INVESTIGATING LONG-RUN STOCK RETURNS AFTER CORPORATE EVENTS: THE UK EVIDENCE

Anupam Dutta*

Abstract

The objective of this paper is to assess the robustness of the existing long-run event study methodologies in the UK stock market. In doing so, the study employs the buy-and-hold abnormal return approach and the calendar time portfolio method to identify the long-term abnormal performance following corporate events. Although many recent studies consider the application of these two widely used approaches, each of the methods is a subject to criticisms. This paper uses the standardized calendar time approach (SCTA) which presents a number of important improvements over the traditional calendar time methodology. The empirical analysis reveals that all the traditional methodologies perform well in the UK security market. Our findings further report that SCTA documents better specification and power than the conventional approaches.

Keywords: Long-run Anomalies, Standardized Abnormal Returns, Specification Issue, Power Issue

***Department of Mathematics and Statistics, University of Vaasa, Wolffintie 34, Vaasa-65200, Finland
Tel: +358458690502
Email: adutta@uwasa.fi**

1 Introduction

Many recent studies investigate the long-term performance of firms after certain corporate events such as IPOs, SEOs, or repurchases. The key articles in this area include Ritter (1991), Barber and Lyon (1997), Kothari and Warner (1997), Fama (1998), Lyon, Barber, and Tsai (1999), Mitchell and Stafford (2000), Loughran and Ritter (2000), Boehme and Sorescu (2002), and Jegadeesh and Karceski (2009). While long-horizon event studies have enjoyed many advances over the years, all the elementary papers focus on United States security markets. Although a number of studies investigate the long-term performance of the UK IPOs (Levis, M., 1993, Espenlaub, Gregory and Tonks, 2000 etc.), not a single simulation study concentrating on this security market is found in the literature. Therefore, the objective of this paper is to conceal such gaps by conducting a simulation study with the UK stock market data. Since the choice of proper methodology plays a key role in investigating the long-run performance, we, like other fundamental studies, employ Buy-and-Hold Abnormal Return (BHAR) approach and Calendar Time Portfolio (CTP) method to measure the long-run anomalies.

However, previous studies document that each of these widely used methods has a number of potential pitfalls. Fama (1998), for example, reports that the BHAR method ignores the issue of potential cross-sectional correlation of event-firm abnormal returns and hence produces misspecified test statistics. Loughran and Ritter (2000), on the other hand, claim that CTP approach has low power to identify the abnormal performance because it gives equal weight to each month, regardless of whether the month has heavy or light event activities. Following the work of Dutta (2014), this paper considers applying the Standardized Calendar Time Approach (SCTA) where we first standardize the abnormal returns for each of the event firms forming the monthly portfolios and then each portfolio is weighted such that periods of heavy event activity receive more weight than periods of low event activity. However, employing standardized abnormal returns is well-documented in the literature. For example, Jaffe (1974) and Mandelker (1974) employ standardized portfolio returns for investigating the long-run abnormal performance. Fama (1998) also suggests to standardize the abnormal returns to resolve the issues raised by Loughran and Ritter (2000). Later, Mitchell and Stafford (2000) use standardized abnormal returns to alleviate the heteroscedasticity problem that often occurs in CTP approach due to the varying portfolio construction.

The empirical analysis reveals that all these methodologies are robust in the UK stock market as well. Our findings also report that the standardized calendar time approach produces reasonably well-specified test statistics in all types of nonrandom samples. The results further show that SCTA documents better power than the existing approaches to identify the long-term abnormal performance. One striking output of our analysis is that the test statistics based on the Fama-French three-factor model are not well specified even in random samples. However, the simulated result shows that the adjusted three-factor model controls well for size and book-to-market ratio biases. The BHAR approach, on the other hand, yields reasonably well-specified test statistics when the control firm approach is employed. While using a reference portfolio as a benchmark, the BHAR methodology does not produce well-specified test statistics. These results are consistent with those reported by Lyon, Barber, and Tsai (1999).

This paper extends the prior literature in three aspects. First, it uses simulated results to assess the performance of the existing long-run event study methodologies using the UK security market data. Second, the study employs a variant of calendar time methodology which yields well-specified test statistics in nonrandom samples. Third, this refined calendar time approach improves the power while inspecting the long-term abnormal performance. However, one major limitation of our proposed approach is that it does not yield well-specified test statistics (not reported in the table) for samples based on pre-event return performance. Lyon, Barber, and Tsai (1999) also report that the buy-and-hold abnormal return approach as well as the traditional calendar time method have this limitation. Therefore, further filtering of the existing methodologies is needed to resolve this problem. Another important drawback of our study is that it does not present the results based on industry-clustered samples due to the non-availability of industry codes.

The remainder of the paper is structured as follows. Section 2 outlines the data and methodology. Section 3 explains the simulation procedure. Section 4 discusses the specification of the tests. Section 5 reports power of the tests and Section 6 concludes the paper.

2 Data and methodology

We obtain stock prices, market value (MV) or size and book-to-market (BM) value data of the UK stock market from DataStream. The sample period ranges from July 1983 to December 2013.

In this paper, we construct 25 size-BM portfolios as expected return benchmarks. In doing so, at the end of June of year *t*, firms are quantiled into five groups on the basis of their market values. Firms are further quantiled into five groups based on their book-to-market ratios. However, we also consider a size-BM-matched control firm to calculate the abnormal returns. Identifying this control firm is a 2-step procedure. First, we identify all the firms with a market value of equity between 70% and 130% of the sample firm at the most recent end of June. Then from this set of firms, we choose the firm with BM closest to that of the sample firm as of the previous December.

2.1 Standardized calendar time approach (SCTA)

The calculation of mean monthly calendar time abnormal return (CTAR) is the following:

$$MMCTAR = \frac{1}{T}\sum_{t=1}^{T} CTAR_t \qquad (1)$$

Where

$$CTAR_t = R_{pt} - E(R_{pt}) \qquad (2)$$

Within this framework, R_{pt} is the monthly return on the portfolio of event firms, $E(R_{pt})$ is the expected return on the event portfolio which is proxied by the raw return on either a reference portfolio or a control firm and T is the total number of months in the sample period.

Following the work of Dutta (2014), this paper uses standardized abnormal returns to compute the monthly CTARs. Dutta argues that since a number of firms in the sample might produce volatile returns, it would cause the distributions of long-run returns to have fat tails. Consequently, test statistics will be seriously misspecified. But standardizing the abnormal returns by their volatility measures is a possible solution to this problem. Although Dutta uses simple return, we consider log return to minimize the skewness problem. Bessembinder and Zhang (2013) and Knif, Kolari and Pynnönen (2013) also document that employing log returns produces better specified test statistics.

The construction of the monthly portfolios in the standardized calendar time approach consists of two steps. We first calculate the standardized abnormal returns for each of the sample firms. In doing so, the abnormal returns for firm *i* are computed as $\varepsilon_{it} = r_{it} - E(r_{it}); t = 1, \ldots, H$, where r_{it} denotes the log return on event firm *i* in the calendar month *t* and $E(r_{it})$ is the expected return which is proxied by the raw return either 25 size-BM reference portfolios or a size-BM matched control firm and H is the holding period which equals 12, 24 or 36 months. The next task is to estimate the event-portfolio residual variances using the H-month residuals computed as monthly differences of *i*-th event firm returns and control firm returns. Dividing ε_{it} by the estimate of its standard deviation yields the corresponding standardized abnormal return, say, z_{it}, for event firm *i* in month *t*. Now let N_t refer to the number of event firms in the calendar month *t*. We then calculate the calendar time abnormal return for portfolio *t* as:

$$CTAR_t = \sum_{t=1}^{N_t} x_{it} z_{it} \quad (3)$$

The weight x_{it} equals $\frac{1}{N_t}$ when the abnormal returns are equally-weighted and $\frac{MV_{it}}{\sum MV_{it}}$ when the abnormal returns are value-weighted by size.

We, like Dutta (2014), also assign weights to each of the monthly CTARs by $1/\sqrt{(\sum_{t=1}^{T} x_{it}^2)}$. For instance, when the abnormal returns are equally weighted i.e., when $x_{it} = \frac{1}{N_t}$, then $1/\sqrt{(\sum_{t=1}^{T} x_{it}^2)} = \sqrt{N_t}$. This weighting scheme is lucrative as it gives more loadings to periods of heavy event activity than the periods of low event activity. Now the grand mean monthly abnormal return, denoted by $\overline{CTAR}$, is calculated as:

$$\overline{CTAR} = \frac{1}{T} \sum_{1}^{T} CTAR_t \quad (4)$$

While finding $\overline{CTAR}$, it might be the case that a number of portfolios do not contain any event firm. In such situations, those months are dropped from the analysis. To test the null hypothesis of no abnormal performance, the *t*-statistic of $\overline{CTAR}$ is computed by using the intertemporal standard deviation of the monthly CTARs defined in equation (3).

2.2 Buy-and-hold abnormal return (BHAR)

An *H*-month BHAR for event firm *i* is defined as:

$$BHAR_{iH} = \prod_{t=1}^{H}(1 + R_{it}) - \prod_{t=1}^{H}(1 + R_{Bt}) \quad (5)$$

R_{it} denotes the return on event firm *i* at time *t* and R_{Bt} indicates the return on a control firm. To test the null hypothesis that the mean buy-and-hold return equals zero, the conventional *t*-statistic is given by:

$$t_{BHAR} = \frac{\overline{BHAR_H}}{\sigma(BHAR_H)/\sqrt{n}} \quad (6)$$

$\overline{BHAR_H}$ implies the sample mean and $\sigma(BHAR_H)$ refers to the cross-sectional sample standard deviation of abnormal returns for the sample containing *n* firms.

However, the earlier studies such as Mitchell and Stafford (2000), Boehme and Sorescu (2002), Jegadeesh and Karceski (2009) report that the BHAR approach does not control well for the cross-sectional correlation among individual firms in nonrandom samples and thus yields misspecified *t*-statistics. Moreover, the test statistics based on BHARs also have this misspecification problem, since the distribution of BHARs is highly skewed. Though bootstrapping corrects for the skewness problem to some extent, it ignores the cross-sectional dependence of abnormal returns.

2.3 Fama-French three-factor model

For each calendar month *t*, we form portfolios consisting of all sample firms that have participated in the event within the last *H* months, where *H* equals 12, 36, or 60 in our study. For each calendar month, the portfolios are rebalanced, i.e., the firms that reach the end of their *H* -month period drop out and new firms that have just executed a transaction are added. We then calculate the portfolio mean monthly abnormal return α_p by regressing its excess return on the three Fama-French factors:

$$R_{pt} - R_{ft} = \alpha_p + \beta_p(R_{mt} - R_{ft}) + s_p SMB_t + h_p HML_t + e_{pt} \quad (7)$$

R_{pt} is the equal or value-weighted return on portfolio *t*, R_{ft} is the risk-free rate, $(R_{mt} - R_{ft})$ is the excess return of the market, *SMB* is the difference between the return on the portfolio of small stocks and big stocks, *HML* is the difference between the return on the portfolio of high and low book-to-market stocks, α_p measures the mean monthly abnormal return of the calendar time portfolio which is zero under the null hypothesis of no abnormal performance and β_p, s_p and $h_p$are sensitivities of the event portfolio to the three factors.

However, since the number of firms changes over the sample period, this may cause the error term to be heteroskedastic and hence the ordinary least squares estimate becomes inefficient. Fama (1998), therefore, suggests to apply the weighted least squares technique instead of ordinary least squares to control for heteroskedasticity. In this study, we estimate regression (7) using weightedleast squares (WLS) procedures. Monthly returns in the WLS model are weighted by $\sqrt{N_t}$, where N_t stands for the number of event firms in month *t*.

2.4 Adjusted Fama-French three-factor model

Fama and French (1993) document that the traditional three-factor model is not able to completely explain the cross section of stock returns. However, Mitchell and Stafford (2000) and later Boehme and Sorescu (2002) refine this three-factor model to deal with the bad model problem. In this paper, we also try to modify the conventional Fama-French three-factor model to moderate the size and book-to-market ratio biases. Our adjusted three-factor model assumes the following form:

$$(R_{event} - R_{control})_{pt} = \alpha_p + \beta_p(R_{mt} - R_{ft}) + s_p SMB_t + h_p HML_t + e_{pt} \quad (8)$$

$(R_{event} - R_{control})_{pt}$ is the equal- or value-weighted monthly portfolio return between the simple returns of each event firm and its size-BM matched control firm. Moreover, for portfolio t, $(R_{event} - R_{control})_{pt}$ contains those firms whose event period includes the month t. In this adjusted model, α_p is a measure of long-term abnormal performance which is zero under the null hypothesis that no abnormal performance exists. Now, to test this null hypothesis, the t-statistic is given as:

$$t = \frac{\hat{\alpha}_p}{s(\hat{\alpha}_p)}$$

$\hat{\alpha}_p$ is an estimator of α_p, and $s(\hat{\alpha}_p)$ is the corresponding standard error of $\hat{\alpha}_p$.

3 Simulation method

Following the work of Lyon, Barber, and Tsai (1999), we randomly select 1000 samples of 200 event months without replacement to assess the specification of the employed methodologies. For each of these 200 event months, we randomly draw one stock from the population of all stocks that are active in the database for that month. For a well-specified test statistic, 1000α tests reject the null hypothesis. A test is conservative if fewer than 1000α null hypotheses are rejected and is anticonservative if more than 1000α null hypotheses are rejected. Based on this procedure, we test the specification of the t-statistic at 5% theoretical levels of significance. A well-specified null hypothesis rejects the null at the theoretical rejection level in favor of the alternative hypothesis of negative (positive) abnormal returns in $1000\alpha/2$ samples.

4 Test specification

In this section, we report the specification of various methodologies under consideration. We first discuss the results in random samples. Later, we consider different types of nonrandom samples based on firm size, book-to-market ratio and overlapping returns.

4.1 Random samples

Table 1 shows the rejection rates in 1000 simulations with a random sample of 200 firms. These results indicate that the BHAR method based on size-BM control firms yields well-specified test statistics in each of the three investment horizons. However, the test statistics are severely missepcified when the BHARs are calculated using the reference portfolios. For example, for a three-year holding period the rejection rates at the 5% level of significance are 3.6% and 0.9% for control firm approach and 4.2% and 0% for reference portfolio method.

The numbers presented in Table 1 further reveal that among the calendar time portfolio methods considered in this paper, our proposed approach is better specified than the rest in each case. One striking finding is that test statistics based on the Fama-French three-factor (henceforth FF3F) model are misspecified regardless of whether equally-weighted or value-weighted portfolios are employed. These findings conform to those reported by Yan (2012) in his empirical research. Moreover, while using the adjusted three-factor model, the size improves. For example, for a five-year holding period and with value-weighted portfolios, the rejection rates at 5% level of significance are 5.2% and 2.0% for FF3F and 2.8% and 1.6% for the adjusted version.

4.2 Nonrandom samples

4.2.1 Firm size

To assess the effect of size-based sampling biases on the employed methods, we randomly choose 1000 samples separately from the largest size decile and smallest size decile. Tables 2A and 2B display these results. The empirical analysis shows that SCTA produces well-specified test statistics for size-based samples. For instance, for a three-year horizon and with large firms and value-weighted portfolios, the rejection rates for SCTA at 5% level of significance are 1.8% and 3.0% when reference portfolios are used and 2.6% and 2.8% when the control firm approach is considered. The BHAR methodology, however, yields either negatively or positively skewed test statistics when the reference portfolio approach is taken into consideration. The size improves if the BHAR is estimated on the basis of control firms. The rejection rates for FF3F, on the other hand, are much higher than the theoretical levels. But, the level of misspecification decreases when the adjusted Fama-French three-factor model is used as an alternative.

Table 1. Specification of tests in random samples

Methods	Benchmark	1 Year		3 Years		5 Years	
		2.5	97.5	2.5	97.5	2.5	97.5
Panel A: Equally-Weighted Portfolios							
Standardized Calendar Time Approach	Size-BM Portfolio	2.4	1.2	2.0	2.8	2.8	2.8
Standardized Calendar Time Approach	Size-BM Control Firm	2.0	0.4	2.8	1.2	1.6	2.4
Buy-and-Hold Abnormal Return Method	Size-BM Portfolio	9.2*	0.0	4.2*	0.0	3.9*	0.0
Buy-and-Hold Abnormal Return Method	Size-BM Control Firm	2.0	1.4	3.6	0.9	3.0	2.3
Fama-French Three-Factor Model	Not Applicable	4.4*	0.0	3.8*	0.0	4.0*	0.0
Adjusted Fama-French Three-Factor Model	Size-BM Control Firm	3.9*	1.7	3.6	2.8	1.2	2.3
Panel B: Value-Weighted Portfolios							
Standardized Calendar Time Approach	Size-BM Portfolio	2.0	2.8	2.8	3.2	2.4	2.0
Standardized Calendar Time Approach	Size-BM Control Firm	2.8	1.6	3.6	2.0	2.8	3.1
Fama-French Three-Factor Model	Not Applicable	0.8	3.0	6.4*	2.6	5.2*	2.0
Adjusted Fama-French Three-Factor Model	Size-BM Control Firm	5.6*	1.2	3.2	0.7	2.8	1.6

Note: This table presents the percentages of 1000 random samples of 200 firms that reject the null hypothesis of no annual, three-year and five-year abnormal returns at 5% level of significance. Panel A and Panel B indicate the specification of tests for equally- and value-weighted portfolios respectively. The numbers marked with * suggest that the empirical size is significantly different from the 5% significance level.

Table 2A. Specification of tests in samples with small firms

Methods	Benchmark	1 Year		3 Years		5 Years	
		2.5	97.5	2.5	97.5	2.5	97.5
Panel A: Equally-Weighted Portfolios							
Standardized Calendar Time Approach	Size-BM Portfolio	0.8	7.2*	1.8	3.0	2.1	3.8*
Standardized Calendar Time Approach	Size-BM Control Firm	1.4	6.1*	2.6	2.8	0.9	3.1
Buy-and-Hold Abnormal Return Method	Size-BM Portfolio	2.8	36.4*	0.0	25.6*	0.0	15.6*
Buy-and-Hold Abnormal Return Method	Size-BM Control Firm	4.8*	1.2	2.0	1.3	3.6	2.1
Fama-French Three-Factor Model	Not Applicable	0.0	42.2*	0.3	34.8*	1.1	22.0*
Adjusted Fama-French Three-Factor Model	Size-BM Control Firm	0.4	3.8*	3.6	1.6	0.8	6.4*
Panel B: Value-Weighted Portfolios							
Standardized Calendar Time Approach	Size-BM Portfolio	0.0	6.8*	2.8	3.4	3.1	4.7*
Standardized Calendar Time Approach	Size-BM Control Firm	1.6	3.4	2.1	2.4	2.3	3.0
Fama-French Three-Factor Model	Not Applicable	0.0	34.8*	0.0	16.4*	0.2	9.2*
Adjusted Fama-French Three-Factor Model	Size-BM Control Firm	2.1	4.8*	1.9	2.8	3.2	3.9*

Note: This table presents the percentages of 1000 samples of 200 large firms that reject the null hypothesis of no annual, three-year and five-year abnormal returns at 5% level of significance. Panel A and Panel B indicate the specification of tests for equally- and value-weighted portfolios respectively. The numbers marked with * suggest that the empirical size is significantly different from the 5% significance level.

Table 2B. Specification of tests in samples with small firms

Methods	Benchmark	1 Year		3 Years		5 Years	
		2.5	97.5	2.5	97.5	2.5	97.5
Panel A: Equally-Weighted Portfolios							
Standardized Calendar Time Approach	Size-BM Portfolio	5.4*	0.4	2.4	2.6	3.6	3.6
Standardized Calendar Time Approach	Size-BM Control Firm	6.8*	1.2	3.2	0.8	1.6	2.8
Buy-and-Hold Abnormal Return Method	Size-BM Portfolio	21.2*	0.0	15.6*	0.0	11.8*	0.3
Buy-and-Hold Abnormal Return Method	Size-BM Control Firm	3.4	1.8	2.4	2.0	4.0*	3.2
Fama-French Three-Factor Model	Not Applicable	8.8*	0.2	8.0*	3.1	12.4*	3.5
Adjusted Fama-French Three-Factor Model	Size-BM Control Firm	4.0*	0.8	6.0*	1.4	5.6*	2.3
Panel B: Value-Weighted Portfolios							
Standardized Calendar Time Approach	Size-BM Portfolio	6.0*	0.0	4.1*	3.1	2.7	3.5
Standardized Calendar Time Approach	Size-BM Control Firm	4.4*	0.2	2.8	1.3	1.6	2.0
Fama-French Three-Factor Model	Not Applicable	9.2*	0.6	7.6*	1.8	5.2*	2.0
Adjusted Fama-French Three-Factor Model	Size-BM Control Firm	4.2*	1.6	3.8*	0.8	2.4	1.2

Note*:* This table presents the percentages of 1000 samples of 200 small firms that reject the null hypothesis of no annual, three-year and five-year abnormal returns at 5% level of significance. Panel A and Panel B indicate the specification of tests for equally- and value-weighted portfolios respectively. The numbers marked with * suggest that the empirical size is significantly different from the 5% significance level.

4.2.2 Book-to-market (BM) ratio

To investigate the specification of the tests under study, we consider drawing firms on the basis of BM values. To do so, firms are deciled into ten groups based on rankings of BM ratio at the end of June each year. We choose the groups with the highest BM ratio and the lowest BM ratio for robustness check. For each group, we select a random sample of 200 firms. We repeat the procedure 1000 times and present the result in Tables 3A and 3B. Inspecting these tables suggests that our proposed calendar time methodology yields reasonably well-specified test statistics for each type of samples based on book-to-market ratios. Our analysis also documents that the BHAR approach produces better specified *t*-statistics when the control firm approach is employed, but the rejection level increases while using reference portfolios. In addition, the conventional Fama-French three-factor model as well as its modified version produce misspecified test statistics.

4.2.3 Overlapping returns

We consider nonrandom samples based on overlapping returns to inspect the behaviour of the employed methods in the presence of cross-sectional correlation of abnormal returns. Selection of these samples involves two steps. The first stage involves a random selection of 100 firms from the population. In the second stage, for each of these 100 firms, we randomly choose a second event month that is within H - 1 periods of the original event month (either before or after), where H equals 12, 36 or 60. Hence we have 200 firms with 200 event months where the same firm appears in the sample twice and this generates the issue of overlapping returns. This procedure is repeated 1000 times and Table 5 presents the results.

The empirical procedure reveals that the BHAR approach produces misspecified test statistics and these results are consistent with those documented in previous studies (e.g. Lyon, Barber, and Tsai (1999) and Mitchell and Stafford (2000)). Such misspecifications are observed, because the BHAR approach assumes that the observations are cross-sectionally uncorrelated. Jegadeesh and Karceski (2009), however, report that this assumption is tenable in random samples of event firms, but it would be violated in nonrandom samples, where the returns for event firms are positively correlated. Our analysis further shows that all the calendar time portfolio methods yield well-specified test statistics with few exceptions occurring in the one year horizon. This result is expected, since Fama (1998) and Mitchell and Stafford (2000) document that by forming monthly calendar time portfolios, all cross-correlations of event-firm abnormal returns are automatically accounted for in the portfolio variance and hence calendar time methodology performs better than BHAR approach in the presence of cross-sectional correlation of event firm anomalies.

Table 3A. Specification of tests in samples of firms with high BM value

Methods	Benchmark	1 Year		3 Years		5 Years	
		2.5	97.5	2.5	97.5	2.5	97.5
Panel A: Equally-Weighted Portfolios							
Standardized Calendar Time Approach	Size-BM Portfolio	0.2	4.8*	2.4	2.8	1.6	3.4
Standardized Calendar Time Approach	Size-BM Control Firm	0.8	6.0*	2.8	2.3	3.2	1.8
Buy-and-Hold Abnormal Return Method	Size-BM Portfolio	0.0	8.4*	0.0	3.9*	0.7	3.4
Buy-and-Hold Abnormal Return Method	Size-BM Control Firm	1.6	2.9	1.8	4.0*	2.8	3.6
Fama-French Three-Factor Model	Not Applicable	1.2	27.4*	0.0	14.2*	0.7	8.2*
Adjusted Fama-French Three-Factor Model	Size-BM Control Firm	1.6	3.8*	2.6	1.4	1.1	5.2*
Panel B: Value-Weighted Portfolios							
Standardized Calendar Time Approach	Size-BM Portfolio	0.3	7.2*	3.2	0.8	1.1	2.4
Standardized Calendar Time Approach	Size-BM Control Firm	0.0	3.9*	3.4	1.9	3.6	2.0
Fama-French Three-Factor Model	Not Applicable	0.0	14.8*	0.0	7.8*	0.8	5.6*
Adjusted Fama-French Three-Factor Model	Size-BM Control Firm	0.8	4.4*	2.9	0.6	3.2	3.8*

Note: This table presents the percentages of 1000 samples of 200 firms with high BM value that reject the null hypothesis of no annual, three-year and five-year abnormal returns at 5% level of significance. Panel A and Panel B indicate the specification of tests for equally- and value-weighted portfolios respectively. The numbers marked with * suggest that the empirical size is significantly different from the 5% significance level.

Table 3B. Specification of tests in samples of firms with low BM value

Methods	Benchmark	1 Year		3 Years		5 Years	
		2.5	97.5	2.5	97.5	2.5	97.5
Panel A: Equally-Weighted Portfolios							
Standardized Calendar Time Approach	Size-BM Portfolio	4.4*	0.6	1.3	2.1	1.2	3.0
Standardized Calendar Time Approach	Size-BM Control Firm	1.3	3.6	1.2	1.6	2.1	3.4
Buy-and-Hold Abnormal Return Method	Size-BM Portfolio	9.5*	0.0	3.2	0.8	4.4*	0.0
Buy-and-Hold Abnormal Return Method	Size-BM Control Firm	1.8	2.4	3.6	1.9	3.6	2.1
Fama-French Three-Factor Model	Not Applicable	17.8*	0.0	12.0*	0.4	3.6	1.3
Adjusted Fama-French Three-Factor Model	Size-BM Control Firm	3.6	0.8	5.4*	1.2	4.2*	0.7
Panel B: Value-Weighted Portfolios							
Standardized Calendar Time Approach	Size-BM Portfolio	1.2	2.8	2.8	3.2	2.4	2.0
Standardized Calendar Time Approach	Size-BM Control Firm	1.1	2.6	3.6	2.0	2.8	3.1
Fama-French Three-Factor Model	Not Applicable	4.8*	0.3	3.9*	0.4	3.1	0.8
Adjusted Fama-French Three-Factor Model	Size-BM Control Firm	3.4	1.6	3.7*	2.7	4.0*	2.8

Note: This table presents the percentages of 1000 samples of 200 firms with low BM value that reject the null hypothesis of no annual, three-year and five-year abnormal returns at 5% level of significance. Panel A and Panel B indicate the specification of tests for equally- and value-weighted portfolios respectively. The numbers marked with * suggest that the empirical size is significantly different from the 5% significance level.

Table 4. Specification of tests in samples of firms with overlapping returns

Methods	Benchmark	1 Year		3 Years		5 Years	
		2.5	97.5	2.5	97.5	2.5	97.5
Panel A: Equally-Weighted Portfolios							
Standardized Calendar Time Approach	Size-BM Portfolio	4.2*	2.3	3.2	0.9	2.8	2.6
Standardized Calendar Time Approach	Size-BM Control Firm	2.8	1.6	2.0	2.8	2.8	1.3
Buy-and-Hold Abnormal Return Method	Size-BM Portfolio	5.8*	1.6	3.6	4.8*	0.9	3.9*
Buy-and-Hold Abnormal Return Method	Size-BM Control Firm	4.0*	0.2	2.0	4.1*	1.7	4.6*
Fama-French Three-Factor Model	Not Applicable	1.4	7.1*	2.6	5.2*	3.7*	0.8
Adjusted Fama-French Three-Factor Model	Size-BM Control Firm	1.6	6.1*	1.1	2.0	1.6	2.8
Panel B: Value-Weighted Portfolios							
Standardized Calendar Time Approach	Size-BM Portfolio	1.2	3.9*	2.3	2.6	1.2	2.4
Standardized Calendar Time Approach	Size-BM Control Firm	1.6	3.4	2.1	2.4	2.3	3.0
Fama-French Three-Factor Model	Not Applicable	1.2	5.2*	1.6	3.8*	2.7	1.1
Adjusted Fama-French Three-Factor Model	Size-BM Control Firm	0.2	4.4*	3.6	2.0	2.1	2.3

Note: This table presents the percentages of 1000 samples of 200 firms with overlapping returns that reject the null hypothesis of no annual, three-year and five-year abnormal returns at 5% level of significance. Panel A and Panel B indicate the specification of tests for equally- and value-weighted portfolios respectively. The numbers marked with * suggest that the empirical size is significantly different from the 5% significance level.

5 Power

This section documents the power of alternative methodologies in random samples. Note that we exclude nonrandom samples from our analysis, since the *t*-tests based on such samples are generally misspecified. To examine the power of test, we introduce a constant level of abnormal return ranging from -20% to 20% at an interval of 5% to event firms. However, we employ only equally-weighted portfolios to make a direct comparison with the BHAR approach. In addition, we consider the estimates based on control firm approach as BHAR estimators based on 25 size-BM reference portfolios are severely skewed in random samples. We also exclude the traditional Fama-French three-factor model from our power analysis as the test statistics based on this model are not well-specified in random samples. Table 6 indicates the percentages of 1000 random samples of 200 firms that reject the null hypothesis of zero abnormal returns over a three-year holding period. Figure 1 also plots power of the tests.

It is evident from Table 5 and Figure 1 that our proposed standardized calendar time approach produces the most powerful *t*-statistic, followed by the BHAR method. The adjusted Fama-French three-factor model, on the other hand, has low power to identify the long-run anomalies. For instance, with 10% (-10%) per year abnormal returns, the rejection rate is 95% (91%) for SCTA, 74%(67%) for the BHAR method and 53% (44%) for the modified three-factor model. We, therefore, conclude that in case of detecting the abnormal performance, the standardized calendar time approach achieves higher power than the BHAR methodology.

6 Conclusions

This paper investigates the performance of the existing long-run event study methodologies with the UK security market data. Doing so employs the buy-and-hold abnormal return approach and the calendar time portfolio method to measure the return anomalies. While numerous recent studies examine the long-term stock price performance by exercising these two popular approaches, none of the methods is free of criticisms. This paper makes the use of a refined calendar time methodology, proposed by Dutta (2014), to resolve the ongoing debates regarding this approach. The empirical analysis indicates that the standardized calendar time approach of Dutta yields reasonably well-specified test statistics in all types of nonrandom samples. The results further show that in case of detecting the abnormal performance, this standardized calendar time methodology has higher power than other empirical procedures used in this study. One of the major findings of this study is that the Fama-French three-factor model produces misspecified test statistics even in random samples. Our simulation also reveals that the adjusted three-factor model performs well after controlling for size and book-to-market ratio biases. In addition, the buy-and-hold abnormal return approach yields reasonably well-specified test statistics when control firm approach is employed. But, the BHAR methodology possesses lower power than the modified calendar time approach. However, we document that all the employed approaches perform well in the UK stock market.

Table 5. Power of alternative methods in random samples

Methods	Induced Level of Abnormal Return (%) over 3 Years								
	-20	-15	-10	-5	0	5	10	15	20
Standardized Calendar Time Approach	1.00	0.91	0.62	0.18	0.04	0.23	0.69	0.95	1.00
Buy-and-Hold Abnormal Return Method	0.89	0.67	0.34	0.11	0.03	0.15	0.43	0.74	0.91
Adjusted Fama-French Three-Factor Model	0.68	0.44	0.21	0.09	0.04	0.12	0.29	0.53	0.78

Note: This table presents the percentages of 1000 random samples of 200 firms that reject the null hypothesis of no abnormal returns over three-year holding period. We add the levels of annual abnormal return indicated in the column heading. In order to make a direct comparison with BHAR approach, only equally-weighted portfolios are considered in our analysis. In addition, we exclude the reference portfolio approach while calculating the power of tests, since the BHAR estimates based on 25 size-BM reference portfolios are generally biased in random samples.

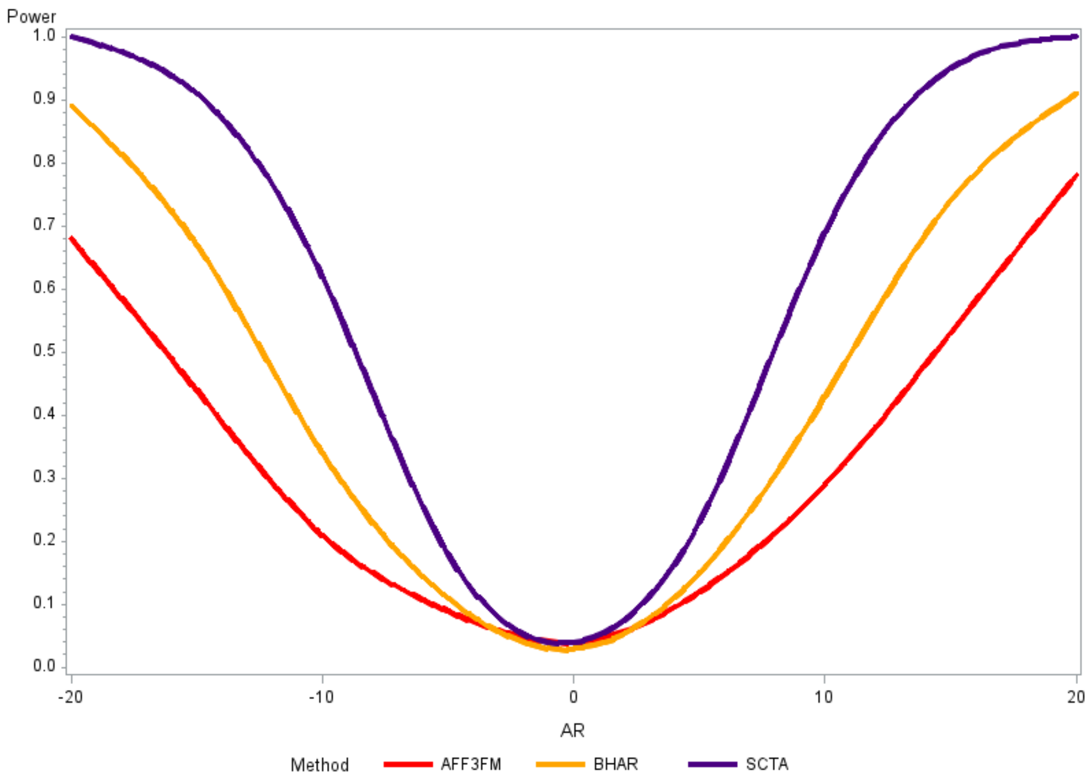

Figure 1. Simulated power of different methods

This figure represents the percentages of 1000 random samples of 200 firms that reject the null hypothesis of no abnormal returns over three-year holding period. We consider equally weighted portfolios to make a direct comparison with BHAR approach. The horizontal axis indicates the induced level of annual abnormal returns (%), while the rejection rates are shown in the vertical axis. In addition, AFF3FM indicates the adjusted Fama-French three-factor model.

References

1. Barber, B. & Lyon, J. 1997. Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. *Journal of Financial Economics*, 43 (3): 341-372.
2. Bessembinder, H. & Zhang, F. 2013. Firm characteristics and long-run stock returns after corporate events. *Journal of Financial Economics*, 109 (1): 83-102.
3. Boehme, D. R. & Sorescu, S. M. 2002. The long-run performance following dividend initiations and

resumptions: Underreaction or Product of Chance?. *The Journal of Finance*, 57 (2): 871-900.
4. Dutta, A. 2014. Does Calendar Time Portfolio Approach Really Lack Power?. *International Journal of Business and Management*, 9 (9): 260-266.
5. Espenlaub, S., Gregory, A. & Tonks, I. 2000). Re-assessing the long-term under-performance of UK initial public offerings. *European Financial Management*, 6: 319-42.
6. Fama, E, & French, K. 1993. Common risk factors in returns on stocks and bonds. *Journal of Financial Economics* 33: 3-56.
7. Fama, E. 1998. Market effciency, long-term returns, and behavioral finance. *Journal of Financial Economics*, 49 (3): 283-306.
8. Huang, Y. 2000. Long-term Abnormal Stock Performance: UK evidence. Ph.D. Thesis. University of Exeter: U.K.
9. Jaffe, J. F. (1974). Special information and insider trading. *Journal of Business*, 47: 410-428.
10. Jegadeesh, N. & Karceski, J. 2009. Long-run performance evaluation: correlationand heteroskedasticity-consistent tests. *Journal of Empirical Finance*, 16 (1): 101-111.
11. Kothari, S. & Warner, J. B. 1997. Measuring long-horizon security price performance. *Journal of Financial Economics*, 43 (3), 301-339.
12. Knif, J., Kolari, J. W. & Pynnönen, S. 2013. A Robust and powerful test of abnormal stock returns in long-horizon event studies. Working Papers Series: Available at SSRN: http://ssrn.com/abstract=2292356 or http://dx.doi.org/10.2139/ssrn.2292356.
13. Levis, M. 1993. The long-run performance of initial public offerings: the UK experience 1980-1988. *Financial Management*, 22: 28-41.
14. Loughran, T. & Ritter, J. 2000. Market efficiency, uniformly least powerful tests of market efficiency. *Journal of Financial Economics*, 55 (3): 361-389.
15. Lyon, J., Barber, B. & Tsai, C. 1999. Improved methods of tests of long-horizon abnormal stock returns. *Journal of Finance*, 54 (1): 165-201.
16. Mandelker, G. 1974. Risk and return: The case of merging firms. *Journal of Financial Economics*, 1: 303-336.
17. Mitchell, M. & Stafford, E. 2000. Managerial decisions and long-term stock price performance. *Journal of Business*, 73 (3): 287-329.
18. Ritter, J. 1991. The long-term performance of initial public offerings. *Journal of Finance*, 46 (1): 3-27.