

This table reports the rejection frequency in 250 samples of 200 firms when applying the Wilcoxon signed-rank test which is a combination of sign test and rank test. It bears the null hypothesis of an equal number of positive and negative returns under a framework of a binomial test. Firstly, the absolute values of 200 BHARs are ranked in each sample. Rank 1 is with firms with the lowest BHARs. Then signs are allocated to observations. Afterwards, the additions of ranks with a positive sign and with a negative sign are computed. Following this, the sum of these two figures is calculated for the purpose of determining whether the null hypothesis of zero median is rejected or not. The process is expressed as follows:

$$\mathrm{SR_N} = \sum_{\mathrm{i}} \mathrm{BHAR_i^+}$$

Where $\mathrm{SR_N} \sim \mathrm{N}(\mathrm{E}(\mathrm{SR_N}), \sigma^2(\mathrm{SR_N}))$

$\mathrm{E}(\mathrm{SR_N}) = \mathrm{N}(\mathrm{N}+1)/4)$

$$\sigma^2(\mathrm{SR_N}) = \mathrm{N}(\mathrm{N}+1)(2\mathrm{N}+1)/24$$

$\mathrm{BHAR_i^+}$ is the positive rank of the absolute value of BHARs. When N is large, test statistics follows a normal distribution. N is denoted as the total number of firms. The test statistics Z are calculated based on mean and variance in order to compare with critical values at a significance level of 5% on both one-sided and two-sided tests. For each sample, I achieve one Z. The number of Z totals 250 over one-, three- and five-year holding periods. The rejection frequency among 250 samples is computed.

5%	**BHARs**		
	α=0	**α<0**	**α>0**
	Panel A: One year		
Market-adjusted model	44.4	60.4	0.0
Market model	14.0	22.8	0.0
CAPM	38.8	55.2	0.0
FF three-factor model	41.2	55.6	0.0
Carhart four-factor model	35.2	43.2	0.0
	Panel B: Three years		
Market-adjusted model	60.4	74.0	0.0
Market model	3.6	7.2	2.0
CAPM	54.4	64.0	0.0
FF three-factor model	58.4	70.4	0.0
Carhart four-factor model	45.2	54.8	0.0
	Panel C: Five years		
Market-adjusted model	77.2	85.2	0.0
Market model	5.6	5.6	6.4
CAPM	65.6	76.4	0.0
FF three-factor model	73.2	80.8	0.0
Carhart four-factor model	57.2	68.0	0.4

Table 3.10 Power of test: market-adjusted model, asset pricing model and Fama-French three-factor model under BHARs

This table reports power of test when abnormal returns are induced with the application of the market-adjusted model, asset pricing model and Fama-French three-factor model under BHARs. Abnormal returns are introduced in the range of -20% to 20% with an interval of 5% to individual stocks over the holding periods. If the test is superior, the percentage of rejection rates is expected to be 100% when there is an existence of abnormal returns.

Abnormal returns	**-20%**	**-15%**	**-10%**	**-5%**	**0%**	**5%**	**10%**	**15%**	**20%**
Models					**Panel A: One year**				
Market-adjusted model	98.8	95.6	65.6	13.2	6.8	41.2	76.0	88.4	94.4
Capital asset pricing model	98.0	96.0	72.4	17.2	4.8	23.2	62.0	82.8	90.8
Fama-French three-factor model	98.0	95.2	65.2	17.2	6.0	19.2	49.6	74.4	87.6
					Panel B: Three years				
Market-adjusted model	95.2	76.0	32.4	6.8	4.4	18.0	39.2	62.8	76.4
Capital asset pricing model	95.6	74.8	37.6	10.4	3.6	12.4	30.0	51.6	67.2
Fama-French three-factor model	92.0	72.8	37.6	14.8	6.0	9.6	24.0	38.8	54.8
					Panel C: Five years				
Market-adjusted model	86.0	53.6	20.0	6.0	8.8	22.0	46.0	65.6	74.8
Capital asset pricing model	85.2	56.0	24.4	9.6	4.4	10.8	24.0	44.0	61.6
Fama-French three-factor model	80.8	49.2	22.8	8.4	3.6	9.6	20.8	37.2	56.0

Table 3.11 Rejection frequency of the null hypothesis of zero mean CARs in small firms - Standard error over the estimation period

This table reports the rejection frequency of the null hypothesis of zero mean cumulative abnormal returns in 250 samples over a period of 1982 to 2008. Both one-tailed and two-tailed tests are examined at significance levels of 1% and 5%. A firm is excluded from the population if it does not have 25 month returns. Then 200 small firms in size decile 1 with the smallest market value are selected randomly. For each firm, I randomly select a random month between October 1982 and October 2007 for one year, October 1982 to October 2005 for three years and October 1982 to October 2003 for five years. The process is repeated 250 times. The estimation period is the previous 24 months before the event month. The test period starts from the event month. The investment horizons are one, three and five years. If a firm does not survive the whole test period, I only track the available returns. The returns over the estimation period are used to estimate the coefficients in the models through regression. I employ four models including the market-adjusted model, market model, Fama-French three-factor model and Carhart four-factor model. The abnormal returns are measured by cumulative abnormal returns (CARs). The standard error in test statistics is estimated based on the abnormal returns over the 24-month estimation period.

CARs are calculated as $\mathrm{CAR}_{p,T} = \sum_{t=0}^{T-1} \mathrm{AR}_{p,t}$ where $\mathrm{AR}_{p,t} = \frac{\sum_{i=1}^{N} \mathrm{AR}_{i,t}}{N}$. The test statistic to test the null hypothesis of zero mean CARs is: $t = \frac{\mathrm{CAR}_{P,T}}{\sigma(\mathrm{CAR}_{p,t}) * \sqrt{T}}$.,

When the standard deviation of CARs is based on estimation period, it is shown as: $\sigma(\mathrm{CAR}_{p,t}) = \sqrt{\sum_{t=-24}^{-1} \left(\mathrm{CAR}_{p,t} - \overline{\mathrm{CAR}_{p,t}}\right)^2 / 23}$

Where $\overline{\mathrm{CAR}_{p,t}} = \frac{1}{24} \sum_{t=-24}^{-1} \mathrm{CAR}_{p,t}$.

	1%			5%						
Model	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**	**Mean CARs**	**SD of CARs**	**Mean Tests**	**SD of Tests**
	Panel A: one year									
Market-adjusted model	84.0	0.0	92.8	98.8	0.0	100.0	0.275	0.067	3.842	1.237
Market model	99.2	0.0	99.2	100.0	0.0	100.0	0.390	0.080	5.758	1.731
CAPM	82.8	0.0	88.0	96.4	0.0	99.2	0.260	0.070	3.840	1.346
FF three-factor model	86.0	0.0	89.2	95.6	0.0	97.2	0.252	0.076	4.046	1.463
Carhart four-factor model	88.8	0.0	92.4	96.0	0.0	97.6	0.279	0.091	4.622	1.770

Table 3.11 continued

Model	1%			5%			Mean CAR	SD of CAR	Mean Tests	SD of Tests
	α=0	α<0	α>0	α=0	α<0	α>0				
					Panel B: three years					
Market-adjusted model	98.0	0.0	98.8	99.2	0.0	100.0	0.547	0.108	4.556	1.252
Market model	100.0	0.0	100.0	100.0	0.0	100.0	0.868	0.167	7.646	2.106
CAPM	98.0	0.0	98.0	98.8	0.0	99.6	0.536	0.120	4.701	1.384
FF three-factor model	98.0	0.0	98.8	99.6	0.0	100.0	0.561	0.130	5.417	1.677
Carhart four-factor model	97.6	0.0	98.4	99.6	0.0	100.0	0.593	0.143	5.908	1.876
					Panel C: five years					
Market-adjusted model	96.4	0.0	98.4	99.6	0.0	100.0	0.681	0.136	4.421	1.186
Market model	100.0	0.0	100.0	100.0	0.0	100.0	1.156	0.270	7.998	2.448
CAPM	97.2	0.0	98.0	99.2	0.0	99.6	0.680	0.154	4.674	1.398
FF three-factor model	97.2	0.0	98.8	99.2	0.0	99.6	0.687	0.168	5.222	1.542
Carhart four-factor model	94.8	0.0	96.8	98.4	0.0	98.8	0.748	0.227	5.891	2.082

Note that the models applied are listed below:

Market-adjusted model: $AR_{i,t} = R_{i,t} - R_{m.t}$;

Market model: $AR_{i,t} = R_{i,t} - (\alpha_i + \beta_i R_{m,t})$;

Capital asset pricing model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$

Fama-French three-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$;

Carhart four-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_{i,1}RPM_t + \beta_{i,2}HML_t + \beta_{i,3}SMB_t\}$

$R_{m.t}$ is denoted as market return which is measured as an equally-weighted market portfolio return, α_i and β_i are achieved by regressing monthly returns for security i on $R_{m,t}$ over the estimation period. $R_{f,t}$ is the risk-free rate which uses three-month market rates of Treasury bills as a proxy. HML_t is the difference between returns on portfolios of high and low book-to-market-ratio securities, whilst SMB_t represents the difference between returns on portfolios of small and large market value securities. RPM_t is the difference between returns on a market index and risk-free investment. β_s are beta factors corresponding to different control factors and are achieved by regression over the estimation period.

Table 3.12 Rejection frequency of the null hypothesis of zero mean BHARs in small firms - Standard error over the test period

This table reports the rejection frequency of the null hypothesis of zero mean buy-and-hold abnormal returns in 250 samples over the period of 1982 to 2008. Both one-tailed and two-tailed tests are examined at significance levels of 1% and 5%. To begin with, a firm is excluded from the population if it does not have 25 month returns. Then 200 small firms in size decile 1 with the smallest market value are selected randomly. For each firm, I randomly select a random month over a period of October 1982 to October 2007 for one year, October 1982 to October 2005 for three years and October 1982 to October 2003 for five years. The process is repeated 250 times. The estimation period is the previous 24 months before the event month. The test period starts from the event month. The investment horizons are one, three and five years. If a firm does not survive the whole test period, I only track the available returns. The returns over the estimation period are used to estimate the coefficients in the models through regression. I employ four models including market-adjusted model, market model, Fama-French three-factor model and Carhart four-factor model. The abnormal returns are measured by buy-and-hold returns (BHARs). The standard error in test statistics are estimated based on the abnormal returns over the estimation period. BHARs are calculated as: $\mathrm{BHAR}_{p,T} = \frac{\sum_{i=1}^{N}\mathrm{BHAR}_{i,T}}{N}$, where $\mathrm{BHAR}_{i,T} = \prod_{t=0}^{T-1}(1+\mathrm{BHAR}_{i,t}) - 1$.

The test statistic to test the null hypothesis of zero mean BHARs is: $t = \frac{\mathrm{BHAR}_{P,T}}{\sigma(\mathrm{BHAR}_{p,t})}$.

The standard deviation of BHARs is based on test period as: as: $\sigma(\mathrm{BHAR}_{p,t}) = \frac{1}{199}\sqrt{\sum_{t=0}^{T-1}\left(\mathrm{BHAR}_{i,T} - \mathrm{BHAR}_{p,T}\right)^2/(T-1)}$.

Model	1%			5%			Mean BHARs	SD of BHARs	Mean Tests	SD of Tests
	α=0	α<0	α>0	α=0	α<0	α>0				
						Panel A: one year				
Market-adjusted model	19.2	0.0	30.4	50.4	0.0	68.8	0.168	0.097	1.991	0.671
Market model	88.4	0.0	93.2	98.4	0.0	99.6	0.433	0.145	3.639	0.810
CAPM	15.2	0.0	30.4	44.8	0.0	62.4	0.191	0.124	1.938	0.683
FF three-factor model	13.6	0.0	25.2	42.0	0.0	62.4	0.192	0.131	1.833	0.706
Carhart four-factor model	20.0	0.0	30.0	54.0	0.0	68.0	0.464	0.523	1.985	0.717

Table 3.12 continued

Model	1% α=0	1% α<0	1% α>0	5% α=0	5% α<0	5% α>0	Mean BHARs	SD of BHARs	Mean Tests	SD of Tests
						Panel B: three years				
Market-adjusted model	10.8	0.0	21.6	40.4	0.4	63.2	0.285	0.180	1.832	0.646
Market model	54.4	0.0	64.0	82.0	0.0	92.4	1.130	12.769	2.739	1.045
CAPM	12.8	0.0	23.6	47.2	0.0	67.6	0.390	0.301	1.917	0.642
FF three-factor model	17.6	0.0	37.6	58.0	0.0	76.4	0.398	0.782	2.066	0.689
Carhart four-factor model	49.6	0.0	63.2	80.0	0.0	88.4	1.142	21.947	2.491	0.773
						Panel C: five years				
Market-adjusted model	7.2	0.0	16.0	30.4	0.0	49.6	0.256	0.177	1.611	0.759
Market model	41.2	0.0	52.0	64.4	0.0	73.6	235998	1599991	2.344	0.874
CAPM	7.2	0.0	12.0	23.6	0.0	45.2	0.518	0.465	1.640	0.627
FF three-factor model	10.0	0.0	17.2	34.0	0.0	55.6	0.510	0.460	1.718	0.708
Carhart four-factor model	20.0	0.0	30.0	48.4	0.0	62.8	39915	653545	1.830	0.961

Note that the models applied are listed below:

Market-adjusted model: $AR_{i,t} = R_{i,t} - R_{m.t}$;

Market model: $AR_{i,t} = R_{i,t} - (\alpha_i + \beta_i R_{m,t})$;

Capital asset pricing model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$

Fama-French three-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$;

Carhart four-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_{i,1}RPM_t + \beta_{i,2}HML_t + \beta_{i,3}SMB_t\}$

$R_{m.t}$ is denoted as market return which is measured as an equally-weighted market portfolio return, α_i and β_i are achieved by regressing monthly returns for security i on $R_{m,t}$ over the estimation period. $R_{f,t}$ is the risk-free rate which uses three-month market rates of Treasury bills as a proxy. HML_t is the difference between returns on portfolios of high and low book-to-market-ratio securities, whilst SMB_t represents the difference between returns on portfolios of small and large market value securities. RPM_t is the difference between returns on a market index and risk-free investment. β_s are beta factors corresponding to different control factors and are achieved by regression over the estimation period.

Table 3.13 Rejection frequency of the null hypothesis of zero mean CARs in large firms- Standard error over the estimation period

This table reports the rejection frequency of the null hypothesis of zero mean cumulative abnormal returns in 250 samples between 1982 and 2008. Both one-tailed and two-tailed tests are examined at significance levels of 1% and 5%. A firm is excluded from the population if it does not have 25-month returns. Then 200 small firms in size decile 10 with the largest market value are selected randomly. For each firm, I randomly select a month over the period of October 1982 to October 2007 for one year, October 1982 to October 2005 for three years and October 1982 to October 2003 for five years. The process is repeated 250 times. The estimation period is the previous 24 months before the event month. The test period starts from the event month. The investment horizons are one, three and five years. If a firm does not survive the whole test period, I only track the available returns. The returns over the estimation period are used to estimate the coefficients in the models through regression. I employ four models including the market-adjusted model, market model, Fama-French three-factor model and Carhart four-factor model. The abnormal returns are measured by cumulative abnormal returns (CARs). The standard error in test statistics are estimated based on the abnormal returns over the 24-month estimation period.

CARs are calculated as $\mathrm{CAR}_{p,T} = \sum_{t=0}^{T-1} \mathrm{AR}_{p,t}$ where $\mathrm{AR}_{p,t} = \frac{\sum_{i=1}^{N} \mathrm{AR}_{i,t}}{N}$. The test statistics to test the null hypothesis of zero mean CARs is: $t = \frac{\mathrm{CAR}_{P,T}}{\sigma(\mathrm{CAR}_{p,t}) * \sqrt{T}}$.,

When the standard deviation of CARs is based on estimation period, it is shown as: $\sigma(\mathrm{CAR}_{p,t}) = \sqrt{\sum_{t=-24}^{-1} \left(\mathrm{CAR}_{p,t} - \overline{\mathrm{CAR}_{p,t}}\right)^2 / 23}$

Where $\overline{\mathrm{CAR}_{p,t}} = \frac{1}{24} \sum_{t=-24}^{-1} \mathrm{CAR}_{p,t}$.

Model	1%			5%			Mean CARs	SD of CARs	Mean Tests	SD of Tests
	α=0	α<0	α>0	α=0	α<0	α>0				
	Panel A: one year									
Market-adjusted model	21.2	27.2	0.0	42.4	56.0	0.0	-0.053	0.031	-1.804	1.033
Market model	90.8	92.4	0.0	94.8	96.4	0.0	-0.118	0.038	-4.394	1.448
CAPM	18.0	23.6	0.0	36.0	48.4	0.8	-0.043	0.032	-1.600	1.211
FF three-factor model	47.6	52.8	0.0	65.2	71.2	0.4	-0.055	0.034	-2.336	1.424
Carhart four-factor model	74.8	79.2	0.0	87.2	90.0	0.0	-0.079	0.035	-3.571	1.584

Table 3.13 continued

Model	1%			5%			Mean CAR	SD of CAR	Mean Tests	SD of Tests
	α=0	α<0	α>0	α=0	α<0	α>0				
	Panel B: three years									
Market-adjusted model	54.4	63.2	0.0	75.2	82.8	0.0	-0.145	0.062	-2.770	1.205
Market model	98.0	98.4	0.0	99.2	100.0	0.0	-0.332	0.096	-6.963	2.080
CAPM	62.4	70.4	0.0	78.4	86.4	0.0	-0.152	0.069	-3.181	1.464
FF three-factor model	74.4	82.0	0.0	87.6	90.4	0.0	-0.159	0.068	-3.827	1.702
Carhart four-factor model	72.4	77.6	0.0	84.0	88.8	0.0	-0.153	0.074	-3.904	1.982
	Panel C: five years									
Market-adjusted model	54.8	66.8	0.0	78.8	89.2	0.0	-0.201	0.065	-2.781	0.967
Market model	99.6	100.0	0.0	100.0	100.0	0.0	-0.540	0.137	-8.280	2.448
CAPM	65.2	73.6	0.0	84.0	90.0	0.0	-0.205	0.077	-3.105	1.187
FF three-factor model	78.8	84.8	0.0	90.8	95.6	0.0	-0.204	0.070	-3.628	1.334
Carhart four-factor model	75.6	80.8	0.0	87.6	92.0	0.0	-0.210	0.086	-3.977	1.796

Note that the models applied are listed below:

Market-adjusted model: $\mathrm{AR}_{i,t} = \mathrm{R}_{i,t} - \mathrm{R}_{m.t}$;

Market model: $\mathrm{AR}_{i,t} = \mathrm{R}_{i,t} - (\alpha_i + \beta_i \mathrm{R}_{m,t})$;

Capital asset pricing model: $\mathrm{AR}_{i,t} = \mathrm{R}_{i,t} - \{\mathrm{R}_{f,t} + \beta_i(\mathrm{R}_{m,t} - \mathrm{R}_{f,t})\}$

Fama-French three-factor model: $\mathrm{AR}_{i,t} = \mathrm{R}_{i,t} - \{\mathrm{R}_{f,t} + \beta_i(\mathrm{R}_{m,t} - \mathrm{R}_{f,t})\}$;

Carhart four-factor model: $\mathrm{AR}_{i,t} = \mathrm{R}_{i,t} - \{\mathrm{R}_{f,t} + \beta_{i,1}\mathrm{RPM}_t + \beta_{i,2}\mathrm{HML}_t + \beta_{i,3}\mathrm{SMB}_t\}$

$\mathrm{R}_{m.t}$ is denoted as market return which is measured as an equally-weighted market portfolio return, α_i and β_i are achieved by regressing monthly returns for security i on $\mathrm{R}_{m,t}$ over the estimation period. $\mathrm{R}_{f,t}$ is the risk-free rate which uses three-month market rates of Treasury bills as a proxy. HML_t is the difference between returns on portfolios of high and low book-to-market-ratio securities, whilst SMB_t represents the difference between returns on portfolios of small and large market value securities. RPM_t is the difference between returns on a market index and risk-free investment. β_s are beta factors corresponding to different control factors and are achieved by regression over the estimation period.

Table 3.14 Rejection frequency of the null hypothesis of zero mean BHARs in large firms - Standard error over the test period

This table reports the rejection frequency of the null hypothesis of zero mean buy-and-hold abnormal returns in 250 samples over the period of 1982 to 2008. Both one-tailed and two-tailed tests are examined at significance levels of 1% and 5%. To begin with, a firm is excluded from the population if it does not have 25 month returns. Then 200 large firms in size decile 10 with the largest market value are selected randomly. For each firm, I randomly select a month over the period of October 1982 to October 2007 for one year, October 1982 to October 2005 for three years and October 1982 to October 2003 for five years. The process is repeated 250 times. The estimation period is the previous 24 months before the event month. The test period starts from the event month. The investment horizons are one, three and five years. If a firm does not survive the whole test period, I only track the available returns. The returns over the estimation period are used to estimate the coefficients in the models through regression. I employ four models, the market-adjusted model, market model, Fama-French three-factor model and Carhart four-factor model. The abnormal returns are measured by buy-and-hold returns (BHARs). The standard error in test statistics are estimated based on the abnormal returns over the estimation period. BHARs are calculated as: $\mathrm{BHAR}_{p,T} = \frac{\sum_{i=1}^{N} \mathrm{BHAR}_{i,T}}{N}$, where $\mathrm{BHAR}_{i,T} = \prod_{t=0}^{T-1}(1 + \mathrm{BHAR}_{i,t}) - 1$.

The test statistics to test the null hypothesis of zero mean BHARs is: $t = \frac{\mathrm{BHAR}_{P,T}}{\sigma(\mathrm{BHAR}_{p,t})}$.

The standard deviation of BHARs is based on test period: $\sigma(\mathrm{BHAR}_{p,t}) = \frac{1}{199}\sqrt{\sum_{t=0}^{T-1}\left(\mathrm{BHAR}_{i,T} - \mathrm{BHAR}_{p,T}\right)^2/(T-1)}$.

Model	**1%**			**5%**			**Mean BHARs**	**SD of BHARs**	**Mean Tests**	**SD of Tests**
	α=0	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**				
	Panel A: one year									
Market-adjusted model	16.4	23.6	0.0	33.6	46.0	0.0	-0.038	0.025	-1.559	1.058
Market model	32.8	39.6	0.0	54.4	66.4	0.0	-0.063	0.030	-2.125	1.080
CAPM	4.0	6.4	0.4	12.8	19.6	1.6	-0.021	0.027	-0.785	1.013
FF three-factor model	4.8	6.0	0.0	12.0	21.2	0.8	-0.025	0.028	-0.883	1.000
Carhart four-factor model	10.4	16.4	0.0	23.6	33.6	0.0	-0.039	0.031	-1.292	1.058

Table 3.14 continued

Model	1% α=0	1% α<0	1% α>0	5% α=0	5% α<0	5% α>0	Mean BHARs	SD of BHARs	Mean Tests	SD of Tests
						Panel B: three years				
Market-adjusted model	51.6	60.4	0.0	71.2	77.2	0.0	-0.095	0.041	-2.629	1.217
Market model	4.0	6.4	0.4	10.4	13.6	3.6	-0.012	0.081	-0.347	1.176
CAPM	30.4	38.0	0.0	51.2	61.6	0.0	-0.077	0.044	-1.989	1.191
FF three-factor model	44.8	50.4	0.0	62.8	74.4	0.0	-0.096	0.045	-2.393	1.184
Carhart four-factor model	20.8	27.2	0.0	35.2	48.4	0.0	-0.072	0.051	-1.659	1.224
						Panel C: five years				
Market-adjusted model	90.0	94.4	0.0	97.2	98.0	0.0	-0.159	0.041	-4.203	1.325
Market model	0.4	0.0	1.2	4.0	0.8	14.8	0	0	0.914	0.784
CAPM	45.2	57.6	0.0	68.8	80.8	0.0	-0.111	0.046	-2.549	1.166
FF three-factor model	72.8	81.2	0.0	87.6	92.0	0.0	-0.151	0.043	-3.403	1.243
Carhart four-factor model	24.0	29.6	0.0	39.2	46.0	0.0	0	0	-1.735	1.257

Note that the models applied are listed below:

Market-adjusted model: $AR_{i,t} = R_{i,t} - R_{m.t}$;

Market model: $AR_{i,t} = R_{i,t} - (\alpha_i + \beta_i R_{m,t})$;

Capital asset pricing model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$

Fama-French three-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$;

Carhart four-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_{i,1}RPM_t + \beta_{i,2}HML_t + \beta_{i,3}SMB_t\}$

$R_{m.t}$ is denoted as market return which is measured as an equally-weighted market portfolio return, α_i and β_i are achieved by regressing monthly returns for security i on $R_{m,t}$ over the estimation period. $R_{f,t}$ is the risk-free rate which uses three-month market rates of Treasury bills as a proxy. HML_t is the difference between returns on portfolios of high and low book-to-market-ratio securities, whilst SMB_t represents the difference between returns on portfolios of small and large market value securities. RPM_t is the difference between returns on a market index and risk-free investment. β_s are beta factors corresponding to different control factors and are achieved by regression over the estimation period.

Table 3.15 Rejection frequency of the null hypothesis of zero mean BHARs in 10 size deciles

This table reports the rejection frequency of the null hypothesis of zero mean buy-and-hold abnormal returns in 250 samples over the period of 1982 to 2008. The two-tailed test is examined with the applications of the market-adjusted model and capital asset pricing model at a significance level of 5%. To begin with, a firm is excluded from the population if it does not have 25 month returns. Then 200 firms are selected randomly. For each firm, I randomly select a month over the period of October 1982 to October 2007 for one year, October 1982 to October 2005 for three years and October 1982 to October 2003 for five years. The process is repeated 250 times. 10 size deciles are investigated. The estimation period is the previous 24 months before the event month. The test period starts from the event month. The investment horizons are one, three and five years. If a firm does not survive the whole test period, I only track the available returns. The returns over the estimation period are used to estimate the coefficients in the models through regression. The abnormal returns are measured by buy-and-hold returns (BHARs). The standard error in test statistics are estimated based on the abnormal returns over the estimation period. BHARs are calculated as: $\mathrm{BHAR}_{p,T} = \frac{\sum_{i=1}^{N} \mathrm{BHAR}_{i,T}}{N}$, where $\mathrm{BHAR}_{i,T} = \prod_{t=0}^{T-1}(1 + \mathrm{BHAR}_{i,t}) - 1$. The test statistic is: $t = \frac{\mathrm{BHAR}_{p,T}}{\sigma(\mathrm{BHAR}_{p,t})}$. The standard deviation of BHARs is based on test period: $\sigma(\mathrm{BHAR}_{p,t}) = \frac{1}{199}\sqrt{\sum_{t=0}^{T-1}\left(\mathrm{BHAR}_{i,T} - \mathrm{BHAR}_{p,T}\right)^2/(T-1)}$.

5%	**One year**					**Three years**					**Five years**				
Size deciles	α=0	α<0	α>0	Mean BHARs	Mean Tests	α=0	α<0	α>0	Mean BHARs	Mean Tests	α=0	α<0	α>0	Mean BHARs	Mean Tests
1	50.4	0.0	68.8	0.168	0.097	40.4	0.4	63.2	0.285	0.180	30.4	0.0	49.6	0.256	0.177
2	23.6	0.4	39.6	0.080	1.373	56.0	0.0	72.0	0.195	2.111	60.0	0.0	80.0	0.268	2.133
3	4.0	6.4	3.6	-0.001	-0.095	12.8	1.2	20.8	0.076	0.948	12.4	0.4	19.2	0.085	0.874
4	8.0	12.0	2.8	-0.008	-0.300	4.4	4.0	6.0	0.026	0.266	7.6	8.0	6.0	0.011	-0.007
5	8.0	8.8	1.6	-0.004	-0.211	4.4	9.2	0.8	0.005	-0.127	7.2	10.4	0.0	-0.013	-0.323
6	6.8	10.8	2.8	-0.008	-0.354	14.4	20.0	0.0	-0.034	-0.782	16.4	22.0	0.8	-0.041	-0.823
7	23.6	32.0	0.0	-0.037	-1.128	28.4	35.6	0.4	-0.064	-1.329	32.0	40.4	0.0	-0.078	-1.431
8	14.0	23.2	0.4	-0.031	-1.014	16.0	24.0	0.0	-0.047	-1.016	42.4	52.0	0.0	-0.095	-1.816
9	40.4	50.4	0.0	-0.047	-1.701	24.4	34.8	0.8	-0.056	-1.342	56.0	63.6	0.0	-0.104	-2.146
10	33.6	46.0	0.0	-0.038	-1.559	71.2	77.2	0.0	-0.095	-2.629	97.2	98.0	0.0	-0.159	-4.203

Table 3.15 continued

5%	One year					Three years					Five years				
Size deciles	α=0	α<0	α>0	Mean BHARs	Mean Tests	α=0	α<0	α>0	Mean BHARs	Mean Tests	α=0	α<0	α>0	Mean BHARs	Mean Tests
1	44.8	0.0	62.4	0.191	0.124	47.2	0.0	67.6	0.390	0.301	23.6	0.0	45.2	0.518	0.465
2	24.4	0.4	42.8	0.092	1.380	65.6	0.0	80.0	0.223	2.278	76.0	0.0	87.2	0.349	2.403
3	5.2	6.8	4.4	0.003	-0.021	13.6	0.0	22.4	0.090	1.047	18.0	0.4	30.8	0.131	1.187
4	9.2	10.8	2.0	-0.008	-0.303	5.6	2.0	9.2	0.036	0.418	6.0	1.6	8.8	0.046	0.410
5	2.8	5.2	2.0	0.008	0.085	4.4	6.4	3.2	0.014	0.027	5.6	7.2	2.0	0.013	0.012
6	6.4	9.2	1.2	-0.009	-0.378	11.2	16.8	0.0	-0.028	-0.670	14.8	19.2	1.2	-0.028	-0.635
7	18.8	28.4	0.4	-0.033	-1.016	24.8	32.4	0.4	-0.061	-1.195	28.0	35.6	0.8	-0.074	-1.297
8	15.6	23.2	0.4	-0.027	-0.889	18.8	26.4	0.0	-0.053	-1.080	24.8	33.6	0.0	-0.069	-1.223
9	17.6	26.0	0.0	-0.030	-1.017	47.6	60.8	0.0	-0.085	-1.948	63.6	69.6	0.0	-0.120	-2.330
10	12.8	19.6	1.6	-0.021	-0.785	51.2	61.6	0.0	-0.077	-1.989	68.8	80.8	0.0	-0.111	-2.549

Note that the models applied are listed below:

Market-adjusted model: $AR_{i,t} = R_{i,t} - R_{m.t}$;

Capital asset pricing model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$

$R_{m.t}$ is denoted as market return which is measured as an equally-weighted market portfolio return, β_i is achieved by regressing monthly returns for security i on $R_{m,t}$ over the estimation period. $R_{f,t}$ is the risk-free rate which uses three-month market rates of Treasury bills as a proxy.

Table 3.16 Rejection frequency of the null hypothesis of zero mean CARs in firms with the lowest BTM- Standard error over the estimation period

This table reports the rejection frequency of the null hypothesis of zero mean cumulative abnormal returns in 250 samples over a period of 1982 to 2008. Both one-tailed and two-tailed tests are examined at significance levels of 1% and 5%. A firm is excluded from the population if it does not have 25 month returns. Then, 200 firms in BTM decile 1 with the lowest BTM are selected randomly. For each firm, I randomly select a random month between October 1982 and October 2007 for one year, October 1982 to October 2005 for three years and October 1982 to October 2003 for five years. The process is repeated 250 times. The estimation period is the previous 24 months before the event month. The test period starts from the event month. The investment horizons are one, three and five years. If a firm does not survive the whole test period, I only track the available returns. The returns over the estimation period are used to estimate the coefficients in the models through regression. I employ four models: the market-adjusted model, market model, Fama-French three-factor model and Carhart four-factor model. The abnormal returns are measured by cumulative abnormal returns (CARs). The standard error in test statistics is estimated based on the abnormal returns over the 24-month estimation period.

CARs are calculated as $\mathrm{CAR}_{p,T} = \sum_{t=0}^{T-1} \mathrm{AR}_{p,t}$ where $\mathrm{AR}_{p,t} = \frac{\sum_{i=1}^{N} \mathrm{AR}_{i,t}}{N}$. The test statistic to test the null hypothesis of zero mean CARs is: $t = \frac{\mathrm{CAR}_{P,T}}{\sigma(\mathrm{CAR}_{p,t}) * \sqrt{T}}$.,

When the standard deviation of CARs is based on estimation period, this is shown as: $\sigma(\mathrm{CAR}_{p,t}) = \sqrt{\sum_{t=-24}^{-1} \left(\mathrm{CAR}_{p,t} - \overline{\mathrm{CAR}_{p,t}}\right)^2 / 23}$

Where $\overline{\mathrm{CAR}_{p,t}} = \frac{1}{24} \sum_{t=-24}^{-1} \mathrm{CAR}_{p,t}$.

	1%			5%						
Model	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**	**Mean CARs**	**SD of CARs**	**Mean Tests**	**SD of Tests**
					Panel A: one year					
Market-adjusted model	32.0	44.4	0.0	65.2	78.8	0.0	-0.145	0.044	-2.263	0.745
Market model	98.8	99.6	0.0	99.6	99.6	0.0	-0.303	0.057	-5.233	1.184
CAPM	31.2	44.8	0.0	68.0	80.0	0.0	-0.133	0.045	-2.287	0.831
FF three-factor model	28.0	36.8	0.0	57.6	69.6	0.0	-0.117	0.050	-2.128	0.959
Carhart four-factor model	54.0	66.8	0.0	76.8	82.8	0.0	-0.145	0.057	-2.734	1.137

Table 3.16 continued

Model	1%			5%			Mean CAR	SD of CAR	Mean Tests	SD of Tests
	α=0	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**				
					Panel B: three years					
Market-adjusted model	32.4	39.6	0.0	65.2	80.0	0.0	-0.249	0.076	-2.281	0.779
Market model	100.0	100.0	0.0	100.0	100.0	0.0	-0.800	0.134	-8.130	1.715
CAPM	61.6	72.4	0.0	84.0	91.2	0.0	-0.288	0.088	-2.921	0.973
FF three-factor model	43.6	51.2	0.0	63.2	75.6	0.0	-0.225	0.099	-2.410	1.099
Carhart four-factor model	58.8	66.4	0.0	74.8	80.4	0.0	-0.263	0.121	-2.908	1.415
					Panel C: five years					
Market-adjusted model	37.6	50.0	0.0	68.4	85.2	0.0	-0.324	0.094	-2.369	0.739
Market model	100.0	100.0	0.0	100.0	100.0	0.0	-1.312	0.195	-10.652	1.976
CAPM	72.0	80.8	0.0	90.4	96.0	0.0	-0.397	0.108	-3.178	0.932
FF three-factor model	41.2	49.6	0.0	62.8	76.8	0.0	-0.284	0.116	-2.379	0.972
Carhart four-factor model	67.2	74.0	0.0	82.0	86.8	0.0	-0.363	0.151	-3.153	1.347

Note that the models applied are listed below:

Market-adjusted model: $AR_{i,t} = R_{i,t} - R_{m.t}$;

Market model: $AR_{i,t} = R_{i,t} - (\alpha_i + \beta_i R_{m,t})$;

Capital asset pricing model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$

Fama-French three-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$;

Carhart four-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_{i,1}RPM_t + \beta_{i,2}HML_t + \beta_{i,3}SMB_t\}$

$R_{m.t}$ is denoted as market return which is measured as an equally-weighted market portfolio return, α_i and β_i are achieved by regressing monthly returns for security i on $R_{m,t}$ over the estimation period. $R_{f,t}$ is the risk-free rate which uses three-month market rates of Treasury bills as a proxy. HML_t is the difference between returns on portfolios of high and low book-to-market-ratio securities, whilst SMB_t represents the difference between returns on portfolios of small and large market value securities. RPM_t is the difference between returns on a market index and risk-free investment. β_s are beta factors corresponding to different control factors and are achieved by regression over the estimation period.

Table 3.17 Rejection frequency of the null hypothesis of zero mean BHARs in firms with the lowest BTM- Standard error over the test period

This table reports the rejection frequency of the null hypothesis of zero mean buy-and-hold abnormal returns in 250 samples over a period of 1982 to 2008. Both one-tailed and two-tailed tests are examined at significance levels of 1% and 5%. To begin with, a firm is excluded from the population if it does not have 25 month returns. Then 200 firms in BTM decile 1 with the lowest BTM are selected randomly. For each firm, I randomly select a random month over from October 1982 to October 2007 for one year, October 1982 to October 2005 for three years and October 1982 to October 2003 for five years. The process is repeated 250 times. The estimation period is the previous 24 months before the event month. The test period starts from the event month. The investment horizons are one, three and five years. If a firm does not survive the whole test period, I only track the available returns. The returns over the estimation period are used to estimate the coefficients in the models through regression. I employ four models: the market-adjusted model, market model, Fama-French three-factor model and Carhart four-factor model. The abnormal returns are measured by buy-and-hold returns (BHARs). The standard error in test statistics is estimated based on the abnormal returns over the estimation period. BHARs are calculated as: $\mathrm{BHAR}_{p,T} = \frac{\sum_{i=1}^{N} \mathrm{BHAR}_{i,T}}{N}$, where $\mathrm{BHAR}_{i,T} = \prod_{t=0}^{T-1}(1 + \mathrm{BHAR}_{i,t}) - 1$.

The test statistic to test the null hypothesis of zero mean BHARs is: $t = \frac{\mathrm{BHAR}_{P,T}}{\sigma(\mathrm{BHAR}_{p,t})}$.

The standard deviation of BHARs is based on test period as: as: $\sigma(\mathrm{BHAR}_{p,t}) = \frac{1}{199}\sqrt{\sum_{t=0}^{T-1}(\mathrm{BHAR}_{i,T} - \mathrm{BHAR}_{p,T})^2/(T-1)}$.

Model	**1%** α=0	**1%** α<0	**1%** α>0	**5%** α=0	**5%** α<0	**5%** α>0	**Mean BHARs**	**SD of BHARs**	**Mean Tests**	**SD of Tests**
						Panel A: one year				
Market-adjusted model	71.2	76.8	0.0	82.4	86.4	0.0	-0.118	0.046	-3.395	1.590
Market model	80.8	82.4	0.0	88.0	91.6	0.0	-0.175	0.081	-4.240	2.016
CAPM	44.4	54.4	0.0	66.0	71.2	0.0	-0.093	0.056	-2.460	1.399
FF three-factor model	22.0	27.6	0.0	36.0	47.6	0.0	-0.061	0.069	-1.546	1.342
Carhart four-factor model	12.8	14.8	0.0	21.2	30.8	0.0	-0.039	0.097	-1.047	1.346

Table 3.17 continued

Model	1% α=0	1% α<0	1% α>0	5% α=0	5% α<0	5% α>0	Mean BHARs	SD of BHARs	Mean Tests	SD of Tests
							Panel B: three years			
Market-adjusted model	91.6	94.0	0.0	97.2	98.4	0.0	-0.235	0.053	-5.027	1.846
Market model	21.2	26.8	0.0	34.0	40.8	0.0	-0.107	0.248	-1.502	1.521
CAPM	84.4	88.4	0.0	92.4	95.6	0.0	-0.233	0.059	-4.507	1.880
FF three-factor model	52.4	56.8	0.0	63.2	67.2	0.0	-0.556	1.717	-2.936	1.951
Carhart four-factor model	19.6	22.0	0.0	29.6	32.0	0.0	-1.656	25.732	-1.069	1.759
							Panel C: five years			
Market-adjusted model	87.2	92.0	0.0	94.4	97.2	0.0	-0.291	0.065	-5.399	2.358
Market model	2.8	3.6	0.0	5.2	6.4	3.2	0	1	0.293	1.171
CAPM	79.2	84.0	0.0	86.4	88.4	0.0	-0.256	0.167	-4.869	2.729
FF three-factor model	55.2	60.0	0.0	64.8	71.2	0.0	-0.127	0.420	-3.226	2.567
Carhart four-factor model	26.4	30.0	0.0	36.0	42.8	0.0	161043	1299378	-1.553	1.975

Note that the models applied are listed below:

Market-adjusted model: $AR_{i,t} = R_{i,t} - R_{m.t}$;

Market model: $AR_{i,t} = R_{i,t} - (\alpha_i + \beta_i R_{m,t})$;

Capital asset pricing model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$

Fama-French three-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$;

Carhart four-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_{i,1}RPM_t + \beta_{i,2}HML_t + \beta_{i,3}SMB_t\}$

$R_{m.t}$ is denoted as market return which is measured as an equally-weighted market portfolio return, α_i and β_i are achieved by regressing monthly returns for security i on $R_{m,t}$ over the estimation period. $R_{f,t}$ is the risk-free rate which uses three-month market rates of Treasury bills as a proxy. HML_t is the difference between returns on portfolios of high and low book-to-market-ratio securities, whilst SMB_t represents the difference between returns on portfolios of small and large market value securities. RPM_t is the difference between returns on a market index and risk-free investment. β_s are beta factors corresponding to different control factors and are achieved by regression over the estimation period.

Table 3.18 Rejection frequency of the null hypothesis of zero mean CARs in firms with the highest BTM- Standard error over the estimation period

This table reports the rejection frequency of the null hypothesis of zero mean cumulative abnormal returns in 250 samples over a period of 1982 to 2008. Both one-tailed and two-tailed tests are examined at significance levels of 1% and 5%. A firm is excluded from the population if it does not have 25 month returns. Then, 200 firms in the BTM decile 10 with the highest BTM are selected randomly. For each firm, I randomly select a random month over a period of October 1982 to October 2007 for one year, October 1982 to October 2005 for three years and October 1982 to October 2003 for five years. The process is repeated 250 times. The estimation period is the previous 24 months before the event month. The test period starts from the event month. The investment horizons are one, three and five years. If a firm does not survive the whole test period, I only track the available returns. The returns over the estimation period are used to estimate the coefficients in the models through regression. I employ four models: the market-adjusted model, market model, Fama-French three-factor model and Carhart four-factor model. The abnormal returns are measured by cumulative abnormal returns (CARs). The standard error in test statistics is estimated based on the abnormal returns over the 24-month estimation period.

CARs are calculated as $\mathrm{CAR}_{p,T} = \sum_{t=0}^{T-1} \mathrm{AR}_{p,t}$ where $\mathrm{AR}_{p,t} = \frac{\sum_{i=1}^{N} \mathrm{AR}_{i,t}}{N}$. The test statistic to test the null hypothesis of zero mean CARs is: $t = \frac{\mathrm{CAR}_{P,T}}{\sigma(\mathrm{CAR}_{p,t}) * \sqrt{T}}$.,

When the standard deviation of CARs is based on estimation period, it is shown as: $\sigma(\mathrm{CAR}_{p,t}) = \sqrt{\sum_{t=-24}^{-1} \left(\mathrm{CAR}_{p,t} - \overline{\mathrm{CAR}_{p,t}}\right)^2 / 23}$

Where $\overline{\mathrm{CAR}_{p,t}} = \frac{1}{24} \sum_{t=-24}^{-1} \mathrm{CAR}_{p,t}$.

Model	1% α=0	1% α<0	1% α>0	5% α=0	5% α<0	5% α>0	Mean CARs	SD of CARs	Mean Tests	SD of Tests
					Panel A: one year					
Market-adjusted model	46.0	0.0	55.6	70.0	0.0	84.0	0.169	0.062	2.638	1.121
Market model	100.0	0.0	100.0	100.0	0.0	100.0	0.378	0.077	6.458	1.852
CAPM	35.2	0.0	46.8	56.8	0.0	70.8	0.140	0.067	2.395	1.287
FF three-factor model	36.8	0.0	43.2	58.8	0.0	68.4	0.127	0.069	2.428	1.478
Carhart four-factor model	70.0	0.0	76.8	86.8	0.0	91.2	0.186	0.073	3.655	1.694

Table 3.18 continued

Model	1%			5%			Mean CAR	SD of CAR	Mean Tests	SD of Tests
	α=0	α<0	α>0	α=0	α<0	α>0				
	Panel B: three years									
Market-adjusted model	84.0	0.0	89.6	96.4	0.0	98.4	0.385	0.093	3.499	0.974
Market model	100.0	0.0	100.0	100.0	0.0	100.0	0.906	0.153	9.183	2.189
CAPM	59.6	0.0	69.2	80.8	0.0	88.8	0.289	0.102	2.912	1.107
FF three-factor model	65.6	0.0	71.2	79.6	0.0	86.4	0.282	0.115	3.179	1.375
Carhart four-factor model	84.0	0.0	87.2	93.2	0.0	94.8	0.382	0.137	4.422	1.769
	Panel C: five years									
Market-adjusted model	82.0	0.0	90.4	96.4	0.0	99.6	0.522	0.106	3.477	0.975
Market model	100.0	0.0	100.0	100.0	0.0	100.0	1.366	0.216	10.327	2.820
CAPM	62.0	0.0	75.6	83.2	0.0	91.2	0.405	0.121	3.047	1.150
FF three-factor model	54.8	0.0	60.4	71.2	0.0	82.0	0.330	0.136	2.806	1.332
Carhart four-factor model	86.4	0.0	89.2	93.2	0.0	95.2	0.510	0.179	4.458	1.914

Note that the models applied are listed below:

Market-adjusted model: $AR_{i,t} = R_{i,t} - R_{m.t}$;

Market model: $AR_{i,t} = R_{i,t} - (\alpha_i + \beta_i R_{m,t})$;

Capital asset pricing model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$

Fama-French three-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$;

Carhart four-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_{i,1}RPM_t + \beta_{i,2}HML_t + \beta_{i,3}SMB_t\}$

$R_{m.t}$ is denoted as market return which is measured as equally-weighted market portfolio return; α_i and β_i are achieved by regressing monthly returns for security i on $R_{m,t}$ over the estimation period. $R_{f,t}$ is the risk-free rate which use three-month market rates of Treasury bill as a proxy. HML_t is the difference between returns on portfolios of high and low book-to-market-ratio securities, whilst SMB_t is the difference between returns on portfolios of small and large market value securities. RPM_t is the difference between returns on a market index and risk-free investment. β_s are beta factors corresponding to different control factors and are achieved by regression over the estimation period.

Table 3.19 Rejection frequency of the null hypothesis of zero mean BHARs in firms with the highest BTM- Standard error over the test period

This table reports the rejection frequency of the null hypothesis of zero mean buy-and-hold abnormal returns in 250 samples over a period of 1982 to 2008. Both one-tailed and two-tailed tests are examined at significance levels of 1% and 5%. To begin with, a firm is excluded from the population if it does not have 25 month returns. Then 200 firms in BTM decile 10 with the highest BTM are selected randomly. For each firm, I randomly select a random month over a period of October 1982 to October 2007 for one year, October 1982 to October 2005 for three years and October 1982 to October 2003 for five years. This process is repeated 250 times. The estimation period is the previous 24 months before the event month. The test period starts from the event month. The investment horizons are one, three and five years. If a firm does not survive the whole test period, I only track the available returns. The returns over the estimation period are used to estimate the coefficients in the models through regression. I employ four models: the market-adjusted model, market model, Fama-French three-factor model and Carhart four-factor model. The abnormal returns are measured by buy-and-hold returns (BHARs). The standard error in test statistics is estimated based on the abnormal returns over the estimation period. BHARs are calculated as: $\mathrm{BHAR_{p,T}} = \frac{\sum_{i=1}^{N}\mathrm{BHAR_{i,T}}}{N}$, where $\mathrm{BHAR_{i,T}} = \prod_{t=0}^{T-1}(1+\mathrm{BHAR_{i,t}})-1$.

The test statistic to test the null hypothesis of zero mean BHARs is: $t = \frac{\mathrm{BHAR_{P,T}}}{\sigma(\mathrm{BHAR_{p,t}})}$.

The standard deviation of BHARs is based on the test period as: as: $\sigma(\mathrm{BHAR_{p,t}}) = \frac{1}{199}\sqrt{\sum_{t=0}^{T-1}\left(\mathrm{BHAR_{i,T}}-\mathrm{BHAR_{p,T}}\right)^2/(T-1)}$.

Model	1%			5%			Mean BHARs	SD of BHARs	Mean Tests	SD of Tests
	α=0	α<0	α>0	α=0	α<0	α>0				
	Panel A: one year									
Market-adjusted model	14.0	0.0	23.6	38.0	0.0	56.0	0.129	0.104	1.763	0.782
Market model	95.6	0.0	96.4	98.0	0.0	99.2	0.460	0.167	4.310	0.967
CAPM	5.2	0.0	10.0	23.6	0.4	39.2	0.134	0.147	1.398	0.807
FF three-factor model	3.2	0.0	6.0	16.8	0.4	30.0	0.124	0.141	1.226	0.809
Carhart four-factor model	12.4	0.0	28.0	48.0	0.0	72.0	0.281	0.203	1.960	0.613

Table 3.19 continued

Model	1% α=0	1% α<0	1% α>0	5% α=0	5% α<0	5% α>0	Mean BHARs	SD of BHARs	Mean Tests	SD of Tests
							Panel B: three years			
Market-adjusted model	19.6	0.0	34.0	57.6	0.0	76.4	0.280	0.160	2.099	0.643
Market model	81.6	0.0	86.8	96.0	0.0	98.4	2.934	1.241	3.326	0.835
CAPM	6.0	0.0	13.2	27.2	0.0	49.2	0.323	0.281	1.660	0.608
FF three-factor model	4.4	0.0	9.6	24.0	0.4	46.0	0.309	0.240	1.577	0.618
Carhart four-factor model	33.2	0.0	49.2	69.2	0.0	85.6	0.994	6.841	2.304	0.611
							Panel C: five years			
Market-adjusted model	25.6	0.0	36.8	62.4	0.0	78.0	0.301	0.165	2.144	0.647
Market model	44.8	0.0	58.4	74.0	0.0	86.0	15	22	2.573	0.830
CAPM	6.4	0.0	13.2	34.4	0.0	52.0	0.568	0.767	1.710	0.618
FF three-factor model	0.8	0.0	4.0	16.8	0.4	36.0	0.335	0.277	1.351	0.704
Carhart four-factor model	17.6	0.0	29.6	49.2	0.0	66.0	1	3	1.826	0.894

Note that the models applied are listed below:

Market-adjusted model: $AR_{i,t} = R_{i,t} - R_{m.t}$;

Market model: $AR_{i,t} = R_{i,t} - (\alpha_i + \beta_i R_{m,t})$;

Capital asset pricing model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$

Fama-French three-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$;

Carhart four-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_{i,1}RPM_t + \beta_{i,2}HML_t + \beta_{i,3}SMB_t\}$

$R_{m.t}$ is denoted as market return which is measured as an equally-weighted market portfolio return, α_i and β_i are achieved by regressing monthly returns for security i on $R_{m,t}$ over the estimation period. $R_{f,t}$ is the risk-free rate which uses three-month market rates of Treasury bills as a proxy. HML_t is the difference between returns on portfolios of high and low book-to-market-ratio securities, whilst SMB_t represents the difference between returns on portfolios of small and large market value securities. RPM_t is the difference between returns on a market index and risk-free investment. β_s are beta factors corresponding to different control factors and are achieved by regression over the estimation period.

Table 3.20 Rejection frequency of the null hypothesis of zero mean BHARs in 10 BTM deciles

This table reports the rejection frequency of the null hypothesis of zero mean buy-and-hold abnormal returns in 250 samples over the period of 1982 to 2008. The two-tailed test is examined with the applications of the market-adjusted model and capital asset pricing model at a significance level of 5%. To begin with, a firm is excluded from the population if it does not have 25 month returns. Then 200 firms are selected randomly. For each firm, I randomly select a month over the period of October 1982 to October 2007 for one year, October 1982 to October 2005 for three years and October 1982 to October 2003 for five years. The process is repeated 250 times. 10 BTM deciles are investigated. The estimation period is the previous 24 months before the event month. The test period starts from the event month. The investment horizons are one, three and five years. If a firm does not survive the whole test period, I only track the available returns. The returns over the estimation period are used to estimate the coefficients in the models through regression. The abnormal returns are measured by buy-and-hold returns (BHARs). The standard error in test statistics are estimated based on the abnormal returns over the estimation period. BHARs are calculated as: $\mathrm{BHAR}_{\mathrm{p,T}} = \frac{\sum_{\mathrm{i=1}}^{\mathrm{N}} \mathrm{BHAR}_{\mathrm{i,T}}}{\mathrm{N}}$, where $\mathrm{BHAR}_{\mathrm{i,T}} = \prod_{\mathrm{t=0}}^{\mathrm{T-1}}(1 + \mathrm{BHAR}_{\mathrm{i,t}}) - 1$. The test statistic is: $\mathrm{t} = \frac{\mathrm{BHAR}_{\mathrm{P,T}}}{\sigma(\mathrm{BHAR}_{\mathrm{p,t}})}$. The standard deviation of BHARs is based on test period: $\sigma(\mathrm{BHAR}_{\mathrm{p,t}}) = \frac{1}{199}\sqrt{\sum_{\mathrm{t=0}}^{\mathrm{T-1}}\left(\mathrm{BHAR}_{\mathrm{i,T}} - \mathrm{BHAR}_{\mathrm{p,T}}\right)^2/(\mathrm{T}-1)}$.

5%	**One year**					**Three years**					**Five years**				
BTM deciles	α=0	α<0	α>0	Mean BHARs	Mean Tests	α=0	α<0	α>0	Mean BHARs	Mean Tests	α=0	α<0	α>0	Mean BHARs	Mean Tests
1	82.4	86.4	0.0	-0.118	-3.395	97.2	98.4	0.0	-0.235	-5.027	94.4	97.2	0.0	-0.291	-5.399
2	60.0	72.0	0.0	-0.079	-2.437	87.6	92.4	0.0	-0.174	-3.702	84.8	88.4	0.0	-0.202	-3.747
3	26.0	35.2	0.0	-0.039	-1.255	44.0	51.6	0.0	-0.085	-1.827	67.6	79.6	0.0	-0.150	-2.836
4	31.2	40.0	0.0	-0.043	-1.393	25.2	34.0	0.8	-0.038	-1.017	30.4	39.6	0.0	-0.070	-1.322
5	6.0	9.2	2.0	0.003	-0.119	4.4	4.8	4.0	0.013	0.105	20.8	28.4	0.4	-0.048	-0.970
6	3.2	6.4	3.6	0.000	-0.072	3.2	1.6	6.0	0.034	0.415	6.4	4.8	6.4	0.021	0.215
7	4.0	2.0	5.6	0.018	0.300	11.6	1.2	23.2	0.075	0.990	4.0	2.4	10.4	0.048	0.562
8	20.4	0.0	34.0	0.049	1.191	32.0	0.0	46.0	0.095	1.557	22.0	0.0	32.0	0.101	1.277
9	28.8	0.0	44.0	0.070	1.540	51.2	0.0	68.0	0.136	1.947	33.2	0.0	51.2	0.145	1.639
10	38.0	0.0	56.0	0.129	1.763	57.6	0.0	76.4	0.280	2.099	62.4	0.0	78.0	0.301	2.144

Table 3.20 continued

5%	One year					Three years					Five years				
BTM deciles	α=0	α<0	α>0	Mean BHARs	Mean Tests	α=0	α<0	α>0	Mean BHARs	Mean Tests	α=0	α<0	α>0	Mean BHARs	Mean Tests
1	66.0	71.2	0.0	-0.093	-2.460	92.4	95.6	0.0	-0.233	-4.507	86.4	88.4	0.0	-0.256	-4.869
2	38.4	47.6	0.0	-0.056	-1.619	76.8	83.2	0.0	-0.160	-3.052	79.6	86.4	0.0	-0.190	-3.239
3	6.8	16.4	0.4	-0.020	-0.633	24.4	36.4	0.0	-0.063	-1.306	44.4	53.2	0.0	-0.104	-1.817
4	20.4	29.2	0.0	-0.033	-1.007	11.2	15.2	1.6	-0.007	-0.436	10.0	15.6	0.0	-0.011	-0.406
5	3.6	5.6	5.2	0.018	0.105	4.8	2.0	7.6	0.034	0.415	6.8	11.2	2.4	-0.006	-0.238
6	3.6	4.0	5.6	0.013	0.207	12.4	0.4	19.2	0.075	0.947	12.0	1.2	22.8	0.088	0.985
7	4.4	2.4	6.0	0.022	0.339	19.2	0.0	33.6	0.102	1.280	20.4	0.0	36.8	0.115	1.275
8	16.8	0.0	24.0	0.045	1.044	29.2	0.0	47.6	0.100	1.552	26.0	0.0	42.8	0.132	1.508
9	20.4	0.0	38.0	0.068	1.390	40.8	0.0	63.2	0.136	1.792	32.4	0.0	54.0	0.159	1.614
10	23.6	0.4	39.2	0.134	1.398	27.2	0.0	49.2	0.323	1.660	34.4	0.0	52.0	0.568	1.710

Note that the models applied are listed below:

Market-adjusted model: $AR_{i,t} = R_{i,t} - R_{m.t}$;

Capital asset pricing model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$

$R_{m.t}$ is denoted as market return which is measured as an equally-weighted market portfolio return, β_i is achieved by regressing monthly returns for security i on $R_{m,t}$ over the estimation period. $R_{f,t}$ is the risk-free rate which uses three-month market rates of Treasury bills as a proxy.

Table 3.21 Distribution of CARs and BHARs

This table reports the distribution of both CARs and BHARs based on different models. In order to address the issue of asymmetric distribution of CARs and BHARs, I examine the distributional properties of both CARs and BHARs with a sample of 50,000 randomly selected firms in one, three and five years over a period of 1982 to 2008. A similar simulation process is conducted as in the previous research. A firm is selected only if it has 25-month consecutive returns. 50,000 long-run CARs and BHARs are eventually achieved to summarize the descriptive statistics over three investment horizons. Abnormal returns are tested based on four models: the market-adjusted model, market model, Fama-French three-factor model and Carhart four-factor model.

	Cumulative Abnormal Returns (CARs)					Buy-and-Hold Returns (BHARs)				
	MA	**MM**	**CAPM**	**FF**	**CARHART**	**MA**	**MM**	**CAPM**	**FF**	**CARHART**
	Panel A: One year									
Mean	0.9	0.5	0.7	1.0	1.1	0.4	4.9	0.9	2.0	4.6
Median	0.1	-0.4	0.6	0.8	0.9	-5.3	-6.0	-4.8	-5.3	-5.6
Standard Deviation	48.6	60.5	51.0	55.3	61.6	61.5	96.6	84.2	87.4	131.0
Minimum	-308.8	-589.6	-369.3	-383.0	-640.3	-108.0	-161.2	-164.4	-510.1	-10365.8
Maximum	1395.4	1396.1	1384.6	1384.4	1370.0	6160.2	12122.9	11749.2	10485.4	16250.2
Kurtosis	12.1	8.0	9.4	9.3	9.9	324.9	1749.4	371.8	41009.6	45032.8
Skewness	1.1	-0.2	0.7	0.7	0.6	10.1	34.5	13.0	-190.2	207.8
	Panel B: Three years									
Mean	2.7	0.2	1.9	1.9	1.6	0.6	53.3	3.0	2.1	23.7
Median	2.0	2.1	3.6	2.7	2.3	-15.4	-15.6	-13.8	-16.9	-18.6
Standard Deviation	78.8	128.1	85.1	92.0	106.1	90.9	529.7	108.8	538.7	1892.5
Minimum	-549.9	-2237.4	-825.2	-993.8	-989.5	-104.2	-149.1	-155.9	-114552.1	-8896.0
Maximum	1308.8	1243.2	1159.9	1304.2	1481.7	5095.4	35807.3	4762.1	25285.1	412180.1
Kurtosis	12.1	8.0	9.4	9.3	9.9	324.9	1749.4	371.8	41009.6	45032.8
Skewness	1.1	-0.2	0.7	0.7	0.6	10.1	34.5	13.0	-190.2	207.8

Table 3.21 continued

	Cumulative Abnormal Returns (CARs)					Buy-and-Hold Returns (BHARs)				
	MA	MM	CAPM	FF	CARHART	MA	MM	CAPM	FF	CARHART
					Panel C: Three years					
Mean	2.9	0.6	3.3	2.3	1.7	-2.0	374.4	4.4	3.6	230.9
Median	3.0	5.2	5.8	4.4	3.5	-22.8	-21.4	-21.1	-26.3	-28.8
Standard Deviation	93.0	180.5	102.0	110.3	135.2	109.3	20082.7	177.6	280.5	30730.2
Minimum	-462.5	-3127.9	-1091.6	-1618.4	-2259.4	-104.5	-1211.4	-4753.0	-4121.7	-336348.6
Maximum	1047.3	1562.6	1136.8	1505.8	2444.8	7336.7	4238245.0	16476.1	48955.4	6100327.0
Kurtosis	7.0	7.6	7.3	9.5	12.1	524.0	39772.6	3143.7	18712.1	32510.6
Skewness	0.7	-0.3	0.5	0.3	0.4	12.6	191.1	40.8	111.5	173.2

Note that the models applied are listed below:

Market-adjusted model: $AR_{i,t} = R_{i,t} - R_{m.t}$;

Market model: $AR_{i,t} = R_{i,t} - (\alpha_i + \beta_i R_{m,t})$;

Capital asset pricing model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$

Fama-French three-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_i(R_{m,t} - R_{f,t})\}$;

Carhart four-factor model: $AR_{i,t} = R_{i,t} - \{R_{f,t} + \beta_{i,1}RPM_t + \beta_{i,2}HML_t + \beta_{i,3}SMB_t\}$

$R_{m.t}$ is denoted as market return which is measured as an equally-weighted market portfolio return, α_i and β_i are achieved by regressing monthly returns for security i on $R_{m,t}$ over the estimation period. $R_{f,t}$ is the risk-free rate which uses three-month market rates of Treasury bills as a proxy. HML_t is the difference between returns on portfolios of high and low book-to-market-ratio securities, whilst SMB_t represents the difference between returns on portfolios of small and large market value securities. RPM_t is the difference between returns on a market index and risk-free investment. β_s are beta factors corresponding to different control factors and are achieved by regression over the estimation period.

Table 3.22 Average cumulative returns and buy-and-hold returns over the test period conditional on the pre-event survival period

This table reports the average cumulative returns and buy-and-hold returns over the test period when there are different requirements of pre-event survival periods. To study the selection bias in the methodology, I compare returns over different test periods conditional on various pre-event survival periods. I start with the case when there is no requirement of pre-event returns. From October 1982, I keep firms with nonmissing return data in that month and track the returns from this month over the test periods of one, three and five years. If a firm does not survive the whole test period, I only track the available return data. I firstly calculate the cumulative return and buy-and-hold return of each firm. And then I take the average of cumulative returns of firms in the portfolio. The month is moved forward to November 1982. A similar process is repeated until October 2007. The number of observations declines with the length of pre-event survival periods.

Pre-event survival period	Observations	Mean returns	Return over test period		
			One year	Three years	Five years
None	303,970	CR	16.75	51.09	81.46
		BH	17.00	57.55	93.21
12 months	287,377	CR	16.60	50.60	80.58
		BH	17.00	58.07	94.26
24 months	272,041	CR	16.48	50.21	79.79
		BH	17.00	58.27	94.49
36 months	267,440	CR	16.47	50.17	79.73
		BH	17.00	58.17	94.58
48 months	263,505	CR	16.51	50.25	79.82
		BH	17.01	58.38	94.95

Chapter 4: Event-time approach: Simulation based on reference portfolios

4.1Research questions and hypotheses

Null hypothesis: the cross-sectional cumulative abnormal return or buy-and-hold abnormal return is zero over the investment horizons of one-, three- and five years.

This chapter follows the studies undertaken by Lyon, Barber and Tsai (1999) who use reference portfolios and the control firm approach which match the event firm with reference firms according to similar characteristics such as size or book-to-market ratio. Rather than using pricing models to estimate the expected return, the returns of reference portfolios or control firm are applied so as to achieve abnormal returns. Both the event month and event are simulated without specific requirements. However, Lyon, Barber and Tsai (1999) conduct a simulation process with 1000 samples of 200 firms. The reason I use 250 samples of 200 firms is because the total number of firms in the UK stock market is fewer than those listed in the US stock market. If 1000 samples of 200 firms are randomly selected to test the null hypothesis of zero abnormal returns, sampling bias could occur when the same firm is selected several times. The reference portfolios are constructed by firstly ranking market value or book-to-market ratio in September of each year. Then firms are deciled into ten groups according to the ranking. An event firm is matched with the reference portfolio as the benchmark which the event firm belongs to. When it comes to the control firm approach, the event firm is matched with only one firm with the closest market value or BTM or both. The skewness-adjusted t test, bootstrapped skewness-adjusted t test, empirical p value from pseudoportfolios are also examined, apart from the student t test. The percentage of rejection rates of 5% at a significance level of 5% indicates the test statistic is well-specified. The power of test is applied in order to identify which test statistic outperforms with the highest power, especially in the long run. An open question exists regarding which characteristic of a firm should be used to match with the event firm.

4.2Data and reference portfolios construction

4.2.1 Data

As with the data I applied in the previous chapter, I keep firms with market value and positive book value as well as having ordinary shares based on Gregory, Tharyan and Huang's (2009) data. The market value in September of year t is derived from the LSPD Archive file with A4 code representing the market value of the company's ordinary shares in £1 million. When a firm has market value less than £1 million which has an A4 denoted as zero, £0.5 million is used to replace the market value of zero. The same treatment is applied for firms with zero book value. The book value requires at least a 6-month lag from September due to firms' different accounting year ends. The market value and book value of firms are used to construct reference portfolios discussed in the next section. Compared with the data applied in event-time approach based on pricing models, there is no requirement of 25-month consecutive returns of a firm when applying even-time approach based on reference portfolios. However, the availability of market value and book value imply the requirement of stock returns in September of each year from 1982 to 2007. Since the last return observation ends in December 2008, the last event month ends in October 2007 for the one-year sample, October 2005 for the three-year sample and October 2003 for the five-year sample. Returns of firms are tracked from event months. The test period starts from October 1982 to September 2008. The investment horizons are one year, three years and three years.

Returns of delisted firms

Since the portfolios are assumed to be held over the investment horizons, the missing returns and returns of delisted firms need to be carefully dealt with. Lyon et al. (1999) propose the replacement of missing values of delisted firms with average monthly returns of reference portfolios, whereas Liu and Strong (2008) suggest replacement with a risk-free rate or zero. This replacement could result in a potential upward bias in BHARs when delisting firms go bankrupt. Since LSPD provide reasons of delisting with a code of g10 and quote returns of delisting firms as missing values, Gregory et al. (2010) suggest the replacement is allowed only when the company still remains in another status such as merger or acquisition. If the company goes bankrupt, the missing value is kept as -1 since it has lost its value. As for comparison with the US data, I apply the replacement of missing values with the mean monthly reference portfolio return introduced by Lyon et al. (1999).

4.2.2 Reference portfolios construction

Reference portfolios are constructed by size, by book-to-market ratio (BTM), by size and BTM following the methods proposed by Lyon and Barber (1997) and Lyon, Barber and Tsai (1999). They are comprised of either a number of firms or a single control firm by matching with different criteria. Returns of reference portfolios are taken as the benchmark returns when detecting abnormal returns of individual stocks.

4.2.2.1 Reference portfolios by size

When grouping reference portfolios by size, the following procedures are applied:

1. I firstly rank market values of firms in September of year t. Since the market value of a firm changes in different years, the rankings of the same firm may differ in different years. Moreover, the composition of reference portfolios is different every time when this procedure is conducted because the market values of firms provided by LSPD are approximated to integers. The imprecision of market values cannot distinguish firms with different market values but with the same figures in LSPD.
2. Firms are then deciled into ten groups based on the ranking in each year. Size 1 is consisted with firms with the smallest market values whereas size 10 has firms with the largest market values.
3. Returns of stocks are tracked from October of each year. As described in the data, the missing returns or returns of delisted firms are replaced with the average returns of reference portfolios.

4.2.2.2 Reference portfolios by book-to-market ratio (BTM)

When grouping reference portfolios by book-to-market ratio, the following procedures are applied:

1. With market values in September of each year derived from LSPD and 6-month lagged book values from data applied by Gregory, Tharyan and Huang (2009), the

book-to-market ratios (BTMs) are calculated by dividing book values by market values.

2. In September of each year, firms are ranked according to their BTMs. I replace zero BTM with 0.5, firms with a book value of 0.5 cannot differentiate themselves in the ranking. Therefore, if the same procedure is repeated to rank firms by BTM, firms may be allocated to different groups.
3. Firms are deciled into ten groups according to their BTM rankings in each year. BTM 1 contains firms with the lowest BTMs whereas BTM 10 has firms with the highest BTMs.
4. Stock returns are tracked from October of each year. The same treatment for missing returns and returns of delisted firms as size deciles is taken.

4.2.2.3 Reference portfolios by size and book-to-market ratio (BTM)

When grouping reference portfolios by size and book-to-market ratio, the following procedures are applied:

1. I firstly decile firms into ten groups according to their market values in September of each year.

2. In each size decile, firms are further quintiled into five groups based on book-to-market ratios. Therefore, in each year, there are 50 groups of reference portfolios.

3. Stock returns are tracked from October of each year. The same treatment for missing returns and returns of delisted firms as size deciles is taken.

4.2.2.4 Control firm approach

The control firm approach is examined based on three criteria. The first is by size. A firm with the closest market value in September of year t is selected to match with the sample firm. If there are multiple firms with the same difference of market value to the sample firm, I randomly select one. This could occur when the market values are the same for multiple control firms because of the imprecision of market values provided by LSPD. The same rule is applied to the control firm approach matched by BTM. However, this method becomes more sophisticated when the matching criterion is both size and BTM. A list of firms in the same year, when the market value of the event firm is based, with market values in the range of 70%-130% of the event firm’s market value is firstly created. Following this, a firm with the closest BTM is chosen as the benchmark. If there are multiple firms in the benchmark, one firm will be randomly selected. Stock returns are tracked from October of each year. The same treatment for missing returns and returns of delisted firms as size deciles is taken.

4.2.2.5 Rebalanced returns vs. Buy-and-hold returns

The compounding of returns has two formats as follows:

$$R_{reb} = \prod_{t=0}^{t=T}\left[1 + \left(\frac{\sum_{i=1}^{n_t} R_{it}}{n_t}\right)\right] - 1 \quad (4.1)$$

$$R_{bh} = \sum_{i=1}^{n_t} \frac{\prod_{t=0}^{t=T}(1+R_{it})-1}{n_t} \quad (4.2)$$

R_{it} is the return of individual stock i at time t and n_t is the number of stocks at time t. In the study, the number of stocks is always 200 when conducting buy-and-hold returns due to the replacement of missing data. R_{reb} is the rebalancing buy-and-hold return which calculates the mean of individual stock returns in different months and then compounds them. Baber, Lyon and Tsai (1999) criticize this measurement on the issues of rebalancing bias which potentially increases portfolio returns and new listing bias which lowers portfolio returns due to underperformance of new listing firms as suggested by Ritter (1991). To mitigate the problems brought by the active rebalancing buy-and-hold strategy, R_{bh} is used to calculate buy-and-hold returns of individual stocks over the investment horizon; then the average of these buy-and-hold stock returns comprising the portfolio is taken. This measurement is a passive investment strategy which assumes investors hold their stocks over the investment horizons even when a stock is delisted or a new stock is listed.

Table 4.1 shows annualized reference portfolio returns for ten size portfolios and ten BTM portfolios when rebalanced returns and buy-and-hold returns are applied. Similar results are documented as studies carried out by Lyon, Barber and Tsai (1999). Regarding the size deciles, size 1 shows the highest annualized returns. When size deciles increase, there is a downward trend for the annualized returns for both rebalanced returns and buy-and-hold returns. For instance, the annualized buy-and-hold return in size 1 is 35.5% in five years and decreases to 12.5% in size 10. The magnitude of the annualized returns is larger in rebalanced returns in size 1, 2 and 3 when compared with the buy-and-hold returns. However, most annualized returns are larger when the measurement is the conventional buy-and-hold returns. With respect to the BTM deciles, firms with low BTM show smaller annualized returns compared with firms with large BTM in most cases. BTM 1 shows an annualized rebalanced return of 11.2% whereas BTM 10 indicates the corresponding figure of 24.6%. Interestingly, the magnitude of the annualized returns is mostly larger when the rebalanced return is applied. It even reaches 10% when the investment horizon is five years in BTM 10 which is consisted with firms with the highest BTM. It can be concluded that when the rebalanced returns, which assumes the portfolios are rebalanced in each month based on an equal weighting scheme, are applied, small firms or firms with the highest BTM generate higher returns compared with the buy-and-hold returns which bear an assumption of holding a portfolio over the investment horizon. Compared with the US data presented by Lyon, Barber and Tsai (1999), the magnitude of returns under both measurements is mostly smaller in the UK stock market. For example, size 5 shows an annualized return of 18.3% in five years when the buy-and-hold returns are applied in the US stock market whereas the corresponding figure in the UK data indicates a return of 14.3%. Moreover, the UK evidence shows that the differences between the rebalanced returns and buy-and-hold returns are generally larger. Apart from establishing reference portfolios by size or BTM separately, I group reference portfolios by both size and BTM. Firstly I decile firms based on the ranking of market values in September of year t. Then, each size decile is quintiled into five groups based on the ranking of firms' BTMs. Therefore, 50 reference portfolios are constructed with criteria of both size and BTM. The last reference portfolio I adopt is equally weighted market return. Since the UK FTSE ALL SHARE index is value-weighted, the equally weighted market return needs to be computed by averaging stock returns across firms in each month of year. To obtain comprehensive market returns, I use the stock return file with all firms traded in London Stock Exchange without any screening from LSPD.

4.3 Research methodology

4.3.1 Models of abnormal returns

This thesis examines monthly cumulative abnormal returns which calculate monthly abnormal returns of individual stocks and then sum them up and buy-and-hold returns which compound individual stock returns over the holding period to compare with the compounding benchmark returns.

Cumulative abnormal returns (CARs)

Cumulative abnormal returns (CARs) calculate monthly abnormal returns for individual stock firstly by taking the difference of monthly returns of the sample firm and the monthly return of benchmark, then summing up the abnormal returns over the investment horizon. The expression is given as:

$$\mathrm{CAR}_{i,T} = \sum_{t=1}^{T}(R_{i,t} - \frac{1}{n_t}\sum_{j=1}^{n_t} R_{jt}) \tag{4.3}$$

R_{jt} is individual stock returns from reference portfolios. The benchmark could be a single firm or a portfolio of firms matching with similar characteristics such as size or BTM. In this study, I have seven types of reference portfolios to examine: 10 size portfolios, 10 BTM portfolios, 50 size/BTM portfolios, portfolios based on equally weighted market portfolio, control firm by size, BTM and both.

Buy-and-hold returns

Buy-and-hold abnormal returns (BHARs) take the differences between the compounded individual stock returns and compounded benchmark returns over the holding period as follows:

$$\mathrm{BHAR}_{i,T} = \prod_{t=1}^{T}(1 + R_{i,t}) - \prod_{t=1}^{T}(1 + R_{b,t}) \tag{4.4}$$

Where $R_{i,t}$ is the monthly individual stock return in month t and $R_{b,t}$ is the benchmark return in month t. T is the holding period of investment including the event month. The benchmark portfolios are the same as applied in cumulative abnormal returns.

Since a portfolio of different firms is studied over a period of one to five years, average cross-sectional returns need to be applied. This could be based on an equally-weighted or value-weighted scheme. Average monthly portfolio returns for buy-and-hold abnormal returns and cumulative abnormal returns are as follows:

$$\overline{\mathrm{CAR}_{p,T}} = \sum_{t=1}^{T} \mathrm{CAR}_{p,t} \tag{4.5}$$

$$\overline{\mathrm{BHAR}_{p,T}} = \sum_{i=1}^{n} w_i * \mathrm{BHAR}_{i,T} \tag{4.6}$$

where $\mathrm{CAR}_{p,t} = \sum_{i=1}^{n} w_i * \mathrm{AR}_{i,t}$ with w_i is the weight of firm i and n is the number of firms in the portfolio in the event month. Mean BHARs take weights into account over a period of time but mean CARs incorporate weights in each month before summation. In this chapter, I only focus on equal weight schemes which suggest that both BHARs and CARs of individual stock are computed first and then the average is taken.

4.3.2 Statistical inferences

Both parametric tests with an assumption of normal distribution of abnormal returns and nonparametric tests which are distribution-free are examined in this chapter. Since there are 250 samples of firms, 250 test statistics are achieved. The percentages of rejection rates of the null hypothesis of zero CARs or BHARs are calculated to identify if the test is well-specified or not. If a test is well specified, the rejection frequency is expected to be 2.5% for each side in two-tailed test and 5% in one-tail test at a significance level of 5%. This indicates approximately 12-13 samples of 250 samples generate test statistics which reject the null hypothesis. The test statistic is conservative when the rejection frequency is less than 5%, whereas the test statistic is anticonservative when the rejection frequency is more than 5%.

Parametric tests

The two parametric tests: conventional student t test and skewness-adjusted test with an assumption of normal distribution of abnormal returns are applied.

Student t test

The parametric tests statistics, following studies by Lyon, Barber and Tsai (1999), are the conventional student t-test statistics and skewness-adjusted test statistics. The nonparametric tests include bootstrapped skewness-adjusted test statistics, p-value from pseudoportfolio and the Wilcoxon signed-rank test. The distribution is assumed to be normal distributed with zero mean and constant variance when conducting the parametric tests. The null hypothesis is zero mean buy-and-hold cross-sectional abnormal returns or zero mean cumulative abnormal returns. Both two-tailed and one-tailed tests are examined at significance level of 5%.

$$\text{student t test} = \frac{\overline{\mathrm{CAR_T}}}{\sigma(\mathrm{CAR_T})/\sqrt{\mathrm{n}}} \tag{4.7}$$

$$\text{student t test} = \frac{\overline{\mathrm{BHAR_T}}}{\sigma(\mathrm{BHAR_T})/\sqrt{\mathrm{n}}} \tag{4.8}$$

n is the number of firms in one sample. Since the missing returns are replaced with the mean equally weighted reference portfolio returns, the size of samples is always 200 firms.

Skewness-adjusted t test

Since most of the long-term stock performance shows right skewness (Lyon and Barber, 1997) with higher rejection rates on the lower tail in two-tailed test and one-tail test, an adjustment is made incorporating the skewness based on the conventional student t test shown as:

$$\text{skewness} - \text{adjusted t} = \text{student t} + \frac{\sqrt{\mathrm{n}}}{3} * \frac{\overline{\mathrm{BHAR_T}}}{\sigma(\mathrm{BHAR_T})} * \text{skewness}^2 + \frac{\sqrt{\mathrm{n}}}{6\mathrm{n}} * \text{skewnss} \tag{4.9}$$

where

$$\text{skewness} = \frac{\sum_{i=1}^{n}(\mathrm{BHAR_i} - \overline{\mathrm{BHAR_T}})^3}{\mathrm{n}\sigma(\mathrm{BHAR_T})^3} \tag{4.10}$$

The same formulas are applied on CARs. This transformed skewness-adjusted test statistic is originally introduced by Johnson (1978) and further developed by Sutton (1993).

Nonparametric tests

Although the skewness adjusted test deals with skewness, the assumption of normal distribution of abnormal returns is easily violated. Therefore, I turn to nonparametric tests such as the bootstrapped skewness-adjusted test, p-value test or Wilcoxon signed-rank test, which do not require normal distribution.

Bootstrapped skewness-adjusted t test

The bootstrapped skewness-adjusted test for long-run abnormal returns is transformed by Lyon, Barber and Tsai (1999). The essence of this test is to formulate critical values for rejection or non-rejection of the null hypothesis by using the endogenous distribution of skewness-adjusted test statistics. To begin with, the skewness-adjusted test of one sample is computed based on 200 cumulative abnormal returns or buy-and-hold returns over the investment horizon. Then the bootstrapping procedure is implemented by randomly drawing

1000 subsamples to construct the bootstrapped distribution. Each subsample consists of 50 firms. Among these 50 firms, a new skewness-adjusted test is calculated as follows:

$$\text{bootstrapped skewness} - \text{adjusted } t_{sub} = \sqrt{n}\left(P + \frac{1}{3} * P * S^2 + \frac{1}{6n} * S\right) \tag{4.11}$$

where

$$P = \frac{\overline{BHAR_T^{sub}} - \overline{BHAR_T}}{\sigma(BHAR_T^{sub})} \tag{4.12}$$

$$\text{skewness} = \frac{\sum_{i=1}^{n_{sub}}(BHAR_{iT}^{sub} - \overline{BHAR_T^{sub}})^3}{n\sigma(BHAR_T^{sub})^3} \tag{4.13}$$

It is worth noting that the first part of the original skewness-adjusted test is supposed to be the student t test with the null hypothesis of zero mean buy-and-hold abnormal returns. However, when conducting this test on the subsamples, the null hypothesis is that the mean buy-and-hold abnormal return of 50 firms is equal to the mean buy-and-hold abnormal returns of the original sample with 200 firms. The computation of skewness still follows the original setup. Eventually I have 1000 skewness-adjusted test statistics for each sample. The critical values to determine whether the null hypothesis is rejected or not are formed by taking the percentiles of these 1000 test statistics. For instance, regarding two-tailed test at a significance level of 5%, I take the percentile of 2.5 as the critical value of the lower tail and 97.5 for the upper tail. The original skewness-adjusted test, but not the student t, for 200 firms in one sample is used to compare with the critical values to decide whether or not the null hypothesis is rejected. This approach not only approximates the empirical distribution of abnormal returns, but also takes skewness into account. The same formulas are applied on CARs.

P-value from pseudoportfolios

An alternative among the group of nonparametric tests is the p value from pseudoportfolios, introduced by Brock et al. (1992) and detailed by Lyon, Barber and Tsai (1999). In this approach, a control firm is randomly selected from the same reference group as the sample firm. Therefore, the reference group can be characterized by size, BTM, size/BTM and equally-weighted portfolio returns. As a result, a pseudoportfolio with 200 new firms is constructed. The same rule of no firms with the same event month is applied. This process continues 1000 times. For one sample, I achieve 1000 pseudoportfolios which are used to construct the distribution of mean abnormal returns. One characteristic which distinguishes the pseudoportfolio approach with other tests is the critical value is mean abnormal return but not test statistics. The null hypothesis results in the fact that long-term abnormal return of one sample with 200 firms is equal to the mean return of 1000 pseudoportfolios. Therefore, the next step is to achieve the mean portfolio abnormal return for 1000 pseudoportfolio in order to construct an empirical distribution of abnormal returns. This distribution is used to determine the critical values to compare with the mean abnormal returns of the original sample. According to the findings demonstrated by Lyon, Barber and Tsai (1999), the critical value is determined by solving the following equation:

$$\Pr\left(\overline{BHAR_T^P} \leq y_l\right) = \Pr\left(\overline{BHAR_T^P} \geq y_u\right) = \frac{\alpha}{2} \tag{4.15}$$

where y_l is the critical value in the low tail and y_u is the critical value on the upper tail. If the mean cross-sectional abnormal return is larger than y_u or smaller than y_l at a significance level of α, the null hypothesis is rejected. The process is carried out for all 250 samples. The percentages of rejection rates are calculated. The same formulas are applied on CARs.

Wilcoxon Signed-rank test

I also examine a popular nonparametric test which is a combination of the sign test and rank test, known as the Wilcoxon Signed-rank test introduced by Wilcoxon (1945). It bears the null hypothesis of an equal numbers of positive and negative returns under a framework of the binomial test. Firstly, the absolute values of 200 CARs or BHARs are ranked in each sample. Rank 1 relates to firms with the lowest CARs or BHARs. Then signs are allocated to observations. Afterwards, the additions of ranks with a positive sign and a negative sign are computed. Following this, the sum of these two figures is calculated for the purpose of determining whether the null hypothesis of zero median is rejected or not. The process is expressed as follows:

$$SR_N = \sum_i BHAR_i^+ \tag{4.16}$$

where

$$SR_N \sim N(E(SR_N), \sigma^2(SR_N)) \tag{4.17}$$

$$E(SR_N) = N(N+1)/4) \tag{4.18}$$

$$\sigma^2(SR_N) = N(N+1)(2N+1)/24 \quad (4.19)$$

CAR_i^+ is the positive rank of the absolute value of BHARs. When N is large, test statistics follow a normal distribution. N is the total number of firms. The test statistics Z is calculated based on mean and variance in order to compare this with critical values at a significance level of 5% on both one-sided and two-sided tests. For each sample, I achieve one Z. The number of Z totals 250 over one-, three- and five-year holding periods. The rejection frequency among 250 samples is computed. The same process is conducted for CARs.

4.3.3 Simulation process

I follow the simulation process on 250 samples of 200 firms as the studies undertaken by Lyon, Barber and Tsai (1999). The choice of 250 samples is due to a smaller sample in the UK compared with the US. I randomly select 200 event dates without replacement from a sample of dates covering October 1982 to October 2007 for a one-year investment horizon, to October 2005 for three-year investment, and October 2003 for five-year investment respectively. Then I randomly select 200 firms. The event firms are possibly the same, which is similar to the practice of when a firm has multiple events over a period of time. Moreover, I assure that no firms appear more than once with the same dates. The 200 firms eventually arrive at different event months; this reduces noise from event time clustering. For each firm in the sample, one reference portfolio including the event firm is created according to different criteria such as size or BTM. The post-event period includes the event month in this study. It demonstrates that the reference portfolios are constructed based on market value and BTM in September each year. Therefore, event months need to be allocated carefully to the appropriate reference portfolio calendar. That is, an event date between October of year t to September of year t+1 should be allocated to the size portfolio which is based on market value in September of year t. For instance, if an event firm experiences an event on January 2000, it should be categorized into a size/BTM portfolio in September 1999. I summarize the rejection frequency in 250 samples. If a test is well specified, a rejection rate of 5% is expected at 5% significance level. Underrejection with less than 5% rejection rate and overrejection with more than 5% rejection rate at 5% significance level suggest the test is misspecified in a conservative way or an anticonservative way.

4.4 Simulation on random samples

4.4.1 Cumulative Abnormal Returns (CARs)

Lower rejection rates of CARs are found in **Table 4.2** compared with BHARs, as consistent with Lyon, Barber and Tsai (1999). Test statistics are mostly negatively biased. However, the benchmark using equally-weighted market return shows positively biased test statistics regardless of investment horizons and test statistics applied. Since the asymmetric pattern of test statistics is shown when applying the conventional test statistics, the skewness-adjusted test and bootstrapped skewness-adjusted test improve the misspecification with higher rejection rates on the upper tail under the measurement of CARs. Although an improvement, the asymmetric pattern of test statistics still exists. No clear trend of CARs over the three investment horizons is exhibited since the results are mixed according to different benchmarks. For instance, when matching reference portfolios by size and BTM, the rejection rate decreases from 6.8% to 3.6% from one year to three years but increases to 4.8% in five years. However, the rejection rate increases to 7.2% in three years but declines to 5.2% in five years with the application of 10 size portfolios. In terms of two-tailed tests, the pseudoportfolio approach does not perform better when comparing this with the application of BHARs, especially when reference portfolios are matched by size in one and three years and by BTM in three years. The rejection rate when the benchmark is 10 size reference portfolios is 7.2% in one year and three years. The control firm approach shows well-specified results as in the pseudoportfolios approach with marginal improvement in skewness. It is worth commenting that the rejection rates in two-tailed test increase in bootstrapped skewness-adjusted test compared with conventional test statistics. Apart from parametric tests, I also test the null hypothesis of zero median abnormal returns with the application of the Wilcoxon signed-rank test which does not require normal distribution. **Table 4.4** presents the results based on CARs and BHARs when benchmarks are reference portfolios and a control matched by size, BTM and size/BTM. Interestingly, the reference portfolios approach and control firm approach yield well-specified results in CARs in most cases. The distributions of abnormal returns based on different benchmarks are mixed. Additionally, the magnitude of rejection rates under CARs is mostly lower than BHARs.

4.4.2 Buy-and-hold Abnormal Returns (BHARs)

The empirical results of rejection frequency are shown in **Table 4.3** at a significance level of 5%. The rebalanced buy-and-hold returns assuming investors rebalance portfolios monthly exhibit severe misspecification compared with other benchmarks when conventional test statistics are applied. The rejection rate with the application of rebalanced buy-and-hold returns in reference portfolios matched by size and BTM is 8.8% in one year, 22% in three years and 27.2% in five years. Moreover, test statistics are negatively biased with higher rejection rates on the lower tail. This is consistent with findings in Lyon et al. (1999) but with slightly higher rejection rates. The conventional BHARs, assuming investors hold the portfolios over the investment horizons, largely improve the misspecification in the rebalanced BHARs. This alleviates the new listing and rebalancing bias. The reference portfolios are matched by size, BTM and size/BTM, and equally-weighted market return. Regarding student t test, the rejection rates increase with the length of investment horizons except for the reference portfolios matching by BTM with a rejection rate of 8.4% in one year, 7.6% in three years and 6.8% in five years. The student test statistics in one-tailed test for all benchmarks have higher rejection rates on the lower tail, indicating a negative bias on test statistics. Therefore, I conduct the skewness-adjusted test and bootstrapped skewness-adjusted test to overcome this issue. Although a significant improvement on asymmetry of test statistics for all benchmarks is demonstrated, the rejection rates in two-tailed test increase in most cases. However, the application of p value from pseudoportfolios yields well-specified tests. Furthermore, the control firm approach in conjunction with conventional test statistics reveals similar results as the empirical p value test. Consequently, it can be concluded that the reference portfolios together with p value and the control firm approach with student t, are appropriate to detect long-term abnormal performance. It is difficult to identify which benchmark is superior since the results are mixed over the investment horizons. For instance, in a one-year investment horizon, the control firm matched by size and BTM has a rejection rate of 4.4% while the control firm matched by BTM indicates a rejection rate of 4% in two-tailed test. The figures are 5.2% and 3.6% respectively in a three-year investment horizon. However, the symmetry of abnormal returns with rejection rates close to 5% is shown when matching firms by BTM.

The Wilcoxon signed-rank test, a nonparametric test, shows that when applying the approach based on reference portfolios, an increased trend of rejection with the length of investment horizons for all benchmarks is evident. Moreover, there is severe misspecification with high rejection rates. For instance, the rejection rate is 58.4% in one year and 62.8% in five years with the application of the benchmark matched by size and BTM. Overrejection clusters on the lower tail, indicating that the median of BHARs is significantly negative. In contrast, the control firm approach presents well-specified results in **Table 4.4**, regardless of benchmarks. For example, the rejection rate is 3.6% in one year, 5.2% in three years and 4.8% in five years when applying the benchmark matched by size.

Power of test when the benchmark is reference portfolios based on size and BTM

To find out which test statistic is superior in terms of power of test, all test statistics are examined under BHARs with negative and positive shocks when the benchmark is reference portfolios based on size and BTM. The power of test is the probability that I reject the null hypothesis when the null is false. To examine it, I introduce abnormal returns ranging from -20% to 20% at an incremental of 5% to stock returns which apply a benchmark matched by size and BTM. If the test is superior, the rejection rates are expected to be 100%, which indicates 250 null hypotheses of zero BHARs are all rejected, when abnormal returns exist. Mixed results are shown in **Table 4.5** and **Figure 4.1** over three investment horizons but with a general conclusion that the bootstrapped skewness-adjusted t test is the most powerful test with positive shocks whereas the student t test is more appropriate with negative shocks. The rejection rates with the application of student t test on induced abnormal return of -20% are 90.4%, 56% and 28.4% when the holding periods are one-, three- and five years. The corresponding figures with the application of bootstrapped skewness-adjusted t test when the induced abnormal return is 20% are 95.6%, 55.2% and 22.8%, respectively. A dramatic decline for all test statistics is evident in rejection rates when the investment horizon lengthens. For example, the skewness-adjusted test shows a decrease in rejection rates from 91.6% to 12.4% from one year to five years when abnormal return of 15% is introduced. Ang and Zhang (2004) also indicate similar findings. Comparison among the bootstrapped skewness-adjusted t test, p-vale from pseudoportfolios and student t test in control firm approach, findings of the bootstrapped skewness-adjusted t test with the highest power when induced abnormal returns are positive and the student t test in control firm approach with the least power when induced abnormal returns are negative are in line with studies over an investment horizon of three years undertaken by Lyon, Barber and Tsai (1999). However, there are some differences. First, Lyon et al. (1999) document the lowest power of test with the application of student t test in control firm approach when induced abnormal returns are positive whereas the results suggest p-value from pseudoportfolios has lowest power when there is a positive shock. Second, the bootstrapped skewness-adjusted t test has the highest power when negative shocks are introduced in the UK data. However, Lyon et al. (1999) shows p-vale from pseudoportfolios is the most appropriate test to be applied based on the US data when there is a negative shock. Last but most importantly, compared with the UK results, the magnitude of rejection rates is much higher when the US data is applied.

4.5 Simulation on non-random samples

4.5.1 Large/Small size

Among all types of reference portfolios based on CARs or BHARs, I investigate the size effect. The issue of small firms is always the focus in empirical studies due to their volatile and abnormal returns. Therefore, I decile firms based on ranking of market value in September each year and choose the one with the highest market value and the one with the lowest market value for robustness check. I repeat the previous process by randomly selecting 250 samples of 200 firms from a large size group and small size group, separately. It is should be noted that although firms belong to different size groups, the benchmark portfolio is still the same as the original setup. **Tables 4.6-4.7** highlight the fact that that CARs and BHARs share similar findings in small firms and large firms. The misspecification is wildly documented together with positive skewness of abnormal returns when using benchmarks matched with BTM or equally weighted market returns in both small and large firms. In contrast, the benchmarks matched with size/BTM and size generate well-specified results for large firms in both the reference portfolios approach and control firm approach. Although the small size effect is reduced significantly compared with approaches based on the model and conventional calendar-time approach, it is still evident when applying the reference portfolio approach in conjunction with conventional test statistics, skewness-adjusted test, bootstrapped skewness adjusted test, on samples filled with small firms. It is important to note that the control firm approach based on benchmarks matched with size and size/BTM yield well-specified results, particularly for small firms. Furthermore, the approach based on reference portfolios with p value from pseudoportfolios slightly overrejects the null hypothesis of zero abnormal returns. In addition, the control firm eliminates the misspecification in large firms with the application of benchmark matched by BTM but fails in small firms. The major difference between CARs and BHARs is that the magnitude of rejection rates under CARs is mostly smaller than those under BHARs.

4.5.2 High/Low book-to-market ratio (BTM)

Firms are categorized into ten groups based on rankings of BTM in September each year. I choose the group with the highest BTM and lowest BTM for robustness check. For each group, 250 samples of 200 firms are randomly selected from the group in order to obtain the statistical inferences. Similar bias to the robustness check on large/small size appears. For instance, the firms in the high BTM group do not fall into the reference portfolios based on Size/BTM. When firms are classified according to size/BTM, initially they are deciled according to ranking in size. For each size, firms are quintiled into five groups based on rankings in BTM. Therefore, the number of firms in the reference portfolio is smaller than the number of firms in the high BTM group. This bias can be reduced by applying a reference portfolio based on BTM as suggested but undone by Lyon, Barber and Tsai (1999). As expected, the benchmark matched with BTM produces more specified test statistics compared with the benchmark matched with size/BTM in most cases, as demonstrated in **Table 4.8-4.9**. BHARs and CARs present similar findings except that the magnitude of rejection rate is slightly higher under BHARs, particularly in the long run. Samples of firms with high BTM represent slight overrejection on the null hypothesis of zero abnormal returns compared with samples of firms with low BTM when applying the approach based on reference portfolios. This may be partially attributed to small size effect since the samples of firms with high BTM potentially include small firms. As in the case of non-random samples based on size, the benchmark is crucial when detecting long-run abnormal returns on firms in different BTM groups. The benchmarks matched by size and equally weighted market returns generate misspecification. Despite the control firm approach eliminating the small size effect, it yields misspecification in the samples of firms with low BTM when the matching criterion of the benchmark is size.

4.5.3 Industry clustering

It is inevitable that most firms in the same industry or sector share the same event date or event period in the stock market. For example, in an economic downturn, a surge of mergers and acquisitions is evident in firms in the same industry. Therefore, I expect to see higher rejection rates with more significant evidence of abnormal returns. I start with G17, which stands for stock exchange industrial classification, in G records from the London Stock Price Database. The industry classification system in the UK follows the Industry Classification Benchmark (ICB) as Dow Jones since Jan, 2006. Two systems were applied, namely, the Stock Exchange Industrial Classification system and the Financial Time Actuaries Index Classification system before 1994. These two systems are merged after 1994 until the ICB appears in 2006. The popular ICB system has 10 industries with 20 supersectors that are separated into 41 sectors with 114 subsectors. The ten industries are: oil and gas (0001), basic materials (1000), industrials (2000), consumer goods (3000), health care (4000), consumer services (5000), telecommunications (6000), utilities (7000), financials (8000), technology (9000). The data provided by Gregory, Tharyan and Huang (2009) reclassifies firms before 2006 to match with the ICB system according to G17. I conduct the simulation process in two steps. The first is to randomly draw an industry from 10 industries. I then randomly select 200 firms with 200 event dates with replacements from this industry. This completes one sample. I repeat the same simulation procedure until there are 250 samples of 200 firms. Note that to avoid overlapping returns, if one firm appears more than once, the next event month is dropped if this falls in both the event period and pre-event period. For instance, if firm A has one event month in January, 2000, its next event month cannot be in any month during the period of January, 1999 to December, 1999 or the period of February, 2000 to January, 2001, when the event period is 12 months. The reference portfolio returns are defined as mean monthly cumulative or buy-and-hold stock returns in the same industry. **Table 4.10** reveals that the different benchmarks on CARs yields well-specified results, particularly in the long run. The benchmark matched with equally-weighted market returns, surprisingly, produces well-specified test statistics. For instance, the rejection rate for this benchmark together with skewness-adjusted test based on CARs is 5.6% in one year, 6.4% in three years and 3.6% in five years. There is little difference when comparing the rejection frequencies of the reference portfolios approach and control firm approach. This indicates that the industry clustering is not a significant issue in long-run event studies with the application of methodologies based on the portfolios approach and control firm approach under BHARs. However, it is important to note that test statistics are negatively biased with higher rejection rates on the lower tail when applying CARs. With respect to BHARs, even the control firm approach generates misspecified results when using the benchmark matched with size/BTM. The reference portfolios approach also yields misspecified results.

4.6 Cross-sectional dependence of returns

4.6.1 Overlapping returns

Suppose Apple experiences an announcement of stock dividends and launch of a new product in February 2006 and April 2006, respectively. Both events affect the stock price not only as far as the event date is concerned; the effect may last for a few days or even months. This suggests cross-sectional returns which mean the stock returns of Apple reflect the impact from both events from April 2006 till December 2006, or even longer. The cross-sectionally related returns potentially result in misspecified tests due to violation of independence of returns in event studies. Since the simulation could generate firms with multiple event months without the special requirement of a gap between event months, the previous results possibly yield misspecified test statistics due to dependence of returns. Therefore, it is worth investigating the level of impact from overlapping returns with special cases according to the procedures designed by Lyon, Barber and Tsai (1999). I firstly select 100 event months without replacement over the event period. Following this, 100 event firms are randomly selected to match with the event months. Firms are not allowed to have more than one event month. For each event firm, I assume it experience another event either within a period of the previous 11 months ending in the event month or in a period over the following 11 months after the event month. The same rules are applied to three- and five-year investment horizons. I eventually arrive at 200 firms with 200 event months where one firm has two event months which generate the issue of overlapping returns. The process is repeated 250 times. All the benchmarks are examined at a significance level of 5% over three event windows. **Table 4.11** illustrates that the reference portfolios approach and control firm approach cannot resolve the issue of overlapping returns; this is one severe type of cross-sectional dependent returns. Both BHARs and CARs with different benchmarks and test statistics indicate severe misspecification as in prior research but with a smaller magnitude of rejection rates. When firms have multiple events in the study period, even the control firm approach, which yields well-specified tests in random samples and non-random samples, loses its advantages. Lyon, Barber and Tsai (1999) advocate applying Loughran and Ritter's studies in 1995 thereby establishing a requirement of pre-event returns to keep a 'clean' dataset for the event studies. However, as discussed in the previous chapter regarding the event-time approach based on models, the stock returns increase when the survival period lengthens.

4.6.2 Calendar clustering

Calendar clustering also creates cross-sectional returns which lead to overstated test statistics (Brown and Warner, 1980; Fama, 1998). Firms sharing the same event month generate returns having impact from events at the same time over the event studies window. This mix-up influence can be overcome by applying the monthly calendar-time portfolio approach with the price of lower power of test (Mitchell and Stafford, 2000). I follow the design as Lyon, Barber and Tsai (1999) by randomly selecting 250 event firms initially without replacement. Subsequently, for each event firm, I randomly select 200 firms without replacement. Eventually I reach 250 samples of 200 firms; the event month in each sample is the same but differs across samples. All benchmarks are examined at a significance level of 5%. Both BHARs and CARs show the ability to resolve the issue of cross-sectional returns with the application of the reference portfolios approach and control firm approach except for the case when the benchmark is equally weighted market portfolio returns. It interesting to note that when applying the equally-weighted market portfolio returns as the benchmark under BHARs, the rejection rate declines with the length of investment horizons. For example, the rejection rate based on student t test decreases from 14.2% in one year to 4.4% in five years. However, the application of CARs shows the opposite results with a higher rejection rate in one year. **Table 4.12** demonstrates some evidence of negatively skewed test statistics, particularly when applying CARs. Compared with the results presented by Lyon, Barber and Tsai (1999), the rejection rates based on the UK data are slightly higher but with similar distribution when using BHARs with a benchmark matched by size and BTM.

4.7 Summary

This chapter follows the methodology of another event-time approach applied by Lyon, Barber and Tsai (1999). This approach is based on reference portfolios or control firm which are matched with the event firm according to similar characteristics. According to the results documented by Agarwal and Taffler (2008), 30th September of each year is a better choice to construct the reference portfolios since 22% of firms' accounting year ends in March whereas 37% of firms release their report in December. Therefore, this study rank market value of firms in September of each year but use the book value of firms with a six-month lag starting from September. Three types of reference portfolios are established. One is by size which firstly ranks market value of firms in September of each year and then decile firms into ten groups according to their rankings. Another is by book-to-market ratio which ranks BTMs in September of each year and then decile firms into ten groups according to their rankings. The last type is by size and book-to-market ratio which firstly decile firms into ten groups based on their rankings by size and then for each size decile, firms are quintiled into five groups based on their rankings by book-to-market ratio, Stock returns are tracked from October of each year. Control firm approach is also found by size, book-to-market ratio or both. When a control firm is matched by size or book-to-market ratio, a firm with the closest market value or book-to-market ratio to the event firm is taken as the control firm. However, when both criteria is used to find the control firm, a list of firms with market values in the range of 70%-130% of the event firm's market value in the same year is firstly created. Following this, a firm with the closest BTM is chosen as the control firm. The null hypothesis of zero cross-sectional cumulative abnormal return or buy-and-hold return over investment horizons of one-, three- and five years is tested. Apart from the traditional student-t test statistic, other parametric tests such as skewness-adjusted t test and nonparametric tests such as bootstrapped skewness-adjusted t test, p-value from pseudoportfolios and Wilcoxon signed-rank test are examined. Since the UK firms are fewer than firms listed in the US, he simulation process is based on 250 samples of 200 firms. If a test is well specified, a rejection rate of 5% is expected at 5% significance level. Underrejection with less than 5% rejection rate and overrejection with more than 5% rejection rate at 5% significance level suggest the test is misspecified in a conservative way or an anticonservative way.

The studies document similar findings as prior research undertaken by Lyon and Barber (1997) and Lyon, Barber and Tsai (1999). Firstly, the magnitude of rejection rates is smaller in CARs, especially with the conventional test statistics, skewness-adjusted test statistics and bootstrapped skewness-adjusted test statistics. Secondly, the negatively skewed test statistics under CARs and BHARs is evidenced with higher rejection rates on the left tail. Thirdly, the skewness-adjusted test and bootstrapped skewness-adjusted test, to some extent, improve the skewness with higher rejection rates on the upper tail, regardless of measurement of abnormal returns. Although an improvement, the asymmetric pattern of test statistics still exists. Moreover, the rejection rates in two-tailed test increase in the bootstrapped skewness-adjusted test compared with conventional test statistics. Fourthly, the control firm approach based on different criteria with conventional test statistics shows superior ability when detecting long-run abnormal returns compared with the reference portfolios approach. Furthermore, similar as findings documented by Ang and Zhang (2004), the Wilcoxon signed-rank tests under CARs mostly generate well-specified test statistics when applying reference portfolios and control firm approach whereas BHARs based on reference portfolios overreject the null hypothesis of zero median. However, the reference portfolio approach in conjunction with empirical p value also generates well-specified test statistics.

Apart from the similarities to prior studies, I also document some differences that may be attributed to the market structure of the UK stock market. To begin with, compared with event-time approach based on models, the reference portfolios and control firm approach yield well-specified results in both CARs and BHARs in random samples. Secondly, although an improvement is indicated regarding the asymmetric pattern of test statistics under CARs and BHARs with the application of skewness-adjusted tests and bootstrapped skewness-adjusted tests, I discover higher rejection rates in two-tailed tests compared with the conventional test statistics. Additionally, the Wilcoxon signed-rank test is well specified when applying the control firm approach regardless of measurement of abnormal returns. However, CARs based on the reference portfolios approach yield well-specified results whereas BHARs show severe misspecification. The issues regarding overlapping returns cannot be resolved based on the application of both reference portfolios approach and control firm approach. To conclude, the reference portfolios approach in conjunction of the p-value from pseudoportfolios and the control firm approach with the conventional student t test statistics are advocated to be applied in detecting long-term abnormal stock performance in both random and non-random samples. It is worth mentioning that the power of test loses its advantage when applying these two approaches. Moreover, when the firms in the sample have overlapping returns, the misspecification cannot be eliminated with the application of these two approaches. Additionally, the matching criteria can be size, BTM or both. It is important to note that when the non-random samples are examined based on size or BTM, the matching criterion of reference portfolios is required to have the same characteristic as the event firms.

In summary, the application of BHARs in conjunction with p-value from pseudoportfolios is appropriate for application in the context of long-run event studies. Furthermore, the control firm approach together with student t-test statistics is proved to yield well-specified test statistics in both random and non-random samples. Firms in reference portfolios and control firms are selected on the basis of size, BTM or both. However, in terms of power of test, these two approaches have the least power whereas the skewness-adjusted test and bootstrapped skewness-adjusted test have the highest power. It is worth noting that when the non-random samples are examined, the benchmark portfolio or control firm needs to share at least one characteristic with the event firm.

Figures and Tables of Chapter 4

Table 4.1 Annualized buy-and-hold returns of reference portfolios

Panel A displays annualized returns of reference portfolios in terms of size. Market value of each stock in September each year is used as a basis to categorize the size group. Size 1 contains 10% of stocks which have the smallest market cap while size 10 has 10% of stock with the largest market cap. The returns of each stock are tracked from October of each year over one, three and five years. Panel B shows the annualized returns of reference portfolios based on book-to-market ratio (BTM). The book-to-market ratios in September each year are ranked to categorize stocks with ten deciles. Decile 1 contains 10% of stocks with the smallest BTM and decile 10 has 10% of stocks with the largest BTM. The remainder of the process is similar to Panel A. Buy-and-hold rebalanced returns compound the mean portfolio returns over the study period

$$\mathrm{PR_{reb}} = \prod_{t=0}^{t=T}\left(1+\frac{\sum_{i=1}^{n}\mathrm{R_{it}}}{n}\right)-1$$

Buy-and-hold returns assume investors keep the portfolio over the holding period. Note that stocks with missing returns or those that are delisted are still taken as a constituent of the portfolio. So the missing returns are replaced by the average return of the portfolio to which the stock belongs. The buy-and-hold return of each stock is computed. Then all buy-and-hold stock returns in the portfolio are taken the average based on the number of stocks.

$$\mathrm{PR_{bnh}} = \sum_{i=1}^{n}\frac{\left[\prod_{t=0}^{t=T}(1+\mathrm{R_{it}})\right]-1}{n}$$

Decile	Rebalanced returns (%)			Buy-and-hold returns (%)			Differences (%)		
	1y	3y	5y	1y	3y	5y	1y	3y	5y
Panel A: Size decile portfolio returns									
1	49.8	39.9	32.6	40.7	30.6	25.5	9.1	9.2	7.2
2	24.8	25.9	24.3	20.4	21.1	19.9	4.4	4.8	4.4
3	18.7	18.8	18.4	16.5	17.5	16.2	2.2	1.4	2.2
4	13.8	15.0	15.4	13.8	15.4	14.3	0.0	-0.4	1.1
5	12.5	13.0	14.4	13.6	14.9	14.3	-1.1	-1.9	0.1
6	13.6	12.8	13.8	15.4	14.6	14.3	-1.8	-1.8	-0.5
7	11.6	11.4	12.3	13.7	13.6	13.2	-2.2	-2.2	-0.8
8	12.1	11.9	12.5	13.9	12.9	13.0	-1.8	-1.0	-0.5
9	12.0	12.8	12.7	13.4	13.2	12.6	-1.4	-0.4	0.1
10	13.1	13.0	12.8	13.8	13.2	12.5	-0.6	-0.2	0.3

Table 4.1 continued

Panel B: Book-to-market decile portfolio returns									
1	11.2	7.7	8.6	12.7	8.6	7.7	-1.5	-0.9	0.9
2	11.2	11.0	11.6	12.1	9.8	8.9	-0.9	1.2	2.7
3	10.4	12.2	12.4	10.8	10.3	9.3	-0.3	1.9	3.1
4	12.9	14.7	14.4	12.8	11.8	10.6	0.1	2.8	3.9
5	13.5	17.3	16.3	13.1	12.3	10.8	0.4	5.0	5.5
6	15.0	17.6	17.8	14.0	12.4	12.4	1.0	5.2	5.5
7	15.0	19.5	18.1	14.3	13.6	11.7	0.7	5.9	6.5
8	15.5	21.2	20.2	14.2	14.0	12.6	1.3	7.2	7.6
9	17.9	24.8	22.3	16.2	16.2	13.6	1.7	8.6	8.7
10	24.6	28.3	26.0	22.7	19.1	16.0	1.9	9.2	10.0

Table 4.2 Rejection frequency using cumulative abnormal returns in random samples

This table shows the rejection rates (as a percentage) of 250 random samples with 200 firms at a significance level of 5% in one-tailed and two-tailed tests over a period of October 1982 to September 2008, under the null hypothesis of zero one-, three-year, five-year cumulative abnormal returns. The test is specified if the overall rejection rate is the same as the significance level together with equal rejection rate on both sides. The abnormal return is defined as the difference between the actual return of a stock and benchmark return. The benchmark is categorized into five types: 50 portfolios by size and book-to-market ratio (BTM) which decile firms by market value in September each year and then quintile firms based on BTM; 10 portfolios ranked by market value; 10 portfolios ranked by BTM, equally weighted market portfolio and control firm which is matched by size, or BTM or size/BTM with sample firm. Apart from conventional student test statistics, skewness-adjusted test statistics, bootstrapped skewness-adjusted test statistics and p-value (Barber, Lyon and Tsai 1999) derived from pseudoportfolios are examined. The descriptive statistics of abnormal returns are summarized as follows.

5%		**One year**			**Three years**			**Five years**		
Test statistics	**Benchmarks**	α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
Conventional t test	**50 size/BTM portfolios**	6.8	8.4	3.6	3.6	6.8	2.4	4.8	3.6	4.8
	10 size portfolios	6.4	8.8	2.0	7.2	7.6	3.6	5.2	6.4	6.4
	10 BTM portfolios	4.4	8.4	4.0	8.0	4.8	8.4	4.8	5.6	4.0
	Equally-weighted market return	6.8	3.2	8.4	5.2	2.0	9.6	7.6	1.6	12.0
Skewness-adjusted test	**50 size/BTM portfolios**	8.0	9.2	4.8	3.6	6.8	2.8	5.2	4.8	6.4
	10 size portfolios	7.2	8.4	4.0	8.4	7.6	4.8	7.6	6.8	8.0
	10 BTM portfolios	5.6	8.0	6.4	9.6	5.2	8.4	6.4	5.6	4.4
	Equally-weighted market return	9.2	2.8	9.6	7.6	2.0	11.6	10.0	1.6	14.0
Bootstrapped skewness-adjusted test	**50 size/BTM portfolios**	7.6	7.6	4.0	3.6	5.6	3.2	6.0	3.6	5.6
	10 size portfolios	7.2	6.8	4.0	7.2	7.2	5.2	5.6	6.0	6.4
	10 BTM portfolios	4.8	5.6	4.8	9.2	4.8	8.8	4.4	5.2	4.4
	Equally-weighted market return	7.6	2.8	9.2	8.0	2.0	12.8	7.6	1.2	12.4

Table 4.2 continued

5%		One year			Three years			Five years		
Test statistics	**Benchmarks**	α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
P value from Pseudoportfolios	**50 size/BTM portfolios**	5.6	6.8	6.0	4.8	8.8	3.6	5.2	3.2	6.0
	10 size portfolios	7.2	7.2	3.2	7.2	6.8	4.0	6.0	6.4	7.6
	10 BTM portfolios	4.8	6.0	5.2	7.2	4.0	8.8	5.2	4.8	4.8
Control firm approach	**by size and BTM**	6.0	8.0	5.2	4.8	6.8	4.0	6.0	6.8	4.8
	by size	5.6	6.4	5.2	6.0	6.4	4.0	5.2	4.8	5.6
	by BTM	4.4	5.2	2.4	6.8	4.4	6.8	4.8	4.8	4.8

Abnormal return is defined as the difference of actual individual i stock return and benchmark return as follows: $\mathrm{AR_{it}} = \mathrm{R_{it}} - \mathrm{BR_{it}}$.
The cumulative abnormal returns (CARs) calculate abnormal returns of individual stocks in month t by taking the difference between the stock return and equally weighted benchmark return. $\mathrm{CAR_{i,T}} = \sum_{t=1}^{T}(\mathrm{R_{i,t}} - \frac{1}{n_t}\sum_{j=1}^{n_t}\mathrm{R_{jt}})$. I replace the missing returns with average portfolio return. The conventional student test statistics is formulated as: $t = \overline{\mathrm{AR}}/(\sigma(\mathrm{AR})/\sqrt{n})$ whereas n is the number of firms in one sample which is 200 in this study. The skewness-adjusted test is calculated as: $\mathrm{skewness-adj.\,t} = \sqrt{n}(S + (1/3) * S^2 * \gamma + (1/6\,n) * \gamma$ whereas γ is skewness of cross-sectional abnormal returns and $\sqrt{n}S$ is student t test. The bootstrapped skewness-adjusted t test bootstraps abnormal returns in each sample 1000 times with size of 50. Then skewness-adjusted test for each subsample of 50 firms is computed in order to construct a distribution which is used to determine the critical value for rejection of the null hypothesis for each sample. Note that skewness-adjusted test calculated in the full sample is used to compare with the critical value. The Pseudo portfolio approach firstly randomly selects a benchmark firm from the same group (by size, BTM or size/BTM) as the sample firm. Then one mean buy-and-hold return for each sample is achieved. With all 250 mean returns, a distribution is constructed for the hypothesis test. Note that the null hypothesis is the equality of sample mean return with critical value.

Table 4.3 Rejection frequency using buy-and-hold returns in random samples

This table shows the rejection rates (as a percentage) of 250 random samples with 200 firms at a significance level of 5% in one-tailed and two-tailed tests over a period of October 1982 to September 2008, under the null hypothesis of zero one-, three-year, five-year buy-and-hold abnormal returns. The test is specified if the overall rejection rate is the same as the significance level together with equal rejection rate on both sides. The abnormal return is defined as the difference between the actual return of a stock and benchmark return. The benchmark is categorized into five types: 50 portfolios by size and book-to-market ratio (BTM) which decile firms by market value in September each year and then quintile firms based on BTM; 10 portfolios ranked by market value; 10 portfolios ranked by BTM, equally weighted market portfolio and control firm which is matched by size, or BTM or size/BTM with sample firm. Apart from conventional student test statistics, skewness-adjusted test statistics, bootstrapped skewness-adjusted test statistics and p-value (Barber, Lyon and Tsai 1999) derived from pseudoportfolios are examined. The descriptive statistics of abnormal returns are summarized as follows.

5%		**One year**			**Three years**			**Five years**		
Test statistics	**Benchmarks**	α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
Conventional t test	**50 size/BTM rebalanced portfolios**	8.8	15.2	0.4	22.0	26.4	0.0	27.2	42.4	0.0
	50 size/BTM buy-and-hold portfolios	6.0	11.6	1.6	7.2	10.8	1.2	8.0	8.8	2.0
	10 size buy-and-hold portfolios	5.6	10.4	2.4	8.8	12.8	1.6	8.0	9.2	1.2
	10 BTM buy-and-hold portfolios	8.4	14.0	1.6	7.6	13.2	3.6	6.8	11.6	0.8
	Equally-weighted market return	6.8	8.4	4.8	8.4	11.6	0.4	8.8	13.6	1.2
Skewness-adjusted test	**50 size/BTM buy-and-hold portfolios**	8.0	10.0	6.0	7.6	10.4	4.4	7.2	10.0	4.0
	10 size buy-and-hold portfolios	6.8	8.8	6.0	8.0	10.8	4.4	7.6	10.0	4.8
	10 BTM buy-and-hold portfolios	8.0	12.0	4.0	8.4	10.0	6.8	7.6	8.8	2.8
	Equally-weighted market return	8.4	6.8	7.6	5.6	9.6	2.4	10.8	11.6	4.0
Bootstrapped skewness-adjusted test	**50 size/BTM buy-and-hold portfolios**	10.8	8.0	8.0	5.6	6.8	6.0	6.4	7.6	4.4
	10 size buy-and-hold portfolios	8.0	6.8	7.2	7.6	8.4	6.0	8.8	7.6	5.6
	10 BTM buy-and-hold portfolios	7.2	9.2	4.0	6.8	6.4	8.0	8.8	6.8	5.2
	Equally-weighted market return	8.8	4.4	8.4	5.6	6.4	4.4	8.4	8.0	4.8

Table 4.3 continued

5%		One year			Three years			Five years		
Test statistics	**Benchmarks**	α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
P value from Pseudoportfolios	**50 size/BTM buy-and-hold portfolios**	7.6	6.0	7.6	4.8	4.8	6.4	5.2	3.6	4.4
	10 size buy-and-hold portfolios	4.4	3.6	4.8	5.2	4.8	4.4	4.4	5.2	5.2
	10 BTM buy-and-hold portfolios	5.2	6.8	3.6	6.0	6.0	4.0	4.4	5.2	4.8
Control firm approach	**by size and BTM**	4.4	6.0	3.6	5.2	7.6	5.2	6.0	4.0	5.2
	by size	4.0	3.6	5.6	3.6	5.2	5.2	4.0	6.0	2.4
	by BTM	3.6	5.2	2.8	6.0	4.0	7.2	6.4	4.8	6.0

Abnormal return is defined as the difference of actual individual i stock return and benchmark return as follows: $\mathrm{AR_{it}} = \mathrm{R_{it}} - \mathrm{BR_{it}}$.
Rebalanced buy-and-hold return assumes investors rebalance portfolios monthly. The monthly portfolio return needs to calculate every month and then compound and average it over the investment horizon. $\mathrm{PR_{reb}} = \prod_{t=0}^{t=T}\left(1 + \frac{\sum_{i=1}^{n} R_{it}}{n}\right) - 1$.The conventional buy-and-hold return assumes rather than rebalancing monthly, investors tend to hold the same portfolio over the investment horizon. Stocks in the portfolios are supposed have return data over the study period. So I replace the missing returns with average portfolio return. $\mathrm{PR_{bnh}} = \sum_{i=1}^{n} \frac{[\prod_{t=0}^{t=T}(1+R_{it})]-1}{n}$. The conventional student test statistics is formulated as: $t = \overline{AR}/(\sigma(AR)/\sqrt{n})$ whereas n is the number of firms in one sample which is 200 in this study. The skewness-adjusted test is calculated as: $\mathrm{skewness-adj.\,t} = \sqrt{n}(S + (1/3) * S^2 * \gamma + (1/6\,n) * \gamma$ whereas γ is skewness of cross-sectional abnormal returns and $\sqrt{n}S$ is student t test. The bootstrapped skewness-adjusted t test bootstraps abnormal returns in each sample 1000 times with size of 50. Then skewness-adjusted test for each subsample of 50 firms is computed in order to construct a distribution which is used to determine the critical value for rejection of the null hypothesis for each sample. Note that skewness-adjusted test calculated in the full sample is used to compare with the critical value. The Pseudo portfolio approach firstly randomly selects a benchmark firm from the same group (by size, BTM or size/BTM) as the sample firm. Then one mean buy-and-hold return for each sample is achieved. With all 250 mean returns, a distribution is constructed for the hypothesis test. Note that the null hypothesis is the equality of sample mean return with critical value.

Table 4.4 Wilcoxon signed-rank test in random samples

This table shows the rejection rates (as a percentage) of 250 random samples with 200 firms at a significance level of 5% in one-tailed and two-tailed tests over a period of October 1982 to September 2008, under the null hypothesis of zero one-, three-year, five-year cumulative abnormal returns and buy-and-hold returns. The test is specified if the overall rejection rate is the same as the significance level together with equal rejection rate on both sides. The abnormal return is defined as the difference between the actual return of a stock and benchmark return. The benchmark is categorized into five types: 50 portfolios by size and book-to-market ratio (BTM) which decile firms by market value in September each year and then quintile firms based on BTM; 10 portfolios ranked by market value; 10 portfolios ranked by BTM, equally weighted market portfolio and control firm which is matched by size, or BTM or size/BTM with sample firm. Apart from conventional student test statistics, skewness-adjusted test statistics, bootstrapped skewness-adjusted test statistics and p-value (Barber, Lyon and Tsai 1999) derived from pseudoportfolios are examined. The descriptive statistics of abnormal returns are summarized as follows. Abnormal return is defined as the difference of actual individual i stock return and benchmark return as follows: $AR_{it} = R_{it} - BR_{it}$.

The conventional buy-and-hold return assumes rather than rebalancing monthly, investors tend to hold the same portfolio over the investment horizon. Stocks in the portfolios are supposed have return data over the study period. So I replace the missing returns with average portfolio return. $PR_{bnh} = \sum_{i=1}^{n} \frac{[\prod_{t=0}^{t=T}(1+R_{it})]-1}{n}$. The cumulative abnormal returns (CARs) calculate abnormal returns of individual stocks in month t by taking the difference between the stock return and equally weighted benchmark return. $CAR_{i,T} = \sum_{t=1}^{T}(R_{i,t} - \frac{1}{n_t}\sum_{j=1}^{n_t} R_{jt})$.

The Wilcoxon signed-rank test is conducted in three steps. Firstly, the absolute values of 200 CARs or BHARs are ranked in each sample. Rank 1 is with firms with the lowest CARs or BHARs. Then signs are allocated to observations. Afterwards, the additions of ranks with a positive sign and with a negative signs are computed. Then the sum of these two figures is calculated for the purpose of determine if the null hypothesis of zero median is rejected or not. The process is given as: $SR_N = \sum_i BHAR_i^+$ where, $SR_N \sim N(E(SR_N), \sigma^2(SR_N))$, $E(SR_N) = N(N+1)/4$, $\sigma^2(SR_N) = N(N+1)(2N+1)/24$. CAR_i^+ is the positive rank of the absolute value of BHARs.

Table 4.4 continued

5%		CARs α=0	CARs α<0	CARs α>0	BHARs α=0	BHARs α<0	BHARs α>0
		Panel A: One year					
Reference portfolios	**by size/BTM**	6.4	7.6	1.6	29.6	40.4	0.0
	by size	3.6	4.8	4.4	28.4	39.2	0.4
	by BTM	7.2	10.4	3.2	43.6	53.6	0.0
Control firm	**by size/BTM**	4.8	5.2	3.2	4.4	6.4	3.2
	by size	3.2	3.6	4.8	3.6	3.2	5.2
	by BTM	4.8	4.8	4.8	3.6	3.6	6.4
		Panel B: Three years					
Reference portfolios	**by size/BTM**	5.2	6.4	2.0	58.4	74.0	0.0
	by size	6.0	6.0	4.8	60.4	71.6	0.0
	by BTM	7.2	6.4	4.8	66.8	79.2	0.0
Control firm	**by size/BTM**	4.8	7.6	3.2	5.6	6.0	3.6
	by size	4.4	6.4	6.0	5.2	6.0	6.8
	by BTM	5.2	3.2	6.8	5.6	4.0	6.4
		Panel C: Five years					
Reference portfolios	**by size/BTM**	5.2	3.2	7.2	62.8	76.0	0.4
	by size	6.8	4.0	9.2	66.0	76.4	0.0
	by BTM	4.4	6.0	2.4	80.0	86.8	0.0
Control firm	**by size/BTM**	6.8	6.4	5.2	5.2	5.2	5.2
	by size	4.8	6.0	5.6	4.8	4.0	4.8
	by BTM	5.6	3.2	5.6	4.4	2.4	4.8

Table 4.5 Power of test in random samples

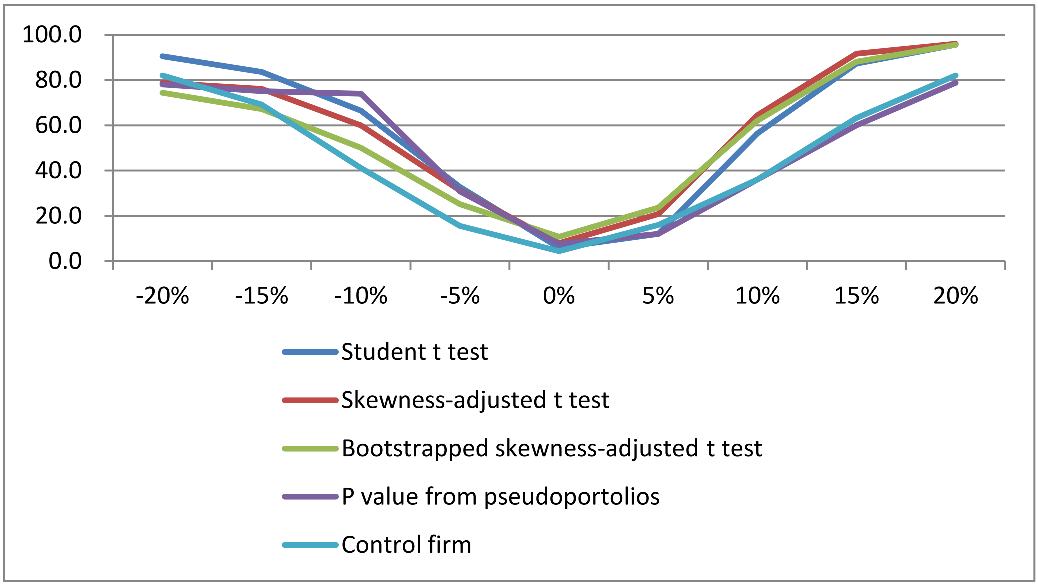
This table shows the rejection rates of null hypothesis of zero abnormal returns when shocks are introduced. To measure the power of test, which examines the ability of the test to reject the null hypothesis of zero mean buy-and-hold returns based on benchmark matching by size and book-to-market ratio when it is false, I introduce abnormal returns range from -20% to 20% at an interval of 5% to each event firm. For example, I add abnormal return of 5% to the buy-and-hold return of an event firm over the holding period. The percentages of rejection rates in 250 samples are computed at a significance level of 5% over three investment horizons (one, three and five years).

	-20%	**-15%**	**-10%**	**-5%**	**0%**	**5%**	**10%**	**15%**	**20%**
					Panel A: One year				
Student t test	90.4	83.6	66.4	32.8	6.0	12.0	56.4	87.2	95.6
Skewness-adjusted t test	78.8	76.0	60.0	31.2	8.0	20.8	64.4	91.6	96.0
Bootstrapped skewness-adjusted t test	74.4	67.2	50.0	25.2	10.8	23.6	62.0	88.0	95.6
P value from pseudoportfolios	78.0	75.0	74.0	30.8	7.6	12.0	36.0	60.0	78.8
Control firm	82.0	69.2	41.2	15.6	4.4	16.0	36.0	63.2	82.0
					Panel B: Three years				
Student t test	56.0	45.6	30.8	16.8	7.2	3.6	8.0	21.6	42.0
Skewness-adjusted t test	49.6	39.2	24.8	14.4	7.6	6.8	16.0	30.4	54.4
Bootstrapped skewness-adjusted t test	36.0	27.6	18.4	10.4	5.6	8.4	17.6	33.2	55.2
P value from pseudoportfolios	33.0	25.0	15.0	11.6	4.8	5.6	9.2	17.2	27.2
Control firm	32.4	24.8	14.4	8.8	5.2	6.8	12.0	20.0	30.8
					Panel C: Five years				
Student t test	28.4	20.8	13.2	9.6	8.0	4.4	3.2	7.2	10.0
Skewness-adjusted t test	25.2	20.4	13.2	11.6	7.2	7.6	8.8	12.4	19.6
Bootstrapped skewness-adjusted t test	20.4	14.8	10.8	8.0	6.4	7.6	10.8	14.8	22.8
P value from pseudoportfolios	27.6	16.8	11.6	7.6	5.2	4.4	6.0	6.8	12.0
Control firm	14.8	9.2	6.4	4.8	6.0	6.0	6.4	9.6	12.4

Figure 4.1 Power of test in random samples

This table shows the specification of test statistics based on reference portfolios approach matching by size and BTM when shocks are introduced. To measure the power of test, which examines the ability of the test to reject the null hypothesis of zero mean buy-and-hold returns based on benchmark matching by size and book-to-market ratio when it is false, I introduce abnormal returns range from -20% to 20% at an interval of 5% to each event firm. For example, I add abnormal return of 5% to the buy-and-hold return of an event firm over the holding period. The percentages of rejection rates in 250 samples are computed at a significance level of 5% over three investment horizons (one, three and five years).

One year



Three years

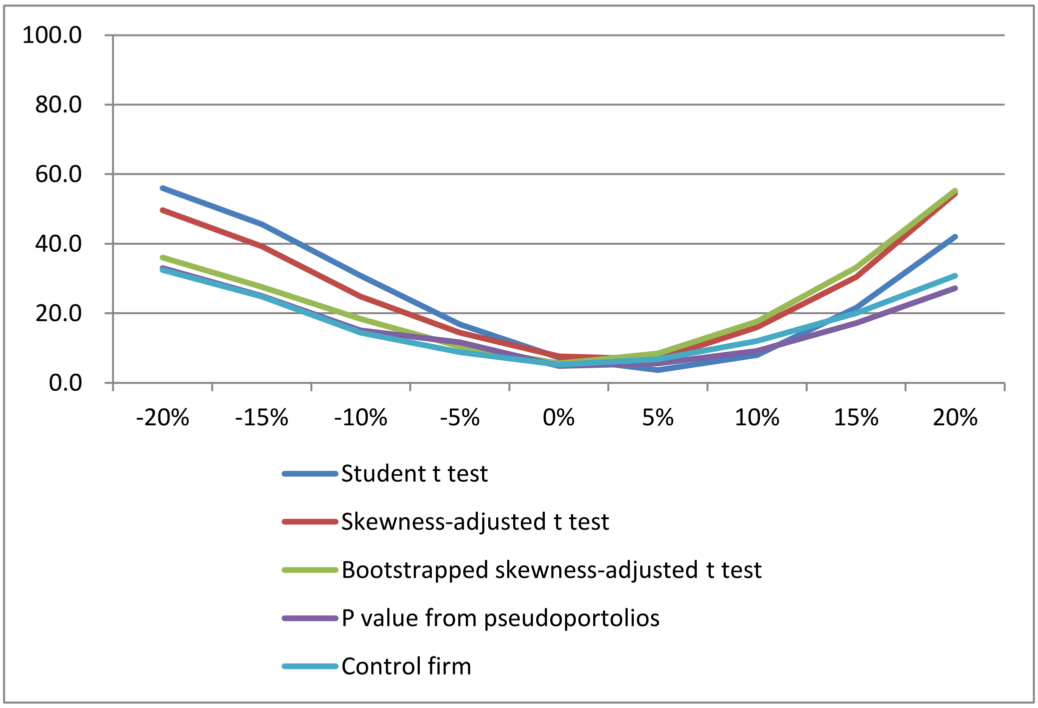


Figure 4.1 continued

Five years

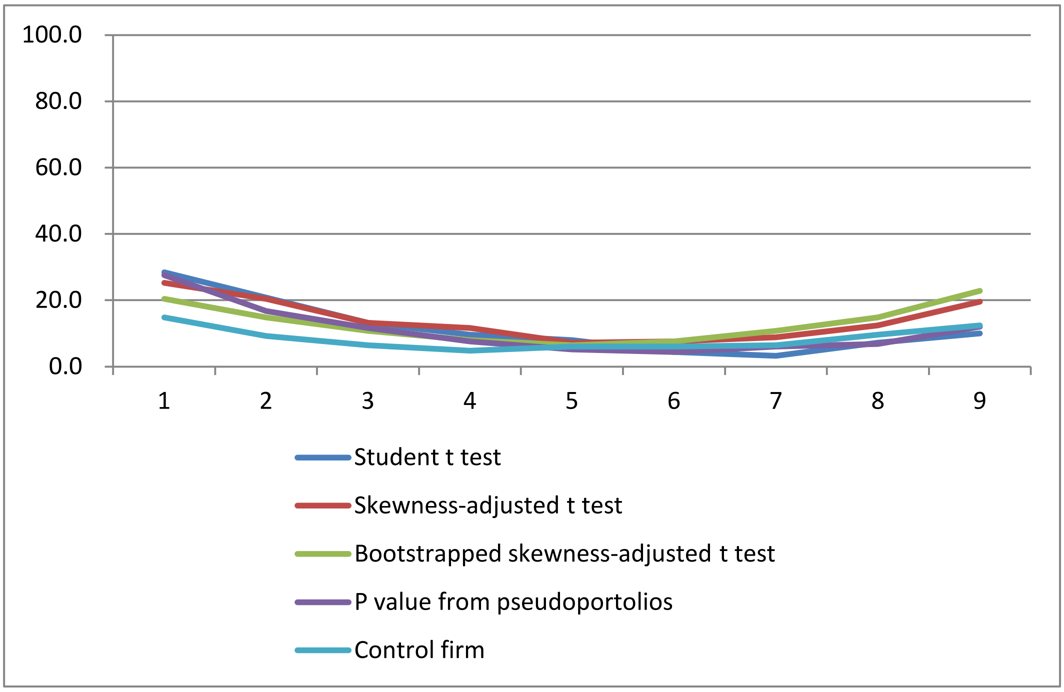


Table 4.6 Rejection frequency using cumulative abnormal returns and buy-and-hold abnormal returns in small firms

This table shows the rejection rates (as a percentage) of 250 samples with 200 small firms in size decile 1 which has firms with the smallest market value in September of each year at a significance level of 5% in one-tailed and two-tailed tests over a period of October 1982 to September 2008, under the null hypothesis of zero one-, three-year, five-year cumulative abnormal returns/buy-and-hold abnormal returns. The test is specified if the overall rejection rate is the same as the significance level together with equal rejection rate on both sides. The abnormal return is defined as the difference between the actual return of a stock and benchmark return. The benchmark is categorized into five types: 50 portfolios by size and book-to-market ratio (BTM) which decile firms by market value in September each year and then quintile firms based on BTM; 10 portfolios ranked by market value; 10 portfolios ranked by BTM, equally weighted market portfolio and control firm which is matched by size, or BTM or size/BTM with sample firm. Apart from conventional student test statistics, skewness-adjusted test statistics, bootstrapped skewness-adjusted test statistics and p-value (Barber, Lyon and Tsai 1999) derived from pseudoportfolios are examined. The event window is one, three and five years. The descriptive statistics of abnormal returns are summarized as follows.

Abnormal return is defined as the difference of actual individual i stock return and benchmark return as follows: $\mathrm{AR}_{\mathrm{it}} = \mathrm{R}_{\mathrm{it}} - \mathrm{BR}_{\mathrm{it}}$.

The cumulative abnormal returns (CARs) calculate abnormal returns of individual stocks in month t by taking the difference between the stock return and equally weighted benchmark return. $\mathrm{CAR}_{\mathrm{i,T}} = \sum_{t=1}^{T}(\mathrm{R}_{\mathrm{i,t}} - \frac{1}{n_t}\sum_{j=1}^{n_t}\mathrm{R}_{\mathrm{jt}})$. The conventional buy-and-hold return assumes investors tend to hold the same portfolio over the investment horizon: $\mathrm{PR}_{\mathrm{bnh}} = \sum_{i=1}^{n}\frac{[\prod_{t=0}^{t=T}(1+\mathrm{R}_{\mathrm{it}})]-1}{n}$. Stocks in the portfolios are supposed have return data over the study period. So I replace the missing returns with average portfolio return.. The conventional student test statistics is formulated as: $\mathrm{t} = \overline{\mathrm{AR}}/(\sigma(\mathrm{AR})/\sqrt{\mathrm{n}})$ whereas n is the number of firms in one sample which is 200 in this study. The skewness-adjusted test is calculated as: $\mathrm{skewness-adj.t} = \sqrt{\mathrm{n}}(\mathrm{S} + (1/3) * \mathrm{S}^2 * \gamma + (1/6\,\mathrm{n}) * \gamma$ whereas γ is skewness of cross-sectional abnormal returns and $\sqrt{\mathrm{n}}\mathrm{S}$ is student t test. The bootstrapped skewness-adjusted t test bootstraps abnormal returns in each sample 1000 times with size of 50. Then skewness-adjusted test for each subsample of 50 firms is computed in order to construct a distribution which is used to determine the critical value for rejection of the null hypothesis for each sample. Note that skewness-adjusted test calculated in the full sample is used to compare with the critical value. The Pseudo portfolio approach firstly randomly selects a benchmark firm from the same group (by size, BTM or size/BTM) as the sample firm. Then one mean buy-and-hold return for each sample is achieved. With all 250 mean returns, a distribution is constructed for the hypothesis test. Note that the null hypothesis is the equality of sample mean return with critical value.

Table 4.6 continued

CARs	5%	One year			Three years			Five years		
		α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
Conventional t test	**50 size/BTM portfolios**	7.2	10.8	2.8	7.2	9.6	2.8	5.2	11.6	2.0
	10 size portfolios	6.4	8.0	2.8	6.0	9.2	2.0	5.6	11.6	2.0
	10 BTM portfolios	62.4	0.4	76.8	75.6	0.0	90.4	76.0	0.0	84.8
	Equally-weighted market return	90.0	0.0	94.4	98.8	0.0	99.6	98.4	0.0	99.6
Skewness-adjusted test	**50 size/BTM portfolios**	8.0	10.0	4.8	8.8	9.6	6.0	6.4	10.4	6.0
	10 size portfolios	7.6	6.8	5.6	6.8	8.4	6.0	6.4	10.4	4.0
	10 BTM portfolios	76.4	0.4	83.2	86.4	0.0	94.0	81.6	0.0	87.6
	Equally-weighted market return	95.6	0.0	97.6	99.6	0.0	99.6	99.2	0.0	99.6
Bootstrapped skewness-adjusted test	**50 size/BTM portfolios**	6.8	5.2	7.6	8.0	6.0	7.2	7.6	7.6	6.4
	10 size portfolios	6.4	4.8	9.2	8.4	6.0	6.8	6.4	7.2	5.2
	10 BTM portfolios	76.0	0.4	83.6	85.2	0.0	92.0	80.4	0.0	87.6
	Equally-weighted market return	95.6	0.0	97.6	99.6	0.0	99.6	99.2	0.0	99.6
P value from Pseudoportfolios	**50 size/BTM portfolios**	5.2	4.4	6.4	8.0	6.8	6.8	6.8	8.0	6.4
	10 size portfolios	4.8	2.4	6.0	5.6	5.6	5.6	5.2	4.8	4.4
	10 BTM portfolios	69.6	0.4	84.4	81.6	0.0	93.2	80.4	0.0	88.8
Control firm approach	**by size and BTM**	4.0	3.2	4.8	6.4	3.6	7.6	2.8	7.2	2.4
	by size	4.0	3.6	6.0	3.2	4.8	6.4	4.8	6.0	4.8
	by BTM	45.2	0.4	58.0	50.8	0.0	69.6	48.8	0.0	62.8

Table 4.6 continued

BHARs	5%	One year			Three years			Five years		
		α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
Conventional t test	**50 size/BTM portfolios**	11.6	14.4	0.4	8.8	13.2	0.8	8.4	15.2	0.8
	10 size portfolios	12.0	16.0	0.0	9.6	13.6	1.2	14.0	21.6	0.4
	10 BTM portfolios	13.2	0.8	27.2	7.2	2.0	12.8	4.4	4.4	7.6
	Equally-weighted market return	32.0	0.4	50.4	22.8	0.0	40.4	16.8	0.0	28.0
Skewness-adjusted test	**50 size/BTM portfolios**	12.4	16.0	3.2	8.4	10.8	2.4	17.6	23.6	2.0
	10 size portfolios	13.2	14.8	3.6	9.2	11.2	4.4	16.8	21.6	2.0
	10 BTM portfolios	37.2	0.4	49.2	19.2	2.0	29.2	8.8	3.2	12.4
	Equally-weighted market return	60.4	0.4	71.2	46.0	0.0	59.2	31.6	0.0	44.4
Bootstrapped skewness-adjusted test	**50 size/BTM portfolios**	12.4	13.2	7.6	8.0	7.2	6.8	17.2	21.6	2.4
	10 size portfolios	12.0	11.2	7.6	9.2	7.6	6.8	12.4	16.0	3.6
	10 BTM portfolios	45.6	0.4	52.8	26.8	1.2	33.6	11.2	2.8	16.0
	Equally-weighted market return	65.6	0.4	74.8	54.8	0.0	64.0	36.8	0.0	48.4
P value from Pseudoportfolios	**50 size/BTM portfolios**	6.8	4.8	4.4	8.0	6.8	7.2	6.4	6.0	5.6
	10 size portfolios	4.4	3.6	5.2	6.4	6.0	7.2	3.2	2.8	2.8
	10 BTM portfolios	23.2	0.8	40.8	14.4	0.4	14.0	6.4	0.8	12.4
Control firm approach	**by size and BTM**	2.4	2.8	6.0	4.0	5.2	6.0	4.8	5.6	6.4
	by size	2.0	1.6	4.0	5.6	2.4	7.6	6.4	4.8	8.8
	by BTM	19.6	1.2	30.8	8.8	1.6	14.8	3.2	4.0	9.6

Table 4.7 Rejection frequency using cumulative abnormal returns and buy-and-hold abnormal returns in large firms

This table shows the rejection rates (as a percentage) of 250 samples with 200 large firms in size decile 10 which has firms with the largest market value in September of each year at a significance level of 5% in one-tailed and two-tailed tests over a period of October 1982 to September 2008, under the null hypothesis of zero one-, three-year, five-year cumulative abnormal returns/buy-and-hold abnormal returns. The test is specified if the overall rejection rate is the same as the significance level together with equal rejection rate on both sides. The abnormal return is defined as the difference between the actual return of a stock and benchmark return. The benchmark is categorized into five types: 50 portfolios by size and book-to-market ratio (BTM) which decile firms by market value in September each year and then quintile firms based on BTM; 10 portfolios ranked by market value; 10 portfolios ranked by BTM, equally weighted market portfolio and control firm which is matched by size, or BTM or size/BTM with sample firm. Apart from conventional student test statistics, skewness-adjusted test statistics, bootstrapped skewness-adjusted test statistics and p-value (Barber, Lyon and Tsai 1999) derived from pseudoportfolios are examined. The event window is one, three and five years. The descriptive statistics of abnormal returns are summarized as follows.

Abnormal return is defined as the difference of actual individual i stock return and benchmark return as follows: $\mathrm{AR_{it}} = \mathrm{R_{it}} - \mathrm{BR_{it}}$.

The cumulative abnormal returns (CARs) calculate abnormal returns of individual stocks in month t by taking the difference between the stock return and equally weighted benchmark return. $\mathrm{CAR_{i,T}} = \sum_{t=1}^{T}(\mathrm{R_{i,t}} - \frac{1}{n_t}\sum_{j=1}^{n_t}\mathrm{R_{jt}})$. The conventional buy-and-hold return assumes investors tend to hold the same portfolio over the investment horizon: $\mathrm{PR_{bnh}} = \sum_{i=1}^{n}\frac{[\prod_{t=0}^{t=T}(1+\mathrm{R_{it}})]-1}{n}$. Stocks in the portfolios are supposed have return data over the study period. So I replace the missing returns with average portfolio return.. The conventional student test statistics is formulated as: $t = \overline{AR}/(\sigma(AR)/\sqrt{n})$ whereas n is the number of firms in one sample which is 200 in this study. The skewness-adjusted test is calculated as: $\mathrm{skewness-adj.}\,t = \sqrt{n}(S + (1/3) * S^2 * \gamma + (1/6\,n) * \gamma$ whereas γ is skewness of cross-sectional abnormal returns and $\sqrt{n}S$ is student t test. The bootstrapped skewness-adjusted t test bootstraps abnormal returns in each sample 1000 times with size of 50. Then skewness-adjusted test for each subsample of 50 firms is computed in order to construct a distribution which is used to determine the critical value for rejection of the null hypothesis for each sample. Note that skewness-adjusted test calculated in the full sample is used to compare with the critical value. The Pseudo portfolio approach firstly randomly selects a benchmark firm from the same group (by size, BTM or size/BTM) as the sample firm. Then one mean buy-and-hold return for each sample is achieved. With all 250 mean returns, a distribution is constructed for the hypothesis test. Note that the null hypothesis is the equality of sample mean return with critical value.

Table 4.7 continued

CARs	5%	One year			Three years			Five years		
		α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
Conventional t test	**50 size/BTM portfolios**	5.2	4.0	6.4	4.0	4.0	5.2	4.0	3.6	6.4
	10 size portfolios	4.8	4.4	7.2	3.6	3.6	4.8	4.0	3.6	4.8
	10 BTM portfolios	10.0	16.8	1.2	20.0	29.6	0.8	34.4	52.4	0.0
	Equally-weighted market return	20.0	29.2	0.4	36.4	47.6	0.0	58.0	68.8	0.0
Skewness-adjusted test	**50 size/BTM portfolios**	6.4	5.6	6.4	4.8	4.0	5.2	4.0	3.6	5.6
	10 size portfolios	6.4	5.2	6.4	4.4	3.6	4.4	4.8	4.4	4.0
	10 BTM portfolios	12.4	19.2	1.2	23.2	31.6	0.8	38.8	55.2	0.0
	Equally-weighted market return	22.8	30.8	0.4	38.4	49.2	0.0	61.6	70.8	0.0
Bootstrapped skewness-adjusted test	**50 size/BTM portfolios**	6.0	5.2	5.2	4.4	4.0	4.8	3.2	3.6	3.6
	10 size portfolios	4.4	6.0	6.4	3.6	3.6	4.0	4.4	4.4	2.4
	10 BTM portfolios	10.8	17.6	0.8	22.0	30.8	0.8	34.8	53.2	0.0
	Equally-weighted market return	21.6	32.0	0.0	35.2	49.6	0.0	59.6	68.4	0.0
P value from Pseudoportfolios	**50 size/BTM portfolios**	6.4	8.8	6.4	5.2	4.4	5.2	5.6	4.4	5.6
	10 size portfolios	6.8	6.8	6.0	4.4	3.2	5.2	4.8	4.4	4.8
	10 BTM portfolios	0.4	3.2	0.0	2.4	8.0	0.0	9.6	18.8	0.0
Control firm approach	**by size and BTM**	5.6	7.2	6.0	3.2	4.0	4.0	6.4	5.2	5.2
	by size	6.8	5.6	5.6	4.8	4.0	6.0	3.6	4.4	4.4
	by BTM	6.0	8.4	4.0	11.6	17.2	0.8	14.8	19.2	0.0

Table 4.7 continued

BHARs	5%	One year			Three years			Five years		
		α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
Conventional t test	**50 size/BTM portfolios**	5.2	6.8	4.4	6.8	7.6	4.8	5.6	4.8	4.0
	10 size portfolios	4.0	5.2	3.6	8.0	6.0	4.8	7.2	6.4	4.0
	10 BTM portfolios	22.0	30.8	0.0	38.8	48.8	0.0	47.6	56.8	0.0
	Equally-weighted market return	43.2	55.6	0.0	80.0	87.2	0.0	90.0	94.4	0.0
Skewness-adjusted test	**50 size/BTM portfolios**	6.0	6.0	5.2	7.2	6.8	6.4	5.2	4.8	5.2
	10 size portfolios	4.4	5.2	6.4	8.0	6.0	6.0	7.2	5.6	4.8
	10 BTM portfolios	22.4	29.6	0.0	36.8	48.8	0.0	44.8	55.6	0.0
	Equally-weighted market return	43.2	54.0	0.0	77.2	86.4	0.0	89.6	93.6	0.0
Bootstrapped skewness-adjusted test	**50 size/BTM portfolios**	6.4	5.6	4.8	6.4	5.6	5.2	7.6	4.4	6.0
	10 size portfolios	5.2	4.8	6.8	7.2	5.6	5.6	6.8	5.2	5.6
	10 BTM portfolios	20.4	28.4	0.0	32.4	45.6	0.0	40.0	50.4	0.0
	Equally-weighted market return	38.4	52.0	0.0	68.8	80.8	0.0	85.2	90.0	0.0
P value from Pseudoportfolios	**50 size/BTM portfolios**	7.6	6.8	7.2	8.0	7.2	6.0	6.8	4.8	6.4
	10 size portfolios	4.8	5.6	6.0	7.6	5.2	6.0	6.8	5.6	5.6
	10 BTM portfolios	2.0	4.0	0.0	6.0	11.6	0.0	5.2	12.4	0.0
Control firm approach	**by size and BTM**	6.8	8.4	6.8	6.8	7.2	6.4	5.2	5.6	5.6
	by size	8.4	6.0	5.2	5.2	7.6	4.8	3.6	5.6	4.0
	by BTM	5.2	10.4	4.0	10.0	16.0	1.2	7.2	14.4	0.8

Table 4.8 Rejection frequency using cumulative abnormal returns and buy-and-hold abnormal returns in firms with low book-to-market ratio

This table shows the rejection rates (as a percentage) of 250 random samples with 200 firms in BTM decile 1 which has firms with the smallest BTM at a significance level of 5% in one-tailed and two-tailed tests over a period of October 1982 to September 2008, under the null hypothesis of zero one-, three-year, five-year cumulative abnormal returns/buy-and-hold abnormal returns. The test is specified if the overall rejection rate is the same as the significance level together with equal rejection rate on both sides. The abnormal return is defined as the difference between the actual return of a stock and benchmark return. The benchmark is categorized into five types: 50 portfolios by size and book-to-market ratio (BTM) which decile firms by market value in September each year and then quintile firms based on BTM; 10 portfolios ranked by market value; 10 portfolios ranked by BTM, equally weighted market portfolio and control firm which is matched by size, or BTM or size/BTM with sample firm. Apart from conventional student test statistics, skewness-adjusted test statistics, bootstrapped skewness-adjusted test statistics and p-value (Barber, Lyon and Tsai 1999) derived from pseudoportfolios are examined. The event window is one, three and five years. The descriptive statistics of abnormal returns are summarized as follows.

Abnormal return is defined as the difference of actual individual i stock return and benchmark return as follows: $\mathrm{AR}_{\mathrm{it}} = \mathrm{R}_{\mathrm{it}} - \mathrm{BR}_{\mathrm{it}}$.

The cumulative abnormal returns (CARs) calculate abnormal returns of individual stocks in month t by taking the difference between the stock return and equally weighted benchmark return. $\mathrm{CAR}_{\mathrm{i,T}} = \sum_{\mathrm{t=1}}^{\mathrm{T}} (\mathrm{R}_{\mathrm{i,t}} - \frac{1}{\mathrm{n_t}} \sum_{\mathrm{j=1}}^{\mathrm{n_t}} \mathrm{R}_{\mathrm{jt}})$. The conventional buy-and-hold return assumes investors tend to hold the same portfolio over the investment horizon: $\mathrm{PR}_{\mathrm{bnh}} = \sum_{\mathrm{i=1}}^{\mathrm{n}} \frac{[\prod_{\mathrm{t=0}}^{\mathrm{t=T}} (1+\mathrm{R}_{\mathrm{it}})]-1}{\mathrm{n}}$. Stocks in the portfolios are supposed have return data over the study period. So I replace the missing returns with average portfolio return.. The conventional student test statistics is formulated as: $\mathrm{t} = \overline{\mathrm{AR}}/(\sigma(\mathrm{AR})/\sqrt{\mathrm{n}})$ whereas n is the number of firms in one sample which is 200 in this study. The skewness-adjusted test is calculated as: $\mathrm{skewness-adj.t} = \sqrt{\mathrm{n}}(\mathrm{S} + (1/3) * \mathrm{S}^2 * \gamma + (1/6\,\mathrm{n}) * \gamma$ whereas γ is skewness of cross-sectional abnormal returns and $\sqrt{\mathrm{n}}\mathrm{S}$ is student t test. The bootstrapped skewness-adjusted t test bootstraps abnormal returns in each sample 1000 times with size of 50. Then skewness-adjusted test for each subsample of 50 firms is computed in order to construct a distribution which is used to determine the critical value for rejection of the null hypothesis for each sample. Note that skewness-adjusted test calculated in the full sample is used to compare with the critical value. The Pseudo portfolio approach firstly randomly selects a benchmark firm from the same group (by size, BTM or size/BTM) as the sample firm. Then one mean buy-and-hold return for each sample is achieved. With all 250 mean returns, a distribution is constructed for the hypothesis test. Note that the null hypothesis is the equality of sample mean return with critical value.

Table 4.8 continued

CARs	5%	One year			Three years			Five years		
		α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
Conventional t test	**50 size/BTM portfolios**	4.8	4.8	6.0	4.4	6.4	3.2	5.2	6.8	2.8
	10 size portfolios	38.0	48.4	0.0	54.8	64.0	0.0	52.4	66.8	0.0
	10 BTM portfolios	5.6	4.8	6.8	4.4	6.4	3.2	5.2	6.4	3.2
	Equally-weighted market return	52.0	59.6	0.0	72.4	78.8	0.0	68.4	78.8	0.0
Skewness-adjusted test	**50 size/BTM portfolios**	6.0	5.2	6.8	4.4	5.6	3.6	6.0	6.8	3.6
	10 size portfolios	36.0	48.0	0.0	53.6	62.4	0.0	52.4	65.6	0.0
	10 BTM portfolios	7.6	4.8	8.0	4.4	6.0	4.0	4.8	6.4	4.0
	Equally-weighted market return	49.6	58.4	0.0	68.0	75.6	0.0	65.6	75.6	0.0
Bootstrapped skewness-adjusted test	**50 size/BTM portfolios**	5.6	5.2	7.2	4.4	5.2	3.6	5.6	6.4	3.6
	10 size portfolios	31.2	42.8	0.0	48.4	57.6	0.0	46.8	61.2	0.0
	10 BTM portfolios	6.8	3.6	7.2	4.8	5.2	3.6	4.0	5.2	4.0
	Equally-weighted market return	42.0	53.6	0.0	61.2	71.2	0.0	60.8	70.8	0.0
P value from Pseudoportfolios	**50 size/BTM portfolios**	9.2	5.6	8.0	6.4	8.8	3.6	7.2	8.8	4.4
	10 size portfolios	52.8	59.2	0.4	69.6	78.8	0.0	70.0	76.8	0.0
	10 BTM portfolios	4.8	4.8	6.0	4.4	6.4	3.6	4.4	6.4	4.8
Control firm approach	**by size and BTM**	38.0	48.4	0.0	4.0	4.8	2.8	5.6	7.6	2.0
	by size	5.6	4.8	6.8	34.4	46.0	0.0	29.6	43.2	0.0
	by BTM	52.0	59.6	0.0	4.4	3.2	4.0	4.4	4.8	4.8

Table 4.8 continued

BHARs	5%	One year			Three years			Five years		
		α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
Conventional t test	**50 size/BTM portfolios**	5.6	6.4	2.0	3.2	5.2	2.8	7.2	10.8	2.0
	10 size portfolios	21.2	29.2	0.4	46.4	54.8	0.0	49.2	55.2	0.0
	10 BTM portfolios	6.0	9.6	2.4	5.6	7.6	2.0	9.2	12.8	1.6
	Equally-weighted market return	39.2	49.6	0.0	68.8	76.8	0.0	68.8	73.6	0.0
Skewness-adjusted test	**50 size/BTM portfolios**	5.6	6.0	4.4	4.4	4.0	4.4	7.2	7.6	7.6
	10 size portfolios	18.4	24.4	1.2	34.0	45.2	0.0	38.4	46.4	0.0
	10 BTM portfolios	7.2	8.8	4.8	4.8	6.4	3.2	6.4	9.6	5.6
	Equally-weighted market return	30.4	41.6	0.4	55.2	64.4	0.0	50.0	59.6	0.0
Bootstrapped skewness-adjusted test	**50 size/BTM portfolios**	8.0	5.2	8.4	4.0	2.0	4.8	9.2	4.0	9.2
	10 size portfolios	15.2	20.4	2.8	23.6	30.8	0.0	23.6	32.4	0.0
	10 BTM portfolios	6.8	6.4	5.2	5.2	2.4	4.0	7.2	5.6	8.0
	Equally-weighted market return	23.6	32.0	0.4	36.8	47.6	0.0	34.0	42.8	0.0
P value from Pseudoportfolios	**50 size/BTM portfolios**	9.2	5.6	8.0	5.6	2.8	6.8	12.0	7.2	11.2
	10 size portfolios	28.8	34.8	4.8	58.4	66.0	0.4	60.0	67.6	0.4
	10 BTM portfolios	5.6	5.6	4.0	5.6	4.4	3.6	8.8	5.6	6.4
Control firm approach	**by size and BTM**	5.6	2.4	6.0	2.0	2.4	3.6	4.8	4.8	4.8
	by size	13.2	19.2	0.8	21.2	29.6	0.0	26.8	33.6	0.0
	by BTM	5.2	5.2	5.2	3.6	6.0	3.6	3.2	4.0	5.2

Table 4.9 Rejection frequency using cumulative abnormal returns and buy-and-hold abnormal returns in firms with high book-to-market ratio

This table shows the rejection rates (as a percentage) of 250 random samples with 200 firms in BTM decile 10 which has firms with the highest BTM at a significance level of 5% in one-tailed and two-tailed tests over a period of October 1982 to September 2008, under the null hypothesis of zero one-, three-year, five-year cumulative abnormal returns/buy-and-hold abnormal returns. The test is specified if the overall rejection rate is the same as the significance level together with equal rejection rate on both sides. The abnormal return is defined as the difference between the actual return of a stock and benchmark return. The benchmark is categorized into five types: 50 portfolios by size and book-to-market ratio (BTM) which decile firms by market value in September each year and then quintile firms based on BTM; 10 portfolios ranked by market value; 10 portfolios ranked by BTM, equally weighted market portfolio and control firm which is matched by size, or BTM or size/BTM with sample firm. Apart from conventional student test statistics, skewness-adjusted test statistics, bootstrapped skewness-adjusted test statistics and p-value (Barber, Lyon and Tsai 1999) derived from pseudoportfolios are examined. The event window is one, three and five years. The descriptive statistics of abnormal returns are summarized as follows.

Abnormal return is defined as the difference of actual individual i stock return and benchmark return as follows: $AR_{it} = R_{it} - BR_{it}$.

The cumulative abnormal returns (CARs) calculate abnormal returns of individual stocks in month t by taking the difference between the stock return and equally weighted benchmark return. $CAR_{i,T} = \sum_{t=1}^{T}(R_{i,t} - \frac{1}{n_t}\sum_{j=1}^{n_t} R_{jt})$. The conventional buy-and-hold return assumes investors tend to hold the same portfolio over the investment horizon: $PR_{bnh} = \sum_{i=1}^{n}\frac{[\prod_{t=0}^{t=T}(1+R_{it})]-1}{n}$. Stocks in the portfolios are supposed have return data over the study period. So I replace the missing returns with average portfolio return.. The conventional student test statistics is formulated as: $t = \overline{AR}/(\sigma(AR)/\sqrt{n})$ whereas n is the number of firms in one sample which is 200 in this study. The skewness-adjusted test is calculated as: $\mathrm{skewness-adj.t} = \sqrt{n}(S + (1/3) * S^2 * \gamma + (1/6\,n) * \gamma$ whereas γ is skewness of cross-sectional abnormal returns and $\sqrt{n}S$ is student t test. The bootstrapped skewness-adjusted t test bootstraps abnormal returns in each sample 1000 times with size of 50. Then skewness-adjusted test for each subsample of 50 firms is computed in order to construct a distribution which is used to determine the critical value for rejection of the null hypothesis for each sample. Note that skewness-adjusted test calculated in the full sample is used to compare with the critical value. The Pseudo portfolio approach firstly randomly selects a benchmark firm from the same group (by size, BTM or size/BTM) as the sample firm. Then one mean buy-and-hold return for each sample is achieved. With all 250 mean returns, a distribution is constructed for the hypothesis test. Note that the null hypothesis is the equality of sample mean return with critical value.

Table 4.9 continued

CARs	5%	One year			Three years			Five years		
		α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
Conventional t test	**50 size/BTM portfolios**	6.0	6.0	4.0	4.8	4.4	3.6	4.0	6.8	4.0
	10 size portfolios	4.4	2.4	9.6	7.2	0.4	14.0	13.2	0.8	18.8
	10 BTM portfolios	8.0	11.2	1.6	5.6	7.6	1.6	3.6	6.0	0.8
	Equally-weighted market return	56.4	0.0	66.8	84.8	0.0	95.6	92.8	0.0	97.6
Skewness-adjusted test	**50 size/BTM portfolios**	7.2	6.0	5.6	5.6	4.4	6.8	6.0	6.8	8.4
	10 size portfolios	8.0	2.4	15.2	12.0	0.4	20.4	18.4	0.8	27.6
	10 BTM portfolios	9.2	10.8	4.4	6.0	6.8	5.2	4.0	4.8	4.4
	Equally-weighted market return	64.4	0.0	73.6	92.8	0.0	96.4	95.6	0.0	98.4
Bootstrapped skewness-adjusted test	**50 size/BTM portfolios**	8.4	6.0	8.0	8.0	4.0	9.2	6.8	4.4	8.0
	10 size portfolios	9.6	1.6	15.2	14.0	0.4	21.2	19.2	0.4	28.0
	10 BTM portfolios	7.6	9.2	6.4	6.4	5.6	7.2	5.2	3.6	6.8
	Equally-weighted market return	64.4	0.0	71.6	91.2	0.0	96.0	93.2	0.0	98.0
P value from Pseudoportfolios	**50 size/BTM portfolios**	7.6	4.8	6.8	10.4	4.4	9.2	8.8	3.2	9.2
	10 size portfolios	4.4	0.4	8.0	8.0	0.4	14.4	11.6	0.4	21.2
	10 BTM portfolios	5.2	6.0	3.6	6.0	5.2	5.6	6.8	3.2	6.4
Control firm approach	**by size and BTM**	4.4	5.2	2.4	5.2	3.6	5.6	3.6	7.6	4.4
	by size	4.4	2.0	9.2	7.6	1.2	12.4	6.4	2.0	11.6
	by BTM	6.8	4.0	6.4	3.2	6.4	6.8	4.0	5.6	5.6

Table 4.9 continued

BHARs	5%	One year			Three years			Five years		
		α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
Conventional t test	**50 size/BTM portfolios**	6.4	10.8	1.6	10.0	14.0	0.8	3.2	9.2	1.2
	10 size portfolios	6.4	4.4	3.6	6.0	0.4	14.0	13.6	0.8	23.2
	10 BTM portfolios	10.8	15.2	1.2	13.6	19.2	0.4	10.0	15.6	0.4
	Equally-weighted market return	16.0	1.2	25.2	29.2	0.0	44.8	41.6	0.0	57.6
Skewness-adjusted test	**50 size/BTM portfolios**	7.6	10.0	4.8	12.4	13.6	4.0	13.6	16.8	4.0
	10 size portfolios	8.0	5.2	6.0	16.4	0.0	28.8	24.8	0.8	35.2
	10 BTM portfolios	10.0	13.6	3.6	11.6	16.8	2.0	12.4	16.4	3.2
	Equally-weighted market return	29.6	1.2	39.2	54.0	0.0	64.8	65.2	0.0	76.0
Bootstrapped skewness-adjusted test	**50 size/BTM portfolios**	6.8	7.6	8.0	10.8	12.0	8.8	18.8	19.6	6.0
	10 size portfolios	10.4	4.8	11.6	24.4	0.0	34.4	29.6	0.8	38.0
	10 BTM portfolios	10.8	10.0	7.2	11.2	10.4	5.6	10.4	12.4	4.0
	Equally-weighted market return	36.8	1.2	45.2	59.2	0.0	68.8	70.0	0.0	77.2
P value from Pseudoportfolios	**50 size/BTM portfolios**	6.4	4.8	4.8	8.4	6.0	8.8	8.0	2.4	8.8
	10 size portfolios	3.2	1.6	4.8	11.2	0.0	19.2	10.0	0.0	21.6
	10 BTM portfolios	5.2	6.4	2.4	4.8	5.2	5.6	6.4	4.0	6.8
Control firm approach	**by size and BTM**	3.6	4.4	2.4	4.4	4.0	3.2	3.2	4.0	4.8
	by size	6.0	3.6	8.4	8.4	1.2	15.2	12.0	0.8	17.6
	by BTM	4.0	4.4	5.2	4.0	2.8	7.2	3.6	3.2	6.4

Table 4.10Rejection frequency using cumulative abnormal returns and buy-and-hold abnormal returns in firms with industry clustering

This table shows the rejection rates (as a percentage) of 250 random samples with 200 firms which are clustered in industry at a significance level of 5% in one-tailed and two-tailed tests over a period of October 1982 to September 2008, under the null hypothesis of zero one-, three-year, five-year cumulative abnormal returns/buy-and-hold abnormal returns. The test is specified if the overall rejection rate is the same as the significance level together with equal rejection rate on both sides. The abnormal return is defined as the difference between the actual return of a stock and benchmark return. The benchmark is categorized into five types: 50 portfolios by size and book-to-market ratio (BTM) which decile firms by market value in September each year and then quintile firms based on BTM; 10 portfolios ranked by market value; 10 portfolios ranked by BTM, equally weighted market portfolio and control firm which is matched by size, or BTM or size/BTM with sample firm. Apart from conventional student test statistics, skewness-adjusted test statistics, bootstrapped skewness-adjusted test statistics and p-value (Barber, Lyon and Tsai 1999) derived from pseudoportfolios are examined. The event window is one, three and five years. The descriptive statistics of abnormal returns are summarized as follows.

Abnormal return is defined as the difference of actual individual i stock return and benchmark return as follows: $AR_{it} = R_{it} - BR_{it}$.

The cumulative abnormal returns (CARs) calculate abnormal returns of individual stocks in month t by taking the difference between the stock return and equally weighted benchmark return. $CAR_{i,T} = \sum_{t=1}^{T}(R_{i,t} - \frac{1}{n_t}\sum_{j=1}^{n_t} R_{jt})$. The conventional buy-and-hold return assumes investors tend to hold the same portfolio over the investment horizon: $PR_{bnh} = \sum_{i=1}^{n} \frac{[\prod_{t=0}^{t=T}(1+R_{it})]-1}{n}$. Stocks in the portfolios are supposed have return data over the study period. So I replace the missing returns with average portfolio return.. The conventional student test statistics is formulated as: $t = \overline{AR}/(\sigma(AR)/\sqrt{n})$ whereas n is the number of firms in one sample which is 200 in this study. The skewness-adjusted test is calculated as: $skewness - adj. t = \sqrt{n}(S + (1/3) * S^2 * \gamma + (1/6\,n) * \gamma$ whereas γ is skewness of cross-sectional abnormal returns and $\sqrt{n}S$ is student t test. The bootstrapped skewness-adjusted t test bootstraps abnormal returns in each sample 1000 times with size of 50. Then skewness-adjusted test for each subsample of 50 firms is computed in order to construct a distribution which is used to determine the critical value for rejection of the null hypothesis for each sample. Note that skewness-adjusted test calculated in the full sample is used to compare with the critical value. The Pseudo portfolio approach firstly randomly selects a benchmark firm from the same group (by size, BTM or size/BTM) as the sample firm. Then one mean buy-and-hold return for each sample is achieved. With all 250 mean returns, a distribution is constructed for the hypothesis test. Note that the null hypothesis is the equality of sample mean return with critical value.

Table 4.10 continued

CARs	5%	One year			Three years			Five years		
		α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
Conventional t test	**50 size/BTM portfolios**	5.6	6.0	3.6	8.0	12.8	2.8	5.6	8.0	4.4
	10 size portfolios	5.6	7.2	5.6	8.4	12.8	1.2	7.2	9.6	1.6
	10 BTM portfolios	5.6	5.6	4.8	8.8	12.8	2.8	6.8	8.4	2.8
	Equally-weighted market return	6.0	4.4	6.8	5.6	7.2	6.8	1.2	2.0	4.0
Skewness-adjusted test	**50 size/BTM portfolios**	6.8	6.4	6.0	8.8	13.6	3.6	6.0	8.0	5.2
	10 size portfolios	6.0	8.0	6.0	9.6	13.2	3.2	7.6	9.6	2.8
	10 BTM portfolios	5.6	6.8	7.2	10.8	12.8	3.6	7.6	8.4	3.6
	Equally-weighted market return	6.4	4.8	9.2	6.4	7.2	8.4	3.6	2.0	6.8
Bootstrapped skewness-adjusted test	**50 size/BTM portfolios**	6.8	6.4	6.0	8.4	12.0	3.2	5.2	6.4	4.0
	10 size portfolios	5.6	6.4	6.0	8.0	12.8	3.2	7.6	9.2	2.4
	10 BTM portfolios	6.8	5.2	7.6	9.6	13.2	4.0	6.8	7.2	4.4
	Equally-weighted market return	9.2	4.0	10.0	5.6	5.6	6.0	4.0	2.0	6.8
P value from Pseudoportfolios	**50 size/BTM portfolios**	6.8	5.6	6.0	8.8	10.0	3.2	7.2	9.6	4.8
	10 size portfolios	5.6	3.6	4.4	7.2	10.8	2.0	8.0	9.2	2.4
	10 BTM portfolios	6.8	5.2	6.4	8.0	10.8	2.8	6.4	7.6	2.8
Control firm approach	**by size and BTM**	4.4	5.2	5.2	6.8	10.0	3.2	6.8	10.0	3.2
	by size	6.8	6.0	6.0	4.4	6.4	2.8	5.2	6.0	2.8
	by BTM	5.2	4.0	5.6	4.4	8.0	4.0	6.4	6.0	4.4

Table 4.10 continued

5%	One year			Three years			Five years		
	α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
50 size/BTM portfolios	8.4	11.2	3.2	13.6	16.8	1.2	6.4	11.6	2.4
10 size portfolios	9.6	13.2	2.8	14.0	20.4	1.6	8.0	12.8	2.0
10 BTM portfolios	10.8	11.6	2.8	14.8	17.6	0.8	8.8	12.4	2.4
Equally-weighted market return	10.0	12.0	2.8	13.2	17.6	0.8	8.8	11.6	2.0
50 size/BTM portfolios	10.0	10.8	5.2	12.4	16.4	2.8	9.2	11.2	6.0
10 size portfolios	11.6	13.2	4.8	13.2	18.0	2.8	9.2	11.6	4.4
10 BTM portfolios	11.6	11.2	3.6	12.8	15.6	2.8	8.0	10.0	4.0
Equally-weighted market return	10.4	10.0	4.8	12.4	16.0	3.6	6.0	8.4	5.6
50 size/BTM portfolios	10.4	11.2	6.0	9.6	13.2	4.0	8.8	8.8	6.4
10 size portfolios	10.8	11.6	7.2	12.0	14.0	4.8	7.6	8.4	4.8
10 BTM portfolios	10.8	10.8	6.4	9.2	11.6	4.4	5.6	6.4	5.2
Equally-weighted market return	10.0	7.2	5.6	11.2	11.6	4.0	6.0	6.0	5.2
50 size/BTM portfolios	6.4	6.8	5.6	8.4	10.0	3.6	4.8	3.2	5.2
10 size portfolios	5.2	8.0	4.4	7.6	9.6	3.6	3.6	4.4	2.8
10 BTM portfolios	5.2	6.8	4.0	14.4	0.4	14.0	4.4	3.6	4.8
by size and BTM	8.4	11.2	3.2	13.6	17.2	1.2	6.8	11.6	2.4
by size	5.2	8.4	4.8	5.6	6.0	5.6	4.4	2.8	5.2
by BTM	4.8	6.0	4.0	7.6	6.8	5.2	4.4	2.8	6.8

Table 4.11Rejection frequency using cumulative abnormal returns and buy-and-hold abnormal returns in firms with overlapping returns

This table shows the rejection rates (as a percentage) of 250 random samples with 200 firms which have multiple events within the study period at a significance level of 5% in one-tailed and two-tailed tests over a period of October 1982 to September 2008, under the null hypothesis of zero one-, three-year, five-year cumulative abnormal returns/buy-and-hold abnormal returns. The test is specified if the overall rejection rate is the same as the significance level together with equal rejection rate on both sides. The abnormal return is defined as the difference between the actual return of a stock and benchmark return. The benchmark is categorized into five types: 50 portfolios by size and book-to-market ratio (BTM) which decile firms by market value in September each year and then quintile firms based on BTM; 10 portfolios ranked by market value; 10 portfolios ranked by BTM, equally weighted market portfolio and control firm which is matched by size, or BTM or size/BTM with sample firm. Apart from conventional student test statistics, skewness-adjusted test statistics, bootstrapped skewness-adjusted test statistics and p-value (Barber, Lyon and Tsai 1999) derived from pseudoportfolios are examined. The event window is one, three and five years. The descriptive statistics of abnormal returns are summarized as follows.

Abnormal return is defined as the difference of actual individual i stock return and benchmark return as follows: $AR_{it} = R_{it} - BR_{it}$.

The cumulative abnormal returns (CARs) calculate abnormal returns of individual stocks in month t by taking the difference between the stock return and equally weighted benchmark return. $CAR_{i,T} = \sum_{t=1}^{T}(R_{i,t} - \frac{1}{n_t}\sum_{j=1}^{n_t} R_{jt})$. The conventional buy-and-hold return assumes investors tend to hold the same portfolio over the investment horizon: $PR_{bnh} = \sum_{i=1}^{n}\frac{[\prod_{t=0}^{t=T}(1+R_{it})]-1}{n}$. Stocks in the portfolios are supposed have return data over the study period. So I replace the missing returns with average portfolio return.. The conventional student test statistics is formulated as: $t = \overline{AR}/(\sigma(AR)/\sqrt{n})$ whereas n is the number of firms in one sample which is 200 in this study. The skewness-adjusted test is calculated as: $\text{skewness} - \text{adj.}\, t = \sqrt{n}(S + (1/3) * S^2 * \gamma + (1/6\, n) * \gamma$ whereas γ is skewness of cross-sectional abnormal returns and $\sqrt{n}S$ is student t test. The bootstrapped skewness-adjusted t test bootstraps abnormal returns in each sample 1000 times with size of 50. Then skewness-adjusted test for each subsample of 50 firms is computed in order to construct a distribution which is used to determine the critical value for rejection of the null hypothesis for each sample. Note that skewness-adjusted test calculated in the full sample is used to compare with the critical value. The Pseudo portfolio approach firstly randomly selects a benchmark firm from the same group (by size, BTM or size/BTM) as the sample firm. Then one mean buy-and-hold return for each sample is achieved. With all 250 mean returns, a distribution is constructed for the hypothesis test. Note that the null hypothesis is the equality of sample mean return with critical value.

Table 4.11 continued

CARs	5%	One year			Three years			Five years		
		α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
Conventional t test	**50 size/BTM portfolios**	18.0	20.4	6.4	22.4	20.0	11.2	18.8	15.6	12.0
	10 size portfolios	16.8	20.8	6.4	23.2	22.4	8.4	16.4	16.4	7.2
	10 BTM portfolios	18.8	20.0	8.0	21.6	18.0	10.4	15.2	10.0	14.8
	Equally-weighted market return	18.0	19.6	6.8	24.0	21.6	9.6	18.0	10.0	14.8
Skewness-adjusted test	**50 size/BTM portfolios**	19.2	20.8	7.2	25.2	20.4	11.6	20.4	14.8	12.8
	10 size portfolios	18.8	21.6	6.4	23.6	23.2	8.8	17.6	16.8	8.4
	10 BTM portfolios	19.2	19.2	8.4	23.2	18.0	10.4	17.2	9.6	15.6
	Equally-weighted market return	19.2	19.6	7.6	24.8	21.2	10.8	19.6	10.4	17.2
Bootstrapped skewness-adjusted test	**50 size/BTM portfolios**	18.4	18.8	6.8	22.0	20.0	11.2	18.8	13.6	12.4
	10 size portfolios	16.4	16.4	5.6	23.2	22.4	8.0	15.6	16.8	7.2
	10 BTM portfolios	17.2	17.2	8.8	22.0	18.0	10.8	16.8	8.8	16.0
	Equally-weighted market return	16.0	17.2	6.8	21.6	17.6	11.2	20.4	9.2	16.0
P value from Pseudoportfolios	**50 size/BTM portfolios**	16.0	16.8	5.6	14.0	13.6	9.6	9.2	7.6	8.4
	10 size portfolios	14.8	17.6	5.2	15.2	17.6	6.8	8.4	9.2	5.2
	10 BTM portfolios	13.6	17.6	6.8	15.2	13.6	9.2	9.6	5.6	10.4
Control firm approach	**by size and BTM**	16.4	15.6	11.2	19.2	14.4	11.6	17.6	13.2	14.4
	by size	19.2	18.4	6.8	20.4	22.4	9.6	13.6	13.2	5.2
	by BTM	18.8	16.0	11.2	22.0	13.6	16.0	20.0	10.0	19.6

Table 4.11 continued

BHARs	5%	One year			Three years			Five years		
		α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
Conventional t test	**50 size/BTM portfolios**	19.6	20.0	5.6	20.0	21.6	9.6	21.2	12.4	16.0
	10 size portfolios	18.4	23.6	5.2	21.2	24.0	7.2	20.4	17.6	10.4
	10 BTM portfolios	20.0	20.0	4.8	21.6	22.4	6.4	17.6	14.4	11.2
	Equally-weighted market return	20.0	23.6	3.6	25.6	28.8	4.8	17.6	14.8	11.2
Skewness-adjusted test	**50 size/BTM portfolios**	20.8	21.2	9.2	24.4	20.8	11.6	23.6	12.8	21.6
	10 size portfolios	19.2	22.0	7.6	22.0	22.0	8.8	22.0	17.2	14.4
	10 BTM portfolios	20.0	20.0	9.6	22.0	20.0	9.6	20.8	13.2	14.0
	Equally-weighted market return	19.2	22.0	6.4	24.4	26.8	6.8	20.8	13.2	14.4
Bootstrapped skewness-adjusted test	**50 size/BTM portfolios**	19.2	18.0	9.2	23.6	18.0	12.4	26.0	12.0	20.4
	10 size portfolios	19.6	15.6	8.0	20.4	18.0	10.0	24.8	15.6	15.6
	10 BTM portfolios	17.6	16.4	10.4	18.4	16.4	10.8	20.8	11.6	15.2
	Equally-weighted market return	17.2	18.0	8.4	21.2	21.2	8.0	21.2	13.6	14.8
P value from Pseudoportfolios	**50 size/BTM portfolios**	13.2	13.2	6.4	5.6	3.6	6.4	10.0	3.6	11.2
	10 size portfolios	9.6	12.0	7.6	4.0	5.2	4.4	5.6	4.0	6.8
	10 BTM portfolios	13.6	14.0	6.0	14.4	0.4	14.0	5.6	2.4	6.8
Control firm approach	**by size and BTM**	20.0	20.4	5.6	20.0	21.6	10.0	21.2	12.0	16.8
	by size	12.8	12.0	9.6	13.2	12.8	8.4	10.8	8.0	12.8
	by BTM	14.0	14.4	12.0	16.8	9.6	16.8	21.6	8.0	26.0

Table 4.12Rejection frequency using cumulative abnormal returns and buy-and-hold abnormal returns in firms with calendar clustering

This table shows the rejection rates (as a percentage) of 250 random samples with 200 firms which have the same event month at a significance level of 5% in one-tailed and two-tailed tests over a period of October 1982 to September 2008, under the null hypothesis of zero one-, three-year, five-year cumulative abnormal returns/buy-and-hold abnormal returns. The test is specified if the overall rejection rate is the same as the significance level together with equal rejection rate on both sides. The abnormal return is defined as the difference between the actual return of a stock and benchmark return. The benchmark is categorized into five types: 50 portfolios by size and book-to-market ratio (BTM) which decile firms by market value in September each year and then quintile firms based on BTM; 10 portfolios ranked by market value; 10 portfolios ranked by BTM, equally weighted market portfolio and control firm which is matched by size, or BTM or size/BTM with sample firm. Apart from conventional student test statistics, skewness-adjusted test statistics, bootstrapped skewness-adjusted test statistics and p-value (Barber, Lyon and Tsai 1999) derived from pseudoportfolios are examined. The event window is one, three and five years. The descriptive statistics of abnormal returns are summarized as follows.

Abnormal return is defined as the difference of actual individual i stock return and benchmark return as follows: $AR_{it} = R_{it} - BR_{it}$.

The cumulative abnormal returns (CARs) calculate abnormal returns of individual stocks in month t by taking the difference between the stock return and equally weighted benchmark return. $CAR_{i,T} = \sum_{t=1}^{T}(R_{i,t} - \frac{1}{n_t}\sum_{j=1}^{n_t} R_{jt})$. The conventional buy-and-hold return assumes investors tend to hold the same portfolio over the investment horizon: $PR_{bnh} = \sum_{i=1}^{n} \frac{[\prod_{t=0}^{t=T}(1+R_{it})]-1}{n}$. Stocks in the portfolios are supposed have return data over the study period. So I replace the missing returns with average portfolio return.. The conventional student test statistics is formulated as: $t = \overline{AR}/(\sigma(AR)/\sqrt{n})$ whereas n is the number of firms in one sample which is 200 in this study. The skewness-adjusted test is calculated as: $skewness - adj.t = \sqrt{n}(S + (1/3) * S^2 * \gamma + (1/6\,n) * \gamma$ whereas γ is skewness of cross-sectional abnormal returns and $\sqrt{n}S$ is student t test. The bootstrapped skewness-adjusted t test bootstraps abnormal returns in each sample 1000 times with size of 50. Then skewness-adjusted test for each subsample of 50 firms is computed in order to construct a distribution which is used to determine the critical value for rejection of the null hypothesis for each sample. Note that skewness-adjusted test calculated in the full sample is used to compare with the critical value. The Pseudo portfolio approach firstly randomly selects a benchmark firm from the same group (by size, BTM or size/BTM) as the sample firm. Then one mean buy-and-hold return for each sample is achieved. With all 250 mean returns, a distribution is constructed for the hypothesis test. Note that the null hypothesis is the equality of sample mean return with critical value.

Table 4.12 continued

CARs	5%	One year			Three years			Five years		
		α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
Conventional t test	**50 size/BTM portfolios**	6.0	5.2	2.8	1.2	1.6	2.4	1.6	2.8	1.6
	10 size portfolios	5.6	6.4	3.6	4.0	4.0	4.0	2.8	2.8	2.0
	10 BTM portfolios	4.8	6.4	3.2	3.2	3.6	4.0	2.8	3.6	3.2
	Equally-weighted market return	8.4	5.6	8.0	12.8	4.4	16.0	17.6	5.6	18.4
Skewness-adjusted test	**50 size/BTM portfolios**	6.0	5.6	2.8	2.0	2.0	4.4	2.0	2.8	2.4
	10 size portfolios	5.6	6.4	4.0	5.2	4.8	5.6	2.8	2.8	4.0
	10 BTM portfolios	5.6	6.8	4.4	4.4	3.6	5.2	3.6	3.6	4.0
	Equally-weighted market return	9.6	5.6	8.8	14.8	4.4	17.6	19.6	5.2	20.0
Bootstrapped skewness-adjusted test	**50 size/BTM portfolios**	3.6	5.2	3.6	2.0	1.2	4.4	2.4	2.0	2.0
	10 size portfolios	4.4	6.0	3.6	2.8	3.6	5.6	2.8	3.2	4.4
	10 BTM portfolios	5.2	5.6	4.0	5.6	3.2	5.2	3.2	2.4	3.6
	Equally-weighted market return	8.4	5.2	8.8	15.2	2.8	18.4	20.4	4.4	22.0
P value from Pseudoportfolios	**50 size/BTM portfolios**	4.8	6.0	4.8	6.0	4.8	8.0	3.6	4.4	5.2
	10 size portfolios	4.4	6.0	4.0	5.2	3.2	6.4	2.0	2.8	5.2
	10 BTM portfolios	4.4	5.6	6.0	4.0	3.6	5.2	2.8	2.4	4.8
Control firm approach	**by size and BTM**	3.2	3.6	3.2	2.4	1.2	4.4	1.6	3.6	2.0
	by size	4.4	2.4	5.2	3.6	3.6	5.6	2.0	4.4	4.0
	by BTM	6.4	6.4	6.0	2.4	4.0	4.4	3.2	2.8	5.2

Table 4.12 continued

BHARs	5%	One year			Three years			Five years		
		α=0	α<0	α>0	α=0	α<0	α>0	α=0	α<0	α>0
Conventional t test	**50 size/BTM portfolios**	6.0	6.8	2.8	2.8	6.4	2.8	3.6	4.4	2.0
	10 size portfolios	5.2	7.2	2.8	5.2	7.2	2.4	3.6	4.4	2.0
	10 BTM portfolios	6.0	7.6	2.8	4.8	6.8	1.2	2.4	4.4	1.6
	Equally-weighted market return	14.2	11.6	6.4	4.8	6.8	0.8	4.4	6.4	3.2
Skewness-adjusted test	**50 size/BTM portfolios**	6.4	6.4	5.6	4.8	5.2	5.2	5.2	4.4	4.8
	10 size portfolios	6.0	6.8	5.6	6.4	5.2	5.6	5.2	4.0	4.0
	10 BTM portfolios	5.2	6.4	6.4	5.2	4.8	4.0	4.0	2.4	3.6
	Equally-weighted market return	13.6	10.8	9.2	4.4	5.2	4.4	5.6	4.8	6.4
Bootstrapped skewness-adjusted test	**50 size/BTM portfolios**	5.2	4.8	5.2	5.2	3.2	8.0	4.4	4.0	5.6
	10 size portfolios	4.8	3.6	6.4	6.4	4.8	8.0	5.2	2.8	6.4
	10 BTM portfolios	4.8	4.4	6.8	4.8	4.0	7.2	3.6	2.8	5.2
	Equally-weighted market return	10.8	8.4	10.0	5.2	2.8	6.0	5.6	4.0	6.4
P value from Pseudoportfolios	**50 size/BTM portfolios**	6.4	4.0	6.4	4.8	4.0	6.4	5.6	3.6	6.4
	10 size portfolios	4.8	5.2	6.8	4.4	4.0	6.8	3.6	3.2	5.6
	10 BTM portfolios	4.0	5.2	5.6	14.4	0.4	14.0	2.4	1.6	4.8
Control firm approach	**by size and BTM**	6.0	6.8	2.8	3.2	6.4	2.8	3.6	4.4	2.0
	by size	6.8	2.4	7.6	1.2	2.4	4.0	4.8	4.0	4.0
	by BTM	5.2	5.2	6.8	2.4	2.0	5.6	3.2	2.4	4.8

Chapter 5: Calendar-time approach

5.1 Research questions and hypotheses

Null hypothesis: the intercept implying the abnormal return of the portfolio is zero.

To resolve the issue of cross-sectional dependence of returns when applying the event-time approach based on reference portfolios and control firm approach, the calendar-time approach proposed by Jaffe (1974) and Mandelker (1974) is employed to detect the long-term abnormal stock performance based on the UK data. The Fama-French three-factor model and the Carhart-four-factor model, the two most popular asset-pricing models with factors of market risk premium, size, book-to-market ratio and momentum, are examined. Both value-weighted and equally-weighted portfolios are expected to achieve zero portfolio abnormal returns. The weighted least square technique advocated by Fama (1998), the Huber/sandwich variance estimators (1982) and the generalized least squares proposed by Gregory, Guermat and Al-Shawawreh (2010) are compared discover the most effective way to deal with heteroskedasticity caused by varying numbers of firms in each calendar month. Therefore, with 250 random samples of 200 firms, a rejection rate of 5% at 5% significance level suggests the test is well-specified. When misspecification exists, this may indicate that the model applied is not an equilibrium model, since the factors cannot fully capture the characteristics of the event firms. Therefore, the mean monthly calendar-time abnormal returns, a variant of the calendar-time portfolio approach, recommended by Lyon, Barber and Tsai (1999), are investigated. The underlying concept is to take abnormal returns of a stock as the difference of observed returns and reference portfolio returns in calendar months. This is similar to the methodology applied in chapter 4 in terms of using reference portfolio returns as benchmark returns. A difference exists in that this chapter uses the calendar month, whereas the previous chapter uses/focuses on the event month. The null hypothesis of this method is zero grand mean monthly abnormal return over the calendar months of T. The percentage of rejection rates determines whether the test is well-specified or not.

5.2 Data

I employ the same data as the studies of the event-time approach based on reference portfolios in the previous chapter. Therefore, stock returns with available market values and book values are applied. Market values of stocks are mainly used to allocate weights to individual stocks so as to measure portfolio returns. Both an equally-weighted scheme which allocate weights evenly to individual stocks in the portfolio and a value-weighted scheme which assign weights based on market values of stocks are applied. However, the missing returns and returns of delisted firms are treated differently in the calendar-time approach when compared with event-time approach based on reference portfolios. Since the number of firms is allowed to differ when applying the Fama-French three-factor model and Carhart four-factor model, missing returns and returns of delisted firms over the test period are left as they are. Therefore, stock returns of a firm are tracked based on its availability. This treatment is the same as the event-time approach based on models. I thus apply average abnormal returns (AAR) to achieve monthly portfolio returns in this approach.

5.3 Research methodology

5.3.1 Models and statistical inferences

5.3.1.1 Fama-French three-factor model

Fama and French (1993) introduce an asset pricing model with three factors of market risk premium, size, and BTM. Although this model is questionable regarding its accuracy to capture the characteristics of an asset, it is still advocated as an alternative to measure and test long-term stock performance due to its elimination of cross-sectional returns. (Fama, 1998). The model is transformed to be utilized in event studies by Jaffe (1974) and Mandelker (1974) as follows:

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t \qquad (5.1)$$

$R_{p,t}$ is the monthly equally-weighted or value-weighted portfolio return. $R_{f,t}$ is the monthly risk-free rate defined as the 90-day treasury bill rate. RPM_t is the difference between value-weighted market return and risk-free rate. HML_t is the difference between returns of value-weighted portfolios of stocks with high BTM and low BTM. SMB_t is the difference between returns of value-weighted portfolios of stocks with small market value and large market value. These factors for the UK stock market are downloaded from the website of Xfi centre of Exeter University. The regression estimates intercept α and other coefficients β. ε stands for the error term which reflects returns which are not captured by the three factors. The intercept α is a proxy to test if the null hypothesis of zero abnormal return is rejected or not at a given significance level. The test statistics for α is assumed to be normally distributed with zero mean and constant variance as follows:

$$t = \frac{\alpha}{\text{Standard Error}(\alpha)} \tag{5.2}$$

Since the error term ε_t could be heteroskedastic due to the change in the number of firms each month, I apply weighted least squares (WLS) in addition to ordinary least squares to correct this. The weights applied in WLS are the number of event firms in the portfolio.

5.3.1.2 Carhart four-factor model

Ang and Zhang (2004) initially conduct simulation on the four-factor model in event studies and conclude that higher rejection rates exist compared with the three-factor model. The four-factor model is developed by Carhart (1997) with incorporation of momentum factor which compares winners and losers. He suggests the persistence in fund performance can be explained by one-year momentum based on studies by Jegadeesh and Titman (1993). The transformed formula for event studies is illustrated as follows:

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp} RPM_t + \beta_{\_hml} HML_t + \beta_{\_smb} SMB_t + \beta_{\_umd} UMD_t + \varepsilon_t \tag{5.3}$$

UMD_t is the zero-investment portfolio which takes the difference between the previous 12-month winners in long position and the previous 12-month losers in short position. Other variables are the same as the three-factor model. Similar test statistics are computed as well. Both OLS and WLS techniques are applied in the model.

5.3.1.3 Mean Monthly Calendar-Time Abnormal Returns

Fama (1998) and Lyon, Barber and Tsai (1999) strongly recommend the application of mean monthly calendar-time abnormal returns which is a variant of the calendar-time portfolio approach aiming at eliminating the issue of cross-sectional returns. The underlying concept is to take abnormal return of a stock as the difference of observed return and reference portfolio returns in a calendar month. The reference portfolios are the same as in the previous chapter.

They include reference portfolios formed by size, BTM, size/BTM; and prevent-event returns. In each calendar month, the benchmark return for an individual stock is the reference portfolio return which the event firm belongs to. The abnormal return is identified as the difference of individual stock return and benchmark return in calendar month t. It is shown as:

$$AR_{i,t} = R_{i,t} - R_{b,t} \tag{5.4}$$

The portfolio return with sample firms is then computed as:

$$CTAR_t = \sum_{i=1}^{n} w_{it} * AR_{it} \tag{5.5}$$

w_{it} is the weight of the sample firm in a calendar month. It can be either equally weighted by simply dividing the total sum of abnormal returns by the number of sample firms in the portfolio or value weighted by taking the percentage of market value of the sample firm relative to the total market value of the portfolio. A grand mean monthly abnormal return over the calendar months of T is computed as:

$$MCTAR = \sum_{t=1}^{T} CTAR_t / T \tag{5.6}$$

The rejection of the null hypothesis of zero abnormal return over a period of time T is determined by the following statistics with time-series of stock returns:

$$t = \frac{MCTAR}{\sigma(MCTAR_t)/\sqrt{T}} \tag{5.7}$$

5.3.1.4 Techniques of regression

The most commonly applied statistical technique for regression is ordinary least squares (OLS). This bears the assumption that stock returns are normally distributed with zero mean and constant variance. It attempts to minimize the sum of the squared error term. Under the Gaus Markov assumptions, the OLS estimators are the best linear unbiased estimators. One key condition under the Gaus Markov assumption is homoskedasticity which assumes constant variance. However, in the calendar-time approach, heteroskedasticity exists because of different numbers of firms in each calendar month (Mitchell and Stafford, 2000). When there is an issue of heteroskedasticity, although the estimators are still unbiased and consistent, they are no longer efficient. This will result in incorrect statistical inferences. Three approaches are suggested to overcome this issue. One is weighted least squares technique (WLS) advocated by Fama (1998) and Ang and Zhang (2004). Instead of distributing equal weight in the OLS, WLS takes the number of firms in the portfolio in each calendar month as weights. Another alternative is to use sandwich variance estimators based on OLS/WLS introduced by Huber (1967) and White (1982) which deals with inconsistent covariance matrix caused by heteroskedasticity of disturbances. The new covariance matrix estimator incorporating the presence of heteroskedasticity is not affected by the model. Therefore, the statistical inferences are correct even when heteroskedasticity cannot be eliminated. The last technique to resolve heteroskedasticity is generalized least squares (GLS) proposed by Gregory, Guermat and Al-Shawawreh (2010). With an assumption of a linear function of the number of firms in portfolios in each calendar month, the authors firstly run basic regression by taking an example of the market model:

$$R_{\tau,t} = \alpha + \beta R_{bt} + u_t \tag{5.8}$$

With the disturbances from the regression, another regression is undertaken as:

$$\log\left(\hat{u}_t^2\right) = \delta_0 + \delta_1 \log(n_t) + \text{error}_t \tag{5.9}$$

The last is to equalize $\widehat{Var}(u_t)$ to $\exp\left(\hat{\delta}_0 + \hat{\delta}_1 \log(n_t)\right)$. The authors claim the GLS provides similar standard errors as in the sandwich variance estimators with OLS but has better adjusted-R square. In this study, I apply all four techniques to find out if there is any difference exists when either ignoring or incorporating heteroskedasticity.

5.3.2 Simulation process

I firstly randomly select 250 samples of 200 event months with replacement from a period of September, 1982 to September 2007 for a one-year investment horizon, to September 2005 for three-year investment horizon and to September 2003 for a five-year investment horizon as Ang and Zhang (2004). Then 200 firms are randomly selected to match with the event months. Event months possibly turn out to be the same with different firms. A firm is not allowed to have multiple events in the same month. But Firms could have more than one different event month. Returns of event firm are tracked following the event month according to the event window, suggesting exclusion of return in the event month. Suppose the event window is one year. For a calendar month, a portfolio consisting of firms which experience events in the previous 12 months ending in the calendar month is constructed. The monthly portfolio return is computed based on mean return of individuals stocks comprising the portfolio in the calendar month. The mean return is measured on the basis of equally-weighted and value-weighted scheme. This mean portfolio monthly return is applied on the left-hand side of the regression formula, but not the individual stock returns as Kothari and Warner (1997). The number of firms in a portfolio varies across calendar months due to different event months of firms. If there is no company in a calendar month, this month is excluded in the regression. Since the calendar-time approach screens out months without event firms, the missing returns are not replaced. If an event firm has a missing return, it is considered as a drop-out from the portfolio in the calendar month. This further indicates that if an event firm is delisted in the test period, the same treatment, which tracks returns of the event firm until it is delisted, is taken as the event-time approach based on models. The regression based on ordinary least squares is conducted according to the Fama-French three-factor model and Carhart four-factor model. To control the bias of heteroskedasticity due to the inconstant number of firms, I implement the weighted least squares technique. The intercept from regression is the ultimate estimate to determine whether absence of long-run abnormal returns exists. Therefore, the conventional test statistics is computed for the intercept. Each sample yields one student-t test statistic. If there is no abnormal return, the intercept is expected to be zero. The rejection frequency among 250 samples is reported at a significance level of 1%, 5% and 10% on both one-tail and two-tailed tests. If the test is well-specified, the rejection frequency is expected to be 5% at a significance level of 5%.

5.3.3 Power of test

A well-specified test needs to control for Type I error which is the probability of rejecting the null hypothesis when it is true and Type II error which is the probability of failing to reject the null when it is false. The power of test is one minus the probability of a Type II error. In other words, it is the probability of rejecting the null when it is false. To examine the power of test, I introduce abnormal returns in a range of -20% to 20% at an interval of 5% to event firms. If this is a 'good model' and a test is well specified, 100% rejection rates are expected with levels of shocks. The shock needs to be spread evenly across the event window for an event firm. For instance, suppose there is a 15% shock. For each event firm in each month over the event period of 12 months, 1.25% of extra return is added to original returns.

5.4 Conventional calendar-time approach

5.4.1 Simulation on random samples

5.4.1.1 Equally-weighted portfolios

Equally-weighted portfolios

Table 5.1 reports rejection frequency of the null hypothesis of zero abnormal returns for both three-factor and four-factor models based on an equally-weighted scheme. The high rejection rates show severe misspecification over three investment horizons at all significance levels. The two factor models exhibit nonzero abnormal returns in 9.2% to 14.4% of all samples over one year, 29.2% to 37.2% of all samples over three years and 36% to 54.8% of all samples over five years at a significance level of 5%. Moreover, the rejection frequency increases with the length of investment horizons. At 5% significance level, the rejection rates are 10%, 29.2% and 36% when applying the three-factor mode under OLS over one-, three- and five years, respectively. Compared with three techniques in the regressions based on two models, OLS shows less misspecified results with lower rejection rates than WLS and GLS. For instance, when applying the three-factor model over an investment horizon of five years, the rejection rates are 48.4%, 55.6% and 58% under OLS, WLS, and GLS, respectively. The one-sided tests show an asymmetric pattern of test statistics with greater rejection rates on the right tail and close to 0% of rejection rate on the left tail. This positively biased test statistics indicates positive abnormal returns are found more often than negative abnormal returns. Compared with the three-factor model, the four-factor model does not improve misspecification since it has higher rejection rates in most cases. For example, when the investment horizon is five years, the rejection rate is 55.2% with the application of the four-factor model but 48.4% when applying the three-factor model under OLS. **Tables 5.2 and 5.3** display the coefficients and R square from regressions based on the three-factor and four-factor models under equal weights. Regarding the three-factor model with all techniques, the mean abnormal return, which is measured by the intercept from the regression, ranges from 0.33% to 0.4% with smaller median over all investment horizons. Among those mean alphas, OLS demonstrates results of 0.33%, 0.4% and 0.38% over one-, three- and five years. The mean test statistics for alpha under OLS increase from 0.7518 to 1.5263 in five years. The same pattern is found in other two techniques. Both coefficients of the market risk premium and size factor with the application of all techniques have a mean close to 0.8 and mean test statistics larger than 5. This indicates a significant correlation between market risk premium or size factor with market portfolio return. However, the coefficient of the book-to-market ratio factor is relatively small with values ranging from -0.0023 to 0.0221. The adjusted R square, that provides the explanatory power of the model, demonstrates higher power when the investment horizon lengthens. Furthermore, the highest value of the adjusted R square when applying WLS, indicates the three-factor model under WLS is the most appropriate model to be applied in term of explanatory power of the model, since the portfolio returns can be better explained by the three factors. 64.4% of independent variables are able to explain dependent variable with the application of WLS in five years. With respect to the four-factor model, the mean alpha increases slightly in all cases ranging from 0.38% to 0.45%. Compared with the three-factor model, the four-factor model shows higher explanatory power with higher adjusted R-squared. For instance, the highest explanatory power is presented by WLS in five year years with an adjusted R square of 64.63% whereas the corresponding figure for the four-factor model is 64.4%. It is interesting to note that the coefficients for the BTM ratio factor are all negative in the four-factor model. The estimates of the momentum factor reveal small correlation with portfolio returns. Additionally, the means of coefficients of the momentum factor are all negative. The four-factor model exhibits similar characteristics as the three-factor model. In terms of the fitness of models, the four-factor model outperforms the three-factor model. However, it generates severer misspecification with higher rejection rates of the null hypothesis of zero abnormal returns. It is important to note that compared with the empirical results in the US data documented by Lyon, Barber and Tsai (1999) and Ang and Zhang (2004), the misspecification is presented with the application of the three-factor model in the studies based on the UK data. I find the rejection rate is 29.2% when applying the three-factor model under OLS over an investment horizon of three years whereas Ang and Zhang (2004) show the corresponding figure of 4%. Moreover, the misspecification is severer when applying the four-factor model under both OLS and WLS in the UK data. The rejection rates are 32.8% and 54% under WLS in the US data and the UK data, respectively.

Winsorized equally-weighted portfolios

With the results of severe misspecification, I cast doubts on one potential cause which is the outliers of raw returns. Therefore, I winsorize 1% of stock raw returns in each tail within observations of 200 firms in each sample. **Tables 5.4-5.6** display the winsorized results based on equal weights. Although the magnitude of rejection rates is smaller compared with the original sample, there is still overrejection. For instance, the rejection rate is 25.6% in five years at a significance level of 10% with the application of the four-factor model under OLS whereas the corresponding figure for the original sample is 55.2%.The positively biased test statistics with overrejection on the right tail are documented as the original sample but with lower rejection rates. WLS only shows little evidence of outperformance. For example, the three-factor model over a five-year investment horizon generates lower rejection rate than OLS. However, GLS show improved performance with lower rejection rates compared to WLS but not as good as OLS. The winsorized mean intercepts ranging from 0.09% to 0.15% are much smaller compared with the original sample. However, in the four-factor model, the coefficients for all variables are positive. The results indicate winsorization does significantly improve the misspecification. In addition, the explanatory power of independent variables to dependent variables is higher when returns are winsorized. The mean adjusted R square is 52.58% with the application of the three-factor model under OLS over five years when the stock returns are unwinsorized whereas the corresponding figure is 58.81% with the winsorized returns. Therefore, the results conclude that the winsorization on returns is applicable to reduce misspecification with lower rejection rates. Moreover, the explanatory power of models is enhanced after the winsorization process. However, it cannot eliminate the misspecification which is caused by the equally-weighted scheme that allocates more weights to small firms with higher returns.

Descriptive statistics of abnormal returns

To further examine the distributional properties of alphas derived from the two-factor models, I compare the descriptive statistics based on equal weights in **Table 5.7-5.8**. Regarding the three-factor model based on the original data, the mean intercepts are in the range of 0.00332 to 0.00403 over three investment horizons. Moreover, the mean intercept shows a pattern of an increase from one year to three years but a decline from three years to five years. The medians are slightly lower than means. This is further proved with positive skewness ranging from 0.19 to 1.57. The kurtosis is above 3 except the case when applying the three-factor model under OLS over three years. The leptokurtic distribution of alphas is more evident when WLS and GLS are applied over three- and five years. For instance, kurtosis is 12.43 with the application of WLS over one year. The winsorized results present smaller magnitude of mean alphas ranging from 0.00088 to 0.00151. However, the medians show mixed results compared with the mean. This is reflected by the distribution of alphas with positive and negative skewness. Negative skewness is exhibited when applying the three-factor model under all techniques in one year and under OLS in five years. Furthermore, skewness close to zero together with kurtosis close to 3 indicates an approximate normal distribution. The winsorization process, therefore, adjusts the asymmetric distribution of abnormal returns to a normal distribution. This results in improvement on misspecification based on the original data. The four-factor model shows similar results as the three-factor model. Compared with the three-factor model, the four-factor model generates higher values of means and medians. Moreover, skewness is mostly higher whereas kurtosis is mostly smaller.

All techniques based on equally-weighted portfolios

Heteroskedasticity, which suggests the changing variance of disturbance, is one concerning issue when conducting the long-run event studies. Apart from the White correction test (1980), WLS taking the number of firms in calendar months as weights (Ang and Zhang, 2004) and GLS with an assumption of a linear function of the number of firms in portfolios in each calendar month (Gregory, Guermat and Al-Shawawreh, 2010) are proposed to handle the issue. **Table 5.9** summarizes all techniques with the application of both the three-factor and four-factor model under an equal-weight scheme. sandwich variance estimators imposed on OLS and WLS based on the two factor models improves misspecification with slightly lower rejection rates in two-tailed test. For instance, the rejection rate is 29.2% when applying the three-factor model under WLS whereas the sandwich variance estimators generates a rejection rate of 26.8%. Since OLS outperforms WLS with lower rejection rates and sandwich variance estimators further reduces misspecification, WLS loses its attractiveness to control heteroskedasticity. The rejection rate is 30.4% when applying the three-factor model under OLS together with sandwich variance estimators whereas WLS yields a rejection rate of 41.2% over five years. Moreover, GLS yields higher rejection rates than OLS in conjunction with sandwich variance estimators. For example, the four-factor model under OLS with sandwich variance estimators generates a rejection rate of 35.2% whereas GLS yields a rejection rate of 54.8% over an investment horizon of five years. Additionally, it is difficult to identify whether GLS outperforms WLS since the results are mixed. Over an investment of three years, the three-factor model shows rejection rates of 34.4% and 31.6% under WLS and GLS whereas the corresponding figures for the four-factor model are 36.4% and 37.2%. The same pattern is found in one-tailed tests. Moreover, test statistics are positively biased. This can be attributed to the asymmetric distribution of abnormal returns. Consequently, OLS together with sandwich variance estimators is the most appropriate technique to be applied in the long run event studies based on the UK data.

Power of test based on equally-weighted portfolios

When conducting null hypothesis tests, type II error occurs when the null hypothesis is false but is not rejected. The statistical power is worth investigating in event studies because it can measure the ability of a test to detect the presence of abnormal returns following an event. In the studies, I introduce shocks ranging from -20% to 20% at an interval of 5% to individual stocks. As explained in the methodology, the stock returns are added in each month with average monthly returns achieved by dividing the shock by the investment months. **Table 5.10** displays the power of test based on equally-weighted portfolios. Regarding negative shocks, in one year, rejection rates decline when abnormal return is -5% but start increasing with higher shocks. The three-factor model generates rejection rates of 10% when there is no abnormal returns introduced and 5.6% when abnormal return of -5% is introduced. But the rejection rate increases to 24.4% when introducing a shock of -10%. It is interesting to note that for longer horizons, for example three years and five years, the power of test reduces significantly. The same patterns are documented when positive shocks are induced. However, with the same absolute value of shocks, the rejection rates are higher with positive shocks.

For instance, the four-factor model under WLS in three years yields a rejection rate of 50.8% when the shock is 5% and 14.8% when the shock is -5%. This can be attributed to the properties of intercepts reported in **Table 5.11** when there is no abnormal return introduced. The mean intercepts are positive in all cases without any introduction of abnormal returns. For instance, the mean intercepts are 0.33%, 0.4% and 0.38% when the three-factor model is applied under OLS over one-, three- and five years, respectively. The overrejection on the right tail with an indication of more positive abnormal returns in the original data makes test statistics more powerful when abnormal returns are positive but less powerful when abnormal returns are negative. Because positive shocks enhance positive intercepts whereas negative shocks offset positive intercepts. Rejection rates close to 5% when abnormal returns are -15% and -20% suggest intercepts are reduced close to zero with the introduction of negative shocks, which satisfy with the null hypothesis of zero abnormal returns. For example, the mean intercept is 0.38% when there is no abnormal return but is still positive when negative shock is introduced. Even when the abnormal return is -20%, the mean intercept is 0.05%, which is close to zero. The one-tailed tests shown in **Table 5.12** further confirm the conclusion that the test loses more power when negative shocks are introduced. The four-factor model under OLS shows more positive abnormal returns with a rejection rate of 55.2% on one-tailed test with the alternative hypothesis of positive abnormal returns and no negative abnormal returns with a rejection rate of 0% on one-tailed test with the alternative hypothesis of negative abnormal returns over five years. When abnormal return is -20%, the rejection rates are 2% under one-tailed test and 5.2% under two-tailed test. However, the rejection rates are 51.6% under one-tailed test and 88.4% under two-tailed test with induced abnormal return of 20%. The reason to focus on the one-tailed test is that the overrejection of the null hypothesis of zero intercept suggests there is presence of abnormal returns when there is no shock is introduced. Moreover, the test loses power in longer time horizons. For instance, based on the three-factor model under OLS, the rejection rates are 99.6%, 95.6% and 80% in one-, three-, and five years when introducing abnormal return of 20%. Studies undertaken by Ang and Zhang (2004) document similar patterns but higher rejection rates, especially in the three-factor model, compared with the results. They advocate the application of WLS with an argument of its higher power of test. Although the results indicates higher power of test when applying WLS, when heteroskedasticity is taken into account, OLS in conjunction with sandwich variance estimators is advocated based on the UK data due to it less severe misspecification.

5.4.1.2 Value-weighted portfolios

Value-weighted portfolios

Table 5.13 reports rejection rates in 250 samples based on the three-factor and four-factor models under a value-weighted scheme. The misspecification is less severe in one year compared with the equally-weighted portfolios. Furthermore, both models yield well-specified results in three- and five years. For instance, the two-sided tests at a significance level of 5% with the application of both two models under OLS have the same rejection frequency of 4.8% in three years whereas five years witness 5.6% in three-factor model and 3.6% in four-factor model. However, in one-year investment horizon, the rejection rates are 10.4% and 7.2% when applying the three-factor and four-factor model, respectively. WLS and GLS improve the misspecification occasionally and marginally. One-tail tests show an asymmetric pattern of test statistics with mixed results. Take one year as an example. At 5% significance level, test statistics are negatively biased with higher rejection rates on the left tail when the three-factor model is applied with all techniques, whilst positively biased test statistics are found when employing the four-factor model with all techniques. However, positively biased test statistics with an indication of more positive abnormal returns are documented over three- and five years. The coefficients displayed in **Tables 5.14-5.15** show similar patterns of coefficients as the equally-weighted portfolios. However, there are some differences. Firstly, a smaller magnitude of intercept, which is a proxy of abnormal return, is found when compared with the results based on the equally-weighted scheme. The mean intercepts range from -0.06% to 0.03% whereas the corresponding figures under equal weights are 0.33% to 0.4%. Secondly, the estimates of market risk premium based on value weights are larger. The means of coefficients of market risk premium are larger than one in the three-factor model, regardless of techniques. The figures become less than one but close to one in the four-factor model. This indicates a larger impact from the market portfolio which is measured by value-weighted market returns. Thirdly, the coefficients of size factor are mostly half of those under equal weights, indicating smaller size effect when conducted under value-weighted scheme. For instance, the mean coefficients of the size factor are 0.259 under value weights and 0.709 under equal weights when applying the four-factor model under OLS in five years. This can be explained with the fact that the value-weighted scheme gives smaller weights to small firms and larger weights to large firms according to firms' market value. Fourthly, although the estimates of BTM factor are larger under value weights, they are still comparably smaller than other factors. The mean coefficients of BTM factor are 0.0012 under equal weights and 0.119 under value weights with the application of the three-factor model in three years. Last but most importantly, smaller adjusted R-squared in three-factor and four-factor model compared with equal weights suggest a lower explanatory power of factors on portfolio returns under value weighting although test statistics are well-specified, especially in three and five years. Interestingly, the means of intercepts are 0% in five years under all techniques when the three-factor model is applied. Similar as the portfolios returns based on an equal weighting scheme, the four-factor model is proved to be a better fit than the three-factor model under the value-weighted schemes.

Winsorized value-weighted portfolios

I also conduct the winsorization process to test if the impact of outliers can be reduced under value weights. **Tables 5.16-5.18** show the empirical results with well-specified test statistics in particularly three- and five years, which are similar as the original data based on value weights. However, the magnitudes of coefficients are mostly smaller compared with the results based on the original data. Moreover, the explanatory power of two factor models is higher when returns are winsorized. Compared with results based on the original data, there is not much difference in the rejection frequency after winsorizing the data. Although there is an improvement on the explanatory power after winsorization, it is better to apply the factor models based on the original data since the value-weighted scheme resolves the issue of misspecification brought by small firms.

Descriptive statistics of abnormal returns

The descriptive statistics of abnormal returns in **Table 5.19-5.20** based on the value-weighted scheme show some similarities as the equally-weighted scheme. The mean intercept increases from one year to three years and declines from three years to five years. Moreover, positive mean intercepts are documented when applying the four-factor models under all techniques. Additionally, the mean intercepts after winsorization are smaller than that based on the original data. However, there are still some differences. First of all, positive and negative mean intercepts are found with the application of the three-factor model. For instance, over an investment horizon of one year, negative mean intercepts are documented based on the original data and winsorized data. Second, the magnitude of the mean intercepts is much smaller based on value weights than that based on equal weights. For instance, the mean intercept is 0.001% under value weights and 0.382% under equal weights when applying the three-factor model under OLS over five years. Last but most important, the distribution of abnormal returns is approximately normal with skewness close to zero and kurtosis close to 3. More specifically, skewness is mostly negative in both the three-factor and four-factor models.

All techniques based on value-weighted portfolios

Table 5.21 lists the specification of tests including the sandwich variance estimators under all techniques. The mean alphas are negative when applying the three-factor model in one year. Moreover, the magnitude is much smaller compared with the results based on equal weights. Similar as the equally-weighted portfolios, sandwich variance estimators controls heteroskedasticity more efficiently with lower rejection rates compared with the original tests. For instance, the four-factor model under OLS yields a rejection rate of 3.6% with the application of the original test and 3.2% when applying sandwich variance estimators. WLS does not perform better than sandwich variance estimators except the case when the three-factor model is applied over five years. The rejection rate under OLS with sandwich variance estimators is 5.6% whereas WLS yields a rejection rate of 5.2%. Different findings are documented when applying two factor models when compared sandwich variance estimators under OL with GLS. Regarding the three-factor model, rejection rates are the same for these two techniques over one- (10.4%) and three year (4.8%) whereas sandwich variance estimators under OLS yields a rejection rate of 5.6% and GLS generates a rejection rate of 4%

over five years. With respect to the four-factor model, GLS yield higher rejection rates than sandwich variance estimators under OLS over all investment horizons. Moreover, GLS outperforms WLS with lower rejection rates except the case when the three-factor model is applied over an investment of one year. The same conclusion of the application of sandwich variance estimators under OLS based on a value-weighted scheme in the long run event studies is drawn as the equally-weighted portfolios.

Power of test-value-weighted portfolios

Similar findings regarding the power of test based on the value-weighted scheme .as equally-weighted portfolios are presented in **Table 5.22** First of all, the rejection rates decrease when investment horizon lengthens under each technique. For instance, the rejection rates are 94.8%, 36.8% and 20% over investment horizons of one-, three- and five years when applying the three-factor model under OLS. Moreover, the power of the approach based on the four-factor model decrease more significantly than the three-factor model. Second, the rejection rates are smaller when negative shocks are introduced compared with the case when positive shocks are introduced when applying the four-factor model over one year and both factor models over the three- and five years. For example, the rejection rate is 36.8% when abnormal return of 20% is introduced and 8.4% when abnormal return of -20% is introduced. This asymmetric power can be explained by positive intercepts derived from the factor models. **Table 5.23** shows that the mean intercepts are mostly positive with some exceptions when there is no shock induced. Positive intercepts are enhanced when positive abnormal returns are introduced whereas intercepts are reduced close to 0 firstly and then reach negative level when negative shock are introduced. For instance, the three-factor model under OLS over five years shows a mean intercept of 0.001% when there is no abnormal return introduced. The corresponding figures are -0.082% when abnormal return of -5% is introduced and 0.084% with induced abnormal return of 5%. The rejection rates under two-tailed test are 7.2% with negative shock and 7.6% with positive shock. The rejection frequency based on two-tailed test with the same magnitude of positive and negative abnormal returns is similar. This is different from the results based on equally-weighted portfolios which suggest the test loses more power when negative shocks are introduced. This difference can be illustrated with the statement of much smaller magnitude of intercepts based on a value-weighted scheme compared with an equally-weighted scheme. To further address the power of test, one-tailed test is conducted and the results are exhibited **Table 5.24**. Since the factor models based on different techniques yield well-specified test statistics in the original samples when there is no shock introduced, one-tailed test confirmed the conclusion of similar pattern and magnitude of rejection rates as the two-tailed test. The exception of higher power with the introduction of negative shocks when applying the three-factor model over one year is attributed to negative intercepts from regressions. Finally, WLS shows the highest power compared with OLS and GLS. This is consistent as studies undertaken by Ang and Zhang (2004). As the previous discussion, although WLS shows superior power, it generates higher rejection rates than OLS.

Summary: equally-weighted portfolios vs. value-weighted portfolios

In summary, the difference of weighting schemes generates widely divergent results. The equally-weighted scheme allocates equal weights to individual stocks while the value-weighted scheme assigns weights according to market value of individual stocks. The following findings regarding the random samples are consistent with prior research (Lyon, Barber and Tsai, 1999, Ang and Zhang, 2004). First of all, the explanatory power of both the three-factor and four-factor models, which is measured by the adjusted R square, increases with investment horizons. In addition, the four-factor model has higher explanatory power compared with the three-factor model. Secondly, the test loses power in longer holding period, especially when the induced abnormal return is negative. Lastly, the magnitude of intercepts is similar when applying the four-factor model under OLS over three investment horizons. However, interesting findings exist which differentiate the study based on the UK market from other studies based on the US market. To begin with, overrejection rates of the null hypothesis of zero abnormal returns are extremely high, especially in the long run, under equal weights. With the adjustment of the winsorization on stock returns, the misspecification still exists over three investment horizons but with lower rejection rates. Value-weighted portfolio returns applied in regression also improved the misspecification in the equally-weighted portfolio returns. Both factor models produce better specified results in the two-sided test but still show an asymmetric pattern of test statistics. Therefore, it can be concluded that the weighting schemes matter in detecting the long-run abnormal stock returns. This finding is inconsistent with the results documented by Lyon, Barber and Tsai (1999). The US data shows well-specified test statistics regardless of weighting schemes. The anomalies in the findings can be attributed to the specific characteristic of the UK market structure which is filled with mostly small firms. However, the finding of misspecification when applying the four-factor model is in line with findings documented by Ang and Zhang (2004). Second, although WLS incorporates the number of firms in calendar months as weights to control heteroskedasticity, the rejection rates are higher compared with sandwich variance estimators under OLS, regardless of weighting schemes. GLS generates lower rejection rates compared with WLS but similar rejection rates compared with sandwich variance estimators under OLS. Therefore, OLS together with sandwich variance estimators is advocated to be applied in the long-run event studies. Last but not the least, in the case of equally-weighted portfolio returns, the power of test shows considerable small rejection rates when negative shock is introduced, especially in three- and five years. Furthermore, the rejection rates decrease significantly with higher magnitude of negative shocks.

5.4.2 Simulation on non-random samples

5.4.2.1 Large/Small size

All size deciles

Firms are deciled into ten groups based on their market value in September each year. Firms with the smallest market cap are classified as size decile 1 while firms with the largest market cap are classified as size decile 10. Small size firms in decile 1 are randomly selected with replacements to achieve 250 samples of 200 firms. A similar process is conducted for large firms. **Table 5.25** shows the rejection frequency over 250 samples for ten sizes under both equal-weighted and value-weighted schemes at a significance level of 5% in both two-tailed and one-tailed tests when applying the three-factor model. Regardless of weighting schemes, size decile 1 presents the highest rejection rates, most of which are larger than 90%, compared with other deciles. The rejection rates of size decile 1 are 99.6% under equal weights and 94.8% under value weights over five years. The rejection rates decline significantly to at least half from size decile 3 to 10. For instance, the rejection rates under equal weights are 96.4% in size decile 3 and 4.8% in size decile 10 when the investment horizon is five years. Winsorization improves the misspecification with slightly lower rejection rates in size decile 1 but does not work better than the original sample in size decile 10. From size 1 to size 6, test statistics are positively biased with higher rejection in the right tail. However, the findings of bias in test statistics are mixed in the remaining deciles. Under equal weights, the findings of rejection rates of 2.8% on the right tail and 2.4% on the left tail in size decile 7 over an investment horizon of five years indicate a positive bias of test statistics while the rest deciles over all investment horizons show negatively biased test statistics. Under value weights, size decile 7 to 9 present negative biased test statistics over three investment horizons whereas size decile 10 shows positively biased test statistics. Compared with the value-weighted scheme, rejection rates are higher in the first 6 deciles under the equally-weighted scheme. Although test statistics are well-specified in the random samples, small firms show misspecification under both weighting schemes. This suggests although the three-factor model under value-weighted scheme yields well-specified results in random samples, it is strongly affected by small firms which generate the small size effect. The findings in the four-factor model are similar as the three-factor model. The results in **Table 5.26** show other differences compared with Table 5.21. The rejection rates are lower in size decile 1 and higher in size decile 10. For example, the three-factor model yields a rejection rate of 94.8% while the four-factor model generates a rejection rate of 82% in size decile 1 over five years under value weights. However, in size decile 10 under value weights, the rejection rates are 6% with the application of the three-factor model and 14% when applying the four-factor model. Additionally, test statistics are positively biased except the case when the size deciles are 7 and 8 with an investment horizon of one year. The overrejection for size 1-3 of firms with small market values indicates small firm effect have significant impact on specification of tests regardless of weighting schemes.

Small size

Small size effect is well-documented in the literature with a summary of inability of three-factor to capture prices of small firms by Fama and French (1993). Long-run abnormal returns have potential issues which are brought in by small firms.[6] Therefore, I randomly select 250 samples of 200 firms from a population of small firms defined as the size decile 1 with the smallest capitalization. **Tables 5.27-5.32** show the empirical results based on firms in decile 1 with the smallest market value. Similar findings are documented as random samples. For instance, the rejection rates increase from investment horizons of one year to three years but decline in five years. Moreover, the positively biased test statistics are similar as random samples but with much higher rejection rates. However, some differences are noted. To begin with, the rejection rates in 250 samples are extremely high, almost reaching 100%. This indicates more severe misspecification in small firms, regardless of weighting schemes. Secondly, the intercepts are in the range of 1.29% to 2.26% in the three-factor model. This is much higher than random samples (0.33%- 0.4%), based on equal weighting. Similar pattern is documented when the value weighting scheme is applied. Thirdly, the coefficients of size factor are closer to 1 whereas the estimates of the BTM are all negative. This suggests a strong relationship between size factor and portfolio returns. Fourthly, although the four-factor model still shows higher explanatory power than the three-factor model in small firms, the adjusted R-squared are much smaller when compared with random samples under equal weights. For instance, the adjusted R square is 24.8% in five years based on the three-factor model with OLS. However, the same figure is 52.58% for random samples. The fitness of models loses its attractiveness when detecting abnormal returns of small firms. However, the explanatory power of the two factor models under value weights is higher in small firms. The adjusted R-squared are 46.32% and 48.23% in random samples and small firms when applying the three-factor model under value weights. Finally, although there is an improvement on misspecification when applying the models based on the value-weighted scheme compared with the equally-weighted scheme, there is still overrejection.

Large size

With respect to large firms in decile 10, **Tables 5.33-5.36** show the portfolios' performance based on an equally-weighted and value weighted schemes over one-, three- and five years. Compared with random samples, large firms produce well-specified test statistics in one year

[6] Fama (1998), Brav, Geczy and Gompers (2000)

regardless of weighting schemes. However, over an investment horizon of three years, well-specified results are indicated when applying the three-factor model based on both weighting schemes under OLS and WLS. GLS yields misspecified test statistics with higher rejection rates. When applying the four-factor model under OLS and WLS, well specified results are generated under value weights. But the rejection rates are 1.6% under WLS, 16.4% under OLS and 16% under GLS with the application of the four-factor model under equal weights. The asymmetry of test statistics with negatively biased test statistics are found when applying the three-factor model and with positively biased test statistics are shown when applying the four-factor model regardless of weighting schemes. The mean intercepts are all negative and close to 0 based on an equally-weighted scheme. However, when the three-factor model based on a value-weighted scheme, the mean intercepts are mostly 0, especially over longer horizons. The distribution of abnormal returns under a value-weighted scheme is normal over three- and five years. There is a significant decrease in the value of coefficients of size factor. It is worth noting that the explanatory power of models increases to the highest level compared with random samples in both weighting schemes and models. However, for large firms, the value-weighted scheme reduces the explanatory power of models.

Summary: small size vs. large size

To conclude, small size effect dominating in the studies cause the misspecification in both random samples and non-random samples categorizing by size of firms when the equally-weighted scheme is applied. Moreover, large firms show misspecified test statistics with higher rejection rates when the four-factor model under equal weights is applied over three and five years. In addition, GLS yields misspecified results with the application of the three-factor model under both weighting schemes. The asymmetric distributions of abnormal returns vary in different size deciles. Although abnormal returns of firms in decile 10 with the largest market value are less misspecified, the power of test reduces with longer investment horizons. Compared with the previous studies based on the US data, different results are documented based on the UK data. Lyon, Barber and Tsai (1999) conduct the calendar-time portfolio approach based on both weighting schemes and find well-specified test statistics for both small firms and large firms. For instance, the rejection rates are 3.3% over an investment horizon of three years when applying the three-factor model under equal weights in small firms based on the US data whereas the corresponding figure is 92.4% based on the UK data. Regarding large firms, the three-factor model under value weights based on the US data yields a rejection rate of 5.4% in three years whereas the corresponding figure is 4% based on the UK data. The findings in small and large firms based on the UK stock market confirm a strong impact from small firms in non-random samples categorized by size. Furthermore, Ang and Zhang (2004) test the abnormal returns in the long run by using the four-factor model under equal weights based on the US data. The misspecification is demonstrated with high rejection rates over three investment horizons as the results based on the UK data, but with a much smaller magnitude. For instance, the rejection rates are 58% when the four-factor model under WLS is applied based equal weights in the US stock market over an investment horizon of five years whereas the corresponding figure based on the UK stock market is 100%.

5.4.2.2 High/Low book-to-market ratio (BTM)

All BTM deciles

To evaluate the impact of firms with different book-to-market ratios, I decile firms into ten groups based on their book-to-market ratios which are achieved according to market value in September of each year and 6-month-lagged book value. For each BTM decile, I randomly select 250 samples of 200 firms to construct portfolio returns for regressions. **Tables 5.37-5.38** list the percentages of rejection rates of intercept test for the two factor models under equally and value weighting schemes. The lowest and highest rejection rates do not always appear in BTM decile 1 and 10. The three-factor model yields the lowest rejection rate of 4.4% in BTM decile 1 and the highest rejection rate of 91.6% in BTM decile 10 when applying the three-factor model under equal weights in five years. However, the lowest rejection rate is in BTM decile 4 under value weights. When the four-factor model under value weights is applied, the lowest rejection rate is found in BTM decile 3. Winsorization improves the misspecification with smaller rejections for BTM decile 10 but worse performs for BTM decile 1. Similar findings are documented in random samples and non-random samples by size. For instance, the four-factor model shows lower rejection rates compared with the three-factor model. Moreover, the equally-weighted scheme yields higher rejection rates than the value-weighted scheme, except for BTM decile 1 with the application of the three-factor model. It should be noted that with an increase value of BTM, the pattern of test statistics shift from a negative bias to a positive bias, mostly from the BTM decile 4 onwards.

Low BTM

Regarding the BTM decile 1 with the lowest BTM, **Tables 5.39-5.40** detail the rejection frequency with all techniques in three-factor and four-factor models over three investment horizons. The misspecification is more severe when applying the value-weighted scheme, especially when the three-factor model is applied. The rejection rate is 14.8% when the three-factor model under OLS is applied with the equally-weighted scheme in one year whereas the corresponding figure is 24% when the value-weighted scheme is used. It is interesting to note that the value-weighted scheme does not improve misspecification as other non-random samples. The rejection rates are 10.8% based on equal weights and 24.4% based on value weights when the three-factor model is applied under OLS in three years. Negatively biased test statistics are shown in most cases. Moreover, the four-factor model improves the misspecification with lower rejection rates under both weighting schemes.

High BTM

However, the results displayed in **Tables 5.41-5.42** show more severe misspecification with high rejection rates, even reaching 100% for BTM decile 10 with the highest BTM. The high rejection rates can be explained by small firms which fall into the BTM decile 10. Moreover, the repetition of firms in one sample generates overlapping returns which yields biased results. Interestingly, although the value-weighted scheme yields misspecified results, it improves the misspecification with lower rejection rates. For instance, the rejection rate is 73.2% when the four-factor model under OLS is applied based on the equal-weighting scheme whereas the corresponding figure under value weights is 21.6%. This is in contrast with the findings in BTM decile 1, which shows an equal weighting scheme is better with lower rejection rates. The test statistics are positively biased with mostly zero rejection rates on the left tail. The four-factor model still shows greater ability to adjust misspecification to achieve lower rejection rates. However, it cannot alleviate the misspecification.

Summary: low BTM vs. high BTM

In summary, the results are in line with findings shown by Lyon, Barber and Tsai (1999) based on the US data. They only examine the three-factor model under OLS based on two weighting schemes. Regarding an equally-weighted scheme, the misspecification is documented with high rejection rates in both firms with the highest and lowest BTM. The US market presents a rejection rate of 16.8% over an investment horizon of five years in BTM decile 1 whereas the corresponding figure based on the UK market shows a rejection rate of 1.6%. However, BTM decile 10 shows rejection rates of 17.1% based on the US data and 82.8% based on the UK data over an investment horizon of three years. The test statistics are negatively biased in firms with the lowest BTM and positively biased in firms with the highest BTM for data from both the US and UK stock markets. With respect to the value-weighted scheme, the improvement of misspecification is shown when the US data is applied. The rejection rates are 10.3% under value weights and 16.8% under equal weights in BTM decile 1 over five years. Moreover, the US stocks illustrate well-specified results in BTM decile 10 whereas the opposite results are found in the UK market.

5.4.2.3 Industry clustering

In practice, it is common to observe that firms in the same industry have similar events in a certain period. The industry clustering may potentially increase the magnitude of abnormal returns. Therefore, I take the extreme case of firms in one sample clustering in one industry with different event months. I randomly choose one industry and then randomly choose 200 firms with event months in this industry. The same process is repeated 250 times. In order to alleviate the effect from overlapping returns, one firm is allowed to appear more than once but is not allowed to have an event month in the previous or latter T-1 months.

When the portfolio return is based on equal weighting, **Table 5.43** shows overrejection from one year to five years. The misspecification increases with higher rejection rates in longer time horizon. In five years, the rejection rate reaches 80.8% with the application of the four-factor model under WLS. Positive biased test statistics are shown with higher rejection rates on the upper tail. The four-factor model does not show superior ability to reduce the misspecification. Furthermore, OLS performs better than WLS and GLS with lower rejection rates in most cases. However, **Table 5.44** shows misspecification is significantly reduced to a level of well-specification when measuring portfolio returns based on a value-weighted scheme. In one year, the rejection frequency among 250 samples is 4.4% and 6.8% in two-tailed test based on three-factor model and four-factor model under OLS at a significance level of 5%. The techniques of WLS and GLS improve marginally on rejection rates. However, one-tailed tests still show an asymmetry of test statistics. It is interesting to note that the three-factor model under WLS and GLS shows negatively biased test statistics while the four-factor model in all techniques shows positively biased test statistics. In three years, the test is underrejected with 3.2% and 2.4% in two-tailed test based on three-factor model and four-factor model under OLS at a significance level of 5%, respectively. The three-factor model under WLS improves two-tailed test with a rejection rate of 5.2% but has negatively biased test statistics. In five years, similar results are documented for three years. To summarize, factor models based on value-weighted scheme shows well-specified tests in one year but underreject the null hypothesis of zero abnormal returns in longer term.

5.4.3 Cross-sectional dependence of returns

5.4.3.1 Calendar time clustering

One of the cross-sectional dependence of abnormal returns is calendar time clustering. I thus take the extreme case when a sample is consisted with firms having the same event month. One firm is not allowed to appear more than once in one sample. 250 samples of 200 firms are randomly selected. **Table 5.45** shows rejection rates based on an equal weighting scheme. The overrejection is enhanced with the length of investment horizons with positively biased test statistics. The four-factor model generates marginally higher rejection rates compared with the three-factor model. The misspecification is strengthened to 42.8% of rejection rates in two-tailed test in five years at a significance level of 5% with the application of the three-factor model under OLS. However, when converting to value-weighted scheme, the misspecification tends to disappear. The results are displayed in **Table 5.46**. The rejection rate in two-tailed test is 6.8% in one year and increases to 8.8% in three years but declines to 6% in five years at a significance level of 5% when employing the three-factor model under OLS.

It indicates the three-factor model works better in longer investment horizon, particularly in five years. The distribution of abnormal returns is still positively skewed but not as severe as the case in equally-weighted scheme. The four-factor model does improve the specification with lower rejection rates, especially in five years. The most specified test appears as rejection rate of 5.2% in two-tailed test at a significance level of 5% with the application of the four-factor model under OLS and GLS in five years.

Compared with the equally-weighted scheme, the value-weighted scheme generates well-specified results in two-tailed test despite of positively biased test statistics. Value-weighted scheme indicates the four-factor model improves the test with lower rejection frequency while equally-weighted scheme suggests marginally change in rejection rates when applying four-factor model. The severe misspecification based on equal weights gives a hint on the weighting scheme matters as other previous non-random samples when detecting the long-term performance of stocks in a case of calendar time clustering. The WLS technique works better in value-weighted scheme rather than equally-weighted scheme while GLS generates mixed results.

5.4.3.2 Overlapping returns

Apart from the calendar time clustering, another case of cross-sectional dependence is caused by overlapping returns which have firms with multiple events clustering in the same event window. From **Table 5.47-5.48**, different weighting schemes generate different results as previous random samples and non-random samples. The equal-weighted scheme still produces higher rejection compared with the value-weighted scheme. Other differences are highlighted as follows. Firstly, the rejection rate of two-tailed test increases from 7.2% in one year to 20.4% in five years with the application of the three-factor model under OLS at a significance level of 5% based on equal weights. However, the corresponding figure under value weights is 6% in one year and increases to 9.6% in three years and declines to 8.4% in five years. Secondly, positively biased test statistics are documented based on equal weights apart from four-factor model under WLS. The pattern of test statistics is mixed under value weights. Test statistics are negatively biased in one year whereas there are positively biased test statistics in both three and five years. Thirdly, the four-factor model generates lower rejection rates under OLS for all investment horizons based on equal weights. However, when applying four-factor model under OLS over a period of five years, the rejection rate increases to 10.4% (8.4% in three-factor model) based on value weights. Fourthly, the WLS technique generates more misspecified results with higher rejection rates compared OLS under both weighting schemes with two exceptions in three years based on equal weights. The rejection rates in two-tailed test at a significance level of 5% are 12.8% and 10.4% when applying three-factor model and four-factor model under OLS based on equal weights. The figures turn to be 8.8% and 10% after taking into consideration of the number of firms in regression as weights. Fifthly, when compared GLS with OLS, the results are mixed with an uncertain answer if GLS outperforms OLS. However, GLS is still superior to WLS. Last but not least, it is worth mentioning that to compare with calendar time clustering under equal weights, which is another cross-section dependence issue, the magnitude of rejection rates is smaller in the case of overlapping returns. However, the magnitude is higher in the case of overlapping returns under value weights. It indicates the value-weighted scheme has stronger positive impact to improve misspecification caused by calendar time clustering.

5.5 Mean Monthly Calendar-Time (MMAR)

5.5.1 Simulation on random samples

Equally-weighted scheme vs. Value-weighted scheme

The analysis of the percentages of rejection rates of the null hypothesis of zero mean monthly calendar-time abnormal return at a significance level of 5% is listed **Table 5.49**. This approach based on different benchmarks produce well-specified test statistics at a significance level of 5% over three investment horizons. The rejection rates over five years are 4.4% and 4.8% when the reference portfolios are matched by size and by BTM, respectively. Regarding the pattern of test statistics under MMARs, there are negatively biased test statistics with higher rejection rates on the lower tail under equally weighted scheme. For instance, the rejection rates over an investment horizon of five years are 6% on the lower tail and 2.4% on the upper tail when the reference portfolios are matched by size and BTM. The underrejection on the upper tail indicates negative abnormal returns are observed more often than positive returns. The mean of abnormal returns does not show a consistent trend for all benchmarks with the length of investment horizons. However, it is all negative over an investment horizon of one year. Moreover, the benchmark matched by BTM shows an upward trend with negative means of abnormal returns from one year to five years. Regarding value-weighted portfolio returns, the results are mixed. However, well-specified test statistics based on different benchmarks are yielded over three investment horizons. Test statistics are also negatively biased as the results based on the equally-weighted scheme except the case when applying the benchmark matched by size and BTM over three years. The mean of abnormal returns is mostly negative apart from the case of matching by size in three years. It is worth noting that portfolios measured based on value weights show smaller rejection returns compared with equal weights in most circumstances. When compared rejection rates across different investment horizons, there is not a clear trend with the length of investment horizons. Different benchmarks yield different trends of rejection rates from one year to five years. For instance, regarding matching by BTM, the rejection rate is 3.6% in one year and declines to 7.2% in three years but increases to 4.8% in five years under equal weights. However, when the value weighted scheme is applied, the rejection rate increases to 6.4% in three years from 6% in one year and decreases to 3.6% in five years. Lyon, Barber and Tsai (1999) report the rejection rates in two-tailed test at a significance level of 5% with a benchmark of size and BTM under equal weights are 4%, 2.5%, 3.6% over one-, three- and five years. I achieve the results of 2.4%, 3.2% and 3.6% over the same investment horizon. With respect to the value-weighted scheme, the rejection rates documented by Lyon et al. (1999) are 3.6%, 4.8% and 4.4% whereas UK market presents 4%, 3.6% and 2.8%. This comparison shows consistent well-specified test statistics in the UK market but different magnitudes. The pattern of test statistics are different in the US and UK market as well. It is difficult to tell which benchmark is superior with well specified results since each benchmark presents inconsistent performance in three periods, including rejection rates and mean abnormal returns. However, if compared with conventional calendar time approach with extremely high rejection rates under equal weights but with reasonable rejection rate under value weights, the mean monthly abnormal returns approach based on different criteria such as size, BTM shows superior ability to detect long-run abnormal returns with lower and stable rejection rates, especially in random samples which potentially include many small firms. This is consistent with the findings shown by Lyon et al. (1999). They attribute this phenomenon to two causes: the violation of assumption of linearity of controlling factors and interactions among factors.

Descriptive statistics of abnormal returns

To further examine the distribution properties of abnormal returns, **Table 5.50** compares the descriptive statistics of abnormal returns based on both an equally-weighted scheme and value-weighted scheme. As discussed in the previous section, most means of abnormal returns are negative with a few exceptions. For instance, over an investment horizon of three years, the mean abnormal return based on equal weights is 0.013% when the reference benchmark is matched by size and BTM and is -0.015% when the reference benchmark is matched by BTM. The magnitude of mean abnormal returns is mostly larger under value weights. For example, the mean abnormal return is -0.005% under equal weights and is -0.135% under value weights when the reference portfolios matched by BTM is applied over three years. Regarding the equally-weighted portfolios, skewness is all positive except when the benchmark is matched by size over five years. Moreover, the value of skewness is very close to zero. Kurtosis is close to 3 in most cases. This approximates normal distribution of abnormal returns. With respect to the value-weighted portfolios, although most of skewness is negative, it has smaller magnitude compared with the equally-weighted portfolios. In addition, kurtosis is more close to 3. Therefore, when a value-weighted scheme is applied, abnormal returns are normally distributed.

Power of test

I also analyze the power of test with all the benchmarks over three investment horizons. **Table 5.51** based on an equal weighting scheme shows a downward trend of rejection rates with the increasing length of investment horizons for all benchmarks. For instance, the rejection rates with abnormal return of -20% induced are 96%, 63.6% and 36% over one-, three- and five year when the reference portfolios are matched by size. Moreover, the power of test increases when higher abnormal returns are introduced except two cases which are when matching by BTM with abnormal return of 5% in three years and when matching by size and BTM with abnormal return of 5% in five years. The exception can be attributed to negative mean abnormal returns. When compared with positive and negative shocks, the power of test is more sensitive with higher rejection rates when the abnormal returns are positive. The most powerful tests are exhibited when the reference portfolios are matched by size and size/BTM. With respect to the value-weighted scheme on portfolio returns in **Table 5.52**, most of the power of test is reduced compared with the equally-weighted portfolios. For instance, the rejection rate based on reference portfolios matched by size is 22% under value weights over five years whereas the corresponding figure under equal weights is 36%. This may be attributed to the mismatch that the portfolios of event firms are value-weighted but the reference portfolios are equally-weighted. The longer time horizon still reduces the power of test as equally weighted portfolio returns. The benchmarks matching by BTM and equally-weighted market returns show the poorest power of test, especially in five years. For instance, the rejection rate is 4.8% with abnormal return of 20% in five years under benchmark of BTM benchmark and 2% under benchmark of equally-weighted market return. MMAR does improve the power of test with higher rejection rates when abnormal returns are induced compared with the conventional calendar time approach. Take an investment horizon of five years as an example, the three-factor model under OLS based on value weights yields a rejection rate of 20% whilst the mean monthly calendar-time portfolio returns under value weights using reference portfolios matched by size produce a rejection of 22%.

Therefore, when calendar-time approach is applied based on the UK data or data with many small firms in the long-run event studies, the approach based on mean monthly calendar-time portfolio returns is more appropriate to be applied in random samples. Because not only the weighting schemes have limited impact on abnormal returns, but also, this approach has higher power compared with the conventional calendar-time approach based on factor models. When compared with other approaches such as event-time based on reference portfolios, this approach outperforms with higher power. In addition, the reference portfolios can be matched by size, BTM or both.

5.5.2 Simulation on non-random samples

5.5.2.1 Large/Small size

Small size

The results based on small firms in size decile 1 with the smallest market value in September year t are reported in **Table 5.53**. When the benchmark is according to book-to-market ratio or equally weighted market return, it yields misspecified test statistics with high rejection rates under both weighting schemes, even compared with random samples. For benchmark matched by BTM under equal weights, the rejection rate increases from 34% to 43.2% from one year to three years but declines to 25.2% in five years at a significance level of 5%. This finding suggests characteristics of the event firm should be matched with characteristics of the reference portfolios when non-random sample is examined. Therefore, in this case, reference portfolios are required to share at least one common characteristic as the event firm in order to generate well-specified results. The other two MMARs under equal weights based on reference portfolios matched by size and size/BTM underrejects the null hypothesis of zero abnormal returns over three investment horizons. For instance, the rejection rates for these two MMARs are 2.8% when the investment horizon is five years. The value-weighted MMARs based on reference portfolios matched by size and size/BTM produce well-specified results in two-tailed test except the case of a rejection of 8% when the reference portfolios are matched by size over an investment horizon of five years. MMARs based on reference portfolios matched by size produce a rejection of 5.2% over three years at a significance level of 5%. Both of these two benchmarks show negatively biased test statistics with higher rejection rates on the lower tail. Take reference portfolios matched by size and BTM as an example, the rejection rates are 0.4% on the upper tail and 10.4% on the lower tail over five years. The value-weighted scheme improves the underrejection of the null hypothesis of zero abnormal returns under equal weights, but it bears higher rejection rates on the lower tail compared with random samples for the benchmarks matched by size and size/BTM. It is important to note that the magnitude of mean abnormal returns in small firms is larger than in random samples. For example, random samples using reference portfolios matched by size and BTM yields mean abnormal return of -0.029% over five years whereas the corresponding figure for small firms is -0.165%.

Large size

Firms in decile 10 with the largest market value generate similar results as random samples in **Table 5.54**. However, the null hypothesis of zero abnormal returns is underrejected when the reference portfolios are matched by BTM. This confirms the finding of matched characteristics of the event firm and reference portfolios in small firms. The other two benchmarks, namely reference portfolios by size and by size/BTM, generate well-specified results over three investment horizons when an equally-weighted scheme is applied. For instance, when reference portfolios matched by size is applied, the rejection rates are 6%, 4.4% and 5.6% over one-, three- and five years, respectively. Similar results are documented when reference portfolios are matched by size and BTM. However, mixed results are indicated when value-weighted scheme is employed. Over an investment horizon of one year, the rejection rates are 5.2% when reference portfolios are matched by size and by size/BTM. However, the corresponding figures are 0.8% and 0% over three years. This indicates severe underrejection of the null hypothesis of zero abnormal returns. When it comes to an investment horizon of five years, the corresponding figures are 3.2% and 2.4%. It is interesting to note that test statistics are mostly negatively biased under both weighting schemes when size is used as a matching criterion whereas reference portfolios matched by size and BTM show positively biased test statistics. For instance, the rejection rates over one year are 5.2% on the lower tail and 4.4% on the upper tail when the reference portfolios are matched by size. However, when size and BTM are used as matching criteria for reference portfolios, the rejection rates are 4.8% on the lower tail and 6% on the upper tail. Compared with random samples, the magnitude of mean abnormal return is smaller in large firms. For instance, when the reference portfolios are matched by size under equal weights over five years, the mean abnormal return is 0.008% in random samples whereas large firms exhibit the mean abnormal return of 0.002%.

Compared with conventional calendar-time approach, MMARs under equal weights based on reference portfolios matched with at least one common characteristic as the event firm yield well-specified test statistics in small or large firms although there is an asymmetry of test statistics.

5.5.2.2 High/Low book-to-market ratio (BTM)

Low BTM

Firms in BTM decile 1 with the lowest BTM prefer benchmark matched by BTM and size/BTM as shown in **Table 5.55**. Similar as non-random samples categorized by size, reference portfolios are required to share at least one common characteristic as the event firm, so as to generate well-specified test statistics. In this case, BTM is one essential criterion when constructing reference portfolios. Therefore, it is not surprised to observe high rejection rate when the reference portfolios are matched by size only. The rejection rates reach 24.4% under equal weights and 13.2% under value weights at a significance level of 5% when applying the reference portfolios based on BTM over an investment horizon of five years. Although the other two benchmarks yield well-specified results in two-tailed test, there is an underrejection when an equally-weighted scheme is applied over three years. The rejection rates are 2.8% and 1.6% when the reference portfolios are matched by BTM and size/BTM, respectively. The asymmetry of test statistics varies when different weighting schemes and benchmarks are employed. For instance, the reference portfolios matched by BTM generate higher rejection rates on the lower tail whereas the reference portfolios based on size and BTM yields higher rejection rates on the upper tail.

High BTM

Table 5.56 exhibits the percentages of rejection rates for firms with highest BTM from BTM decile 10. It shows similar results as firms with lowest BTM but with smaller magnitude of rejection rates when applying benchmark matched by size and larger magnitude of rejection rates when matching by BTM and size/BTM under equal weights. For instance, under an equally-weighted scheme, the rejection rate is 13.6% in one year when matching by size for firms with the highest BTM and 25.6% for firms with the smallest BTM. However, the rejection increases from 6.8% to 11.2% in one year for firms with the lowest BTM and firms with the largest BTM when matching the reference portfolios by BTM only. Moreover, the magnitude of means of abnormal returns is higher when compared with firms with the smallest BTM. For instance, under a value-weighted scheme, the mean abnormal return is 0.061% for firms with the highest BTM and 0.041% for firms with the lowest BTM when the reference portfolios are matched by size and BTM over an investment horizon of five years. Most of benchmarks over the three investment horizons show positively biased statistics with two exceptions. One is when matching by BTM in one-, three- and five years under both weighting schemes. The other is when matching by size/BTM in three and five years under value weights.

Compared with conventional calendar-time approach based on factor models which generates misspecification, MMARs with reference portfolios matched by BTM or size/BTM yield well-specified results in most circumstances for firms with the lowest and highest BTM.

5.6 Summary

This chapter focuses on the conventional calendar-time approach which is based on the Fama-French three-factor model and Carhart four-factor model, and mean monthly calendar time portfolio return which is a combination of reference portfolios and calendar months. The methodology applied in this chapter follows the studies undertaken by Lyon, Barber and Tsai (1999) and Ang and Zhang (2004). Both an equally-weighted scheme which allocate weights evenly to individual stocks in the portfolio and a value-weighted scheme which assign weights based on market values of stocks are applied. The numbers of firms are different in different calendar months due to missing returns of firms and delisted firms. However, the previous chapter applying the event-time approach based on reference portfolios and control firm replace the missing returns with the benchmark returns. Since the ordinary least squares bears the assumption of normally distributed returns, the weighted least squares advocated by Fama (1998), sandwich variance estimators (1982) and the generalized least squares proposed by Gregory, Guermat and Al-Shawawreh (2010) in the regression are examined to deal with heteroskedasticity caused by different number of firms in each calendar month. The null hypothesis is zero intercept which indicates abnormal return of a portfolio over T period when applying the calendar-time approach based on models. However, the null hypothesis is zero grand mean monthly abnormal return over the calendar months of T when applying the mean monthly calendar-time abnormal returns. 250 samples of 200 event months with replacement from a period of September, 1982 to September 2007 are firstly randomly selected. Then 200 firms are randomly selected to match with the event months. Event months possibly turn out to be the same with different firms. A firm is not allowed to have multiple events in the same month but firms could have more than one different event months. If the test is well-specified, the rejection frequency is expected to be 5% at a significance level of 5%.

There are a few conclusions to be drawn from the empirical results. Firstly, the weighting schemes play an important role when detecting long-run abnormal returns. This finding is inconsistent with the results documented by Lyon, Barber and Tsai (1999). The US data shows well-specified test statistics regardless of weighting schemes. When applying an equally-weighted scheme, there is severe misspecification with high rejection rates, particularly in longer investment horizons with the application of the conventional calendar-time approach. The winsorization does improve the misspecification but still cannot turn around the whole situation. However, when applying the value-weighted scheme, the misspecification is significantly improved and disappears in longer investment horizons. The anomalies in this study can be attributed to the specific characteristic of the UK market structure which is filled with mostly small firms. It is worth mentioning that the magnitude of rejection rates under equal weights is much higher when compared with the results documented by Lyon, Barber and Tsai (1999). Secondly, the pattern of test statistics is asymmetric regardless of the approaches or weighting schemes over all investment horizons. Moreover, these patterns vary on random samples and non-random samples. Thirdly, small size effect dominates in both random and non-random samples when conventional calendar-time approach is applied, especially under an equal-weighted scheme. Lyon, Barber and Tsai (1999) document a rejection rate of 3.3% over an investment horizon of three years when applying the three-factor model under equal weights in small firms based on the US data whereas the corresponding figure is 92.4% based on the UK data in this study. Since the UK market is filled with mostly firms with small capitalization, the results are consistent with the UK market structure. Fourthly, the power of test decreases with the length of investment horizons, which is consistent as prior research (Ang and Zhang, 2004). Moreover, in the case of equally-weighted portfolio returns, the power of test shows considerable small rejection rates when negative shock is introduced, especially in three- and five years. Additionally, the rejection rates decrease significantly with higher magnitude of negative shocks. Fifthly, the explanatory power of both the three-factor and four-four factor models, which is measured by the adjusted R square, increases with investment horizons. In addition, the four-factor model generates results with lower rejection rates and with higher explanatory power of the independent variables to dependent variables compared with the three-factor model. However, the four-factor model is not favourable as suggested by Ang and Zhang (2004). The misspecification is demonstrated with high rejection rates over three investment horizons by Ang and Zhang (2004) as the results based on the UK data in this study, but with a much smaller magnitude. For instance, the rejection rates are 58% when the four-factor model under WLS is applied based equal weights in the UK stock market over an investment horizon of five years whereas the corresponding figure based on the UK stock market is 100%. Sixthly, the three-factor model under OLS together with sandwich variance estimators outperforms in the conventional calendar-time approach. Although WLS incorporates the number of firms in calendar months as weights to control heteroskedasticity, the rejection rates are higher compared with the White estimator under OLS, regardless of weighting schemes. GLS generates lower rejection rates compared with WLS but similar rejection rates compared with the White estimator under OLS. Therefore, OLS with White estimator is advocated to be applied in the long-run event studies. Last but not least, the mean monthly abnormal portfolio return largely improves the misspecification regardless of the weighting scheme. This is consistent as Lyon, Barber and Tsai (1999). Moreover, this method has the highest power of test compared with all other approaches. It is more robust in non-random samples when the matching criteria are similar as the characteristics of the sample firms.

In summary, the Fama-French three-factor model under a value-weighted scheme is more appropriate to apply in the long-run event studies to detect the long-term abnormal stock performance in random samples because of its well-specified test statistics and higher power of test especially in the long run. MMAR is also suitable, especially in non-random samples. When the power of test and non-random samples are taken into consideration, MMAR outperforms the Fama-French three-factor model under OLS with White estimator.

Figures and Tables of Chapter 5

Table 5.1 Rejection frequency of intercept tests from calendar-time approach-Equally weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp} RPM_t + \beta_{\_hml} HML_t + \beta_{\_smb} SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp} RPM_t + \beta_{\_hml} HML_t + \beta_{\_smb} SMB_t + \beta_{\_umd} UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor (SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	2.0	10.0	16.4	0.4	3.2	0.8	15.6	2.4	28
	WLS	2.4	13.2	21.2	0.0	4.0	0.8	20.4	2.8	31.2
	GLS	2.4	13.6	20.0	0.0	4.4	1.2	18.8	2.0	31.6
Four-factor model	**OLS**	1.6	9.2	20.0	0.0	3.6	1.2	18.8	1.2	32.4
	WLS	2.8	14.4	24.4	0.4	5.6	0.4	24.0	1.6	36.0
	GLS	2.4	13.6	26.4	0.0	4.8	1.2	25.2	1.2	34.4

Table 5.1 continued

		1%	5%	10%	1%		5%		10%	
		α=0	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
		Panel B: Three years								
Three-factor model	**OLS**	9.2	29.2	42.0	0.0	15.6	0.0	42.0	0.0	55.6
	WLS	13.6	34.4	44.4	0.0	19.6	0.0	44.4	0.0	62.0
	GLS	18.8	36.4	54.4	0.0	28.0	0.0	54.4	0.0	72.4
Four-factor model	**OLS**	11.6	34.0	45.6	0.0	17.6	0.0	45.6	0.0	60.8
	WLS	14.4	31.6	42.8	0.0	21.6	0.0	42.8	0.0	62.4
	GLS	18.8	37.2	52.4	0.0	26.8	0.0	52.4	0.0	72.4
		Panel C: Five years								
Three-factor model	**OLS**	10.0	36.0	48.4	0.0	18.4	0.0	48.4	0.4	60.4
	WLS	14.8	41.2	55.6	0.0	23.6	0.0	55.6	0.0	73.2
	GLS	15.2	40.4	58.0	0.0	24.0	0.0	58.0	0.0	72.8
Four-factor model	**OLS**	16.0	40.0	55.2	0.0	25.2	0.0	55.2	0.4	67.6
	WLS	24.8	54.0	70.0	0.0	39.6	0.0	70.0	0.0	84.8
	GLS	25.6	54.8	70.4	0.0	38.4	0.0	70.4	0.0	85.2

Table 5.2 Coefficients and R-squared from three-factor model -Equally weighted portfolios

This table reports coefficients from regression based on Fama-French three-factor model to test the null hypothesis of zero abnormal return which is represented by α when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp} RPM_t + \beta_{\_hml} HML_t + \beta_{\_smb} SMB_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), and market risk premium(RPM_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. The means and medians of coefficients among 250 samples are calculated for comparison.

		One year				**Three years**				**Five years**			
		Coefficients		**Test statistics**		**Coefficients**		**Test statistics**		**Coefficients**		**Test statistics**	
		Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
	α	0.0033	0.0032	0.7518	0.8151	0.0040	0.0038	1.4513	1.3854	0.0038	0.0036	1.5263	1.6085
	β_rp	0.8670	0.8655	9.8495	10.0521	0.8658	0.8611	14.5507	14.5634	0.8601	0.8569	16.2497	16.6023
OLS	**β_smb**	0.7285	0.7113	6.0942	6.1592	0.7223	0.7208	9.0591	9.0680	0.7041	0.6958	9.8525	10.0030
	β_hml	0.0103	0.0225	0.1360	0.2376	0.0012	0.0108	0.0475	0.1499	0.0221	0.0287	0.4002	0.4928
	R^2	0.3074	0.3124			0.4835	0.4948			0.5305	0.5477		
	Adj. R^2	0.3006	0.3056			0.4784	0.4898			0.5258	0.5432		

Table 5.2 continued

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
WLS	α	0.0033	0.0032	0.8517	0.8791	0.0038	0.0036	1.6057	1.5505	0.0036	0.0035	1.7720	1.7397
	β_rp	0.8566	0.8578	10.8511	11.0784	0.8583	0.8566	16.9641	17.0374	0.8529	0.8574	20.3404	20.3220
	β_smb	0.7408	0.7258	6.7892	6.8394	0.7470	0.7441	10.9797	11.0588	0.7325	0.7280	13.0400	13.1563
	β_hml	-0.0032	0.0068	0.0267	0.0695	-0.0023	0.0026	-0.0273	0.0437	0.0207	0.0338	0.5026	0.7102
	R^2	0.3494	0.3541			0.5650	0.5738			0.6474	0.6549		
	Adj. R^2	0.3430	0.3478			0.5607	0.5696			0.6440	0.6515		
GLS	α	0.0033	0.0030	0.8362	0.8299	0.0038	0.0036	1.6007	1.5417	0.0037	0.0035	1.7746	1.7767
	β_rp	0.8571	0.8585	10.7416	10.9466	0.8589	0.8613	16.8131	16.9871	0.8541	0.8580	19.8105	19.8971
	β_smb	0.7399	0.7210	6.7148	6.7313	0.7432	0.7385	10.8159	10.8742	0.7263	0.7234	12.5512	12.6129
	β_hml	-0.0023	0.0090	0.0443	0.0926	-0.0023	0.0032	-0.0272	0.0583	0.0197	0.0313	0.4697	0.6262
	R^2	0.3445	0.3494			0.5596	0.5705			0.6338	0.6427		
	Adj. R^2	0.3381	0.3430			0.5553	0.5663			0.6302	0.6392		

Table 5.3 Coefficients and R-squared from four-factor model -Equally weighted portfolios

This table reports coefficients from regression based on Carhart four-factor model to test the null hypothesis of zero abnormal return which is represented by α when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. The means and medians of coefficients among 250 samples are calculated for comparison.

		One year				**Three years**				**Five years**			
		Coefficients		**Test statistics**		**Coefficients**		**Test statistics**		**Coefficients**		**Test statistics**	
		Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
OLS	**α**	0.0038	0.0038	0.8145	0.8935	0.0044	0.0045	1.5294	1.5153	0.0045	0.0044	1.7043	1.7497
	β_rp	0.8627	0.8551	9.7182	9.9052	0.8619	0.8550	14.3579	14.4627	0.8537	0.8477	16.0019	16.3456
	β_smb	0.7318	0.7175	6.1069	6.1997	0.7253	0.7257	9.0823	9.0590	0.7091	0.6982	9.9072	10.0489
	β_hml	-0.0089	0.0088	-0.0346	0.0757	-0.0161	-0.0025	-0.1888	-0.0348	-0.0065	0.0008	-0.0469	0.0100
	β_umd	-0.0405	-0.0337	-0.3191	-0.2734	-0.0366	-0.0398	-0.4811	-0.5152	-0.0602	-0.0593	-0.8196	-0.8607
	R^2	0.3114	0.3147			0.4866	0.4980			0.5340	0.5509		
	Adj. R^2	0.3024	0.3057			0.4799	0.4914			0.5279	0.5451		

Table 5.3 continued

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
WLS	**α**	0.0039	0.0036	0.9548	0.9428	0.0045	0.0041	1.8130	1.7501	0.0044	0.0042	2.0996	2.1354
	β_rp	0.8510	0.8533	10.6781	10.8952	0.8507	0.8497	16.6376	16.7521	0.8425	0.8461	19.8663	19.8785
	β_smb	0.7448	0.7294	6.8074	6.8217	0.7526	0.7509	11.0315	11.0727	0.7395	0.7352	13.1507	13.2487
	β_hml	-0.0285	-0.0158	-0.2052	-0.1552	-0.0330	-0.0223	-0.4806	-0.3400	-0.0186	-0.0130	-0.2654	-0.2367
	β_umd	-0.0546	-0.0512	-0.4668	-0.4497	-0.0657	-0.0653	-0.9256	-0.8692	-0.0811	-0.0743	-1.3705	-1.3386
	$\mathbf{R^2}$	0.3529	0.3587			0.5679	0.5752			0.6509	0.6571		
	$\mathbf{Adj.\,R^2}$	0.3444	0.3503			0.5622	0.5697			0.6463	0.6526		
GLS	**α**	0.0038	0.0035	0.9327	0.9083	0.0044	0.0042	1.7969	1.6860	0.0044	0.0042	2.0726	2.0979
	β_rp	0.8519	0.8526	10.5731	10.7748	0.8516	0.8507	16.5025	16.6740	0.8445	0.8466	19.3750	19.3890
	β_smb	0.7440	0.7274	6.7333	6.7682	0.7487	0.7461	10.8664	10.9081	0.7333	0.7295	12.6548	12.6710
	β_hml	-0.0262	-0.0087	-0.1775	-0.0766	-0.0319	-0.0227	-0.4583	-0.3484	-0.0174	-0.0144	-0.2378	-0.2537
	β_umd	-0.0528	-0.0520	-0.4519	-0.4777	-0.0635	-0.0632	-0.8890	-0.8637	-0.0771	-0.0741	-1.2713	-1.2317
	$\mathbf{R^2}$	0.3480	0.3547			0.5624	0.5723			0.6372	0.6444		
	$\mathbf{Adj.\,R^2}$	0.3394	0.3463			0.5567	0.5667			0.6324	0.6397		

Table 5.4 Rejection frequency of intercept tests from calendar-time approach-Winsorized equally weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$
$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected. Winsorization process is taken as winsorizing 1% of stock raw returns of each sample in each tail.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	2.0	5.6	13.2	0.8	2.0	3.6	9.6	8.4	17.2
	WLS	1.2	8.4	16.0	0.8	3.2	4.4	11.6	8.4	18.4
	GLS	2.0	8.0	15.2	0.8	3.2	4.0	11.2	7.6	16.8
Four-factor model	**OLS**	2.0	4.8	12.8	0.8	2.0	3.6	9.2	6.8	16.4
	WLS	2.8	9.2	15.2	0.8	4.0	4.0	11.2	5.2	19.6
	GLS	2.4	7.2	13.6	1.2	3.6	3.2	10.4	5.6	19.6

Table 5.4 Continued

		1%	5%	10%	1%		5%		10%	
		α=0	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
						Panel B: Three years				
Three-factor model	**OLS**	2.4	8.8	15.6	0.0	4.4	0.4	15.2	2.4	29.2
	WLS	2.8	10.4	18.4	0.0	6.0	0.4	18.0	2.8	30.0
	GLS	4.4	17.2	23.6	0.0	7.6	0.0	23.6	0.8	34.4
Four-factor model	**OLS**	2.4	12.8	20.0	0.0	6.8	0.8	19.2	3.2	31.6
	WLS	2.8	10.0	19.2	0.0	5.6	0.4	18.8	2.0	28.0
	GLS	4.4	17.2	23.6	0.0	7.6	0.0	23.6	0.8	34.4
						Panel C: Five years				
Three-factor model	**OLS**	2.4	10.0	18.8	0.0	4.0	0.4	18.4	2.4	29.2
	WLS	0.4	6.8	15.6	0.0	2.4	0.0	15.6	0.4	28.0
	GLS	0.8	7.6	17.6	0.0	2.4	0.0	17.6	0.0	27.2
Four-factor model	**OLS**	3.6	13.6	25.6	0.0	6.0	1.6	24.0	1.6	37.2
	WLS	2.4	14.8	28.4	0.0	6.0	0.0	28.4	0.0	40.0
	GLS	2.8	15.2	28.0	0.0	5.6	0.0	28.0	0.0	41.2

Table 5.5 Coefficients and R-squared from three-factor model –Winsorized equally weighted portfolios

This table reports coefficients from regression based on Fama-French three-factor model to test the null hypothesis of zero abnormal return which is represented by α when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), and market risk premium(RPM_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. The means and medians of coefficients among 250 samples are calculated for comparison. Winsorization process is taken as winsorizing 1% of stock raw returns of each sample in each tail.

		One year				**Three years**				**Five years**			
		Coefficients		**Test statistics**		**Coefficients**		**Test statistics**		**Coefficients**		**Test statistics**	
		Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
OLS	**α**	0.0009	0.0010	0.2522	0.3080	0.0015	0.0015	0.6653	0.6548	0.0014	0.0016	0.6974	0.7418
	β_rp	0.8333	0.8318	11.3441	11.2568	0.8282	0.8247	16.7769	16.8392	0.8289	0.8264	18.6002	18.5863
	β_smb	0.6559	0.6558	6.6538	6.6392	0.6554	0.6578	9.9070	9.8466	0.6376	0.6316	10.6759	10.6467
	β_hml	0.0324	0.0293	0.3856	0.3613	0.0316	0.0349	0.5618	0.5919	0.0468	0.0457	0.9083	0.9134
	R^2	0.3606	0.3581			0.5471	0.5560			0.5921	0.6000		
	Adj. R^2	0.3544	0.3518			0.5426	0.5516			0.5881	0.5961		

Table 5.5 continued

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
WLS	**α**	0.0009	0.0010	0.3083	0.3165	0.0013	0.0013	0.7092	0.6809	0.0012	0.0011	0.7252	0.6980
	β_rp	0.8279	0.8256	12.6141	12.6148	0.8234	0.8194	19.8493	19.7478	0.8215	0.8234	23.5538	23.5398
	β_smb	0.6663	0.6635	7.4704	7.4414	0.6781	0.6752	12.1568	12.1518	0.6618	0.6597	14.2934	14.1661
	β_hml	0.0215	0.0235	0.2890	0.3296	0.0276	0.0283	0.5657	0.6070	0.0455	0.0456	1.1534	1.0909
	$\mathbf{R^2}$	0.4105	0.4120			0.6333	0.6330			0.7070	0.7088		
	$\mathbf{Adj.\,R^2}$	0.4047	0.4063			0.6297	0.6294			0.7041	0.7060		
GLS	**α**	0.0009	0.0009	0.3023	0.2963	0.0014	0.0013	0.7158	0.6650	0.0012	0.0012	0.7623	0.7133
	β_rp	0.8272	0.8242	12.5095	12.5132	0.8238	0.8213	19.6854	19.5517	0.8226	0.8236	23.0602	22.9488
	β_smb	0.6653	0.6610	7.4102	7.4063	0.6760	0.6753	12.0142	12.0425	0.6574	0.6559	13.8644	13.7715
	β_hml	0.0225	0.0288	0.2971	0.3582	0.0274	0.0291	0.5545	0.6094	0.0449	0.0456	1.1105	1.0957
	$\mathbf{R^2}$	0.4062	0.4089			0.6287	0.6292			0.6971	0.6973		
	$\mathbf{Adj.\,R^2}$	0.4004	0.4031			0.6251	0.6256			0.6941	0.6944		

Table 5.6 Coefficients and R-squared from four-factor model –Winsorized equally weighted portfolios

This table reports coefficients from regression based on Carhart four-factor model to test the null hypothesis of zero abnormal return which is represented by α when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp} RPM_t + \beta_{\_hml} HML_t + \beta_{\_smb} SMB_t + \beta_{\_umd} UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. The means and medians of coefficients among 250 samples are calculated for comparison. Winsorization process is taken as winsorizing 1% of stock raw returns of each sample in each tail.

		One year				**Three years**				**Five years**			
		Coefficients		**Test statistics**		**Coefficients**		**Test statistics**		**Coefficients**		**Test statistics**	
		Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
OLS	**α**	0.0012	0.0013	0.3221	0.3628	0.0018	0.0018	0.7697	0.7486	0.0019	0.0020	0.8916	0.9245
	β_rp	0.8303	0.8282	11.2067	11.1710	0.8251	0.8218	16.5661	16.6156	0.8240	0.8203	18.3363	18.3163
	β_smb	0.6581	0.6618	6.6624	6.6895	0.6578	0.6592	9.9260	9.8665	0.6414	0.6373	10.7217	10.7165
	β_hml	0.0194	0.0342	0.2069	0.3288	0.0180	0.0180	0.2738	0.2790	0.0252	0.0220	0.4289	0.3799
	β_umd	-0.0273	-0.0376	-0.2707	-0.3595	-0.0290	-0.0314	-0.4535	-0.4499	-0.0454	-0.0472	-0.7533	-0.8205
	R^2	0.3643	0.3628			0.5498	0.5597			0.5950	0.6018		
	Adj. R^2	0.3559	0.3544			0.5439	0.5539			0.5897	0.5965		

Table 5.6 Continued

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
WLS	**α**	0.0013	0.0013	0.4171	0.4105	0.0019	0.0018	0.9487	0.8656	0.0018	0.0018	1.0927	1.0855
	β_rp	0.8241	0.8212	12.4335	12.4562	0.8174	0.8147	19.4895	19.3562	0.8131	0.8144	23.0374	23.0113
	β_smb	0.6693	0.6686	7.4839	7.4617	0.6826	0.6784	12.2009	12.1857	0.6674	0.6676	14.3996	14.2589
	β_hml	0.0036	0.0099	0.0487	0.1080	0.0031	0.0004	0.0458	0.0072	0.0138	0.0122	0.2980	0.2558
	β_umd	-0.0381	-0.0413	-0.4051	-0.4329	-0.0523	-0.0491	-0.9003	-0.7979	-0.0655	-0.0655	-1.3609	-1.3448
	$\mathbf{R^2}$	0.4136	0.4193			0.6355	0.6356			0.7096	0.7110		
	$\mathbf{Adj.\ R^2}$	0.4058	0.4117			0.6307	0.6308			0.7058	0.7071		
GLS	**α**	0.0013	0.0012	0.4117	0.3840	0.0019	0.0018	0.9464	0.8934	0.0019	0.0019	1.1021	1.0680
	β_rp	0.8236	0.8215	12.3390	12.3514	0.8180	0.8160	19.3228	19.1261	0.8150	0.8152	22.5826	22.4392
	β_smb	0.6686	0.6662	7.4238	7.4179	0.6800	0.6784	12.0422	12.0478	0.6629	0.6595	13.9589	13.8563
	β_hml	0.0049	0.0079	0.0627	0.0976	0.0036	0.0018	0.0531	0.0356	0.0148	0.0150	0.3124	0.3194
	β_umd	-0.0381	-0.0411	-0.4051	-0.4479	-0.0514	-0.0490	-0.8781	-0.8485	-0.0623	-0.0632	-1.2717	-1.2910
	$\mathbf{R^2}$	0.4094	0.4120			0.6305	0.6309			0.6995	0.6992		
	$\mathbf{Adj.\ R^2}$	0.4016	0.4042			0.6256	0.6261			0.6956	0.6953		

Table 5.7 Descriptive statistics of alphas from the Fama-French three-factor model based on equally-weighted portfolios

This table reports descriptive statistics of alphas derived from the Fama-French three-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), and market risk premium(RPM_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms.

Three-factor model	**Equally-weighted portfolios**				**Winsorized equally-weighted portfolios**			
One year	**Mean**	**Median**	**Skewness**	**Kurtosis**	**Mean**	**Median**	**Skewness**	**Kurtosis**
OLS	0.00333	0.00316	0.98	8.97	0.00088	0.00101	-0.28	3.53
WLS	0.00332	0.00316	1.41	12.43	0.00094	0.00098	-0.16	3.23
GLS	0.00328	0.00304	1.36	11.95	0.00092	0.00092	-0.13	3.24
Three years								
OLS	0.00403	0.00380	0.19	2.65	0.00151	0.00147	0.14	2.78
WLS	0.00377	0.00361	0.53	3.82	0.00134	0.00131	0.21	3.44
GLS	0.00379	0.00357	0.56	3.65	0.00136	0.00129	0.24	3.41
Five years								
OLS	0.00382	0.00364	0.22	3.81	0.00142	0.00157	-0.10	2.96
WLS	0.00356	0.00345	1.57	10.44	0.00116	0.00111	0.08	2.54
GLS	0.00365	0.00351	1.45	9.66	0.00125	0.00118	0.03	2.44

Table 5.8 Descriptive statistics of alpha from the Carhart four-factor model based on equally-weighted portfolios

This table reports descriptive statistics of alpha derived from the Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms.

Four-factor model	**Equally-weighted portfolios**				**Winsorized equally-weighted portfolios**			
One year	**Mean**	**Median**	**Skewness**	**Kurtosis**	**Mean**	**Median**	**Skewness**	**Kurtosis**
OLS	0.00376	0.00384	1.05	8.44	0.00117	0.00126	-0.34	3.31
WLS	0.00388	0.00358	1.32	10.81	0.00133	0.00127	-0.14	3.21
GLS	0.00383	0.00345	1.26	10.43	0.00132	0.00124	-0.17	3.24
Three years								
OLS	0.00441	0.00450	-0.01	2.79	0.00181	0.00176	-0.01	2.99
WLS	0.00445	0.00412	0.57	3.48	0.00188	0.00177	0.29	3.39
GLS	0.00445	0.00416	0.55	3.37	0.00188	0.00180	0.29	3.33
Five years								
OLS	0.00446	0.00445	0.28	4.40	0.00190	0.00196	-0.23	3.18
WLS	0.00439	0.00415	1.89	11.45	0.00183	0.00183	0.10	2.46
GLS	0.00444	0.00423	1.85	11.19	0.00188	0.00189	0.04	2.37

Table 5.9 Rejection frequency of all tests in calendar-time approach-Equally-weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. 250 samples of 200 firms are random selected. Conventional ordinary least squares and three approaches correcting heteroskedasticity including weighted least squares (WLS), sandwich variance estimators (1980) combing OLS and WLS, and generalized least squares (GLS), are applied in both models over three investment horizons. Both one-tail and two-tailed tests are examined at a significance level of 5%. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

				5%			**Mean Alpha**
Holding period	**Model**	**Technique**		**α=0**	**α<0**	**α>0**	
One year	**Three-factor model**	OLS	Original test	10.0	0.8	15.6	0.0033328
			sandwich variance estimators	9.2	1.2	16.0	0.0033328
		WLS	Original test	13.2	0.8	20.4	0.0033168
			sandwich variance estimators	13.2	1.2	19.6	0.0033168
		GLS	Original test	13.6	1.2	18.8	0.0032835
	Four-factor model	OLS	Original test	9.2	1.2	18.8	0.0037594
			sandwich variance estimators	7.6	1.2	16.0	0.0037594
		WLS	Original test	14.4	0.4	24.0	0.0038844
			sandwich variance estimators	12.4	0.8	20.8	0.0038844
		GLS	Original test	13.6	1.2	25.2	0.0038269

Table 5.9 continued

Holding period	Model	Technique		5% α=0	5% α<0	5% α>0	Mean Alpha
Three years	Three-factor model	OLS	Original test	29.2	0.0	42.0	0.0040342
			sandwich variance estimators	26.8	0.0	41.6	0.0040342
		WLS	Original test	34.4	0.0	44.4	0.0037734
			sandwich variance estimators	33.6	0.0	44.8	0.0037734
		GLS	Original test	31.6	0.0	42.8	0.0037921
	Four-factor model	OLS	Original test	34.0	0.0	45.6	0.0044149
			sandwich variance estimators	28.8	0.0	41.2	0.0044149
		WLS	Original test	36.4	0.0	54.4	0.0044528
			sandwich variance estimators	36.0	0.0	54.4	0.0044528
		GLS	Original test	37.2	0.0	52.4	0.0044492
Five years	Three-factor model	OLS	Original test	36.0	0.0	48.4	0.0038232
			sandwich variance estimators	30.4	0.0	46.0	0.0038232
		WLS	Original test	41.2	0.0	55.6	0.0035599
			sandwich variance estimators	37.2	0.0	53.2	0.0035599
		GLS	Original test	40.4	0.0	58.0	0.0036509
	Four-factor model	OLS	Original test	40.0	0.0	55.2	0.0044571
			sandwich variance estimators	35.2	0.0	48.4	0.0044571
		WLS	Original test	54.0	0.0	70.0	0.0043857
			sandwich variance estimators	48.8	0.0	65.6	0.0043857
		GLS	Original test	54.8	0.0	70.4	0.0044366

Table 5.10 Power of test-Equally-weighted portfolios (two-tailed test)

This table reports percentages of rejection rates of the null hypothesis of zero intercept on two-tailed test. The intercepts are derived from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is the study period (12-. 36- and 60 months). The independent variables such as size factor (SMB_t), book-to-market ratio factor (HML_t), market risk premium (RPM_t), and momentum factors (UMD_t) are obtained from the website of Exeter University based on studies done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. The two-tailed tests are examined at significance level of 5%. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected. Abnormal returns in the range of -20% to 20% at an interval of 5% are introduced to individual stocks in order to test the power of test.

		Induced abnormal returns								
		-20%	**-15%**	**-10%**	**-5%**	**0%**	**5%**	**10%**	**15%**	**20%**
						Panel A: one year				
	OLS	89.2	65.2	24.4	5.6	10.0	42.0	84.0	97.6	99.6
Three-factor model	**WLS**	92.8	72.0	30.0	6.0	13.2	51.6	90.0	99.2	100
	GLS	93.6	73.2	30.4	5.2	13.6	50.0	88.8	99.2	100
	OLS	87.2	57.6	19.6	5.6	9.2	45.6	80.8	97.6	99.2
Four-factor model	**WLS**	90.4	63.2	23.6	4.4	14.4	53.2	90.4	98.4	100
	GLS	91.2	65.2	22.4	4.4	13.6	52.8	90.0	98.4	100

Table 5.10 continued

		Induced abnormal returns								
		-20%	**-15%**	**-10%**	**-5%**	**0%**	**5%**	**10%**	**15%**	**20%**
		Panel B: three years								
	OLS	6.8	2.0	5.2	12.4	29.2	48.0	70.0	87.6	95.6
Three-factor model	**WLS**	8.8	2.8	3.6	14.4	34.4	60.0	82.4	94.4	100.0
	GLS	9.6	2.8	3.6	15.6	36.4	58.8	82.4	95.2	100.0
	OLS	5.6	2.8	5.6	14.8	34.0	50.8	73.2	85.2	93.6
Four-factor model	**WLS**	4.4	2.0	6.4	21.2	31.6	64.4	90.0	97.6	99.6
	GLS	4.0	1.6	6.4	20.8	37.2	64.8	88.8	98.4	99.2
		Panel C: five years								
	OLS	4.0	6.0	10.4	20.4	36.0	49.6	62.0	70.8	80.0
Three-factor model	**WLS**	0.8	3.6	8.8	22.8	41.2	62.4	79.6	93.2	97.2
	GLS	1.6	3.2	9.6	23.2	40.4	62.0	78.4	92.8	96.8
	OLS	5.2	9.2	14.4	27.2	40.0	54.8	68.0	78.4	88.4
Four-factor model	**WLS**	2.8	7.6	19.6	39.2	54.0	74.4	89.6	97.6	99.6
	GLS	2.8	6.8	19.2	36.0	54.8	72.8	88.4	97.2	99.2

Table 5.11 Mean intercepts in power of test-Equally-weighted portfolios

This table reports mean intercepts in power of test based on two-tailed test under equal weights. The intercepts are derived from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp} RPM_t + \beta_{\_hml} HML_t + \beta_{\_smb} SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp} RPM_t + \beta_{\_hml} HML_t + \beta_{\_smb} SMB_t + \beta_{\_umd} UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is the study period (12-. 36- and 60 months). The independent variables such as size factor (SMB_t), book-to-market ratio factor (HML_t), market risk premium (RPM_t), and momentum factors (UMD_t) are obtained from the website of Exeter University based on studies done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. The two-tailed tests are examined at significance level of 5%. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected. Abnormal returns in the range of -20% to 20% at an interval of 5% are introduced to individual stocks in order to test the power of test.

		Induced abnormal returns								
		-20%	**-15%**	**-10%**	**-5%**	**0%**	**5%**	**10%**	**15%**	**20%**
		Panel A: one year								
Three-factor model	**OLS**	-0.01333	-0.00917	-0.00500	-0.00083	0.00333	0.00750	0.01167	0.01583	0.02000
	WLS	-0.01335	-0.00918	-0.00502	-0.00085	0.00332	0.00748	0.01165	0.01582	0.01998
	GLS	-0.01338	-0.00922	-0.00505	-0.00088	0.00328	0.00745	0.01162	0.01578	0.01995
Four-factor model	**OLS**	-0.01291	-0.00874	-0.00457	-0.00041	0.00376	0.00793	0.01209	0.01626	0.02043
	WLS	-0.01278	-0.00862	-0.00445	-0.00028	0.00388	0.00805	0.01222	0.01638	0.02055
	GLS	-0.01284	-0.00867	-0.00451	-0.00034	0.00383	0.00799	0.01216	0.01633	0.02049

Table 5.11 continued

		Induced abnormal returns								
		-20%	**-15%**	**-10%**	**-5%**	**0%**	**5%**	**10%**	**15%**	**20%**
		Panel B: three years								
Three-factor model	**OLS**	-0.00152	-0.00013	0.00126	0.00265	0.00403	0.00542	0.00681	0.00820	0.00959
	WLS	-0.00178	-0.00039	0.00100	0.00238	0.00377	0.00516	0.00655	0.00794	0.00933
	GLS	-0.00176	-0.00037	0.00101	0.00240	0.00379	0.00518	0.00657	0.00796	0.00935
Four-factor model	**OLS**	-0.00114	0.00025	0.00164	0.00303	0.00441	0.00580	0.00719	0.00858	0.00997
	WLS	-0.00110	0.00029	0.00168	0.00306	0.00445	0.00584	0.00723	0.00862	0.01001
	GLS	-0.00111	0.00028	0.00167	0.00306	0.00445	0.00584	0.00723	0.00862	0.01000
		Panel C: five years								
Three-factor model	**OLS**	0.00049	0.00132	0.00216	0.00299	0.00382	0.00466	0.00549	0.00632	0.00716
	WLS	0.00023	0.00106	0.00189	0.00273	0.00356	0.00439	0.00523	0.00606	0.00689
	GLS	0.00032	0.00115	0.00198	0.00282	0.00365	0.00448	0.00532	0.00615	0.00698
Four-factor model	**OLS**	0.00112	0.00196	0.00279	0.00362	0.00446	0.00529	0.00612	0.00696	0.00779
	WLS	0.00105	0.00189	0.00272	0.00355	0.00439	0.00522	0.00605	0.00689	0.00772
	GLS	0.00110	0.00194	0.00277	0.00360	0.00444	0.00527	0.00610	0.00694	0.00777

Table 5.12 Power of test-Equally-weighted portfolios (one-tailed test)

This table reports percentages of rejection rates on one-tailed test. The alternative hypotheses are positive intercept when positive abnormal returns are induced and negative intercept when negative abnormal returns are induced. The intercepts are derived from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp} RPM_t + \beta_{\_hml} HML_t + \beta_{\_smb} SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp} RPM_t + \beta_{\_hml} HML_t + \beta_{\_smb} SMB_t + \beta_{\_umd} UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is the study period (12-. 36- and 60 months). The independent variables such as size factor (SMB_t), book-to-market ratio factor (HML_t), market risk premium (RPM_t), and momentum factors (UMD_t) are obtained from the website of Exeter University based on studies done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. The one-tailed tests are examined at significance level of 5%. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected. Abnormal returns in the range of -20% to 20% at an interval of 5% are introduced to individual stocks in order to test the power of test.

		-20%	**-15%**	**-10%**	**-5%**	**0%**		**5%**	**10%**	**15%**	**20%**
				α<0		**α<0**	**α>0**		**α>0**		
						Panel A: one year					
Three-factor model	**OLS**	91.2	73.6	36.0	7.2	0.8	15.6	60.0	88.8	99.2	99.6
	WLS	94.8	79.6	41.6	8.8	0.8	20.4	67.6	93.6	99.6	100.0
	GLS	94.8	80.4	42.4	7.6	1.2	18.8	66.0	93.6	99.6	100.0
Four-factor model	**OLS**	92.0	67.6	29.6	8.4	1.2	18.8	59.2	88.0	98.4	100.0
	WLS	94.0	73.2	33.6	6.4	0.4	24.0	68.4	94.4	100.0	100.0
	GLS	94.0	72.0	32.4	7.6	1.2	25.2	70.0	94.0	99.2	100.0

Table 5.12 continued

		-20%	-15%	-10%	-5%	0%		5%	10%	15%	20%
		α<0				α<0	α>0	α>0			
		Panel B: three years									
Three-factor model	**OLS**	12.4	4.4	0.4	0.0	0.0	42.0	62.4	82.8	93.2	97.6
	WLS	18.0	4.8	0.4	0.0	0.0	44.4	74.0	92.0	98.0	100.0
	GLS	16.4	4.4	0.0	0.0	0.0	54.4	73.6	91.6	99.2	100.0
Four-factor model	**OLS**	10.8	4.4	0.8	0.0	0.0	45.6	67.2	82.8	92.8	95.6
	WLS	7.2	2.0	0.0	0.0	0.0	42.8	81.6	96.4	98.8	100.0
	GLS	7.2	1.6	0.0	0.0	0.0	52.4	81.6	95.2	98.4	100.0
		Panel C: five years									
Three-factor model	**OLS**	3.6	2.0	1.2	0.4	0.0	48.4	60.8	69.2	78.8	91.2
	WLS	2.0	0.8	0.0	0.0	0.0	55.6	75.6	89.2	96.8	99.6
	GLS	0.8	0.8	0.0	0.0	0.0	58.0	75.6	89.6	96.4	99.2
Four-factor model	**OLS**	2.0	1.2	0.8	0.4	0.0	55.2	66.4	78.4	87.6	92.0
	WLS	0.0	0.0	0.0	0.0	0.0	70.0	86.8	96.4	99.6	100.0
	GLS	0.0	0.0	0.0	0.0	0.0	70.4	86.4	95.6	99.2	100.0

Table 5.13 Rejection frequency of intercept tests from calendar-time approach-Value weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is the study period (12-. 36- and 60 months). The independent variables such as size factor (SMB_t), book-to-market ratio factor (HML_t), market risk premium (RPM_t), and momentum factors (UMD_t) are obtained from the website of Exeter University based on studies done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. The two-tailed tests are examined at significance level of 5%. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected. Abnormal returns in the range of -20% to 20% at an interval of 5% are introduced to individual stocks in order to test the power of test.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**					**Panel A: One year**				
Three-factor model	**OLS**	2.4	10.4	15.6	1.6	2.4	10.0	5.6	16.8	9.6
	WLS	4.0	8.8	16.4	2.8	2.4	9.6	6.8	18.0	12.4
	GLS	3.2	10.4	17.2	2.0	2.8	10.4	6.8	17.2	12.0
Four-factor model	**OLS**	2.8	7.2	14.4	0.4	2.4	5.6	8.8	8.8	15.2
	WLS	2.4	8.0	12.4	0.8	3.6	4.0	8.4	8.4	17.2
	GLS	2.4	7.2	12.0	0.4	3.2	3.6	8.4	9.6	16.0

Table 5.13 continued

		1%	5%	10%	1%		5%		10%	
		α=0	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
		Panel B: Three years								
Three-factor model	**OLS**	0.0	4.8	8.8	0.4	0.4	4.0	4.8	8.8	10.4
	WLS	0.4	6.8	12.8	0.4	2.0	5.2	7.6	12.8	11.6
	GLS	1.6	8.4	11.2	0.0	3.2	0.4	10.8	2.8	21.2
Four-factor model	**OLS**	0.8	4.8	14.4	0.0	2.0	2.0	12.4	4.0	20.8
	WLS	0.4	4.8	11.2	0.0	1.2	4.4	6.8	11.2	12.0
	GLS	1.6	6.8	10.8	0.0	2.8	0.4	10.4	2.4	22.0
		Panel C: Five years								
Three-factor model	**OLS**	0.0	5.6	12.8	0.4	1.2	6.0	6.8	13.2	10.0
	WLS	0.4	5.2	10.8	0.8	0.8	5.6	5.2	12.0	11.2
	GLS	0.0	4.0	11.6	0.8	0.8	6.4	5.2	10.8	9.2
Four-factor model	**OLS**	0.8	3.6	8.4	0.0	2.8	0.8	7.6	2.8	19.2
	WLS	0.8	5.2	12.4	0.0	2.0	0.8	11.6	2.8	21.6
	GLS	0.8	4.4	10.8	0.0	2.0	0.4	10.4	1.6	19.2

Table 5.14 Coefficients and R-squared from three-factor model -Value weighted portfolios

This table reports coefficients from regression based on Fama-French three-factor model to test the null hypothesis of zero abnormal return which is represented by α when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp} RPM_t + \beta_{\_hml} HML_t + \beta_{\_smb} SMB_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), and market risk premium(RPM_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. The means and medians of coefficients among 250 samples are calculated for comparison.

		One year				**Three years**				**Five years**			
		Coefficients		**Test statistics**		**Coefficients**		**Test statistics**		**Coefficients**		**Test statistics**	
		Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
OLS	**α**	-0.0004	-0.0002	-0.1060	-0.0396	0.0003	0.0004	0.1138	0.1318	0.0000	0.0003	0.0312	0.0883
	β_rp	1.0121	1.0069	10.8534	10.8551	1.0097	1.0049	14.5236	14.5256	1.0057	1.0002	15.9314	16.2787
	β_smb	0.4708	0.4752	3.7311	3.5951	0.3011	0.3106	3.1469	3.3514	0.2483	0.2454	2.7924	2.8481
	β_hml	0.0747	0.0971	0.6975	0.8549	0.1190	0.1216	1.4661	1.5699	0.1358	0.1396	1.8255	1.9139
	$\mathbf{R^2}$	0.3070	0.3062			0.4213	0.4293			0.4617	0.4789		
	$\mathbf{Adj. R^2}$	0.3002	0.2994			0.4156	0.4237			0.4564	0.4737		

Table 5.14 continued

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
WLS	**α**	-0.0006	-0.0005	-0.1429	-0.1197	0.0000	0.0004	0.0674	0.1429	0.0000	0.0003	0.0344	0.1094
	β_rp	1.0166	1.0171	11.8401	11.8077	1.0148	1.0142	16.3129	16.2936	1.0103	1.0045	18.8260	19.0332
	β_smb	0.4552	0.4631	3.8649	3.8605	0.2900	0.2997	3.3667	3.4361	0.2284	0.2135	3.0592	3.2248
	β_hml	0.0682	0.0754	0.6704	0.7832	0.1236	0.1195	1.6763	1.6859	0.1458	0.1549	2.3561	2.6270
	$\mathbf{R^2}$	0.3413	0.3393			0.4766	0.4807			0.5431	0.5535		
	$\mathbf{Adj.\,R^2}$	0.3348	0.3328			0.4715	0.4756			0.5386	0.5492		
GLS	**α**	-0.0005	-0.0006	-0.1195	-0.1440	0.0001	0.0004	0.0675	0.1514	0.0000	0.0003	0.0490	0.1165
	β_rp	1.0139	1.0138	11.6254	11.6827	1.0135	1.0132	16.0489	16.0091	1.0089	1.0016	18.1704	18.3234
	β_smb	0.4601	0.4630	3.8613	3.7858	0.2927	0.3009	3.3482	3.4540	0.2317	0.2141	2.9977	3.1106
	β_hml	0.0686	0.0852	0.6707	0.8067	0.1228	0.1215	1.6390	1.6976	0.1418	0.1557	2.2029	2.4286
	$\mathbf{R^2}$	0.3339	0.3335			0.4687	0.4707			0.5260	0.5372		
	$\mathbf{Adj.\,R^2}$	0.3273	0.3269			0.4635	0.4655			0.5213	0.5327		

Table 5.15 Coefficients and R-squared from four-factor model -Value weighted portfolios

This table reports coefficients from regression based on Carhart four-factor model to test the null hypothesis of zero abnormal return which is represented by α when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. The means and medians of coefficients among 250 samples are calculated for comparison.

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
OLS	**α**	0.0008	0.0008	0.1791	0.1785	0.0016	0.0013	0.4929	0.4060	0.0014	0.0016	0.4726	0.5345
	β_rp	0.9990	0.9989	10.6599	10.7271	0.9961	0.9898	14.2809	14.3085	0.9918	0.9914	15.6652	15.9594
	β_smb	0.4807	0.4824	3.8147	3.7289	0.3115	0.3223	3.2660	3.5002	0.2590	0.2567	2.9297	3.0563
	β_hml	0.0168	0.0231	0.1526	0.1777	0.0588	0.0667	0.6583	0.6835	0.0741	0.0736	0.8772	0.8361
	β_umd	-0.1226	-0.1146	-0.9545	-0.9401	-0.1275	-0.1124	-1.3044	-1.2726	-0.1303	-0.1280	-1.5013	-1.4756
	R^2	0.3154	0.3190			0.4303	0.4402			0.4701	0.4871		
	Adj. R^2	0.3064	0.3100			0.4228	0.4329			0.4632	0.4803		

Table 5.15 continued

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
WLS	**α**	0.0009	0.0008	0.2115	0.2120	0.0016	0.0017	0.5664	0.5701	0.0014	0.0015	0.5705	0.6094
	β_rp	1.0016	1.0055	11.6099	11.6157	0.9977	0.9933	15.9775	16.0808	0.9926	0.9902	18.3831	18.5709
	β_smb	0.4655	0.4723	3.9599	3.9696	0.3032	0.3076	3.5329	3.5847	0.2406	0.2261	3.2454	3.3951
	β_hml	0.0017	0.0066	0.0330	0.0590	0.0528	0.0485	0.6476	0.6069	0.0771	0.0729	1.0914	1.0702
	β_umd	-0.1432	-0.1237	-1.1485	-1.0666	-0.1515	-0.1377	-1.6656	-1.6585	-0.1419	-0.1348	-1.8511	-1.9310
	$\mathbf{R^2}$	0.3504	0.3498			0.4859	0.4894			0.5516	0.5626		
	$\mathbf{Adj.\ R^2}$	0.3418	0.3412			0.4792	0.4827			0.5457	0.5569		
GLS	**α**	0.0009	0.0011	0.2101	0.2668	0.0016	0.0016	0.5571	0.5636	0.0014	0.0016	0.5627	0.6205
	β_rp	0.9999	1.0009	11.4107	11.4159	0.9968	0.9934	15.7438	15.7611	0.9919	0.9861	17.7422	17.8802
	β_smb	0.4702	0.4678	3.9529	3.9059	0.3055	0.3113	3.5077	3.6189	0.2438	0.2327	3.1645	3.3795
	β_hml	0.0051	0.0117	0.0642	0.1029	0.0523	0.0478	0.6329	0.6120	0.0747	0.0753	1.0241	1.1032
	β_umd	-0.1351	-0.1254	-1.0813	-1.0774	-0.1504	-0.1365	-1.6399	-1.5615	-0.1396	-0.1322	-1.7692	-1.7761
	$\mathbf{R^2}$	0.3424	0.3414			0.4782	0.4822			0.5331	0.5448		
	$\mathbf{Adj.\ R^2}$	0.3338	0.3328			0.4713	0.4755			0.5270	0.5388		

Table 5.16 Rejection frequency of intercept tests from calendar-time approach-Winsorized value weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp} RPM_t + \beta_{\_hml} HML_t + \beta_{\_smb} SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp} RPM_t + \beta_{\_hml} HML_t + \beta_{\_smb} SMB_t + \beta_{\_umd} UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected. Winsorization process is taken as winsorizing 1% of stock raw returns of each sample in each tail.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	3.6	9.6	15.2	3.6	2.4	9.6	5.6	15.6	10.4
	WLS	3.2	10.8	17.2	3.2	2.8	10.4	6.8	18.8	11.2
	GLS	3.2	12.4	17.2	3.2	2.8	10.4	6.8	17.2	13.2
Four-factor model	**OLS**	3.2	9.6	15.6	0.8	2.4	6.8	8.8	11.2	14.4
	WLS	2.8	8.8	14.4	1.6	3.2	5.6	8.8	11.6	14.4
	GLS	2.4	8.8	14.0	1.2	3.2	5.2	8.8	10.4	14.4

Table 5.16 continued

		1%	5%	10%	1%		5%		10%	
		α=0	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
		Panel B: Three years								
Three-factor model	**OLS**	0.0	4.8	8.8	0.8	0.4	4.4	4.4	7.6	10.0
	WLS	0.4	7.6	12.0	0.8	1.6	4.8	7.2	11.2	10.0
	GLS	2.0	7.2	10.4	0.0	3.2	0.4	10.0	3.2	22.4
Four-factor model	**OLS**	1.2	5.2	12.4	0.0	1.6	1.6	10.8	3.6	19.6
	WLS	0.4	4.8	11.6	0.4	1.2	5.6	6.0	11.2	10.8
	GLS	1.6	6.0	10.8	0.0	2.8	0.4	10.4	4.0	22.8
		Panel C: Five years								
Three-factor model	**OLS**	0.0	4.8	14.4	1.2	1.2	7.6	6.8	13.6	10.0
	WLS	1.6	6.0	12.8	1.2	0.8	7.2	5.6	12.8	10.4
	GLS	0.8	5.2	11.6	0.8	0.4	6.8	4.8	12.0	10.0
Four-factor model	**OLS**	0.8	4.0	8.8	0.0	2.0	0.8	8.0	4.4	17.6
	WLS	0.0	5.2	12.4	0.0	2.0	0.8	11.6	4.4	21.6
	GLS	0.0	3.6	11.2	0.0	2.0	0.4	10.8	3.2	19.6

Table 5.17 Coefficients and R-squared from three-factor model –Winsorized value weighted portfolios

This table reports coefficients from regression based on Fama-French three-factor model to test the null hypothesis of zero abnormal return which is represented by α when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), and market risk premium(RPM_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. The means and medians of coefficients among 250 samples are calculated for comparison. Winsorization process is taken as winsorizing 1% of stock raw returns of each sample in each tail.

		One year				**Three years**				**Five years**			
		Coefficients		**Test statistics**		**Coefficients**		**Test statistics**		**Coefficients**		**Test statistics**	
		Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
OLS	**α**	-0.0006	-0.0006	-0.1506	-0.1584	0.0001	0.0003	0.0662	0.1201	-0.0001	0.0001	-0.0066	0.0406
	β_rp	0.9906	0.9916	11.3034	11.3181	0.9909	0.9880	15.0820	15.0457	0.9878	0.9845	16.5483	16.7122
	β_smb	0.4466	0.4503	3.7764	3.7406	0.2823	0.2965	3.1589	3.2767	0.2294	0.2265	2.7790	2.9133
	β_hml	0.0767	0.0937	0.7617	0.9212	0.1182	0.1190	1.5386	1.5655	0.1317	0.1409	1.8935	2.1138
	R^2	0.3225	0.3217			0.4392	0.4473			0.4810	0.4915		
	Adj. R^2	0.3159	0.3150			0.4336	0.4418			0.4759	0.4865		

Table 5.17 continued

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
WLS	**α**	-0.0007	-0.0007	-0.1862	-0.1813	-0.0001	0.0002	0.0173	0.0785	-0.0001	0.0001	-0.0093	0.0564
	β_rp	0.9978	0.9999	12.2964	12.3700	0.9975	1.0000	16.8602	16.6551	0.9921	0.9891	19.3884	19.4121
	β_smb	0.4312	0.4419	3.8860	3.8856	0.2708	0.2817	3.3497	3.4295	0.2092	0.2009	2.9963	3.0974
	β_hml	0.0708	0.0784	0.7403	0.8558	0.1224	0.1204	1.7596	1.8406	0.1417	0.1533	2.4281	2.6269
	$\mathbf{R^2}$	0.3567	0.3581			0.4925	0.4904			0.5579	0.5625		
	$\mathbf{Adj.\,R^2}$	0.3503	0.3518			0.4875	0.4854			0.5535	0.5582		
GLS	**α**	-0.0007	-0.0006	-0.1655	-0.1832	-0.0001	0.0002	0.0191	0.0884	-0.0001	0.0001	0.0084	0.0598
	β_rp	0.9950	0.9951	12.0741	12.1400	0.9960	0.9989	16.5629	16.3929	0.9907	0.9856	18.7391	18.7578
	β_smb	0.4358	0.4375	3.8854	3.8500	0.2733	0.2861	3.3289	3.3886	0.2127	0.2092	2.9413	3.1671
	β_hml	0.0711	0.0823	0.7362	0.8765	0.1217	0.1220	1.7206	1.8216	0.1378	0.1532	2.2742	2.4504
	$\mathbf{R^2}$	0.3490	0.3495			0.4838	0.4858			0.5414	0.5450		
	$\mathbf{Adj.\,R^2}$	0.3426	0.3431			0.4787	0.4807			0.5368	0.5405		

Table 5.18 Coefficients and R-squared from four-factor model –Winsorized value weighted portfolios

This table reports coefficients from regression based on Carhart four-factor model to test the null hypothesis of zero abnormal return which is represented by α when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. The means and medians of coefficients among 250 samples are calculated for comparison. Winsorization process is taken as winsorizing 1% of stock raw returns of each sample in each tail.

		One year				**Three years**				**Five years**			
		Coefficients		**Test statistics**		**Coefficients**		**Test statistics**		**Coefficients**		**Test statistics**	
		Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
OLS	**α**	0.0005	0.0005	0.1283	0.1150	0.0014	0.0012	0.4518	0.4003	0.0012	0.0013	0.4326	0.4711
	β_rp	0.9787	0.9866	11.1100	11.1247	0.9780	0.9753	14.8332	14.8613	0.9747	0.9752	16.2764	16.3314
	β_smb	0.4556	0.4605	3.8573	3.8289	0.2921	0.3072	3.2782	3.3945	0.2395	0.2415	2.9140	3.1469
	β_hml	0.0239	0.0467	0.2225	0.3746	0.0613	0.0634	0.7152	0.7442	0.0741	0.0796	0.9429	0.9720
	β_umd	-0.1118	-0.1043	-0.9266	-0.9381	-0.1204	-0.1205	-1.3198	-1.3351	-0.1216	-0.1216	-1.4886	-1.4912
	R^2	0.3305	0.3319			0.4477	0.4522			0.4890	0.5008		
	Adj. R^2	0.3217	0.3231			0.4405	0.4449			0.4823	0.4943		

Table 5.18 continued

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
WLS	**α**	0.0006	0.0007	0.1595	0.1868	0.0014	0.0014	0.5129	0.4912	0.0012	0.0012	0.5242	0.5126
	β_rp	0.9841	0.9847	12.0679	12.1100	0.9817	0.9805	16.5258	16.4218	0.9754	0.9743	18.9414	18.9236
	β_smb	0.4407	0.4468	3.9796	3.9717	0.2830	0.2896	3.5144	3.5133	0.2206	0.2168	3.1806	3.3154
	β_hml	0.0102	0.0204	0.1081	0.1913	0.0569	0.0569	0.7292	0.7556	0.0773	0.0742	1.1610	1.1361
	β_umd	-0.1302	-0.1130	-1.1153	-1.0098	-0.1400	-0.1276	-1.6483	-1.6088	-0.1332	-0.1285	-1.8371	-1.8336
	$\mathbf{R^2}$	0.3654	0.3645			0.5014	0.5024			0.5660	0.5702		
	Adj. $\mathbf{R^2}$	0.3570	0.3561			0.4949	0.4959			0.5603	0.5646		
GLS	**α**	0.0006	0.0010	0.1629	0.2257	0.0014	0.0013	0.5039	0.4408	0.0013	0.0013	0.5175	0.5648
	β_rp	0.9814	0.9883	11.8563	11.9513	0.9808	0.9799	16.2756	16.2499	0.9752	0.9706	18.3364	18.4494
	β_smb	0.4455	0.4423	3.9763	3.9383	0.2852	0.3028	3.4911	3.5987	0.2237	0.2223	3.1096	3.3841
	β_hml	0.0126	0.0262	0.1312	0.2341	0.0563	0.0580	0.7123	0.7721	0.0758	0.0759	1.1016	1.1088
	β_umd	-0.1237	-0.1152	-1.0547	-1.0159	-0.1392	-0.1265	-1.6235	-1.5393	-0.1295	-0.1191	-1.7419	-1.6902
	$\mathbf{R^2}$	0.3572	0.3530			0.4935	0.4981			0.5490	0.5520		
	Adj. $\mathbf{R^2}$	0.3488	0.3446			0.4868	0.4914			0.5431	0.5461		

Table 5.19 Descriptive statistics of alphas from the Fama-French three-factor model based on value-weighted portfolios

This table reports descriptive statistics of alphas derived from the Fama-French three-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp} RPM_t + \beta_{\_hml} HML_t + \beta_{\_smb} SMB_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor (SMB_t), book-to-market ratio factor(HML_t), and market risk premium(RPM_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms.

Three-factor model	**Value-weighted portfolios**				**Winsorized value-weighted portfolios**			
One year	**Mean**	**Median**	**Skewness**	**Kurtosis**	**Mean**	**Median**	**Skewness**	**Kurtosis**
OLS	-0.00045	-0.00017	-0.15	2.87	-0.00064	-0.00060	-0.14	2.98
WLS	-0.00063	-0.00050	-0.19	3.08	-0.00075	-0.00068	-0.10	3.14
GLS	-0.00054	-0.00058	-0.11	2.98	-0.00068	-0.00063	-0.04	3.04
Three years								
OLS	0.00030	0.00040	-0.43	3.40	0.00012	0.00032	-0.34	3.18
WLS	0.00004	0.00042	-0.31	3.44	-0.00007	0.00022	-0.15	3.14
GLS	0.00006	0.00044	-0.33	3.39	-0.00005	0.00023	-0.19	3.14
Five years								
OLS	0.00001	0.00027	-0.24	3.52	-0.00009	0.00011	-0.15	3.05
WLS	-0.00005	0.00028	-0.55	3.84	-0.00011	0.00013	-0.32	2.99
GLS	-0.00001	0.00032	-0.59	3.86	-0.00007	0.00015	-0.39	3.08

Table 5.20 Descriptive statistics of alpha from the Carhart four-factor model based on value-weighted portfolios

This table reports descriptive statistics of alpha derived from the Carhart four-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms.

Four-factor model	**Value-weighted portfolios**				**Winsorized value weighted portfolios**			
One year	**Mean**	**Median**	**Skewness**	**Kurtosis**	**Mean**	**Median**	**Skewness**	**Kurtosis**
OLS	0.00084	0.00085	0.05	3.09	0.00053	0.00050	-0.02	3.21
WLS	0.00086	0.00085	-0.06	2.91	0.00061	0.00068	-0.09	2.96
GLS	0.00087	0.00105	0.05	3.01	0.00063	0.00097	0.00	3.08
Three years								
OLS	0.00163	0.00134	-0.03	3.12	0.00138	0.00124	0.00	3.03
WLS	0.00161	0.00169	-0.07	2.93	0.00138	0.00143	-0.04	2.89
GLS	0.00162	0.00157	-0.06	2.93	0.00138	0.00133	-0.02	2.94
Five years								
OLS	0.00138	0.00157	-0.28	3.04	0.00119	0.00127	-0.20	2.82
WLS	0.00139	0.00149	-0.27	2.72	0.00124	0.00123	-0.22	2.64
GLS	0.00143	0.00158	-0.19	3.44	0.00126	0.00134	-0.28	2.78

Table 5.21 Rejection frequency of all tests in calendar-time approach-Value-weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. 250 samples of 200 firms are randomly selected. Conventional ordinary least squares and three approaches correcting heteroskedasticity including weighted least squares (WLS), sandwich variance estimators (1980) combing OLS and WLS, and generalized least squares (GLS), are applied in both models over three investment horizons. Both one-tail and two-tailed tests are examined at a significance level of 5%. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

				5%			**Mean Alpha**
Holding period	**Model**	**Technique**		**α=0**	**α<0**	**α>0**	
One year	**Three-factor model**	OLS	Original test	10.4	10.0	5.6	-0.0004497
			sandwich variance estimators	10.4	10.0	5.2	-0.0004497
		WLS	Original test	8.8	9.6	6.8	-0.0006334
			sandwich variance estimators	8.0	8.4	6.4	-0.0006334
		GLS	Original test	10.4	10.4	6.8	-0.0005352
	Four-factor model	OLS	Original test	7.2	5.6	8.8	0.0008357
			sandwich variance estimators	6.0	5.2	6.0	0.0008357
		WLS	Original test	8.0	4.0	8.4	0.0008586
			sandwich variance estimators	7.2	3.2	6.8	0.0008586
		GLS	Original test	7.2	3.6	8.4	0.0008721

Table 5.21 continued

				5%			Mean Alpha
Holding period	**Model**	**Technique**		**α=0**	**α<0**	**α>0**	
Three years	**Three-factor model**	OLS	Original test	4.8	4.0	4.8	0.0002991
			sandwich variance estimators	4.8	4.4	4.4	0.0002991
		WLS	Original test	6.8	5.2	7.6	0.0000429
			sandwich variance estimators	6.0	4.0	6.4	0.0000429
		GLS	Original test	4.8	4.4	6.8	0.000061
	Four-factor model	OLS	Original test	4.8	2.0	12.4	0.0016313
			sandwich variance estimators	4.4	1.2	10.4	0.0016313
		WLS	Original test	8.4	0.4	10.8	0.0016109
			sandwich variance estimators	7.2	0.4	10.8	0.0016109
		GLS	Original test	6.8	0.4	10.4	0.0016232
Five years	**Three-factor model**	OLS	Original test	5.6	6.0	6.8	0.0000106
			sandwich variance estimators	5.6	6.0	5.6	0.0000106
		WLS	Original test	5.2	5.6	5.2	-0.0000495
			sandwich variance estimators	5.2	4.4	5.2	-0.0000495
		GLS	Original test	4.0	6.4	5.2	-0.0000113
	Four-factor model	OLS	Original test	3.6	0.8	7.6	0.001376
			sandwich variance estimators	3.2	0.4	7.2	0.001376
		WLS	Original test	5.2	0.8	11.6	0.0013931
			sandwich variance estimators	3.2	0.8	9.2	0.0013931
		GLS	Original test	4.4	0.4	10.4	0.0014268

Table 5.22 Power of test-Value-weighted portfolios (two-tailed test)

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is the study period (12-. 36- and 60 months). The independent variables such as size factor (SMB_t), book-to-market ratio factor (HML_t), market risk premium (RPM_t), and momentum factors (UMD_t) are obtained from the website of Exeter University based on studies done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. The two-tailed tests are examined at significance level of 5%. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected. Abnormal returns in the range of -20% to 20% at an interval of 5% are introduced to individual stocks in order to test the power of test.

		Induced abnormal returns								
		-20%	**-15%**	**-10%**	**-5%**	**0%**	**5%**	**10%**	**15%**	**20%**
		Panel A: one year								
	OLS	94.8	80.8	51.2	22.8	10.4	16.0	46.8	77.2	92.4
Three-factor model	**WLS**	97.6	85.6	59.2	30.0	8.8	21.6	51.6	82.8	96.0
	GLS	96.8	85.2	59.2	26.4	10.4	20.4	50.8	81.6	95.2
	OLS	92.4	71.2	41.6	16.0	7.2	18.8	53.2	82.8	94.4
Four-factor model	**WLS**	94.8	78.4	45.6	17.6	8.0	24.8	63.6	88.0	97.2
	GLS	94.4	78.0	44.4	18.8	7.2	24.0	61.6	86.4	97.6

Table 5.22 continued

		Induced abnormal returns								
		-20%	**-15%**	**-10%**	**-5%**	**0%**	**5%**	**10%**	**15%**	**20%**
		Panel B: three years								
	OLS	36.8	22.8	13.2	5.2	4.8	6.4	16.0	31.6	49.6
Three-factor model	**WLS**	47.2	28.8	15.2	7.6	6.8	9.2	17.6	36.8	53.2
	GLS	46.8	29.2	15.6	7.6	8.4	10.0	17.6	35.6	52.8
	OLS	19.6	10.0	4.8	4.0	4.8	14.0	27.2	42.0	56.4
Four-factor model	**WLS**	24.0	11.2	5.6	4.4	4.8	13.6	32.8	51.6	70.8
	GLS	23.2	10.8	5.2	4.0	6.8	14.8	30.4	51.6	70.4
		Panel C: five years								
	OLS	20.0	16.4	9.6	7.2	5.6	7.6	8.4	13.6	20.8
Three-factor model	**WLS**	28.4	16.4	10.8	6.8	5.2	6.4	11.6	18.8	30.4
	GLS	24.0	14.4	9.6	7.2	4.0	6.4	10.8	17.6	28.4
	OLS	8.4	4.4	2.8	3.6	3.6	8.0	14.8	24.0	36.8
Four-factor model	**WLS**	8.0	4.0	2.8	2.8	5.2	12.0	20.8	36.4	49.6
	GLS	7.6	4.0	2.0	3.2	4.4	10.4	18.0	32.0	46.4

Table 5.23 Mean intercepts in power of test-Value-weighted portfolios

This table reports mean intercepts in power of test based on two-tailed test under value weights. The intercepts are derived from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is the study period (12-. 36- and 60 months). The independent variables such as size factor (SMB_t), book-to-market ratio factor (HML_t), market risk premium (RPM_t), and momentum factors (UMD_t) are obtained from the website of Exeter University based on studies done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. The two-tailed tests are examined at significance level of 5%. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected. Abnormal returns in the range of -20% to 20% at an interval of 5% are introduced to individual stocks in order to test the power of test.

		Induced abnormal returns								
		-20%	**-15%**	**-10%**	**-5%**	**0%**	**5%**	**10%**	**15%**	**20%**
		Panel A: one year								
Three-factor model	**OLS**	-0.01712	-0.01295	-0.00878	-0.00462	-0.00045	0.00372	0.00788	0.01205	0.01622
	WLS	-0.01730	-0.01313	-0.00897	-0.00480	-0.00063	0.00353	0.00770	0.01187	0.01603
	GLS	-0.01720	-0.01304	-0.00887	-0.00470	-0.00054	0.00363	0.00780	0.01196	0.01613
Four-factor model	**OLS**	-0.01583	-0.01166	-0.00750	-0.00333	0.00084	0.00500	0.00917	0.01334	0.01750
	WLS	-0.01581	-0.01164	-0.00747	-0.00331	0.00086	0.00503	0.00919	0.01336	0.01753
	GLS	-0.01579	-0.01163	-0.00746	-0.00329	0.00087	0.00504	0.00921	0.01337	0.01754

Table 5.23 continued

		Induced abnormal returns								
		-20%	**-15%**	**-10%**	**-5%**	**0%**	**5%**	**10%**	**15%**	**20%**
		Panel B: three years								
Three-factor model	**OLS**	-0.00526	-0.00387	-0.00248	-0.00109	0.00030	0.00169	0.00308	0.00447	0.00585
	WLS	-0.00551	-0.00412	-0.00273	-0.00135	0.00004	0.00143	0.00282	0.00421	0.00560
	GLS	-0.00549	-0.00411	-0.00272	-0.00133	0.00006	0.00145	0.00284	0.00423	0.00562
Four-factor model	**OLS**	-0.00392	-0.00254	-0.00115	0.00024	0.00163	0.00302	0.00441	0.00580	0.00719
	WLS	-0.00394	-0.00256	-0.00117	0.00022	0.00161	0.00300	0.00439	0.00578	0.00717
	GLS	-0.00393	-0.00254	-0.00115	0.00023	0.00162	0.00301	0.00440	0.00579	0.00718
		Panel C: five years								
Three-factor model	**OLS**	-0.00332	-0.00249	-0.00166	-0.00082	0.00001	0.00084	0.00168	0.00251	0.00334
	WLS	-0.00338	-0.00255	-0.00172	-0.00088	-0.00005	0.00078	0.00162	0.00245	0.00328
	GLS	-0.00334	-0.00251	-0.00168	-0.00084	-0.00001	0.00082	0.00166	0.00249	0.00332
Four-factor model	**OLS**	-0.00196	-0.00112	-0.00029	0.00054	0.00138	0.00221	0.00304	0.00388	0.00471
	WLS	-0.00194	-0.00111	-0.00027	0.00056	0.00139	0.00223	0.00306	0.00389	0.00473
	GLS	-0.00191	-0.00107	-0.00024	0.00059	0.00143	0.00226	0.00309	0.00393	0.00476

Table 5.24 Power of test-Value-weighted portfolios (one-tailed test)

This table reports percentages of rejection rates on one-tailed test. The alternative hypotheses are positive intercept when positive abnormal returns are induced and negative intercept when negative abnormal returns are induced. The intercepts are derived from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$
$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is the study period (12-. 36- and 60 months). The independent variables such as size factor (SMB_t), book-to-market ratio factor (HML_t), market risk premium (RPM_t), and momentum factors (UMD_t) are obtained from the website of Exeter University based on studies done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including conventional ordinary least squares, weighted least squares which considers different number of firms in each calendar month and generalized least squares which is an alternative to correct heteroskedasticity, are applied in both models over three investment horizons. The two-tailed tests are examined at significance level of 5%. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected. Abnormal returns in the range of -20% to 20% at an interval of 5% are introduced to individual stocks in order to test the power of test.

		-20%	**-15%**	**-10%**	**-5%**	**0%**		**5%**	**10%**	**15%**	**20%**
		α<0				**α<0**	**α>0**	**α>0**			
		Panel A: one year									
Three-factor model	**OLS**	98.4	87.2	62.0	33.6	10.0	5.6	25.2	56.8	85.2	94.8
	WLS	99.2	90.8	68.8	37.6	9.6	6.8	28.8	61.2	87.2	96.8
	GLS	98.4	90.0	70.0	34.8	10.4	6.8	28.4	61.2	87.2	96.8
Four-factor model	**OLS**	95.6	80.8	51.6	22.8	5.6	8.8	30.4	66.4	86.4	96.4
	WLS	96.8	86.4	57.6	22.0	4.0	8.4	36.0	71.6	90.8	98.8
	GLS	96.8	84.8	54.4	20.8	3.6	8.4	33.6	70.8	90.4	98.0

Table 5.24 continued

		-20%	-15%	-10%	-5%	0%		5%	10%	15%	20%
		α<0				α<0	α>0	α>0			
		Panel B: three years									
Three-factor model	**OLS**	49.2	32.4	19.2	10.0	4.0	4.8	12.4	27.6	44.0	61.6
	WLS	58.8	39.2	23.6	13.2	5.2	7.6	13.2	27.2	48.0	64.0
	GLS	58.8	37.6	22.8	12.4	0.4	10.8	13.2	26.0	46.8	64.4
Four-factor model	**OLS**	31.2	16.0	8.0	4.0	2.0	12.4	22.4	38.0	54.0	71.6
	WLS	37.2	18.4	7.2	3.6	4.4	6.8	26.0	46.0	64.8	76.0
	GLS	37.6	17.6	7.6	3.2	0.4	10.4	24.4	45.2	64.8	78.4
		Panel C: five years									
Three-factor model	**OLS**	29.6	20.4	16.8	11.2	6.0	6.8	8.8	13.6	20.8	35.6
	WLS	36.8	27.2	16.4	11.2	5.6	5.2	10.4	18.4	30.4	42.8
	GLS	33.6	23.2	15.6	10.4	6.4	5.2	10.0	16.8	28.0	42.0
Four-factor model	**OLS**	15.6	10.4	5.2	1.6	0.8	7.6	16.0	25.6	38.8	51.6
	WLS	15.6	8.4	4.0	2.8	0.8	11.6	20.4	36.0	49.2	62.4
	GLS	13.2	8.0	4.4	1.6	0.4	10.4	18.8	33.2	45.6	62.4

Table 5.25 Rejection frequency of intercept tests from three-factor model by size

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Both one-tail and two-tailed tests are examined at significance level of 5%. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected. Winsorization process is taken as winsorizing 1% of stock raw returns of each sample in each tail. Firm sizes are deciled according to size rank in September each year. Size 1 has firms with the smallest size whereas size 10 has firms with the largest size.

	One year			**Three years**			**Five years**		
5%	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**
EW sizes									
1	93.2	0.0	96.4	99.6	0.0	100.0	98.8	0.0	99.6
1*	85.6	0.0	93.2	94.4	0.0	98.0	87.2	0.0	93.2
2	57.2	0.0	69.6	94.0	0.0	97.6	96.4	0.0	97.6
3	15.2	0.4	26.4	40.8	0.0	57.6	54.0	0.0	68.4
4	4.4	1.6	10.0	17.2	0.0	24.8	20.8	0.0	40.0
5	4.0	3.6	6.0	6.4	2.0	10.8	12.8	0.0	22.8
6	4.8	4.8	4.8	4.4	2.0	6.4	10.0	0.0	16.8
7	9.2	12.8	2.4	5.2	6.8	2.8	2.4	2.4	2.8
8	8.0	14.4	2.4	2.8	4.0	2.8	3.2	6.0	1.6
9	5.6	8.8	2.0	4.4	7.6	2.8	1.6	3.6	2.8
10	6.4	8.0	2.8	3.6	4.4	3.6	2.8	4.8	1.2
10*	5.6	6.4	4.0	4.4	6.0	2.8	3.6	6.4	2.4

1*winsorized size 1 with the smallest market cap

10*winsorized size 10 with the largest market cap

Table 5.25 continued

	One year			Three years			Five years		
VW sizes	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**
1	80.4	0.0	92.4	98.4	0.0	100.0	94.8	0.0	98.4
1*	73.2	0.0	84.4	92.0	0.0	95.6	76.0	0.0	89.6
2	52.8	0.0	68.0	93.2	0.0	97.2	95.2	0.0	98.0
3	15.2	0.4	24.4	38.0	0.0	52.0	50.4	0.0	67.2
4	3.2	2.0	7.2	14.4	0.4	23.6	19.6	0.0	32.4
5	4.8	3.6	7.2	5.2	2.8	8.0	10.4	0.0	17.6
6	5.6	4.8	5.2	4.0	2.4	4.0	6.0	0.4	14.8
7	7.6	13.2	2.8	2.8	4.8	3.6	2.8	4.4	2.8
8	8.0	15.2	2.4	3.6	5.6	2.0	4.4	6.8	0.4
9	6.0	9.2	2.4	5.2	8.0	3.2	1.6	3.2	3.2
10	8.4	7.6	6.8	3.6	2.8	4.4	6.0	5.2	6.4
10*	7.6	7.6	4.8	6.4	5.6	6.0	7.2	4.8	8.0

1*winsorized size 1 with the smallest market cap

10*winsorized size 10 with the largest market cap

Table 5.26 Rejection frequency of intercept tests from four-factor model by size

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French four-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t) , and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Both one-tail and two-tailed tests are examined at significance level of 5%. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected. Winsorization process is taken as winsorizing 1% of stock raw returns of each sample in each tail. Firm sizes are deciled according to size rank in September each year. Size 1 has firms with the smallest size whereas size 10 has firms with the largest size.

	One year			**Three years**			**Five years**		
5%	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**
EW sizes									
1	81.2	0.0	90.8	94.8	0.0	98.8	91.6	0.0	98.0
1*	70.0	0.0	82.8	87.6	0.0	92.0	66.8	0.0	79.2
2	48.0	0.0	59.6	86.4	0.0	92.4	93.2	0.0	96.0
3	14.8	0.4	22.8	34.4	0.0	45.6	41.6	0.0	57.2
4	3.6	2.0	6.4	16.4	0.0	24.4	24.8	0.0	39.6
5	6.8	2.0	10.0	10.0	0.8	14.0	17.6	0.0	28.8
6	6.0	3.2	7.6	5.6	0.8	11.2	16.0	0.0	24.0
7	8.4	10.4	3.2	3.2	2.4	6.0	4.4	1.6	8.0
8	3.6	7.2	2.8	5.6	2.8	6.8	6.0	0.0	11.6
9	4.4	2.0	9.6	8.8	2.4	15.6	10.0	0.0	20.8
10	9.2	0.8	12.0	13.2	0.0	24.0	18.0	0.0	29.2
10*	10.0	2.0	15.2	14.4	0.8	24.0	14.0	0.4	25.6

1*winsorized size 1 with the smallest market cap

10*winsorized size 10 with the largest market cap

Table 5.26 continued

	One year			Three years			Five years		
5%	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**
VW sizes									
1	70.4	0.0	79.2	88.4	0.0	95.6	82.0	0.0	90.4
1*	54.4	0.0	69.6	79.6	0.0	86.8	50.0	0.0	66.8
2	44.8	0.0	58.8	85.2	0.0	92.4	91.2	0.0	96.0
3	13.6	0.4	22.8	29.2	0.0	42.0	39.2	0.0	53.6
4	2.8	2.8	4.4	13.6	0.4	22.4	22.4	0.0	33.2
5	6.8	1.6	10.4	10.0	1.6	13.2	12.4	0.0	20.8
6	5.6	4.0	7.2	3.6	1.2	9.6	12.4	0.0	21.6
7	6.4	9.6	4.0	4.4	2.0	6.8	4.0	2.4	7.2
8	3.2	6.0	2.8	5.6	2.8	6.8	4.4	1.2	10.4
9	5.6	2.0	8.8	10.0	1.2	15.6	11.6	0.0	24.8
10	8.8	0.8	13.2	9.6	0.0	14.8	14.0	0.0	25.2
10*	7.6	2.0	12.8	11.6	0.0	19.2	13.2	1.2	23.6

1*winsorized size 1 with the smallest market cap

10*winsorized size 10 with the largest market cap

Table 5.27 Rejection frequency of intercept tests for small firms-Equally weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Ten size deciles are grouped according to their market value in September of each year. Small firms are those in size decile 1 which has the smallest market value. 250 samples of 100 events are randomly selected from small firms with replacement. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	70.8	92.4	96.4	0.0	80.4	0.0	96.4	0.0	99.6
	WLS	82.8	95.6	98.4	0.0	88.4	0.0	98.4	0.0	99.2
	GLS	80.4	95.2	97.6	0.0	88.0	0.0	97.6	0.0	99.2
Four-factor model	**OLS**	51.2	82.8	89.2	0.0	64.8	0.0	89.2	0.0	97.2
	WLS	58.8	91.6	96.8	0.0	76.4	0.0	96.8	0.0	98.8
	GLS	53.2	89.6	95.6	0.0	72.4	0.0	95.6	0.0	98.8

Table 5.27 continued

		1%	5%	10%	1%		5%		10%	
		α=0	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
		Panel B: Three years								
Three-factor model	**OLS**	94.4	99.6	100.0	0.0	97.6	0.0	100.0	0.0	100.0
	WLS	95.6	100.0	100.0	0.0	98.8	0.0	100.0	0.0	100.0
	GLS	89.6	99.2	100.0	0.0	94.4	0.0	100.0	0.0	100.0
Four-factor model	**OLS**	81.2	97.2	99.6	0.0	92.4	0.0	99.6	0.0	100.0
	WLS	96.0	100.0	100.0	0.0	98.4	0.0	100.0	0.0	100.0
	GLS	90.4	99.2	100.0	0.0	94.8	0.0	100.0	0.0	100.0
		Panel C: Five years								
Three-factor model	**OLS**	84.0	97.6	99.2	0.0	90.4	0.0	99.2	0.0	100.0
	WLS	98.0	100.0	100.0	0.0	100.0	0.0	100.0	0.0	100.0
	GLS	97.6	100.0	100.0	0.0	100.0	0.0	100.0	0.0	100.0
Four-factor model	**OLS**	68.4	89.6	94.8	0.0	76.4	0.0	94.8	0.0	99.6
	WLS	91.6	100.0	100.0	0.0	96.8	0.0	100.0	0.0	100.0
	GLS	89.2	100.0	100.0	0.0	96.0	0.0	100.0	0.0	100.0

Table 5.28 Coefficients and R-squared from three-factor model for small firms-Equally weighted portfolios

This table reports coefficients from regression based on Fama-French three-factor model to test the null hypothesis of zero abnormal return which is represented by α when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), and market risk premium(RPM_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Ten size deciles are grouped according to their market value in September of each year. Small firms are those in size decile 1 which has the smallest market value. 250 samples of 100 events are randomly selected from small firms with replacement. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

		One year				**Three years**				**Five years**			
		Coefficients		**Test statistics**		**Coefficients**		**Test statistics**		**Coefficients**		**Test statistics**	
Three-factor model		**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
OLS	**α**	0.0226	0.0214	2.6325	2.6302	0.0175	0.0165	3.3386	3.3221	0.0134	0.0134	3.0673	3.1154
	β_rp	0.6437	0.6645	3.6617	3.6959	0.6612	0.6664	5.9297	6.1425	0.6909	0.7017	7.2777	7.4285
	β_smb	1.0502	0.9920	4.1761	4.2013	0.9887	0.9497	6.4654	6.5366	0.8742	0.8616	6.8535	6.9595
	β_hml	-0.2995	-0.2006	-1.2123	-1.0026	-0.3070	-0.2773	-2.2801	-2.3036	-0.2163	-0.1972	-1.9633	-1.8369
	R^2	0.1078	0.1069			0.2164	0.2183			0.2553	0.2597		
	Adj. R^2	0.0990	0.0982			0.2087	0.2106			0.2480	0.2524		

Table 5.28 continued

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
WLS	**α**	0.0216	0.0205	2.8450	2.8396	0.0166	0.0159	3.5861	3.5494	0.0129	0.0127	3.6156	3.6187
	β_rp	0.6259	0.6367	3.9476	3.9569	0.6175	0.6312	6.2918	6.5827	0.6409	0.6523	8.3650	8.4816
	β_smb	1.0433	0.9969	4.6127	4.6868	0.9920	0.9532	7.3931	7.4667	0.8953	0.8829	8.8994	8.9348
	β_hml	-0.2963	-0.2165	-1.3164	-1.1576	-0.3277	-0.2892	-2.7589	-2.7577	-0.2451	-0.2249	-2.8397	-2.7384
	$\mathbf{R^2}$	0.1235	0.1222			0.2548	0.2621			0.3413	0.3426		
	$\mathbf{Adj.R^2}$	0.1148	0.1136			0.2475	0.2548			0.3348	0.3361		
GLS	**α**	0.0217	0.0203	2.7805	2.7813	0.0169	0.0161	3.5801	3.5374	0.0131	0.0131	3.5679	3.5759
	β_rp	0.6279	0.6402	3.8724	3.9353	0.6246	0.6389	6.2465	6.5256	0.6499	0.6590	8.1993	8.3254
	β_smb	1.0440	0.9937	4.4993	4.6168	0.9944	0.9535	7.2639	7.3198	0.8931	0.8816	8.5378	8.5917
	β_hml	-0.2958	-0.2072	-1.2836	-1.1300	-0.3257	-0.2853	-2.6882	-2.6782	-0.2420	-0.2251	-2.6894	-2.6309
	$\mathbf{R^2}$	0.1187	0.1201			0.2490	0.2548			0.3266	0.3312		
	$\mathbf{Adj.R^2}$	0.1100	0.1113			0.2416	0.2473			0.3200	0.3246		

Table 5.29 Coefficients and R-squared from four-factor model for small firms-Equally weighted portfolios

This table reports coefficients from regression based on Carhart four-factor model to test the null hypothesis of zero abnormal return which is represented by α when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), and market risk premium(RPM_t) and momentum factors(UMD_t)are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Ten size deciles are grouped according to their market value in September of each year. Small firms are those in size decile 1 which has the smallest market value. 250 samples of 100 events are randomly selected from small firms with replacement. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
OLS	**α**	0.0245	0.0226	2.6031	2.6146	0.0183	0.0174	3.2207	3.1771	0.0142	0.0139	2.9607	2.9747
	β_rp	0.6424	0.6869	3.5689	3.7335	0.6507	0.6773	5.7813	6.1542	0.6917	0.7022	7.0785	7.4067
	β_smb	1.0348	0.9622	3.9278	3.9735	1.0206	0.9604	6.3776	6.4268	0.9034	0.8870	6.8335	6.9202
	β_hml	-0.2721	-0.1396	-0.8367	-0.6155	-0.2770	-0.2233	-1.6253	-1.6139	-0.1739	-0.1411	-1.2812	-1.1841
	β_umd	0.1439	0.1642	0.6468	0.6878	0.1467	0.1502	0.9753	1.0266	0.1337	0.1341	1.0481	1.0116
	R^2	0.1075	0.1040			0.2181	0.2246			0.2560	0.2640		
	Adj. R^2	0.0957	0.0922			0.2078	0.2144			0.2462	0.2542		

Table 5.29 continued

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
WLS	**α**	0.0234	0.0211	2.7988	2.7314	0.0175	0.0167	3.4727	3.4540	0.0136	0.0133	3.5627	3.5775
	β_rp	0.6189	0.6526	3.8105	3.9562	0.6045	0.6380	6.0984	6.4986	0.6339	0.6483	8.1002	8.3805
	β_smb	1.0267	0.9506	4.3164	4.3343	1.0309	0.9622	7.3195	7.3616	0.9335	0.9144	9.0428	9.1517
	β_hml	-0.2823	-0.1681	-0.9707	-0.7805	-0.3185	-0.2696	-2.1223	-2.1138	-0.2345	-0.2107	-2.2274	-2.1559
	β_umd	0.1315	0.1589	0.6360	0.7155	0.1096	0.1059	0.8267	0.7485	0.0879	0.0957	0.8571	0.8937
	$\mathbf{R^2}$	0.1210	0.1162			0.2567	0.2633			0.3466	0.3535		
	Adj. $\mathbf{R^2}$	0.1094	0.1046			0.2469	0.2536			0.3380	0.3450		
GLS	**α**	0.0236	0.0216	2.7405	2.6671	0.0178	0.0170	3.4497	3.4295	0.0138	0.0135	3.4950	3.5349
	β_rp	0.6242	0.6564	3.7665	3.9145	0.6112	0.6437	6.0568	6.4955	0.6439	0.6582	7.9458	8.2494
	β_smb	1.0260	0.9548	4.2174	4.2557	1.0321	0.9618	7.1752	7.2248	0.9304	0.9115	8.6446	8.6857
	β_hml	-0.2763	-0.1662	-0.9255	-0.7545	-0.3128	-0.2562	-2.0331	-2.0027	-0.2266	-0.2015	-2.0593	-1.9908
	β_umd	0.1369	0.1661	0.6557	0.7076	0.1182	0.1146	0.8721	0.8169	0.0961	0.1040	0.8946	0.9650
	$\mathbf{R^2}$	0.1170	0.1116			0.2503	0.2602			0.3307	0.3363		
	Adj. $\mathbf{R^2}$	0.1054	0.0998			0.2404	0.2504			0.3218	0.3276		

Table 5.30 Rejection frequency of intercept tests for small firms-Value weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Ten size deciles are grouped according to their market value in September of each year. Small firms are those in size decile 1 which has the smallest market value. 250 samples of 100 events are randomly selected from small firms with replacement. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	51.2	84.4	90.4	0.0	66.8	0.0	90.4	0.0	96.4
	WLS	66.0	88.4	93.6	0.0	78.0	0.0	93.6	0.0	98.4
	GLS	62.0	88.0	94.0	0.0	74.8	0.0	94.0	0.0	97.6
Four-factor model	**OLS**	36.0	70.0	81.2	0.0	48.0	0.0	81.2	0.0	91.2
	WLS	43.2	75.6	88.4	0.0	58.8	0.0	88.4	0.0	95.2
	GLS	41.6	73.2	87.2	0.0	56.0	0.0	87.2	0.0	94.4

Table 5.30 continued

		1%	5%	10%	1%		5%		10%	
		α=0	α=0	α=0	α<0	α>0	α<0	α>0	α<0	α>0
		Panel B: Three years								
Three-factor model	**OLS**	86.8	98.4	99.6	0.0	92.4	0.0	99.6	0.0	100.0
	WLS	93.2	99.2	100.0	0.0	96.0	0.0	100.0	0.0	100.0
	GLS	75.6	96.0	98.8	0.0	84.8	0.0	98.8	0.0	99.2
Four-factor model	**OLS**	63.6	91.2	97.2	0.0	78.0	0.0	97.2	0.0	99.2
	WLS	94.0	99.6	100.0	0.0	96.8	0.0	100.0	0.0	100.0
	GLS	76.8	95.6	98.8	0.0	89.6	0.0	98.8	0.0	99.2
		Panel C: Five years								
Three-factor model	**OLS**	74.0	90.8	97.2	0.0	82.0	0.0	97.2	0.0	99.2
	WLS	94.4	100.0	100.0	0.0	99.2	0.0	100.0	0.0	100.0
	GLS	93.2	99.6	100.0	0.0	99.2	0.0	100.0	0.0	100.0
Four-factor model	**OLS**	50.8	80.4	87.6	0.0	62.8	0.0	87.6	0.0	96.8
	WLS	80.0	98.0	100.0	0.0	86.4	0.0	100.0	0.0	100.0
	GLS	76.8	97.2	100.0	0.0	86.0	0.0	100.0	0.0	100.0

Table 5.31 Coefficients and R-squared from three-factor model for small firms-Value weighted portfolios

This table reports coefficients from regression based on Fama-French three-factor model to test the null hypothesis of zero abnormal return which is represented by α when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), and market risk premium(RPM_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Ten size deciles are grouped according to their market value in September of each year. Small firms are those in size decile 1 which has the smallest market value. 250 samples of 100 events are randomly selected from small firms with replacement. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

		One year				**Three years**				**Five years**			
		Coefficients		**Test statistics**		**Coefficients**		**Test statistics**		**Coefficients**		**Test statistics**	
		Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
OLS	**α**	0.0226	0.0214	2.6325	2.6302	0.0175	0.0165	3.3386	3.3221	0.0134	0.0134	3.0673	3.1154
	β_rp	0.6437	0.6645	3.6617	3.6959	0.6612	0.6664	5.9297	6.1425	0.6909	0.7017	7.2777	7.4285
	β_smb	1.0502	0.9920	4.1761	4.2013	0.9887	0.9497	6.4654	6.5366	0.8742	0.8616	6.8535	6.9595
	β_hml	-0.2995	-0.2006	-1.2123	-1.0026	-0.3070	-0.2773	-2.2801	-2.3036	-0.2163	-0.1972	-1.9633	-1.8369
	R^2	0.1078	0.1069			0.2164	0.2183			0.2553	0.2597		
	Adj. R^2	0.0990	0.0982			0.2087	0.2106			0.2480	0.2524		

Table 5.31 continued

		One year				Three years				Five years			
		Coefficients		**Test statistics**		**Coefficients**		**Test statistics**		**Coefficients**		**Test statistics**	
		Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
WLS	**α**	0.0216	0.0205	2.8450	2.8396	0.0166	0.0159	3.5861	3.5494	0.0129	0.0127	3.6156	3.6187
	β_rp	0.6259	0.6367	3.9476	3.9569	0.6175	0.6312	6.2918	6.5827	0.6409	0.6523	8.3650	8.4816
	β_smb	1.0433	0.9969	4.6127	4.6868	0.9920	0.9532	7.3931	7.4667	0.8953	0.8829	8.8994	8.9348
	β_hml	-0.2963	-0.2165	-1.3164	-1.1576	-0.3277	-0.2892	-2.7589	-2.7577	-0.2451	-0.2249	-2.8397	-2.7384
	$\mathbf{R^2}$	0.1235	0.1222			0.2548	0.2621			0.3413	0.3426		
	$\mathbf{Adj.\,R^2}$	0.1148	0.1136			0.2475	0.2548			0.3348	0.3361		
GLS	**α**	0.0217	0.0203	2.7805	2.7813	0.0169	0.0161	3.5801	3.5374	0.0131	0.0131	3.5679	3.5759
	β_rp	0.6279	0.6402	3.8724	3.9353	0.6246	0.6389	6.2465	6.5256	0.6499	0.6590	8.1993	8.3254
	β_smb	1.0440	0.9937	4.4993	4.6168	0.9944	0.9535	7.2639	7.3198	0.8931	0.8816	8.5378	8.5917
	β_hml	-0.2958	-0.2072	-1.2836	-1.1300	-0.3257	-0.2853	-2.6882	-2.6782	-0.2420	-0.2251	-2.6894	-2.6309
	$\mathbf{R^2}$	0.1187	0.1201			0.2490	0.2548			0.3266	0.3312		
	$\mathbf{Adj.\,R^2}$	0.1100	0.1113			0.2416	0.2473			0.3200	0.3246		

Table 5.32 Coefficients and R-squared from four-factor model for small firms-Value weighted portfolios

This table reports coefficients from regression based on Carhart four-factor model to test the null hypothesis of zero abnormal return which is represented by α when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), and market risk premium(RPM_t) and momentum factors(UMD_t)are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Ten size deciles are grouped according to their market value in September of each year. Small firms are those in size decile 1 which has the smallest market value. 250 samples of 100 events are randomly selected from small firms with replacement. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

		One year				**Three years**				**Five years**			
		Coefficients		**Test statistics**		**Coefficients**		**Test statistics**		**Coefficients**		**Test statistics**	
		Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
OLS	**α**	0.0005	0.0005	0.1283	0.1150	0.0014	0.0012	0.4518	0.4003	0.0012	0.0013	0.4326	0.4711
	β_rp	0.9787	0.9866	11.1100	11.1247	0.9780	0.9753	14.8332	14.8613	0.9747	0.9752	16.2764	16.3314
	β_smb	0.4556	0.4605	3.8573	3.8289	0.2921	0.3072	3.2782	3.3945	0.2395	0.2415	2.9140	3.1469
	β_hml	0.0239	0.0467	0.2225	0.3746	0.0613	0.0634	0.7152	0.7442	0.0741	0.0796	0.9429	0.9720
	β_umd	-0.1118	-0.1043	-0.9266	-0.9381	-0.1204	-0.1205	-1.3198	-1.3351	-0.1216	-0.1216	-1.4886	-1.4912
	$\mathbf{R^2}$	0.3305	0.3319			0.4477	0.4522			0.4890	0.5008		
	$\mathbf{Adj. R^2}$	0.3217	0.3231			0.4405	0.4449			0.4823	0.4943		

Table 5.32 continued

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
WLS	**α**	0.0006	0.0007	0.1595	0.1868	0.0014	0.0014	0.5129	0.4912	0.0012	0.0012	0.5242	0.5126
	β_rp	0.9841	0.9847	12.0679	12.1100	0.9817	0.9805	16.5258	16.4218	0.9754	0.9743	18.9414	18.9236
	β_smb	0.4407	0.4468	3.9796	3.9717	0.2830	0.2896	3.5144	3.5133	0.2206	0.2168	3.1806	3.3154
	β_hml	0.0102	0.0204	0.1081	0.1913	0.0569	0.0569	0.7292	0.7556	0.0773	0.0742	1.1610	1.1361
	β_umd	-0.1302	-0.1130	-1.1153	-1.0098	-0.1400	-0.1276	-1.6483	-1.6088	-0.1332	-0.1285	-1.8371	-1.8336
	$\mathbf{R^2}$	0.3654	0.3645			0.5014	0.5024			0.5660	0.5702		
	Adj. $\mathbf{R^2}$	0.3570	0.3561			0.4949	0.4959			0.5603	0.5646		
GLS	**α**	0.0006	0.0010	0.1629	0.2257	0.0014	0.0013	0.5039	0.4408	0.0013	0.0013	0.5175	0.5648
	β_rp	0.9814	0.9883	11.8563	11.9513	0.9808	0.9799	16.2756	16.2499	0.9752	0.9706	18.3364	18.4494
	β_smb	0.4455	0.4423	3.9763	3.9383	0.2852	0.3028	3.4911	3.5987	0.2237	0.2223	3.1096	3.3841
	β_hml	0.0126	0.0262	0.1312	0.2341	0.0563	0.0580	0.7123	0.7721	0.0758	0.0759	1.1016	1.1088
	β_umd	-0.1237	-0.1152	-1.0547	-1.0159	-0.1392	-0.1265	-1.6235	-1.5393	-0.1295	-0.1191	-1.7419	-1.6902
	$\mathbf{R^2}$	0.3572	0.3530			0.4935	0.4981			0.5490	0.5520		
	Adj. $\mathbf{R^2}$	0.3488	0.3446			0.4868	0.4914			0.5431	0.5461		

Table 5.33 Rejection frequency of intercept tests for large firms-Equally weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Ten size deciles are grouped according to their market value in September of each year. Large firms are those in size decile 10 which has the largest market value. 250 samples of 100 events are randomly selected from small firms with replacement. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**					**Panel A: One year**				
Three-factor model	**OLS**	0.8	4.4	10.0	1.2	0.8	8.0	2.0	12.0	6.0
	WLS	0.4	4.0	6.8	0.0	0.4	4.8	2.0	9.6	6.4
	GLS	0.4	3.2	6.8	0.4	0.4	4.8	2.0	9.2	7.2
Four-factor model	**OLS**	1.2	6.0	13.2	0.0	2.4	1.6	11.6	4.0	24.0
	WLS	1.6	6.8	12.4	0.0	2.0	0.4	12.0	0.8	25.6
	GLS	1.6	6.8	12.8	0.0	2.0	0.4	12.4	1.2	24.0

Table 5.33 continued

		1%	5%	10%	1%		5%		10%	
		α=0	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
		Panel B: Three years								
Three-factor model	**OLS**	0.4	2.8	8.4	0.0	0.4	4.8	3.6	9.6	8.4
	WLS	0.0	2.0	6.0	0.0	0.0	4.0	2.0	7.2	6.8
	GLS	4.4	16.8	23.2	0.0	8.8	0.0	23.2	0.0	42.0
Four-factor model	**OLS**	4.0	16.4	29.6	0.0	6.4	0.0	29.6	0.8	43.6
	WLS	0.0	1.6	5.6	0.0	0.0	3.6	2.0	6.8	6.4
	GLS	4.0	16.0	23.6	0.0	8.0	0.0	23.6	0.0	42.0
		Panel C: Five years								
Three-factor model	**OLS**	0.0	3.2	7.2	0.4	0.0	4.0	3.2	10.4	6.0
	WLS	0.0	1.6	3.2	0.4	0.4	2.0	1.2	8.0	3.6
	GLS	0.0	1.2	2.8	0.0	0.4	2.0	0.8	7.2	3.6
Four-factor model	**OLS**	3.6	18.4	32.0	0.0	7.2	0.4	31.6	1.2	48.4
	WLS	2.0	14.0	24.8	0.0	6.4	0.0	24.8	0.0	46.0
	GLS	1.6	12.8	25.2	0.0	4.8	0.0	25.2	0.0	45.6

Table 5.34 Coefficients and R-squared from three-factor model for large firms-Equally weighted portfolios

This table reports coefficients from regression based on Fama-French three-factor model to test the null hypothesis of zero abnormal return which is represented by α when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), and market risk premium(RPM_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Ten size deciles are grouped according to their market value in September of each year. Large firms are those in size decile 10 which has the largest market value. 250 samples of 100 events are randomly selected from small firms with replacement. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
OLS	**α**	-0.0004	-0.0003	-0.1368	-0.1338	-0.0001	-0.0001	-0.0420	-0.0818	-0.0002	-0.0001	-0.0816	-0.0890
	β_rp	1.0613	1.0582	21.6231	21.6127	1.0601	1.0538	30.2943	30.3624	1.0534	1.0518	32.7270	33.2999
	β_smb	0.1848	0.1836	2.7466	2.7361	0.1927	0.1914	4.0748	4.1257	0.1963	0.1936	4.4936	4.4737
	β_hml	0.0823	0.0941	1.4504	1.6337	0.1364	0.1358	3.3688	3.3001	0.1531	0.1537	4.1112	4.1473
	$\mathbf{R^2}$	0.6073	0.6111			0.7499	0.7563			0.7765	0.7890		
	$\mathbf{Adj.\,R^2}$	0.6035	0.6073			0.7475	0.7539			0.7743	0.7869		

Table 5.34 continued

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
WLS	**α**	-0.0002	-0.0002	-0.0907	-0.1033	-0.0001	-0.0002	-0.0934	-0.1443	-0.0002	-0.0002	-0.1394	-0.1571
	β_rp	1.0604	1.0584	24.1636	24.2258	1.0700	1.0705	35.6013	35.5151	1.0615	1.0616	40.1831	40.3660
	β_smb	0.1786	0.1779	2.9471	3.0693	0.1854	0.1830	4.5578	4.5899	0.1782	0.1749	5.0284	4.9838
	β_hml	0.0841	0.0889	1.6158	1.7611	0.1460	0.1471	4.2111	4.2062	0.1667	0.1658	5.5364	5.5030
	$\mathbf{R^2}$	0.6580	0.6611			0.8070	0.8076			0.8421	0.8440		
	$\mathbf{Adj.\,R^2}$	0.6547	0.6578			0.8051	0.8057			0.8406	0.8424		
GLS	**α**	-0.0002	-0.0003	-0.0742	-0.1364	-0.0001	-0.0002	-0.0840	-0.1369	-0.0002	-0.0001	-0.1137	-0.1201
	β_rp	1.0596	1.0576	24.1748	24.3713	1.0685	1.0661	35.2241	35.2404	1.0597	1.0597	39.0030	39.0654
	β_smb	0.1786	0.1780	2.9479	3.0665	0.1852	0.1856	4.5067	4.5114	0.1807	0.1784	4.9494	4.8538
	β_hml	0.0826	0.0923	1.5813	1.7793	0.1453	0.1462	4.1374	4.1449	0.1639	0.1640	5.2825	5.2835
	$\mathbf{R^2}$	0.6579	0.6637			0.8034	0.8053			0.8339	0.8365		
	$\mathbf{Adj.\,R^2}$	0.6545	0.6604			0.8014	0.8034			0.8323	0.8349		

Table 5.35 Rejection frequency of intercept tests for large firms-Value weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Ten size deciles are grouped according to their market value in September of each year. Large firms are those in size decile 10 which has the largest market value. 250 samples of 100 events are randomly selected from small firms with replacement. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	0.8	6.4	12.4	0.8	1.2	6.4	6.0	11.6	9.6
	WLS	1.2	8.0	13.2	0.4	2.0	6.8	6.4	8.8	10.4
	GLS	0.8	7.6	12.4	0.8	1.6	5.6	6.8	8.4	10.4
Four-factor model	**OLS**	1.6	8.0	13.2	1.2	3.6	2.0	11.2	4.0	19.2
	WLS	2.4	6.8	10.0	0.4	4.4	1.2	8.8	3.2	17.6
	GLS	2.4	6.4	9.6	0.0	4.8	1.2	8.4	3.6	17.2

Table 5.35 continued

		1%	5%	10%	1%		5%		10%	
		α=0	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
						Panel B: Three years				
Three-factor model	**OLS**	0.0	4.0	9.2	0.0	0.4	5.6	3.6	12.0	10.4
	WLS	0.0	3.2	7.6	0.4	0.4	4.0	3.6	10.8	10.8
	GLS	1.6	11.6	17.2	0.0	5.2	0.0	17.2	0.4	27.2
Four-factor model	**OLS**	2.0	7.6	18.0	0.0	4.4	0.4	17.6	0.8	31.2
	WLS	0.0	3.6	7.2	0.4	0.0	3.2	4.0	11.2	11.6
	GLS	1.2	11.6	16.8	0.0	6.0	0.0	16.8	0.4	28.4
						Panel C: Five years				
Three-factor model	**OLS**	0.0	3.6	9.6	0.0	1.2	4.4	5.2	9.2	10.0
	WLS	0.0	3.2	6.0	0.0	0.4	2.0	4.0	5.6	8.8
	GLS	0.0	2.0	5.6	0.0	0.4	1.6	4.0	4.8	9.2
Four-factor model	**OLS**	2.4	9.6	18.0	0.0	3.6	0.0	18.0	0.4	36.0
	WLS	0.8	8.0	16.4	0.0	2.4	0.0	16.4	0.0	34.4
	GLS	0.8	7.6	14.8	0.0	3.6	0.0	14.8	0.0	36.0

Table 5.36 Coefficients and R-squared from three-factor model for large firms-Value weighted portfolios

This table reports coefficients from regression based on Fama-French three-factor model to test the null hypothesis of zero abnormal return which is represented by α when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), and market risk premium(RPM_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Ten size deciles are grouped according to their market value in September of each year. Large firms are those in size decile 10 which has the largest market value. 250 samples of 100 events are randomly selected from small firms with replacement. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

		One year				**Three years**				**Five years**			
		Coefficients		**Test statistics**		**Coefficients**		**Test statistics**		**Coefficients**		**Test statistics**	
		Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**
OLS	**α**	-0.0002	0.0000	-0.0379	-0.0152	0.0000	0.0001	0.0419	0.0487	0.0001	0.0002	0.1027	0.1029
	β_rp	1.0045	1.0000	18.4093	18.4239	0.9947	0.9899	24.8008	24.9224	0.9893	0.9837	28.3022	28.4193
	β_smb	0.0282	0.0201	0.2839	0.2761	-0.0066	-0.0130	-0.1921	-0.2429	-0.0087	-0.0123	-0.2936	-0.2910
	β_hml	0.0537	0.0625	0.8698	0.9736	0.1188	0.1131	2.5509	2.3839	0.1196	0.1148	2.9445	2.9654
	$\mathbf{R^2}$	0.5277	0.5315			0.6660	0.6728			0.7190	0.7272		
	Adj. $\mathbf{R^2}$	0.5231	0.5270			0.6627	0.6696			0.7163	0.7246		

Table 5.36 continued

		One year				Three years				Five years			
		Coefficients		Test statistics		Coefficients		Test statistics		Coefficients		Test statistics	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
WLS	**α**	0.0000	0.0000	0.0252	-0.0113	0.0000	0.0000	0.0223	0.0088	0.0000	0.0001	0.0760	0.1005
	β_rp	0.9979	0.9925	20.1382	20.2110	1.0002	0.9974	28.5298	28.8111	0.9934	0.9893	35.0890	35.4853
	β_smb	0.0131	0.0018	0.1076	0.0274	-0.0183	-0.0244	-0.4602	-0.5045	-0.0218	-0.0213	-0.6895	-0.6174
	β_hml	0.0566	0.0636	0.9776	1.1355	0.1264	0.1188	3.1024	3.0602	0.1290	0.1241	3.9672	3.8485
	$\mathbf{R^2}$	0.5715	0.5768			0.7255	0.7317			0.7980	0.8055		
	$\mathbf{Adj.\,R^2}$	0.5673	0.5727			0.7228	0.7290			0.7961	0.8036		
GLS	**α**	0.0000	0.0000	0.0428	-0.0180	0.0000	0.0000	0.0344	0.0178	0.0001	0.0002	0.0931	0.1124
	β_rp	0.9972	0.9899	20.0239	20.0916	0.9988	0.9960	28.3128	28.4868	0.9926	0.9880	34.3228	34.5439
	β_smb	0.0138	0.0015	0.1201	0.0201	-0.0177	-0.0263	-0.4452	-0.5379	-0.0200	-0.0188	-0.6389	-0.4933
	β_hml	0.0549	0.0607	0.9452	1.1072	0.1254	0.1153	3.0453	3.0567	0.1275	0.1226	3.8334	3.7231
	$\mathbf{R^2}$	0.5686	0.5748			0.7222	0.7291			0.7907	0.7976		
	$\mathbf{Adj.\,R^2}$	0.5643	0.5707			0.7194	0.7264			0.7887	0.7956		

Table 5.37 Rejection frequency of intercept tests from three-factor model by book-to-market ratio

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Both one-tail and two-tailed tests are examined at significance level of 5%. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected. Winsorization process is taken as winsorizing 1% of stock raw returns of each sample in each tail. Book-to-market ratios (BTM) of firms are deciled according to BTM rank in September each year. BTM 1 has firms with the lowest BTM whereas BTM 10 has firms with the highest BTM.

	One year			**Three years**			**Five years**		
5%	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**
EW sizes									
1	14.8	21.2	1.2	10.0	14.4	0.0	4.4	7.2	0.8
1*	42.8	51.6	0.0	42.0	56.4	0.0	30.4	42.0	0.0
2	9.6	13.6	1.2	4.4	6.4	3.2	4.8	1.6	7.2
3	3.6	6.4	4.4	3.2	1.6	4.8	5.2	1.2	12.4
4	10.4	4.8	9.6	11.6	0.8	15.2	8.0	0.0	17.2
5	14.0	0.4	22.4	26.0	0.0	42.0	23.6	0.4	39.6
6	10.0	0.8	18.8	26.4	0.0	40.4	44.8	0.0	57.6
7	18.0	0.0	25.6	45.6	0.0	59.2	46.0	0.0	64.0
8	43.2	0.0	60.4	68.4	0.0	76.4	83.2	0.0	88.8
9	48.8	0.0	64.0	95.2	0.0	97.6	88.4	0.0	93.2
10	48.0	0.0	66.8	88.0	0.0	94.4	91.6	0.0	95.2
10*	43.6	0.4	55.2	64.0	0.0	76.8	78.0	0.0	83.2

1*winsorized size 1 with the lowest book-to-market ratio

10*winsorized size 10 with the highest book-to-market ratio

Table 5.37 continued

	One year			Three years			Five years		
5%	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**
VW sizes									
1	22.0	30.4	0.8	17.6	28.4	0.8	10.0	18.8	0.8
1*	22.4	35.6	0.8	16.8	25.2	0.0	12.8	21.2	0.8
2	15.2	22.8	0.0	10.8	16.8	1.2	9.6	17.2	2.4
3	10.4	13.6	3.6	5.2	9.6	0.4	5.6	12.0	0.4
4	7.6	8.8	4.8	2.4	2.0	7.6	3.6	2.0	6.8
5	5.2	4.4	6.4	7.2	4.4	10.8	5.6	2.4	12.0
6	5.2	3.2	7.6	9.6	0.8	18.0	13.2	0.4	22.0
7	12.0	2.0	16	16.8	0.4	23.2	8.0	0.8	15.6
8	11.2	0.4	17.6	13.2	0.0	22.0	14.8	0.0	28.0
9	16.8	0.0	24.0	27.2	0.0	41.6	26.0	0.0	34.4
10	23.2	0.0	33.2	34.8	0.4	46.0	30.4	0.0	45.6
10*	22.4	0.0	35.2	32.4	0.0	42.8	39.2	0.0	53.6

1*winsorized size 1 with the lowest book-to-market ratio

10*winsorized size 10 with the highest book-to-market ratio

Table 5.38 Rejection frequency of intercept tests from four-factor model by book-to-market ratio

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French four-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t) , and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is a proxy of return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. 200 event months are randomly selected with replacement firstly and then 200 firms are randomly selected based on event months. The process is repeated 250 times to achieve 250 samples of 200 firms. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Both one-tail and two-tailed tests are examined at significance level of 5%. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected. Winsorization process is taken as winsorizing 1% of stock raw returns of each sample in each tail. Book-to-market ratios (BTM) of firms are deciled according to BTM rank in September each year. BTM 1 has firms with the lowest BTM whereas BTM 10 has firms with the highest BTM.

	One year			**Three years**			**Five years**		
5%	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**
EW sizes									
1	8.4	13.6	2.0	3.6	7.6	0.8	4.4	2.8	5.2
1*	32.8	41.6	0.0	26.4	36.4	0.0	10.8	18.8	0.0
2	6.0	9.6	3.6	7.6	3.6	10.8	9.6	0.4	16.4
3	3.2	2.4	5.2	4.8	1.2	11.6	12.8	1.2	19.6
4	8.8	4.0	11.6	14.4	0.8	22.4	19.6	0.8	29.2
5	12.8	0.0	24.8	34.4	0.0	49.6	31.6	0.0	51.6
6	9.6	1.2	16.8	29.6	0.0	43.2	48.8	0.0	63.2
7	17.6	0.0	25.6	44.8	0.0	59.2	48.4	0.0	62.0
8	43.6	0.0	54.8	64.4	0.0	76.4	75.6	0.0	85.6
9	42.4	0.0	57.2	90.4	0.0	94.4	80.0	0.0	90.8
10	44.8	0.0	58.8	73.6	0.0	88.8	79.6	0.0	88.4
10*	32.8	0.4	48.4	48.4	0.0	64.0	60.8	0.0	72.0

1*winsorized size 1 with the lowest book-to-market ratio

10*winsorized size 10 with the highest book-to-market ratio

Table 5.38 continued

	One year			Three years			Five years		
5%	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**	**α=0**	**α<0**	**α>0**
VW sizes									
1	8.0	12.4	2.0	6.4	8.4	2.0	4.4	6.8	5.2
1*	11.6	18.0	1.6	2.8	6.8	1.2	3.6	5.2	3.6
2	6.0	10.4	1.6	5.2	5.2	7.6	7.2	4.0	8.4
3	6.4	6.0	8.4	2.8	4.0	5.2	2.0	1.6	2.8
4	6.0	5.6	8.0	7.2	1.6	12.4	5.6	0.8	9.2
5	6.8	2.8	8.4	7.6	0.4	16.8	12.8	0.8	18.8
6	4.0	1.6	5.6	13.6	0.4	21.2	12.8	0.4	24.0
7	10.0	0.8	18.0	17.6	0.0	32.4	12.8	0.0	22.8
8	11.2	0.4	18.8	12.0	0.0	21.6	16.4	0.0	26.4
9	10.8	0.0	18.4	25.2	0.0	37.2	20.4	0.0	30.0
10	19.2	0.0	31.6	20.0	0.4	35.2	21.2	0.0	35.2
10*	16.8	0.0	27.2	19.6	0.0	29.6	23.6	0.0	35.6

1*winsorized size 1 with the lowest book-to-market ratio

10*winsorized size 10 with the highest book-to-market ratio

Table 5.39 Rejection frequency of intercept tests for firms with low book-to-market ratio (BTM)-Equally weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Book-to-market ratios (BTM) of firms are deciled according to BTM rank in September each year. Decile 1, which has firms with the lowest BTM are examined. 250 samples of 100 events are randomly selected from small firms with replacement. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	3.2	14.8	28.0	6.0	0.0	27.6	0.4	40.0	1.2
	WLS	2.8	16.8	28.0	6.4	0.0	28.0	0.0	39.6	1.6
	GLS	2.8	16.4	28.8	6.0	0.0	28.8	0.0	42.4	0.4
Four-factor model	**OLS**	2.4	9.2	15.6	4.8	0.0	14.8	0.8	26.0	1.2
	WLS	1.6	5.6	11.6	2.8	0.0	10.8	0.8	22.0	4.4
	GLS	1.6	4.4	12.4	2.4	0.0	12.0	0.4	22.8	3.2

Table 5.39 continued

		1%	5%	10%	1%		5%		10%	
		α=0	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
						Panel B: Three years				
Three-factor model	**OLS**	4.8	10.8	18.0	6.0	0.0	17.2	0.8	28.4	1.2
	WLS	2.8	14.0	21.2	6.0	0.0	21.2	0.0	32.0	0.8
	GLS	0.8	2.8	6.4	0.8	0.4	3.6	2.8	12.0	6.0
Four-factor model	**OLS**	1.2	7.6	14.4	2.4	0.4	12.4	2.0	15.6	3.6
	WLS	3.2	12.0	20.0	5.6	0.0	20.0	0.0	31.2	0.8
	GLS	1.2	2.8	6.0	0.8	0.4	4.4	1.6	11.6	5.2
						Panel C: Five years				
Three-factor model	**OLS**	0.0	1.6	5.2	0.4	0.0	4.0	1.2	11.2	4.0
	WLS	0.4	2.0	3.6	0.8	0.0	3.2	0.4	8.4	1.6
	GLS	0.4	1.2	3.6	0.4	0.0	3.2	0.4	8.4	1.6
Four-factor model	**OLS**	0.0	2.8	9.2	0.0	0.0	0.8	8.4	1.6	14.4
	WLS	0.4	4.0	8.0	0.0	0.8	0.4	7.6	2.0	17.2
	GLS	0.0	3.2	6.8	0.0	0.8	0.4	6.4	1.2	14.0

Table 5.40 Rejection frequency of intercept tests for firms with low book-to-market ratio (BTM)-Value weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp} RPM_t + \beta_{\_hml} HML_t + \beta_{\_smb} SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp} RPM_t + \beta_{\_hml} HML_t + \beta_{\_smb} SMB_t + \beta_{\_umd} UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Book-to-market ratios (BTM) of firms are deciled according to BTM rank in September each year. Decile 1, which has firms with the lowest BTM are examined. 250 samples of 100 events are randomly selected from small firms with replacement. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	10.8	24.0	37.2	17.6	0.0	37.2	0.0	55.2	0.4
	WLS	13.2	26.8	39.6	18.0	0.0	39.6	0.0	51.6	0.0
	GLS	12.8	25.6	38.4	17.6	0.0	38.4	0.0	52.4	0.4
Four-factor model	**OLS**	4.0	12.0	19.6	6.8	0.0	18.8	0.8	30.8	2.4
	WLS	2.8	12.0	17.2	4.4	0.4	16.4	0.8	28.0	2.0
	GLS	2.8	11.2	17.6	2.8	0.4	17.2	0.4	29.2	1.6

Table 5.40 continued

		1%	5%	10%	1%		5%		10%	
		α=0	α=0	α=0	α<0	α>0	α<0	α>0	α<0	α>0
		Panel B: Three years								
Three-factor model	**OLS**	10.4	24.4	33.2	16.4	0.0	33.2	0.0	49.2	0.0
	WLS	9.6	24.4	34.8	14.8	0.0	34.4	0.4	47.6	1.2
	GLS	0.4	4.4	10.4	1.6	0.0	7.2	3.2	16.4	7.6
Four-factor model	**OLS**	2.0	8.4	15.6	4.0	0.0	13.6	2.0	22.0	4.4
	WLS	9.6	26.4	32.8	14.8	0.0	32.8	0.0	46.4	0.4
	GLS	0.8	4.0	10.8	1.6	0.0	8.4	2.4	17.6	7.2
		Panel C: Five years								
Three-factor model	**OLS**	2.4	10.8	18.8	4.0	0.0	18.4	0.4	32.0	0.4
	WLS	2.0	13.6	22.4	6.8	0.4	22.0	0.4	31.6	1.2
	GLS	1.6	11.6	21.2	4.8	0.0	20.8	0.4	30.4	0.4
Four-factor model	**OLS**	0.0	2.4	6.0	0.4	0.0	3.6	2.4	8.0	6.4
	WLS	0.0	3.6	5.2	0.4	0.8	2.0	3.2	5.2	8.8
	GLS	0.0	3.2	5.2	0.0	0.0	2.4	2.8	4.8	6.4

Table 5.41 Rejection frequency of intercept tests for firms with high book-to-market ratio (BTM)-Equally weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Book-to-market ratios (BTM) of firms are deciled according to BTM rank in September each year. Decile 10, which has firms with the highest BTM are examined. 250 samples of 100 events are randomly selected from small firms with replacement. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	23.6	53.2	67.2	0.0	34.0	0.0	67.2	0.0	83.6
	WLS	29.6	61.6	78.4	0.0	41.2	0.0	78.4	0.0	88.8
	GLS	28.8	59.2	76.4	0.0	41.2	0.0	76.4	0.0	85.6
Four-factor model	**OLS**	14.8	45.6	59.6	0.0	25.6	0.0	59.6	0.0	74.0
	WLS	18.4	52.4	64.8	0.0	31.6	0.0	64.8	0.0	82.8
	GLS	20.0	50.4	67.2	0.0	29.6	0.0	67.2	0.0	80.4

Table 5.41 continued

		1%	5%	10%	1%		5%		10%	
		α=0	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
		Panel B: Three years								
Three-factor model	**OLS**	54.8	82.8	89.6	0.0	69.6	0.0	89.6	0.0	98.0
	WLS	73.6	95.6	98.8	0.0	88.0	0.0	98.8	0.0	99.2
	GLS	51.2	85.2	94.8	0.0	69.2	0.0	94.8	0.0	96.8
Four-factor model	**OLS**	38.4	73.2	82.0	0.0	53.6	0.0	82.0	0.0	90.8
	WLS	73.2	95.6	99.2	0.0	86.8	0.0	99.2	0.0	99.6
	GLS	50.4	86.0	94.4	0.0	70.0	0.0	94.4	0.0	97.2
		Panel C: Five years								
Three-factor model	**OLS**	71.6	91.2	95.6	0.0	80.0	0.0	95.6	0.0	99.2
	WLS	92.4	99.6	100.0	0.0	97.6	0.0	100.0	0.0	100.0
	GLS	91.6	99.6	100.0	0.0	96.4	0.0	100.0	0.0	100.0
Four-factor model	**OLS**	49.6	79.6	88.4	0.0	61.2	0.0	88.4	0.0	95.2
	WLS	82.0	97.6	99.6	0.0	89.6	0.0	99.6	0.0	100.0
	GLS	78.4	95.6	99.2	0.0	88.4	0.0	99.2	0.0	100.0

Table 5.42 Rejection frequency of intercept tests for firms with high book-to-market ratio (BTM)-Value weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Book-to-market ratios (BTM) of firms are deciled according to BTM rank in September each year. Decile 10, which has firms with the highest BTM are examined. 250 samples of 100 events are randomly selected from small firms with replacement. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level. Moreover, symmetrical rejection rates for both sides are expected.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	9.6	26.0	38.0	0.0	11.2	0.0	38	0.4	53.2
	WLS	11.2	28.0	38.8	0.0	19.2	0.0	38.8	0.8	54.8
	GLS	10.8	27.6	40.4	0.0	17.6	0.0	40.4	0.4	54.4
Four-factor model	**OLS**	7.2	18.0	28.8	0.0	10.8	0.0	28.8	0.0	46.4
	WLS	7.6	20.4	29.2	0.0	13.6	0.0	29.2	0.0	45.6
	GLS	7.2	21.6	28.4	0.0	12.8	0.0	28.4	0.0	46.8

Table 5.42 continued

		1%	5%	10%	1%		5%		10%	
		α=0	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
					Panel B: Three years					
Three-factor model	**OLS**	12.0	28.8	44.4	0.0	16.0	0.0	44.4	0.0	65.2
	WLS	13.2	33.6	48.0	0.0	23.6	0.0	48.0	0.0	66.4
	GLS	5.2	21.6	30.8	0.0	11.6	0.0	30.8	0.0	49.2
Four-factor model	**OLS**	5.2	16.8	28.0	0.0	10.0	0.0	28.0	0.0	45.6
	WLS	13.2	33.6	48.0	0.0	22.0	0.0	48.0	0.0	66.4
	GLS	4.4	20.4	28.8	0.0	11.2	0.0	28.8	0.0	48.0
					Panel C: Five years					
Three-factor model	**OLS**	12.4	34.4	48.4	0.0	18.0	0.0	48.4	0.0	66.0
	WLS	16.8	37.6	50.0	0.0	23.2	0.0	50.0	0.0	64.0
	GLS	16.8	36.8	52.4	0.0	24.4	0.0	52.4	0.0	66.4
Four-factor model	**OLS**	5.6	21.6	32.0	0.0	8.8	0.0	32.0	0.0	49.2
	WLS	8.0	21.6	34.4	0.0	11.6	0.0	34.4	0.0	49.2
	GLS	8.0	22.0	34.8	0.0	11.6	0.0	34.8	0.0	52.4

Table 5.43 Rejection frequency of intercept tests for firms clustering in industry-Equal weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Firms are categorized into ten industries. Firstly, one industry is selected randomly out of ten. Then 200 firms in the same industry are selected randomly with 200 event months. To avoid overlapping issue, one firm is not allowed to have a multiple event in the last or the latter T-1 months. T is denoted as the event window. The same process is repeated for 250 times. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	1.2	10.0	16.4	0.0	4.0	0.8	15.6	2.0	25.2
	WLS	2.0	10.4	17.2	0.0	4.4	1.6	15.6	3.6	27.2
	GLS	2.0	10.4	15.6	0.0	3.2	1.2	14.4	2.8	27.6
Four-factor model	**OLS**	2.8	8.0	18	0.0	3.6	0.8	17.2	1.6	28.8
	WLS	3.2	11.6	19.2	0.0	4.8	1.6	17.6	2.4	32.4
	GLS	2.8	10.0	20	0.0	4.4	1.2	18.8	2.8	32.4

Table 5.43 continued

		1%	5%	10%	1%		5%		10%	
		α=0	α=0	α=0	α<0	α>0	α<0	α>0	α<0	α>0
		Panel B: Three years								
Three-factor model	**OLS**	7.2	16.4	28.8	0.4	9.6	0.4	28.4	0.4	47.6
	WLS	7.2	20.4	35.6	0.0	13.2	0.0	35.6	0.4	49.6
	GLS	12.8	33.2	51.6	0.0	18.4	0.0	51.6	0.0	67.2
Four-factor model	**OLS**	6.4	19.2	29.2	0.4	10.8	0.4	28.8	0.8	50.0
	WLS	6.4	20.0	34.8	0.0	10.4	0.4	34.4	0.4	50.8
	GLS	8.8	27.6	45.2	0.0	16.4	0.0	45.2	0.4	62.4
		Panel C: Five years								
Three-factor model	**OLS**	8.4	29.6	42.8	0.0	13.6	0.0	42.8	0.4	61.6
	WLS	19.2	40.4	58.4	0.0	25.6	0.0	58.4	0.0	76.0
	GLS	12.4	36.0	53.6	0.0	19.2	0.0	53.6	0.0	72.0
Four-factor model	**OLS**	9.6	35.6	50.4	0.0	18.4	0.0	50.4	0.8	68.0
	WLS	32.8	64.4	80.8	0.0	43.6	0.0	80.8	0.0	90.4
	GLS	20.0	52.0	72.0	0.0	32.8	0.0	72.0	0.0	87.6

Table 5.44 Rejection frequency of intercept tests for firms clustering in industry-Value weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. Firms are categorized into ten industries. Firstly, one industry is selected randomly out of ten. Then 200 firms in the same industry are selected randomly with 200 event months. To avoid overlapping issue, one firm is not allowed to have a multiple event in the last or the latter T-1 months. T is denoted as the event window. The same process is repeated for 250 times. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	1.6	4.4	11.6	0.8	2.4	5.2	6.4	8.4	11.6
	WLS	0.8	6.0	10.4	1.2	1.2	6.0	4.4	13.6	9.6
	GLS	1.2	6.0	9.2	0.4	1.6	4.8	4.4	10.4	10.8
Four-factor model	**OLS**	2.4	6.8	12.4	0.4	2.4	2.0	10.4	4.8	16.4
	WLS	1.2	6.4	12.8	0.8	2.0	2.4	10.4	4.8	17.2
	GLS	1.2	7.2	11.2	0.4	2.0	2.0	9.2	5.2	15.6

Table 5.44 continued

		1%	5%	10%	1%		5%		10%	
		α=0	α=0	α=0	α<0	α>0	α<0	α>0	α<0	α>0
		Panel B: Three years								
Three-factor model	**OLS**	1.2	3.2	7.2	1.2	0.4	5.6	1.6	12.0	4.8
	WLS	0.0	5.2	12.8	1.6	0.0	12.4	0.4	20.8	3.2
	GLS	0.0	2.0	5.6	0.0	0.8	0.0	5.6	2.0	13.2
Four-factor model	**OLS**	0.4	2.4	6.0	0.0	0.8	1.2	4.8	4.0	14.0
	WLS	0.4	2.0	6.8	0.8	0.0	6.4	0.4	14.0	3.6
	GLS	0.0	1.6	5.2	0.0	0.0	0.0	5.2	1.2.0	11.2
		Panel C: Five years								
Three-factor model	**OLS**	0.8	2.4	7.6	0.4	0.4	5.6	2.0	10.0	5.2
	WLS	0.4	6.8	11.6	2.0	0.4	10.0	1.6	22.0	2.8
	GLS	0.0	1.2	5.2	0.4	0.0	3.2	2.0	10.0	3.6
Four-factor model	**OLS**	0.0	3.2	9.6	0.4	0.4	0.8	8.8	2.8	14.0
	WLS	0.0	3.2	6.8	0.0	0.4	0.0	6.8	2.0	14.8
	GLS	0.0	2.8	5.6	0.0	0.0	0.8	4.8	0.8	15.2

Table 5.45 Rejection frequency of intercept tests for firms clustering in calendar month-Equally weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. The simulation process is to firstly choose one month randomly from event window. Then 200 firms with the same event month are selected randomly without replacement. The same process is repeated for 250 times. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	12.0	23.2	32.8	2.4	13.6	5.2	27.6	7.6	35.6
	WLS	12.0	24.4	32.8	2.4	13.6	5.2	27.6	6.8	35.6
	GLS	17.6	28.4	37.6	3.6	18.4	6.8	30.8	10.0	36.4
Four-factor model	**OLS**	12.4	23.6	31.2	2.8	12.0	6.8	24.4	9.6	32.0
	WLS	12.8	23.2	31.6	2.8	12.0	7.2	24.4	9.6	32.4
	GLS	15.2	26.4	36.4	4.0	14.4	9.6	26.8	12.4	34.0

Table 5.45 continued

		1%	5%	10%	1%		5%		10%	
		α=0	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
					Panel B: Three years					
Three-factor model	**OLS**	20.0	38.0	51.2	0.4	26.0	3.2	48.0	4.4	57.6
	WLS	20.0	38.8	50.8	0.8	26.0	2.8	48.0	4.4	58.0
	GLS	19.2	39.6	49.2	0.4	26.0	3.2	46.0	4.4	57.2
Four-factor model	**OLS**	18.8	39.6	49.2	0.4	26.0	3.2	46.0	4.0	57.2
	WLS	22.8	40.0	53.6	0.8	28.0	4.4	49.2	6.8	59.2
	GLS	19.6	38.8	50.8	1.2	27.2	3.6	47.2	5.6	58.0
					Panel C: Five years					
Three-factor model	**OLS**	21.6	42.8	56	0.0	32.4	0.4	55.6	0.8	65.2
	WLS	23.2	40.8	56.8	0.0	31.2	0.4	56.4	0.8	65.2
	GLS	24.8	41.2	54.4	0.0	31.6	0.4	54.0	2.0	66.4
Four-factor model	**OLS**	28.4	46.0	56.8	0.0	35.2	0.4	56.4	1.2	65.6
	WLS	28.0	45.6	55.6	0.0	35.2	0.4	55.2	1.2	66.4
	GLS	27.6	44.8	59.2	0.0	32.8	0.8	58.4	1.6	66.0

Table 5.46 Rejection frequency of intercept tests for firms clustering in calendar month-Value weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. The simulation process is to firstly choose one month randomly from event window. Then 200 firms with the same event month are selected randomly without replacement. The same process is repeated for 250 times. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	1.6	6.8	11.2	1.2	1.2	3.2	8.0	6.4	13.2
	WLS	1.6	6.4	11.2	1.2	1.2	3.2	8.0	6.0	13.2
	GLS	4.0	10.4	17.2	1.2	3.6	4.0	13.2	7.6	17.2
Four-factor model	**OLS**	2.0	6.8	12.4	0.4	2.0	3.2	9.2	5.6	13.6
	WLS	1.6	6.4	12.0	0.4	2.0	3.2	8.8	5.6	14.0
	GLS	4.4	11.6	18.4	2.4	4.8	4.4	14.0	7.6	19.6

Table 5.46 continued

		1%	5%	10%	1%		5%		10%	
		α=0	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
		Panel B: Three years								
Three-factor model	**OLS**	2.8	8.8	14.8	0.0	4.0	2.4	12.4	4.8	22.4
	WLS	2.8	8.8	15.2	0.0	4.0	2.4	12.8	4.8	22.0
	GLS	2.4	7.6	14.4	0.4	4.0	0.8	13.6	2.4	25.2
Four-factor model	**OLS**	2.4	7.2	15.6	0.4	3.6	0.8	14.8	2.4	23.6
	WLS	3.6	11.6	16.4	0.0	4.8	2.8	13.6	7.6	22.0
	GLS	3.2	10.0	16.0	0.4	5.2	0.8	15.2	4.0	24.8
		Panel C: Five years								
Three-factor model	**OLS**	1.6	6.0	9.2	0.0	2.0	2.4	6.8	6.4	15.2
	WLS	1.6	4.8	8.8	0.0	2.0	2.4	6.4	5.6	14.0
	GLS	1.6	5.6	9.2	0.4	2.8	3.2	6.0	6.8	15.2
Four-factor model	**OLS**	1.6	5.2	10.0	0.0	2.0	0.0	10.0	0.4	22.0
	WLS	1.6	4.0	8.0	0.0	2.0	0.0	8.0	0.4	21.2
	GLS	1.6	5.2	9.6	0.0	2.8	0.0	9.6	0.8	23.6

Table 5.47 Rejection frequency of intercept tests for firms with overlapping returns-Equally weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are equally weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. The simulation process is to firstly choose 100 months with 100 event firms randomly from event window. Then with the same firm, another event month within the last or latter T-1 is selected randomly. The same process is repeated for 250 times. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	0.4	7.2	11.6	0.8	1.6	2.4	9.2	6.8	18
	WLS	1.6	7.2	13.2	2.0	0.4	6.4	6.8	13.6	10.8
	GLS	0.4	6.8	12.0	0.8	0.8	3.2	8.8	6.8	14.4
Four-factor model	**OLS**	0.4	5.6	10.8	0.8	0.8	2.0	8.8	4.4	16.8
	WLS	2.8	6.0	12.4	3.2	0.4	9.6	2.8	18.0	8.0
	GLS	0.8	5.6	9.6	1.2	0.4	3.2	6.4	6.4	12.8

Table 5.47 continued

		1%	5%	10%	1%		5%		10%	
		α=0	α=0	α=0	α<0	α>0	α<0	α>0	α<0	α>0
		Panel B: Three years								
Three-factor model	**OLS**	1.2	12.8	20.4	0.4	5.2	0.8	19.6	2.0	33.6
	WLS	1.2	8.8	17.6	0.4	3.2	2.0	15.6	3.2	24.8
	GLS	0.8	9.2	15.6	0.4	3.6	0.4	15.2	1.6	25.2
Four-factor model	**OLS**	0.8	10.4	20.4	0.0	4.4	0.4	20.0	0.8	30.4
	WLS	1.6	10.0	20.0	0.4	4.0	0.8	19.2	2.0	30.4
	GLS	1.6	10.0	16.8	0.4	3.6	0.4	16.4	0.8	28.8
		Panel C: Five years								
Three-factor model	**OLS**	4.4	20.4	28.4	0.0	9.6	0.0	28.4	0.0	45.6
	WLS	11.6	31.2	45.6	0.0	16.0	0.0	45.6	0.0	65.2
	GLS	6.8	24.0	38.8	0.0	11.6	0.0	38.8	0.0	54.4
Four-factor model	**OLS**	5.2	16.0	28.0	0.0	8.8	0.0	28.0	0.4	45.2
	WLS	14.8	36.0	54.0	0.0	24.4	0.0	54.0	0.0	70.0
	GLS	8.8	24.4	40.0	0.0	13.6	0.0	40.0	0.0	58.4

Table 5.48 Rejection frequency of intercept tests for with overlapping returns -Value weighted portfolios

This table reports percentages of rejection rates of the null hypothesis of zero intercept, which also means zero abnormal return, from regression based on Fama-French three-factor model and Carhart four-factor model when portfolio returns are value weighted.

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \varepsilon_t$$

$$R_{p,t} - R_{f,t} = \alpha + \beta_{\_rp}RPM_t + \beta_{\_hml}HML_t + \beta_{\_smb}SMB_t + \beta_{\_umd}UMD_t + \varepsilon_t$$

Portfolio returns in calendar months from October 1982 to September 2008 are dependent variables. A firm including in the portfolio in a calendar month is required to experience an event in the last n month prior to the calendar month. N is denoted as the study period which is 12-. 36- and 60 months. The independent variables such as size factor(SMB_t), book-to-market ratio factor(HML_t), market risk premium(RPM_t), and momentum factors(UMD_t) are obtained directly from website of Exeter University based on the work done by Gregory, Tharyan and Huang (2009). The risk free rate $R_{f,t}$ is return of three-month Treasury bill. Missing returns and returns of delisted firms are kept as it is. If there is no firm in a calendar month, the month is dropped out from the time series. Three techniques including ordinary least squares, weighted least squares and generalized least squares are applied in both models over three investment horizons. The simulation process is to firstly choose 100 months with 100 event firms randomly from event window. Then with the same firm, another event month within the last or latter T-1 is selected randomly. The same process is repeated for 250 times. Both one-tail and two-tailed tests are examined at significance level of 1%, 5% and 10%, respectively. The overall rejection rates among 250 samples are accumulated to compute the rejection frequency. If the model is well specified, this figure is expected to be the same as the significance level.

Significance level		**1%**	**5%**	**10%**	**1%**		**5%**		**10%**	
Null hypothesis		**α=0**	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
Models	**Techniques**	**Panel A: One year**								
Three-factor model	**OLS**	2.8	6.0	12.4	2.4	0.8	7.2	5.2	8.8	10.4
	WLS	5.6	11.2	19.6	5.6	2.0	15.2	4.4	22.0	10.0
	GLS	3.2	5.2	11.6	2.8	1.6	8.4	3.2	10.8	8.0
Four-factor model	**OLS**	2.4	5.6	11.2	2.0	0.8	5.6	5.6	10.4	10.4
	WLS	6.0	16.4	23.2	7.2	1.2	18.4	4.8	27.2	8.0
	GLS	2.4	4.8	14.0	2.4	0.4	9.2	4.8	12.8	10.0

Table 5.48 continued

		1%	5%	10%	1%		5%		10%	
		α=0	**α=0**	**α=0**	**α<0**	**α>0**	**α<0**	**α>0**	**α<0**	**α>0**
		Panel B: Three years								
Three-factor model	**OLS**	5.2	9.6	14.8	0.8	5.2	3.6	11.2	8.4	18.0
	WLS	4.4	11.6	17.6	1.6	4.8	5.6	12.0	10.0	18.4
	GLS	3.6	10.4	14.0	0.8	4.8	2.8	11.2	6.0	19.6
Four-factor model	**OLS**	3.2	9.6	15.6	0.4	4.4	2.4	13.2	3.6	20.0
	WLS	4.4	10.0	16.0	0.8	4.8	4.0	12.0	9.6	16.0
	GLS	3.2	8.4	14.8	0.4	4.4	2.0	12.8	4.4	19.2
		Panel C: Five years								
Three-factor model	**OLS**	2.4	8.4	14.4	0.8	3.6	2.4	12	4.4	21.2
	WLS	4.8	13.2	20.0	0.4	7.2	1.2	18.8	4.4	27.6
	GLS	3.2	10.8	17.6	0.4	5.6	2.0	15.6	4.8	25.2
Four-factor model	**OLS**	1.6	10.4	17.2	0.0	2.4	1.2	16.0	2.8	27.2
	WLS	4.4	14.4	23.6	0.0	7.6	0.4	23.2	1.6	37.2
	GLS	2.0	11.6	21.6	0.0	4.8	0.8	20.8	2.0	32.8

Table 5.49 Rejection frequency of the null hypothesis of zero mean monthly calendar-time abnormal return (MMAR)

This table reports rejection frequency of 250 samples with the null hypothesis of zero mean monthly calendar-time abnormal return (MMAR) based on both equal weighting and value-weighted schemes. MMAR is another variant of calendar-time approach which combines reference portfolios and calendar time. Instead of running regression based on factor models, abnormal return of an individual stock in each calendar month is defined as the difference of stock return and benchmark return as $AR_{i,t} = R_{i,t} - R_{b,t}$. The benchmark is based on reference portfolios which are matched with the event firm by similar characteristics such as size, book-to-market ratio, size and book-to-market ratio, equally-weighted market return. In each calendar month, the monthly portfolio return is cumulated either with equally-weighted scheme or value-weighted scheme: $CTAR_t = \sum_{i=1}^{n} w_{it} * AR_{it}$. Eventually, a time series of monthly portfolio returns is achieved over the study period. The grand mean abnormal return is computed as $MCTAR = \sum_{t=1}^{T} CTAR_t / T$. The conventional cross-sectional test statistics is calculated as: $t = \frac{MCTAR}{\sigma(MCTAR_t)/\sqrt{T}}$. 250 samples of 200 event months are randomly selected with replacement. Both two-tailed and one-tailed tests are examined at a significance level of 5%. The missing returns are kept as they are without replacement. The rejection frequency in 250 samples is computed. If the test is well-specified, the rejection frequency is expected to 5% at a significance level of 5%. Furthermore, the distribution of abnormal returns is expected to be symmetric with equal rejection rates of 5% on both sides.

	Equally-weighted portfolio return				Value-weighted portfolio return			
Benchmark criteria	5%			Mean AR	5%			Mean AR
	α=0	α<0	α>0		α=0	α<0	α>0	
	Panel A: One year							
By size	6.0	7.2	2.8	-0.00039	6.4	6.4	4.8	-0.00021
By BTM	3.6	5.6	1.2	-0.00018	6.4	10.8	2.0	-0.00178
By size and BTM	2.4	5.2	1.2	-0.00001	4.0	6.8	4.0	-0.00025
	Panel B: Three years							
By size	3.6	6.4	2.0	-0.00004	4.0	4.4	4.0	0.00010
By BTM	7.2	8.0	4.0	-0.00015	6.0	10.4	2.0	-0.00135
By size and BTM	3.2	4.4	3.2	0.00013	3.6	3.6	5.2	-0.00011

Table 5.49 continued

Benchmark criteria	Equally-weighted portfolio return				Value-weighted portfolio return			
	5%			Mean AR	5%			Mean AR
	α=0	α<0	α>0		α=0	α<0	α>0	
	Panel C: Five years							
By size	4.4	6.0	4.4	0.00008	4.4	6.4	3.2	-0.00015
By BTM	4.8	4.4	2.8	-0.00005	3.6	10.4	0.0	-0.00178
By size and BTM	3.6	6.0	2.4	0.00005	2.8	5.2	3.2	-0.00029

Table 5.50 Descriptive statistics of abnormal returns

This table reports descriptive statistics of abnormal returns when applying mean monthly calendar-time abnormal return (MMAR) based on both equal weighting and value-weighted schemes. MMAR is another variant of calendar-time approach which combines reference portfolios and calendar time. Instead of running regression based on factor models, abnormal return of an individual stock in each calendar month is defined as the difference of stock return and benchmark return as $AR_{i,t} = R_{i,t} - R_{b,t}$. The benchmark is based on reference portfolios which are matched with the event firm by similar characteristics such as size, book-to-market ratio, size and book-to-market ratio, equally-weighted market return. In each calendar month, the monthly portfolio return is cumulated either with equally-weighted scheme or value-weighted scheme: $CTAR_t = \sum_{i=1}^{n} w_{it} * AR_{it}$. Eventually, a time series of monthly portfolio returns is achieved over the study period. The grand mean abnormal return is computed as $MCTAR = \sum_{t=1}^{T} CTAR_t / T$. The conventional cross-sectional test statistics is calculated as: $t = \frac{MCTAR}{\sigma(MCTAR_t)/\sqrt{T}}$. 250 samples of 200 event months are randomly selected with replacement. Both two-tailed and one-tailed tests are examined at a significance level of 5%. The missing returns are kept as they are without replacement. The rejection frequency in 250 samples is computed. If the test is well-specified, the rejection frequency is expected to 5% at a significance level of 5%. Furthermore, the distribution of abnormal returns is expected to be symmetric with equal rejection rates of 5% on both sides.

MMARs	**Equally-weighted portfolios**				**Value-weighted portfolios**			
One year	**Mean**	**Median**	**Skewness**	**Kurtosis**	**Mean**	**Median**	**Skewness**	**Kurtosis**
By size	-0.00039	-0.00072	0.35	4.25	-0.00021	-0.00047	0.06	2.94
By BTM	-0.00018	-0.00051	0.14	2.76	-0.00178	-0.00196	-0.02	3.54
By size and BTM	-0.00001	-0.00005	0.47	5.78	-0.00025	-0.00042	-0.08	2.80
Three years								
By size	-0.00004	-0.00007	0.84	5.68	0.00010	0.00013	0.04	2.94
By BTM	-0.00015	-0.00016	0.30	3.73	-0.00135	-0.00140	0.03	3.11
By size and BTM	0.00013	0.00012	0.24	3.14	-0.00011	-0.00015	-0.11	3.21
Five years								
By size	0.00008	0.00006	-0.03	3.22	-0.00015	-0.00007	-0.23	3.76
By BTM	-0.00005	-0.00015	0.49	4.75	-0.00178	-0.00165	-0.19	3.09
By size and BTM	0.00005	-0.00009	0.08	3.71	-0.00029	-0.00025	-0.24	3.06

Table 5.51 Power of test of mean monthly calendar-time portfolio returns-Equally weighted portfolio returns

This table reports percentages of rejection rates of the null hypothesis of zero mean monthly portfolio returns in 250 samples of 200 firms, which are randomly selected, over a period of 1982 to 2008 when applying equal weighting scheme. For each individual stock, the abnormal returns in the range of -20% to 20% at an incremental of 5% are introduced to each individual over one-, three- and five years. This abnormal return needs to be evenly allocated to each stock's monthly returns. For instance, if the abnormal return is 10%, the monthly returns of a stock are added with 0.83% over one year, 0.27% over three years and 0.17% over five years. The rejection frequency in 250 samples is computed. If this approach is well specified, the rejection rates in all levels are expected to be 100% since there is abnormal return. I study the power of test based on four benchmarks with matching criteria as size, BTM, size/BTM and equally weighted market return.

Benchmark criteria	-20%	-15%	-10%	-5%	0%	5%	10%	15%	20%
					Panel A: One year				
By size	96.0	86.4	60.4	23.6	6.0	11.2	48.8	86.0	98.0
By BTM	96.4	83.6	60.4	24.4	3.6	14.0	51.2	88.4	98.0
By size and BTM	96.4	87.2	55.2	16.4	2.4	15.6	52.8	86.0	98.0
					Panel B: Three year				
By size	63.6	40.8	25.6	10.4	3.6	4.4	14.4	34.4	60.0
By BTM	62.4	41.2	24.8	12.0	7.2	5.6	15.6	33.2	56.0
By size and BTM	56.4	36.4	21.6	8.4	3.2	5.6	16.8	38.8	58.0
					Panel C: Five year				
By size	36.0	24.4	12.8	7.6	4.4	5.6	10.0	20.4	32.0
By BTM	33.6	18.8	10.4	4.8	4.8	4.0	9.2	18.0	30.8
By size and BTM	32.4	19.6	12.0	6.0	3.6	3.2	8.0	20.8	32.0

Table 5.52 Power of test of mean monthly calendar-time portfolio returns-Value weighted portfolio returns

This table reports percentages of rejection rates of the null hypothesis of zero mean monthly portfolio returns in 250 samples of 200 firms, which are randomly selected, over a period of 1982 to 2008 when applying value-weighted scheme. For each individual stock, the abnormal returns in the range of -20% to 20% at an incremental of 5% are introduced to each individual over one-, three- and five years. This abnormal return needs to be evenly allocated to each stock's monthly returns. For instance, if the abnormal return is 10%, the monthly returns of a stock are added with 0.83% over one year, 0.27% over three years and 0.17% over five years. The rejection frequency in 250 samples is computed. If this approach is well specified, the rejection rates in all levels are expected to be 100% since there is abnormal return. I study the power of test based on four benchmarks with matching criteria as size, BTM, size/BTM and equally weighted market return.

Benchmark criteria	-20%	-15%	-10%	-5%	0%	5%	10%	15%	20%
					Panel A: One year				
By size	95.2	82.8	55.2	18.0	6.4	16.0	46.0	80.0	96.0
By BTM	97.6	87.6	64.8	28.8	6.4	11.2	30.8	70.8	92.0
By size and BTM	94.8	85.2	56.4	15.6	4.0	16.8	46.4	83.2	94.0
					Panel B: Three year				
By size	42.0	28.4	12.4	6.0	4.0	5.2	14.8	30.4	44.0
By BTM	53.2	38.8	21.6	13.2	6.0	4.0	6.4	12.4	23.2
By size and BTM	44.4	25.6	14.4	6.4	3.6	6.8	12.4	22.4	40.0
					Panel C: Five year				
By size	22.0	14.4	8.4	6.0	4.4	5.2	8.0	16.8	26.0
By BTM	28.8	16.4	10.4	6.8	3.6	1.6	0.8	2.0	4.8
By size and BTM	20.4	11.2	7.2	4.0	2.8	4.8	8.0	14.4	23.6

Table 5.53 Rejection frequency of the null hypothesis of zero mean monthly calendar-time abnormal return-Small firms

This table reports percentages of rejection rates of 250 samples with the null hypothesis of zero mean monthly calendar-time abnormal return (MMAR) based on both equal weighting and value-weighted schemes. MMAR is another variant of calendar-time approach which combines reference portfolios and calendar time. Instead of running regression based on factor models, abnormal return of an individual stock in each calendar month is defined as the difference of stock return and benchmark return as $AR_{i,t} = R_{i,t} - R_{b,t}$. The benchmark is based on reference portfolios which are matched with the event firm by similar characteristics such as size, book-to-market ratio, size and book-to-market ratio, equally-weighted market return. In each calendar month, the monthly portfolio return is cumulated either with equally-weighted scheme or value-weighted scheme: $CTAR_t = \sum_{i=1}^{n} w_{it} * AR_{it}$. Eventually, a time series of monthly portfolio returns is achieved over the study period. The grand mean abnormal return is computed as $MCTAR = \sum_{t=1}^{T} CTAR_t / T$. Then the conventional cross-sectional test statistics is calculated as: $t = \frac{MCTAR}{\sigma(MCTAR_t)/\sqrt{T}}$. 250 samples of 200 event months are randomly selected with replacement. Then 250 samples of 200 small firms from size decile 1 with the smallest market value in September year t are randomly selected to match the event months. Both two-tailed and one-tailed tests are examined at a significance level of 5%. The missing returns are kept as they are without replacement. The rejection frequency in 250 samples is computed. If the test is well-specified, the rejection frequency is expected to 5% at a significance level of 5%. Furthermore, the distribution of abnormal returns is expected to be symmetric with equal rejection rates of 5% on both sides.

	Equally-weighted portfolio return				**Value-weighted portfolio return**			
Benchmark criteria	**5%**			**Mean AR**	**5%**			**Mean AR**
	α=0	**α<0**	**α>0**		**α=0**	**α<0**	**α>0**	
				Panel A: One year				
By size	3.6	6.0	1.2	-0.00007	5.2	12.8	0.4	-0.00324
By BTM	34.0	0.0	54.4	0.01587	25.2	0.0	39.6	0.01336
By size and BTM	3.2	7.2	1.2	-0.00024	5.2	9.6	0.4	-0.00309
				Panel B: Three years				
By size	1.2	3.2	1.2	0.00001	5.2	9.2	0.4	-0.00185
By BTM	43.2	0.0	67.6	0.01044	35.2	0.0	51.2	0.00919
By size and BTM	1.2	4.0	2.4	-0.00003	3.6	8.0	0.4	-0.00139

Table 5.53 continued

Benchmark criteria	Equally-weighted portfolio return				Value-weighted portfolio return			
		5%				5%		
	α=0	α<0	α>0	Mean AR	α=0	α<0	α>0	Mean AR
				Panel C: Five years				
By size	2.8	8.0	0.8	-0.00053	8.0	14.8	0.4	-0.00225
By BTM	25.2	0.0	43.6	0.00639	14.0	0.0	28.8	0.00524
By size and BTM	2.8	6.4	1.2	-0.00046	4.0	10.4	0.4	-0.00165

Table 5.54 Rejection frequency of the null hypothesis of zero mean monthly calendar-time abnormal return-Large firms

This table reports percentages of rejection rates of 250 samples with the null hypothesis of zero mean monthly calendar-time abnormal return (MMAR) based on both equal weighting and value-weighted schemes. MMAR is another variant of calendar-time approach which combines reference portfolios and calendar time. Instead of running regression based on factor models, abnormal return of an individual stock in each calendar month is defined as the difference of stock return and benchmark return as $\mathrm{AR}_{i,t} = \mathrm{R}_{i,t} - \mathrm{R}_{b,t}$. The benchmark is based on reference portfolios which are matched with the event firm by similar characteristics such as size, book-to-market ratio, size and book-to-market ratio, equally-weighted market return. In each calendar month, the monthly portfolio return is cumulated either with equally-weighted scheme or value-weighted scheme: $\mathrm{CTAR}_t = \sum_{i=1}^{n} w_{it} * \mathrm{AR}_{it}$. Eventually, a time series of monthly portfolio returns is achieved over the study period. The grand mean abnormal return is computed as $\mathrm{MCTAR} = \sum_{t=1}^{T} \mathrm{CTAR}_t / \mathrm{T}$. Then the conventional cross-sectional test statistics is calculated as: $t = \frac{\mathrm{MCTAR}}{\sigma(\mathrm{MCTAR}_t)/\sqrt{T}}$. 250 samples of 200 event months are randomly selected with replacement. Then 250 samples of 200 large firms from size decile 10 with the largest market value in September year t are randomly selected to match the event months. Both two-tailed and one-tailed tests are examined at a significance level of 5%. The missing returns are kept as they are without replacement. The rejection frequency in 250 samples is computed. If the test is well-specified, the rejection frequency is expected to 5% at a significance level of 5%. Furthermore, the distribution of abnormal returns is expected to be symmetric with equal rejection rates of 5% on both sides.

	Equally-weighted portfolio return				Value-weighted portfolio return			
Benchmark criteria	**5%**			**Mean AR**	**5%**			**Mean AR**
	α=0	**α<0**	**α>0**		**α=0**	**α<0**	**α>0**	
	Panel A: One year							
By size	6.0	5.2	4.4	0.00002	5.2	7.2	6.0	-0.00007
By BTM	3.6	5.2	0.0	-0.00110	2.0	5.6	0.8	-0.00083
By size and BTM	4.0	4.8	6.0	0.00001	5.2	4.8	8.4	0.00029
	Panel B: Three years							
By size	4.4	4.4	5.6	0.00001	0.8	3.6	1.6	-0.00039
By BTM	1.2	3.2	0.0	-0.00111	0.0	1.2	0.0	-0.00114
By size and BTM	3.2	2.8	5.6	0.00002	0.0	1.2	2.0	-0.00004

Table 5.54 continued

Benchmark criteria	Equally-weighted portfolio return				Value-weighted portfolio return			
		5%		Mean AR		5%		Mean AR
	α=0	α<0	α>0		α=0	α<0	α>0	
	Panel C: Five years							
By size	5.6	6.0	4.4	0.00000	3.2	5.6	2.8	-0.00019
By BTM	3.2	6.0	0.0	-0.00129	1.2	3.6	0.4	-0.00120
By size and BTM	4.4	4.4	4.8	0.00002	2.4	2.4	3.2	0.00009

Table 5.55 Rejection frequency of the null hypothesis of zero mean monthly calendar-time abnormal return-Firms with lowest BTM

This table reports rejection frequency of 250 samples with the null hypothesis of zero mean monthly calendar-time abnormal return (MMAR) based on both equal weighting and value-weighted schemes. MMAR is another variant of calendar-time approach which combines reference portfolios and calendar time. Instead of running regression based on factor models, abnormal return of an individual stock in each calendar month is defined as the difference of stock return and benchmark return as $AR_{i,t} = R_{i,t} - R_{b,t}$. The benchmark is based on reference portfolios which are matched with the event firm by similar characteristics such as size, book-to-market ratio, size and book-to-market ratio, equally-weighted market return. In each calendar month, the monthly portfolio return is cumulated either with equally-weighted scheme or value-weighted scheme: $CTAR_t = \sum_{i=1}^{n} w_{it} * AR_{it}$. Eventually, a time series of monthly portfolio returns is achieved over the study period. The grand mean abnormal return is computed as $MCTAR = \sum_{t=1}^{T} CTAR_t / T$. Then the conventional cross-sectional test statistics is calculated as: $t = \frac{MCTAR}{\sigma(MCTAR_t)/\sqrt{T}}$. 250 samples of 200 event months are randomly selected with replacement. Then 250 samples of 200 firms from BTM decile 1 with the lowest book-to-market ratio are randomly selected to match the event months. Both two-tailed and one-tailed tests are examined at a significance level of 5%. The missing returns are kept as they are without replacement. The rejection frequency in 250 samples is computed. If the test is well-specified, the rejection frequency is expected to 5% at a significance level of 5%. Furthermore, the distribution of abnormal returns is expected to be symmetric with equal rejection rates of 5% on both sides.

	Equally-weighted portfolio return				**Value-weighted portfolio return**			
	5%				**5%**			
Benchmark criteria	**α=0**	**α<0**	**α>0**	**Mean AR**	**α=0**	**α<0**	**α>0**	**Mean AR**
	Panel A: One year							
By size	25.6	36.4	0.4	-0.00488	21.6	30.4	0.4	-0.00485
By BTM	6.8	5.2	6.4	0.00049	4.8	6.4	2.8	-0.00098
By size and BTM	5.2	4.8	3.6	0.00004	5.2	2.8	4.4	0.00046
	Panel B: Three years							
By size	28.4	41.6	0.0	-0.00420	17.2	29.6	0.0	-0.00379
By BTM	2.8	6.4	0.8	-0.00024	4.4	8.0	2.0	-0.00105
By size and BTM	1.6	4.8	1.6	-0.00029	4.8	3.2	8.0	0.00066

Table 5.55 continued

Benchmark criteria	Equally-weighted portfolio return				Value-weighted portfolio return			
	5%			Mean AR	5%			Mean AR
	α=0	α<0	α>0		α=0	α<0	α>0	
	Panel C: Five years							
By size	24.4	37.2	0.0	-0.00337	13.2	24.8	0.0	-0.00324
By BTM	3.6	5.2	3.6	-0.00007	2.8	5.6	1.6	-0.00108
By size and BTM	4.4	6.4	2.8	-0.00029	3.2	1.2	4.8	0.00041

Table 5.56 Rejection frequency of the null hypothesis of zero mean monthly calendar-time abnormal return-Firms with highest BTM

This table reports percentages of rejection rates of 250 samples with the null hypothesis of zero mean monthly calendar-time abnormal return (MMAR) based on both equal weighting and value-weighted schemes. MMAR is another variant of calendar-time approach which combines reference portfolios and calendar time. Instead of running regression based on factor models, abnormal return of an individual stock in each calendar month is defined as the difference of stock return and benchmark return as $AR_{i,t} = R_{i,t} - R_{b,t}$. The benchmark is based on reference portfolios which are matched with the event firm by similar characteristics such as size, book-to-market ratio, size and book-to-market ratio, equally-weighted market return. In each calendar month, the monthly portfolio return is cumulated either with equally-weighted scheme or value-weighted scheme: $CTAR_t = \sum_{i=1}^{n} w_{it} * AR_{it}$. Eventually, a time series of monthly portfolio returns is achieved over the study period. The grand mean abnormal return is computed as $MCTAR = \sum_{t=1}^{T} CTAR_t / T$. Then the conventional cross-sectional test statistics is calculated as: $t = \frac{MCTAR}{\sigma(MCTAR_t)/\sqrt{T}}$. 250 samples of 200 event months are randomly selected with replacement. Then 250 samples of 200 firms from BTM decile 10 with the highest book-to-market ratio are randomly selected to match the event months. Both two-tailed and one-tailed tests are examined at a significance level of 5%. The missing returns are kept as they are without replacement. The rejection frequency in 250 samples is computed. If the test is well-specified, the rejection frequency is expected to 5% at a significance level of 5%. Furthermore, the distribution of abnormal returns is expected to be symmetric with equal rejection rates of 5% on both sides.

	Equally-weighted portfolio return				Value-weighted portfolio return			
Benchmark criteria	**5%**			**Mean AR**	**5%**			**Mean AR**
	α=0	**α<0**	**α>0**		**α=0**	**α<0**	**α>0**	
	Panel A: One year							
By size	13.6	0.4	18.4	0.00639	2.4	2.4	5.6	0.00288
By BTM	11.2	16.8	0.8	-0.00309	4.0	6.0	1.6	-0.00035
By size and BTM	6.0	3.6	8.4	0.00225	3.2	4.8	2.4	0.00053
	Panel B: Three years							
By size	15.6	0.4	26.4	0.00579	2.4	0.8	8.0	0.00265
By BTM	5.2	11.2	0.0	-0.00235	2.4	7.6	0.8	0.00013
By size and BTM	2.8	1.6	8.4	0.00230	2.4	5.2	1.2	0.00042

Table 5.56 continued

	Equally-weighted portfolio return				Value-weighted portfolio return			
Benchmark criteria		**5%**		**Mean AR**		**5%**		**Mean AR**
	α=0	**α<0**	**α>0**		**α=0**	**α<0**	**α>0**	
	Panel C: Five years							
By size	15.6	0.0	28.0	0.00482	4.8	0.0	12.0	0.00274
By BTM	8.8	16.0	0.0	-0.00235	3.6	8.0	2.0	0.00012
By size and BTM	5.2	2.0	9.2	0.00194	3.2	6.4	3.2	0.00061

Chapter 6 Conclusions

This thesis is consisted with three main empirical chapters based on the UK data which illustrate three most popular approaches: event-time approach based on models, event-time approach based on reference portfolios and calendar-time approach. In comparing the three methodologies on the UK data, the empirical results suggest the following methods are appropriately applied to detect long-term abnormal stock performance. When the event-time approach is applied based on models, although the measurement of BHARs together with the market-adjusted model, capital asset pricing model and Fama-French three-factor model generate well-specified results, the test statistics derived from these models are not reliable because BHARs show severe positively skewed and leptokurtic distribution. Moreover, the reference portfolios in conjunction with p-value from pseudoportfolios and the control firm approach with student t test in the event-time approach are advocated although with lower power of test. When it comes to the calendar-time approach, the three-factor model under OLS together with the sandwich variance estimators using the value-weighted scheme and the mean monthly calendar-time abnormal returns under equal weights are proved to be the most appropriate methods. When power of test is taken into account, the method based on mean monthly calendar-time abnormal returns outperforms.

Since the UK market structure is filled mostly with firms with small market value, I fill in the gap in the literature to explore if the market structure has an impact on the results when conducting a simulation process. **Chapter 3** starts with benchmarks measured by using different models including asset pricing models. Apart from the four models applied by Kothari and Warner (1997), I add four-factor model, introduced by Carhart (1997) with factors downloaded from the University of Exeter. To test the model specification not only in random samples but also in non-random samples, I use the grouping system as Lyon, Barber and Tsai (1999) to categorize small/large firms and firms with low/high book-to-market ratios. The findings demonstrate the differences of the choice of standard error of abnormal returns from estimation period or test period in the UK stock market. Moreover, the buy-and-hold abnormal returns show well-specified results except for the market model while the cumulative abnormal returns show overrejection in most cases. The small rejection rates in BHARs suggest insignificant compounding effect from the measurement of abnormal returns. although the measurement of BHARs together with the market-adjusted model, capital asset pricing model and Fama-French three-factor model generate well-specified results, the test statistics derived from these models are not reliable because BHARs show severe positively skewed and leptokurtic distribution. The misspecification of the models, especially under CARs, confirms the conclusion from prior research which indicates the study of long-term abnormal return is not only a test of market efficiency but also to identify an equilibrium asset pricing model (Fama, 1998). Therefore, I attempt to explore another approach in event-time study, which is to use reference portfolios or a control firm as the benchmark.

The methodology applied in **Chapter 4** is based on studies conducted by Lyon and Barber (1997), Lyon, Barber and Tsai (1999) and Ang and Zhang (2004). To identify the benchmarks as reference portfolios based on size, book-to-market ratio (BTM), size/BTM and equally weighted market returns, I apply both parametric tests with an assumption of normal distribution with zero mean and constant variance and nonparametric tests with no requirement on distribution. The parametric tests have conventional student t test and skewness-adjusted test. The nonparametric tests are bootstrapped skewness-adjusted test, empirical p value from pseudoportfolios and Wilcoxon signed-rank test. Both CARs and BHARs are examined in this chapter. Similar findings are documented with the UK evidence as the US. Both CARs and BHARs produce well-specified test statistics in random samples. However, it is important to note that the choice of benchmark yields different results. Moreover, BHARs generates higher rejection rates compared with CARs, although it takes investors' experience into account. The reference portfolios with p value from pseudoportfolios and the control firm approach with student t test outperform.

In order to solve the issue of cross-sectional dependent returns, the calendar-time approach proposed by Jaffe (1974) and Mandelker (1974) is examined in **Chapter 5**. The conventional calendar-time approach based on the Fama-French three-factor model and Carhart four-factor model, and the mean monthly calendar-time abnormal returns which combines reference portfolios and calendar time, are applied. To comprehensively cover the topic of heteroskedasticity in the regression, three techniques are adopted to overcome the weakness of ordinary least squares. These include White's (1980) test, weighted least squares using the number of firms as weights and generalized least squares. The approach of mean monthly calendar-time abnormal returns applies benchmarks matching by size, book-to-market ratio and size/book-to-market ratio. The main findings indicate the importance of weighting schemes when applying the conventional calendar time approach based on the UK data which has many small firms. Value-weighted scheme shows superior ability to reduce the misspecification. Moreover, the mean monthly calendar-time abnormal returns generate well-specified results in both random samples and non-random samples with the highest power of test compared with all other approaches, regardless of weighting schemes.

6.1 Conclusion: Event-time approach based on models

Chapter 3 shows the empirical results based on the UK stock market with the application of five models to measure the benchmark returns: market-adjusted model, market model, capital asset pricing model, Fama-French three-factor model, and Carhart four-factor model. The findings are mostly consistent as prior research undertaken by Kothari and Warner (1997) based on the US stock market. To begin with, CARs shows misspecification with overrejections of the null hypothesis of zero mean CARs at significance levels of 1% and 5% in one-tailed and two-tailed tests, in all models over one-, three- and five-year investment horizons. In particular, the market model presents the most severe misspecification with the highest rejection rates compared with other models, even when BHARs is applied. Second, test statistics of CARs and BHARs are asymmetric. Positively biased test statistics are shown for random samples, small firms and firms with high book-to-market ratios under BHARs and CARs in 250 samples of 200 firms. However, when the sample is consisted with large firms or firms with low book-to-market ratios, test statistics are negatively biased with higher rejection rates on the lower tail. To further study whether the distributional properties of CARs and BHARs cause the misspecification, I examine a sample of 50,000 firms which are randomly selected without replacement. The magnitude of BHARs is larger than CARs. Additionally, the distribution of BHARs is severely right-skewed and fat-tailed with large values of skewness and kurtosis. Moreover, the medians of BHARs in all models are negative whereas the medians of CARs in all models are positive except for the case when applying the market model in one year. This is further proved by Wilcoxon signed-rank test. Four, the standard errors based on the estimation period or test period mark a difference in the results. Regarding CARs, the misspecification is significantly improved when using standard errors based on the test period. Nevertheless, the number of firms varies over the test period. This potentially leads to biased statistical inferences. This finding indicates even the application of the market-adjusted model, three-factor model and capital asset pricing model yield well-specified results, the test statistics are not reliable because of severely asymmetric distribution of abnormal returns. Last but not least, the sample selection bias in this methodology, which requires at least 25-month consecutive returns is more significant in the long run. The average stock returns with requirements of pre-event survival periods reach 79.79% in CARs and 94.49% in BHARs when the test period is five years. The longer the pre-event survival periods, the higher the magnitudes of CARs and BHARs.

The UK data exhibits distinguished results from prior studies carried out by Kothari and Warner (1997) as follow. Firstly, the magnitude of rejection rates under CARs and BHARs in most cases is larger than the case when applying the methodology based on the US data. Secondly, the distribution of BHARs in a sample of 50,000 firms shows more misspecified results with extremely large value in kurtosis and skewness. This may be caused by outliers of raw returns and small size effect. Thirdly, it is surprising to find out that BHARs improves the misspecification in CARs with a dramatic decrease in rejection rates when applying on all models based on 250 samples of 200 firms. However, Kothari and Warner (1997) document similar rejection rates in BHARs compared with CARs except for the market model. Moreover, when applying standard errors based on the test period, the misspecification is much less than the case when using standard errors based on the estimation period. The decline in rejection rates is larger than the results in prior research. Fourthly, apart from the three-factor model, I employ the Carhart four-factor model which introduces a momentum factor to capture the difference between returns of winners and losers. The four-factor model generates misspecified test statistics with high rejection rates in the long run. Fifthly, to take into account the skewness bias, I adopt Wilcoxon signed-rank test, which does not have an assumption of normal distribution to detect the long-term abnormal performance. Both CARs and BHARs overreject the null hypothesis of zero median in two-tailed test at a significance level of 5%. Furthermore, CARs displays higher rejection rates on the upper tail which indicates the median of CARs is positive. This is consistent with the analysis of distributional properties of CARs on a sample of 50,000 firms. However, BHARs shows overrejection on the lower tail which is in line with the findings of negative medians in all models over three investment horizons. It is interesting to note that the rejection rate is 5.6% in two-tailed test, 5.6% on the lower-tailed test and 6.4% on the upper-tailed test when applying the market model in five years. Sixthly, Kothari and Warner (1997) find large firms outperform small firms with larger CARs. However, the UK stock market exhibits an opposite case. The higher rejection rates in large firms compared with small firms suggest either the pricing models do not generate reasonable normal returns for large firms or large firms with large returns are resampled. Moreover, small firms tend to show positive mean CARs and BHARs whereas large firms have negative mean CARs and BHARs. This indicates positive stock performance for small firms over the period of 1982 to 2008. Small firms and large firms, regardless of models or measurement of abnormal returns, show overrejection of the null hypothesis of zero mean abnormal returns. Lastly, although there is a decline in rejection rates for firms with low book-to-market ratios compared with firms with high lower book-to-market ratios, both firms based on book-to-market ratios show severe misspecification with overrejection in three investment horizons.

Ever since the research based on models in event-time approach undertaken by Kothari and Warner (1997), the following studies shift focuses to other approaches. The findings support this shift with evidence of severe misspecification in random samples and non-random samples, particularly with the measurement of CARs. Although BHARs improves the misspecification significantly, it is subject to severe asymmetric distribution with high value of kurtosis. This methodology, based on models in event time, has some limitations. The first is sample selection bias. The number of firms declines from 4,977 to 3,978 after incorporating the requirement of at least 25-month returns available when the test period is one year. This exclusion of firms results in higher raw returns, particularly in the long run. It is inevitable to have requirement of pre-event data since the coefficients in regressions need a series of continuous data. Even applying the reference portfolios, the matching criteria such as book value, market value, still requires data availability in a specific period. Another limitation is the models which predict the normal returns with the absence of events. The market model shows severe misspecification whereas the four-factor model generates higher rejection rates compared with the three-factor model. Different models exhibit different rejection rates under BHARs and CARs. This indicates the evidence of abnormal returns depends on models applied. Therefore, there is still an open question on which model is superior to reflect all the characteristics of stock returns. As suggested by Fama (1998), the long-run event studies involves the test of market efficiency but also the choice of an equilibrium asset pricing model. The last limitation is the missing returns. Kothari and Warner (1997) track stock returns with available data. This suggests if there is missing returns over the test period, the stock returns are replaced with zero if there are still returns following the missing returns. This may cause the biased in the statistical inferences. Moreover, the anomalies with higher rejection rates in large firms indicate high raw returns and longer survival horizons of large firms.

6.2 Conclusion: Event-time approach based on reference portfolios

With the application of both parametric tests and nonparametric tests on all benchmarks, I achieve similar conclusions with the UK evidence as prior research as Lyon and Barber (1997), Lyon, Barber and Tsai (1999) and Ang and Zhang (2004). First of all, both CARs and BHARs based on reference portfolios and a control firm in event time produce well-specified test statistics in two-tailed test at a significance level of 5%. The magnitude of rejection rates is smaller in CARs. Second, in one-tailed test, asymmetry of test statistics is implied. All benchmarks under both CARs and BHARs show negatively biased test statistics with higher rejection rates on the lower tail but there is an exception when the benchmark is equally-weighted portfolio returns. The application of skewness-adjusted test, the bootstrapped skewness-adjusted test and p value from pseudoportfolios improve the skewness with higher rejection rates on the upper tail. Third, there is no particular pattern of rejection frequency of the null hypothesis of mean abnormal returns over the investment horizons. In some case, the rejection rates decline in a five-year investment horizon. Four, the control firm approach in conjunction with student t test generate well-specified test statistics when detecting long-term abnormal stock performance, especially in two-tailed test. However, it has the poorest power of test. Five, the Wilcoxon signed-rank test, a nonparametric test, shows severe misspecification when the benchmark is reference portfolios under BHARs. Six, compared the power of test of different test statistics based on the benchmark matched by size and BTM, the bootstrapped skewness-adjusted test illustrates consistent outperformance when compared with p value from pseudoportfolios. Moreover, there is a dramatic decline in rejection rates when the investment horizon lengthens. Additionally, control firm approach shows the lowest power of test when negative returns are induced. Seven, even the control firm approach lose the advantage when the stock returns overlap in the study period. However, the reference portfolio approach and control firm approach yield well-specified results when resolving the issue of calendar time clustering, which is one type of cross-sectional dependent returns. At last, the reference portfolios approach and control firm approach are proved to be appropriate to apply under CARs when there is industry clustering.

Interestingly, some UK evidence is distinguished from prior research which is primarily based on the US market. At first, with comparison among all the benchmarks, it is difficult to identify which matching criterion is superior when detecting long-term abnormal stock performance since the empirical findings are mixed. However, the benchmark with equally-weighted market portfolio returns is clearly not a good choice when conducting research on non-random samples. It is applicable in random samples but has positively based test statistics which is opposite when applying other benchmarks. Secondly, although there is an improvement with the application of the skewness-adjusted test and bootstrapped skewness-adjusted test, which aim to resolve the issue of skewness, the rejection rates in the two-tailed tail increase. This indicates misspecification in two-tailed test. Therefore, the UK evidence is not in favour of the application of these two test statistics. However, the p value from pseudoportfolios exhibits satisfactory results without sacrifice on two-tailed test. Thirdly, with the extension of power of test over all three investment horizons based on the benchmark matched by size and BTM, it is noticeable that the parametric tests, especially skewness-adjusted test, show consistent high power with either positive shocks or negative shocks. However, these tests are not advocated in the long-run event studies because of misspecification where there is no abnormal return induced. Fourthly, prior research undertaken by Lyon (1997) examines different benchmarks in random samples whereas Lyon, Barber and Tsai (1999) conduct robustness checks on non-random samples with the application of benchmark matched by size/BTM only. Ang and Zhang (2004) further investigate control firm approach on small firms. The extensive coverage of benchmarks on non-random samples states a fact that the benchmark, which does not share the characteristics of the non-random samples, generates misspecifications. For instance, overrejection of the null hypothesis of zero abnormal returns is documented in small firms when the benchmark is matched by BTM only. Fifthly, it is interesting to note that when conducting the Wilcoxon signed-rank test, CARs produce well-specified results regardless of reference portfolios approach or control firm approach. Last but most importantly, the magnitude of rejection rates is higher when the cross-sectional dependence of returns is caused by overlapping returns and calendar time clustering. Moreover, the application of equally-weighted market return yield misspecified tests in both cases.

The reference portfolios in conjunction with p-value from pseudoportfolios and the control firm approach with student t test are favourable when detecting long-term abnormal stock performance. Even with the UK data which is filled with mostly small firms, these approaches yield well-specified results, especially in two-tailed tests. However, the power of test in both approaches is the worst compared with other tests. Moreover, this approach in this thesis is subject to some limitations. Firstly, the reference portfolios established according to different characteristics of firms include the event firm. This potentially results in biased statistical inferences. It is worth to examine the abnormal performance based on cleaner benchmark returns which exclude the event firm. Secondly, the missing returns are replaced with the average returns of reference portfolios as Lyon, Barber and Tsai (1999) suggested. However, if the sample has a few small firms, the returns will be overstated. Furthermore, the replacement of missing returns has strong impact on the accuracy and reliability of test statistics if there are a significant number of delisted event firms. Thirdly, when applying the event-time approach based on models, I discover that the selection bias with requirement of pre-event data could result in misspecification since the average stock returns increase when the survival period lengthens. Similarly, with the necessity to use market value and book value to categorize firms into reference portfolios, requirement of market value in September of year t and book value in March of year t is applied. It is inevitable to place on requirement on pre-event stock returns in the event-time studies. However, when applying calendar-time approach based on equal-weighted scheme in random samples, the sample selection bias can be eliminated. Fourthly, since the UK market is filled with small firms, the application of equally-weighted market returns as a benchmark is expected to yield misspecified results. It is worth to examine the rejection frequency when applying the value-weighted scheme which allocates reasonable weights to small firms according to their market value. Lastly, the monthly market values of firms provided by the London Stock Price Database (LSPD) are not precise. The rounding system in the LSPD marks firms with market value below £1000 as zero. To make it feasible to calculate the book-to-market ratio, I replace the value of zero with 0.5. However, when firms are deciled according to their market values, it is difficult to accurately allocate them into the right groups.

6.3 Conclusion: Calendar-time approach

Chapter 5 analyzes the calendar-time approach based on both equal weighting scheme and value-weighted scheme. Similar finding as prior research undertaken by Lyon, Barber and Tsai (1999) and Ang and Zhang (2004) are documented as follows. Regarding the conventional calendar-time approach, it generates well-specified test statistics when applying a value-weighted scheme. Moreover, non-random samples including small/large firms, firms with low/high BTM, yield misspecified results regardless of weighting schemes. In addition, the industry clustering with cross-sectional returns is resolved when running regressions based on value-weighted scheme. Compared with the three-factor model, the four-factor with a momentum factor improves the misspecification, especially under value weights, and shows higher explanatory power of independent variables to the dependent variables. The power of test increases with the magnitude of abnormal returns introduced but decreases with the length of investment horizons. Furthermore, the test shows stronger power to capture the positive shocks compared with negative shocks. With respect to mean monthly calendar-time returns, the magnitude of rejection rates is smaller compared with that when applying the conventional approach. This approach generates well-specified results regardless of weighting schemes in random samples and non-random samples.

Although the results show similarities as the prior studies, there are a few differences worth discussing. Regarding the conventional calendar-time approach, the application of different weighting schemes generates different results. The value-weighted scheme yields well-specified results in random samples whereas the equally-weighted scheme generates severe misspecification. I attribute this to the UK market structure which is filled mostly with small firms. When the equal weighting scheme is adopted, with an equal weight allocated to each firm, small firms are given more weights than they deserve. However, the value-weighted scheme allocates weights to firms according to their market value. The prior research based the US data does not show significant difference from the studies when applying different weighting schemes. To reduce the misspecification under equal weights, I find the winsorization process which deals with the outliers of stock returns can effectively reduce the misspecification but fail to eliminate it. Therefore, the value-weighted scheme is a better choice when studying a sample with many small firms. Secondly, small firms and firms with high BTM generate misspecified results regardless of weighting schemes. Small size effect is used to explain this misspecification. Regarding the misspecification in large firms and firms with low BTM, the causes are attributed to resampling of large firms with large returns. Thirdly, to correct the heteroskedasticity, I compare the three techniques. Compared with the ordinary least squares, the weighted least squares does not outperform based on the UK data. This is in contrast to the conclusion drawn by Ang and Zhang (2004). sandwich variance estimators under OLS outperform WLS and GLS with lower rejection rates. Therefore, it is advocated to be applied in the long-run event studies, especially for data containing many small firms. Fourthly, industry clustering and calendar clustering both generate well-specified results under value-weighted scheme. However, the conventional approach yields misspecified test statistics for firms with overlapping returns.

Regarding the mean monthly calendar-time returns, the weighting scheme does not matter with an exception when the benchmark is the equally weighted market portfolio returns under value weights. Second, the power of test decreases as the investment horizons lengthen under both weighting schemes as other approaches. However, the equally weighted scheme shows higher power with positive abnormal returns whereas the value weighted scheme represents higher power with negative abnormal returns. It is worth noting that the power of test is the highest comparing with other approaches. Last but not least, when conducting the robustness checks on small/large firms, weighting schemes do not play an important role. Both small and large firms generate well-specified test statistics when applying the benchmark matched by size and size/BTM. However, misspecification appears with higher rejection rates with the application of benchmarks matched with equally-weighted market portfolio and by BTM. Similar results are found for firms with low/high BTM. The overrejection is documented when applying the benchmarks matched by size and equally-weighted market portfolios. This indicates the meaninglessness to conduct robustness check with benchmarks matched with characteristics which are totally different from sample firms. When compared with the conventional calendar-time approach based on factor models, MMAR outperforms in terms of power of test and non-random samples.

In the studies, the conventional calendar time approach experiences difficulties when the sample is filled with mostly small firms, particularly when applying equally-weighted scheme. It still cannot resolve the issue of overlapping returns with cross-sectional dependent returns, especially when the sample contains a number of small firms. Moreover, the robustness check for non-random samples generates severe misspecification. In addition, the three-factor model is subject to criticism on its ability to capture the characteristics of stock returns. This ‘bad model’ issue proposed by Fama (1998) still exists based on the UK data. The application of four-factor model only lifts up the power of explanatory power of independent variables to dependent variables but generates more misspecified results. Lastly, the correct of heteroskedasticity requires more attention since there is some improvement when applying other techniques compared with ordinary least squares technique. Although the mean monthly calendar-time approach is favourable for its well-specified results regardless of weighting schemes in random sample or non-random samples, the benchmark needs to be carefully constructed. Moreover, when applying value-weighted scheme, there is an open question which market value of the event firm should apply. In the studies, the market value of event firms in the month when the calendar portfolio is constructed is used. This results in bias in the portfolio returns since the market value is not the most updated. Therefore, some studies suggest applying the market value of event firms in the previous month. However, if a firm is newly listed, the market value in the previous month is missing but there is return observation in the current month. In addition, when conducting robustness check on non-random samples, the matching criteria in the benchmarks cannot be overlooked. The empirical results shown in the analysis suggest the benchmarks applied with characteristics similar as the non-random samples yield meaningful results. This is to say when studying non-random samples biased in size, the benchmark needs to be matched with at least with one criterion as size.

6.4 Limitations and Suggestions for future research

Although I cover the UK data excluding financials over a period of 1982 to 2008, there are still some issues regarding the data. First of all, the raw returns show a few outliers which could potentially affect the results, particularly when applying buy-and-hold returns. There are 27 monthly returns larger than 500%. One observation even reaches 4500%. I do not exclude those outliers in the studies but the winsorization process is carried out in the calendar time approach for 250 samples. It is worth investigating the effect of winsorization in different approaches. Second, the data screen to include only firms with market value and book value, shows a significant decline in the number of firms, especially in the year of 2007 with drop-out firms of 545. The 423 firms with missing book value peak in 2007. Third, since the market value is needed for categorizing firms into reference portfolios, it is important to have the precise value. However, the London Stock Price Database (LSPD) provides monthly market value with 0 denoted for firms with market value smaller than £1 million. Therefore, when I use market value as matching criteria, small firms cannot differentiate from each other with the same value of market value. The rounding system in LSPD also makes it difficult to precisely classify firms on the basis of market value. Four, since the UK market consists of main market, alternative market and unlisted securities market, it is interesting to test the long-term performance in the submarket separately. However, this results in a reduction of the number of firms?, especially after the exclusion of financials. The simulation process needs to shift to a smaller sample size in order to avoid the issue of data mining. Lastly, the sub-period long-term stock performance is worth examining since I cover a long period from 1982 to 2008. The large number of dropouts in 2007 may be due to the credit crunch; therefore, the results may be different when separating the periods into decades: 1980s, 1990s and 2000s.

With respect to the event time approach based on models, severe misspecification appears when applying CARs, regardless of standard deviation based on estimation period or test period. The market model is the most misspecified with the highest rejection rates in both CARs and BHARs. This is in line with the argument regarding the “bad mode” issue in event studies by Fama (1998). The asset pricing models are considered to lack the ability to fully capture the characteristics of stock returns. Even the Carhart four-factor model, which incorporates the momentum factor, cannot improve the misspecification. Future research is encouraged to identify a well-specified model to reflect more detailed characteristics of stock returns. Since the reference portfolios approach and control firm approach show superior power when detecting long-term stock performance, the benchmark portfolio is a good starting point to improve the model. In the factor models, additional factors which relate to the specific characteristics of the event firm may be added as a dummy variable when meeting certain requirements. Apart from the model issue, the distribution of CARs and BHARs requires greater attention. BHARs shows higher mean BHARs with high value in kurtosis and skewness. Although extensive literature exists regarding the improvement on skewness through nonparametric tests such as bootstrapped procedures (Lyon, Barber and Tsai, 1999), there is limited discussion regarding resolutions for the “fat-tail” problem.

Regarding the reference portfolios, the empirical results, which require a careful selection of benchmarks, are satisfactory in both random samples and non-random samples. The criteria of benchmarks needs to be more specified, especially when applying the control firm approach. As with the methodology by Lyon, Barber and Tsai (1999), a control firm is the one with the closest BTM to the event firm after screening based on the market value in the range of 70% to 130% of the event firm. However, in most cases, multiple firms exist after the first screening. There is no specific criterion to choose among these control firms. Therefore, apart from size and book-to-market ratio, more studies are encouraged to focus on other criteria, such as P/E ratio, industry. Since CARs and BHARs in the studies are based on an equally-weighted scheme, it is worth examining whether any difference exists when using the value-weighted scheme to measure the portfolio return.

With regards to the calendar time approach, this still concerns the ‘bad model’ issue. Although the Carhart four-factor model shows higher explanatory power of independent variables to dependent variable, it generates misspecified results. The heteroskedasticity is not improved significantly with weighted least squares, generalized least squares and sandwich variance estimators. Therefore, a well-specified model and techniques is the focus of future research.

6.5 Implications for the finance industry

The empirical results documented in this study provide a comprehensive framework for researchers and practitioners who are interested in studying the long-term stock performance following unknown events. The results based on the UK stock market which is mostly filled with small firms suggest the most appropriate method to detect long-term abnormal stock performance is the mean monthly calendar-time abnormal returns under either equal weights or value weights. This can be also extended to studies relating to specific events such as IPOs, SEOs, or Mergers and Acquisitions, etc., Moreover, the relevant test statistics including non-parametric tests can be applied in the short-term event studies. The market efficiency hypothesis is under debate in the context of event studies, since the methodology of event studies not only tests the null hypothesis of zero abnormal returns but also tests the equilibrium model applied. It is important to identify which benchmark should be employed to compare with the actual returns so as to achieve abnormal returns. The performance of stock in the long run will provide a useful guideline to value investors when making investment decisions as well as establishing benchmarks.

Bibliography

1. Abad, D. and Rubia, A., 1999. Evaluation of the fixing trading system in the Spanish market. Working paper, series EC 99-17, Instituto Valenciano de Investigaciones Económicas.
2. Anderson, C., Mandelker, G., 1993. Long run return anomalies and the book-to-market effect: evidence on mergers and IPOs. Unpublished Working Paper, University of Pittsburgh.
3. Andre, P., Kooli, M. and L'Her, J., 2004. The long-run performance of mergers and acquisitions: Evidence from the Canadian stock market. *Financial Management*, Winter, 27-43.
4. Ang, A., Gu, L. and Hochberg, Y.V., 2007. Is IPO underperformance a peso problem? *Journal of Financial and Quantitative Analysis, 42*, 565-594.
5. Aggarwal, R. and Rivoli, P., 1990. Fads in the initial public offering market? *Financial Management,* 19, 45-57.
6. Agrawal, A., Jaffe, J.F. and Mandelker, G., 1992. The post-merger performance of acquiring firms: a re-examination of an anomaly. *Journal of Finance,* 47, 1605-1621.
7. Agrawal, A. and Mandelker, G., 1990. Large shareholders and the monitoring of managers: The case of anti-takeover charter amendments. *Journal of Financial and Quantitative Analysis*, 25, 143-161.
8. Agarwal, V. and Taffler R.J., 2008. Compared the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking and Finance,* 32, 1541-1551.
9. Ang, S. and Zhang, S., 2004. An evaluation of testing procedures for long horizon event studies. *Review of Quantitative Finance and Accounting*, 23, 251-274.
10. Ashley, J., 1962. Stock prices and changes in earnings and dividends: some empirical results. *The Journal of Political Economy,* 70(1), 82-85.
11. Asquith, P., Bruner, R. and Mullins, D., 1983. The Gains to Bidding Firms from Merger. *Journal of Financial Economics,* 11, 121-139.
12. Asquith, P. and Wizman, T.A., 1990. Event risk, covenants, and bondholder returns in leveraged buyouts. *Journal of Financial Economics*, 27, 195-213.

13. Baek J.S., Kang, J.K. and Park K.S., 2004. Corporate governance and firm value: evidence from the Korean financial crisis. *Journal of Financial Economics,* 71, N2.

14. Banz, R., 1981. The relationship between returns and market value of common stock. *Journal of Financial Economics,* 9: 3-18.

15. Basu. S., 1977. Investment performance of common stocks in relation to their Price-Earnings ratio: A test of the efficient market hypothesis. *Journal of Finance,* 32, 663-682.

16. Ball, R. and Brown P., 1968. An empirical evaluation of accounting income numbers. *Journal of Accounting Research,* 6, 159-178.

17. Barberis, N.C., Shleifer, A. and Vishny, R.W., 1998. A model of investor sentiment. *Journal of Financial Economics,* 49, 309-343.

18. Barker, A., 1956. Effective stock splits. *Harvard Business Review,* 34 (1), 101-06.

19. Ball, R. & Brown, P., 1980. Risk and return from equity investments in the Australian mining industry: January 1958-February 1979. *Australian Journal of Management*, 5 (1 & 2), 45-66.

20. Beaver, W. H., 1968. The information content of annual announcements. *Journal of Accounting research Supplement,* 6, 6792.

21. Bernard, V., 1987. Cross-sectional dependence and problems in inference in market-based accounting research. *Journal of Accounting Research,* 25, 1-48.

22. Bhagat, S. and Frost, P.A., 1986. Issuing costs to existing shareholders in competitive and negotiated underwritten public utility equity offerings. *Journal of Financial Economics,* 15, 233-259.

23. Binder, J. J., 1998. The event studies methodology since 1969. *Review of Quantitative Finance and Accounting,* 11, 111-137.

24. Black, F., 1972. Capital market equilibrium with restricted borrowing. *Journal of Business,* 45, 444-455.

25. Blume, M. and Stambaugh, R., 1983. Biases in computed returns: An application to the size effect. *Journal of Financial Economics,* 12, 387-404.

26. Boehmer, E., Musumeci, J. and Poulsen, A., 1991. Event-study methodology under conditions of event-induced variance. *Journal of Financial Economics,* 30, 253-272.

27. Brav, A., Geczy, C., and Gompers, P.A., 1995. Underperformance of seasoned equity offerings revisited. Working paper, Harvard University.

28. Brav, A., and Gompers, P.A., 1997. Myth or reality? The long-run underperformance of initial public offerings: evidence from venture and nonventure capital-backed companies. *Journal of Finance, 52*, 1791-1821.
29. Brav, A., Geczy C. and Gompers, P.A., 2000. Is the abnormal return following equity issuances anomalous? *Journal of Financial Economics,* 56, 209-249.
30. Brock, W. A., Lakonishok, J. and LeBaron, B., 1992. Simple technical trading rules and the stochastic properties of stock returns. *Journal of Finance,* 47, 1731-1764.
31. Brown, S. and Warner, J., 1980. Measuring security price performance. *Journal of Financial Economics*, 8, 205-258.
32. Brown, S. and Warner, J., 1985. Using daily stock returns: the case of event studies. *Journal of Financial Economics*, 14, 3-31.
33. Bruner, R.F, 1999. An analysis of value destruction and recovery in the alliance and proposed merger of Volvo and Renaults *Journal of Financial Economics*, 51, 125-166.
34. Butler, R.W. and Frost, C.A, 1992. An analysis of "exactly distributed" test statistics for multivariate event tests. *Review of Quantitative Finance and Accounting,* 2, 327-341.
35. Buser, S.A. and Chan K.C., 1987. NASDAQ/NMS qualification standards, Ohio registration experience and the price performance of initial public offerings, Columbus, Ohio Department of Commerce and National Association of Securities Dealers, Inc.
36. Byun, J. and Rozeff, M., 2003. Long-run performance after stock splits: 1927-1996, *Journal of Finance*, 58 (3), 1063-1085.
37. Campbell, J.Y., Andrew, W.L. and MacKinlay A.C., 1997. The econometrics of financial markets. Princeton: Princeton University Press.
38. Campbell, C. and Wasley, C., 1993. Measuring security price performance using daily NASDAQ returns. *Journal of Financial Economics,* 33 (1) 73-92.
39. Carter, R.B., Frederick, H.D. and A.K. Singh, 1998. Underwriter reputation, initial returns, and the long-run performance of IPO stocks. *Journal of Finance,* 53 (1), 285-311.
40. Carhart, M.M., 1997. On persistence in mutual fund performance. *Journal of Finance,* 52, 57-82.
41. Chan, L.K.C., Chen, H. and Lakonishok, J., 2002. On mutual fund investment styles. *Review of Financial Studies,* 15, 1407-1437.

42. Chou, P., 2004. Bootstrap tests for multivariate event studies. *Review of Quantitative Finance and Accounting,* 23, 275-290.
43. Collins, D., and Dent W., 1984. A comparison of alternative testing models used in capital market research. *Journal of Accounting Research,* 22, 48-84.
44. Connolly, R.A. and McMillan, H., 1989. Time conditional variances and event studies: The case of capital structure changes. Working paper, University of California.
45. Conrad, J., Cooper, M. and Kaul, G., 2003. Value versus glamour. *Journal of Finance,* 58, 1969-95.
46. Conrad, J. and Kaul, G., 1993. Long-term market overreaction or biases in computed returns? *Journal of Finance,* 48, 39-63.
47. Corrado, C.J, 1989. A nonparametric test for abnormal security performance in event studies. *Journal of Financial Economics*, 23, 385-395.
48. Corrado, C.J. and Truong, C., 2008. Conducting event studies with Asia-Pacific security market data. *Pacific-Basin Finance Journal*, 16 (5), 493-521.
49. Corrado, C.J. and Zivney, T.L., 1992. The specification and power of the sign test in event studies hypothesis tests using daily stock returns. *Journal of Financial and Quantitative Analysis*, 27 (3), 465-478.
50. Cowan, A.R., 1992. Nonparametric event studies tests. *Review of Quantitative Finance and Accounting,* 2, 343-358.
51. Cowen, A.R. and Sergeant, A.M.A., 1996. Trading frequency and event studies test Specification. *Journal of Banking and Finance,* 20, 1731-1757.
52. Cowan, A. R. and Sergeant, A.M.A., 2001. Interacting biases, non-normal return distributions and the performance of tests for long-horizon event studies. *Journal of banking and Finance,* 25, 741-765.
53. Daniel, K., Grinblatt, M., Titman, S. and Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of Finance,* 52(3), 1035-1058.
54. DeBondt, W. F. M. and Thaler, R., 1985. Does the stock market overreact? *Journal of Finance,* 40, 793-805.

55. DeFond, M., Hann, R. N. and Hu, X., 2005. Does the market value financial expertise on audit committees of boards of directors? *Journal of Accounting Research, 43*(2), 153-193.

56. DeFusco, R.A., Johnson, R.R. and Zorn, T.S., 1990. The effect of executive stock option plans on stockholders and bondholders. *Journal of Finance,* 45, 617-627.

57. Desai, H. and Jain, P., 1997. Long-run common stock returns following splits and reverse splits. *Journal of Business,* 70, 409-433.

58. Dodd, P. and Ruback, R., 1977. Tender offers and stockholder returns: An empirical analysis. *Journal of Financial Economics,* 5 (3), 351-373.

59. Dodd,P., 1980. Merger proposals, management discretion and stockholder wealth. *Journal of Financial Economics,* 8, 105-37.

60. Dodd, P. and Warner, J.B., 1983. On corporate governance: A study of proxy contests. *Journal of Financial Economics,* 11 (1), 401-438.

61. Dharan, B. and Ikenberry, D., 1995. The long-run negative drift of post-listing stock returns. *Journal of Finance,* 50, 1547-1574.

62. Dhrymes, P., Friend, F. and Gultekin, M., 1984. A critical re-examination of the empirical evidence on the arbitrage pricing theory. *Journal of Finance*, 39 (2), 323-346.

63. D'Mello, R., Tawatnuntacha, O. and Yaman, D., 2002. Does the sequence of seasoned equity offerings matter? *Financial Management,* Vol. 32, No. 4

64. Dutta, S. and Jog, V., 2009. The long-term performance of acquiring firms: A re-examination of an anomaly. *Journal of Banking & Finance,* 33, 1400-1412.

65. Eckbo, B. E., Masulis, R.W. and Norli, Ø., 2000. Seasoned Public Of- ferings: Resolution of the New Issues Puzzle. *Journal of Financial Economics*, 56, 251-291.

66. Edgington, E.S., 1995. Randomization tests (Third Edition). New York: Marcel Dekker.

67. Efron, B., 1982. The jackknife, the bootstrap, and other resampling plans, Monograph 38. *CBMS-NSF Regional Conference Series in Applied Mathematics,* Philadelphia.

68. Efron, B. and Tibshirani, R. J., 1993. An introduction to the bootstrap. New York: Chapman and Hall.

69. Espenlaub, S., Gregory, A. and Tonks, I. 1998.Testing the robustness of long-term under-performance of UK initial public offerings. *Discussion Paper in Business and Management,* 98/07.

70. Fama, E., 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 25 (2), 383-417.

71. Fama, E.F., 1976. Foundations of Finance. New York: Basic Books, Inc.

72. Fama, E.F., 1991. Efficient capital markets: II. *Journal of Finance,* 46: 1575-1617.

73. Fama, E.F., 1998. Market efficiency, long-run returns, and behavioural finance. *Journal of Financial Economics*, 49(3), 283-306.

74. Fama, E.F., Fisher, L, Jensen, M. and Roll, R., 1969. The adjustment of stock prices to new information. *International Economic Review*, 10, 1-21.

75. Fama, E. and French K., 1992. The cross-section of expected stock returns. *Journal of Finance* 32, 427-465.

76. Fama, E.F. and French, K.R., 1993. Common risk factors in the returns on stocks and bond. *Journal of Financial Economics,* 39, 3-56.

77. Fama, E.F. and French, K.R., 1998. Value versus growth: The international evidence. *Journal of Finance 53,* 1975-1999.

78. Fama, E.F. and MacBeth, J., 1973. Risk, return and equilibrium: Empirical tests. *Journal of Political Economy*, 81, 607-636.

79. Firth, M., 1997. An analysis of the stock market performance of new issues in New Zealand. *Pacific-Basin Finance Journal,* 5, 63-85.

80. Fisher, R.A. (1921). On the "probable error" of a coefficient of correlation deduced from a small sample. *Metron, 1,* 3-32.

81. Franks, J., Harris, R. and Titman, S., 1991. The postmerger share-price performance of acquiring firms. *Journal of Financial Economics*, 29, 81-96.

82. Freidlin, B. and Gastwirth, J. L., 2000. Should the median test be retired from general use? *The American Statistician,* 54(3), 161-64.

83. Giammarino, R.M., Lewis, T., 1989. A theory of negotiated equity financing. *Review of Financial Studies 1,* 265-288.

84. Gompers, P. A. and Lemer, J., 2003. The really long-run performance of Initial Public Offerings: The pre-NASDAQ evidence. *Journal of Finance,* 58, 1355-1392.

85. Gregory, A., 1997. An examination of the long run performance of UK acquiring firms. *Journal of Business Finance and Accounting,* 24, 971-1002.

86. Gregory A., Guermat, C. and Shawawreh, F., 2010, UK IPOs: Long Run Returns, Behavioural Timing and Pseudo Timing. *Journal of Business Finance & Accounting*, 37 (5-6), 612-647.

87. Gregory, A., Tharyan R. and Huang A., 2009. The Fama-French and momentum portfolios and factors in the UK. *University of Exeter Business School, Xfi Centre for Finance and Investment Paper* No. 09/05.

88. Gugler, K., Mueller, D.C. and Yurtoglu, B.B., 2004. Marginal q, Tobin's q, Cash Flow, and Investment. *Southern Economic Journal*, 70, 512-531.

89. Hall, P., 1992. On the removal of skewness by transformation. *Journal of the Royal*

90. *Statistical Society,* Series B 54, 221-228.

91. Hall, P. and Horowitz, J.L., 1996. Bootstrap critical values for tests based on generalized method of-moments estimators. *Econometrica,* 64, 891-916.

92. Haw, I., Pastena, V.S. and Lilien, S.B., 1990. Market manifestation of nonpublic information prior to mergers: The effect of capital structure. *Accounting Review,* 65 (2) 432-451.

93. Hein, S.E. and Westfall P., 2004. Improving tests of abnormal returns by bootstrapping the multivariate regression model with event parameters. *Journal of Financial Econometrics,* 2, 451-471.

94. Higson, C. and Elliott, J., 1998. Post-takeover returns - the UK evidence. *Journal of Empirical Finance,* 5, 27-46.

95. Hou, K., Robinson, D.T. and Olsson, T., 2000. Do takeovers increase stockholder value? Working paper, University of Chicago, IL, Chicago.

96. Huber, P. J. 1947. The behavior of maximum likelihood estimation under nonstandard conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1,* Lecam, L. M. and Neyman, J. editors. University of California Press, p 221-233

97. Ikenberry, D., Lakonishok, J. and Vermaelen, T., 1995. Market underreaction to open market share repurchases. *Journal of Financial Economics,* 39, 181-208.

98. Ikenberry, D., Lakonishok, J. and Vermaelen, T., 2000. Stock repurchases in Canada: performance and strategic trading. *The Journal of Finance*, 55, 2373-2397.

99. Ikenberry, D. L., Rankine, G. and Stice, E.K., 1996. What do stock splits really signal? *Journal of Financial and Quantitative Analysis,* 31, 357-375.

100. Ibbotson, R.G., 1975. Price performance of common stock new issues. *Journal of Financial Economics*, 3, 235-272.

101. Jaffe, J.F., 1974. Special information and insider trading. *Journal of Business,* 7, 410-428.

102. Jain, P., 1986. Analyses of the distribution of market model prediction errors for daily returns data. *Journal of Accounting Research,* 24 (1), 76-96.

103. Jayaraman, N. and Shastri, K., 1993. The effects of the announcements t s o f dividend increases on stock return volatility: the evidence from the options market. *Journal of Business Finance & Accounting,* 20, 673-685.

104. Jegadeesh, N., 2000. Long term performance of seasoned equity offerings: Benchmark errors and biases in expectation. *Financial Management,* 29, 3-50.

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

106. Jensen, M.C., & Ruback, R.S., 1983. The market for corporate control. *Journal of Financial Economics*, 11, 5-50.

107. Johnson, N.J., 1978. Modified t tests and confidence intervals for asymmetrical populations. *Journal of the American Statistical Association,* 73, 536 - 544.

108. Jones, C.P. and Litzenberger, R.H., 1970. Quarterly earnings reports and intermediate stock price trends. *Journal of Finance*, 25, 143-148.

109. Karafiath, I., 2009. Detecting cumulative abnormal volume: a comparison of event studies methods. *Applied Economics Letters*, 16, 797-802.

110. Karafiath, I. and Spencer, D.E., 1991. Statistical inference in multi-period event studies: An assessment. *Review of Quantitative Finance and Accounting*, 1, 353-371.

111. Keloharju, M., 1993. The winner’s curse, legal liability, and the long-run performance of initial public offerings in Finland. *Journal of Financial Economics,* 34, 251-277.

112. Kim, E.H. and Schatzberg, J.D., 1987. Voluntary corporate liquidations. *Journal of Financial Economics,* 19 (2), 311-328.

113. Kothari, S.P. and Warner, J.B., 2008. Econometrics of event studies, in: Eckbo, B. Espen (ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol. 1, Elsevier/NorthHolland, 3-36.

114. Kothari, S.P., Lewellen, J.W. and Warner, J.B., 2006. Stock returns, aggregate earnings surprises, and behavioural finance. *Journal of Financial Economics,* 79, 537-568.

115. Kothari, S.P. and Warner, J.B., 1997. Measuring long-horizon security price performance. *Journal of Financial Economics,* 43, 301-339.

116. Kramer, L. A., 2001. Alternative methods for robust analysis in event studies applications, in Cheng-Few Lee, ed.: Advances in Investment Analysis and Portfolio Management, vol. 8, pp. 109-132 (Elsevier: North-Holland).

117. Lakonishok, J., and Vermaelen, T., 1990. Anomalous price behaviour around repurchase tender offers. *Journal of Finance*, 45, 455-477.

118. Lakonishok, J., Shleifer, A. and Vishny, R., 1994. Contrarian investment, extrapolation and risk." *Journal of Finance,* 49 (5), 1541-1578.

119. Langetieg T., 1978. An application of a three factor performance index to measure stockholder gains from merger. *Journal of Financial Economics,* 6, 365-383.

120. Lee, I., 1997. Do firms knowingly sell overvalued equity? *Journal of Finance, 52*, 1439-1466.

121. Lee, P., Taylor, S. and Walter, T., 1996. Australian IPO underpricing in the short- and long-run. *Journal of Banking and Finance,* 20, 1189-1210.

122. Levis, M., 1993. The long-run performance of initial public offerings: The U.K. experience 1980-88. *Financial Management*, 22, 28-41.

123. Liu, W. and Strong, N., 2008. Biases in decomposing holding-period portfolio returns. *Review of Financial Studies*, 21, 2243-2274.

124. Lintner, J., 1965. The valuation of risk assets and selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47, 13-37.

125. Loderer, C. and Martin, K., 1992. Post Acquisition Performance of Acquiring Firms. *Financial Management,* 21(3), 69-79.

126. Lougharan, T., 1993. NYSE vs NASDAQ returns: Market microstructure or the poor performance of initial public offerings. *Journal of Financial Economics,* 33, 241-260.

127. Loughran, T and Ritter, R.J., 1995. The new issues puzzle. *Journal of Finance,* 1, 23-51.

128. Loughran T., Ritter, J.R, 1997. The operating performance of firms conducting seasoned equity offerings. *Journal of Finance,* 52, 1823-1850.

129. Loughran, T. and Ritter, J.R., 2000. Uniformly least powerful tests of market efficiency. *Journal of Financial Economics,* 55, 361-389.

130. Loughran, T. and Vijh, A.M., 1997. Do long-term shareholders benefit from corporate acquisitions? *Journal of Finance,* 52, 1765-1790.

131. Lyon, J. and., 1997. Detecting long-run abnormal stock returns: the empirical power and specification of test statistics. *Journal of Financial Economics,* 43, 341-372.

132. Lyon, J., Barber, B., and Tsai, C., 1999. Improved methods of tests of long-horizon abnormal stock returns. *Journal of Finance*, *54*, 165-201.

133. MacKinlay, A.C., 1997. Event studies in economics and finance. *Journal of Economic Literature,* 35, 13-39.

134. Magenheim, E.B. and Mueller, D.C., 1988. Are Acquiring-Firm Shareholders Better off after an Acquisition, in John C. Coffee, Jr., Louis Lowenstein, and Susan Rose-Ackerman, editors, Knights, Raiders and Targets: The Impact of the Hostile Takeover, New York: Oxford University Press.

135. Mandelker, G., 1974. Risk and return: the case of merging firms. *Journal of Financial Economics,* 1, 303-335.

136. Markowitz, H., 1959. Portfolio selection: efficient diversification of investments. New York: John Wiley & Sons, Inc.

137. Marais, M., 1984.An application of the bootstrap method to the analysis of squared standardized market model prediction errors. *Journal of Accounting Research*, 22 Supplement, 34-54.

138. Masulis, R.W., 1980. The effects of capital structure changes on security prices: a study of exchange offers. *Journal of Financial Economics,* 8, 139-178.

139. Maynes, E. and Rumsey, J., 1993. Conducting event studies with thinly traded stocks. *Journal of Banking and Finance,* 17, 145-157.

140. McWilliams, V.B., 1990. Managerial share ownership and the stock price effects of anti-takeover amendment proposals. *Journal of Finance,* 45(5), 1627-1640.

141. Michaely, R., Thaler, R.H. and Womack, K.L, 1995. Price reactions to dividend initiations and omissions: overreaction or drift? *Journal of Finance*, 50, 573-608.

142. Mikkelson, W. and Partch, M., 1986. Valuation effects of security offerings and the issuance process. *Journal of Financial Economics,* 15, 31-60.

143. Mitchell, M.L. and Lehn, K., 1990. Do bad bidders become good targets? *Journal of Political Economy,* 98, 372-398.

144. Mitchell, M.L and Stafford, E., 1997. Managerial decisions and long-term stock price performance. CRSP working papers 453, Centre for Research in Security Prices, Graduate School of Business, University of Chicago.

145. Mitchell, M.L and Stafford, E., 1999. Managerial decisions and long term stock price performance. Working Paper, Harvard University.

146. Myers, J. and Bakay, A., 1948. Influence of stock split-ups on market price. *Harvard Business Review,* 26, 251-55.

147. McGuinness, P., 1993. The market valuation of initial public offering in Hong Kong. *Applied Financial Economics,* 3, 267-281.

148. Mitchell, M. and Stafford, E., 2000. Managerial decisions and long-term stock price performance. *Journal of Business,* 73, 287-329.

149. Moeller, S., Schlingemann, F.P. and Stulz, R.M., 2003. Do shareholders of acquiring firms gain from acquisitions? Working paper, Ohio State University.

150. Parsons, J.E. and Raviv, A., 1985.Underpricing of seasoned issues. *Journal of Financial Economics*, 14, 377-397.

151. Patell, J., 1976. Corporate forecasts of earnings per share and stock price behaviour: Empirical tests. *Journal of Accounting Research*, 14, 246-276.

152. Rajan, R. and Servaes, H., 1997. Analyst following of Initial Public Offerings. *Journal of Finance*, 52, 507-529.Rau, P.R. and Vermaelen, T., 1998. Glamour, value and the post-acquisition performance of acquiring firms. *Journal of Financial Economics,* 49, 223-253.

153. Ritter, J.R., 1991. The Long-run performance of initial public offerings. *Journal of Finance,* 46, 3-27.

154. Roll, R., 1977. A critique of the asset pricing theory's tests part I: On past and potential testability of the theory. *Journal of Financial Economics,* 4, 129-176.

155. Roll, R., 1983. On computing mean returns and the small firm premium. *Journal of Financial Economics*, 12, 371-386.

156. Ross, S.A., 1976. The arbitrage theory of capital asset pricing. *Journal of Economic Theory,* 13, 341-360.

157. Schultz, P., 2003. Pseudo market timing and the long-run underperformance of IPOs. *Journal of Finance,* 58, 483-517.

158. Schumann, L., 1988. State regulation of takeovers and shareholder wealth: The case of New York's 1985 takeover statutes. *Rand Journal of Economics,* 19, 557-567.

159. Serra, A.P., 2002. Event studies tests: A brief survey. Working Paper FEP 117, University of Porto.

160. Sharpe,W.F., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance*, 19, 425-442.

161. Slovin, M.B., Shushka, M.E. and Polonchek, J.A., 1991. The information content of multiple seasoned common stock offerings by bank holding companies. *Journal of Banking and Finance,* 15, 633-646.

162. Smith, C.W., 1977. Alternative methods for raising capital: Rights versus underwritten offerings. *Journal of Financial Economics,* 5, 273-307.

163. Spiess, D. K., and Affleck-Graves J., 1995. Underperformance in long-run stock returns following seasoned equity offering. *Journal of Financial Economics,* 38, 243-267.

164. Stattman, D., 1980. Book values and stock returns. *Chicago MBA: A Journal of Selected Papers,* 5, 25-45.

165. Stoll, H.R. and Curley, A.J, 1970. Small business and the new issues market for equities. *Journal of Financial and Quantitative Analysis,* 5, 309-322.

166. Sutton, C.D., 1993. Computer-intensive methods for tests about the mean of an asymmetrical distribution. *Journal of the American Statistical Association,* 88, 802-810.

167. Tuch, C. and O'Sullivan, N., 2007. The impact of acquisitions on firm performance: A review of the evidence. *International Journal of Management Reviews,* 9 (2), 141-170.

168. Wermers, R., 2000. Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses. *Journal of Finance*, 55, 1655-1695.

169. Weber, J., Willenborg, M. and Zhang. J., 2008. Does auditor reputation matter? The case of KPMG Germany and ComROAD AG. *Journal of Accounting Research*, 46 (4), 941-972.

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

171. White, H., 1982. Maximum likelihood estimation of misspecified models. *Econometrica 50*, 1-26

172. Yeo, G. H. H., Chen, S. S., Ho, K. W. and Lee, C. F., 1999. Effects of executive share option plans on shareholder wealth and firm performance: The Singapore evidence. *The Financial Review*, 34 (2), 1-20.