Does calendar time portfolio approach really lack power?

Author(s): Dutta, Anupam

Title: Does calendar time portfolio approach really lack power?

Year: 2014

Version: Publisher's PDF

Please cite the original version:

Does Calendar Time Portfolio Approach Really Lack Power?

Anupam Dutta[1]

[1] Department of Mathematics & Statistics, University of Vaasa, Vaasa, Finland

Correspondence: Anupam Dutta, Department of Mathematics & Statistics, University of Vaasa, Vaasa, Finland. E-mail: adutta@uwasa.fi

Abstract

This paper investigates whether the calendar time methodology lacks power in detecting the long-run abnormal performance of the firms after major corporate events. In addition, the study proposes a variant of calendar time approach by standardizing the abnormal returns of the event firms forming the monthly portfolios. To assess the robustness of the modified method, the results from buy-and-hold abnormal return approach and the mean monthly calendar time abnormal return method are also reported. The empirical analysis documents that the proposed approach improves the power in random samples and in samples with small firms and with calendar clustering.

Keywords: event study, long-run anomalies, standardized abnormal returns, specification issue, power issue

1. Introduction

A large number of recent studies examine the price behavior of equity for periods of one to five years following significant corporate events (e.g., IPOs, SEOs, repurchases, or bond rating changes). Although there have been many advances in long-run event study methodology over the years, the elementary papers 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), and Jegadeesh and Karceski (2009).

While investigating the long-term abnormal returns of the event firms, two important issues are taken into account. These are the power issue and the specification issue. Prior literature suggests that these two issues still remain unsolved and further filtering of the existing methods is required for solving such issues. Kothari and Warner (2007), for instance, conclude that whether calendar time portfolio (CTP) method or buy-and-hold abnormal return (BHAR) approach can best address these long-horizon issues remains an open question.

However, although each of these two issues is important in inspecting the long-run abnormal performance, only a few studies focus on the power issue. Ang and Zhang (2004), for instance, is the solo study in the literature that extensively reports the power of several empirical procedures at three different investment horizons. In order to extend this limited literature, the present study makes a modest attempt to compare the power of alternative methodologies. In doing so, we analyze the power of buy-and-hold abnormal return approach and the mean monthly calendar time abnormal return (CTAR) methodology for random samples and samples with small firms and with calendar clustering.

Loughran and Ritter (2000), however, criticize the calendar time methodology claiming that it has low power. Ang and Zhang (2004) also document that the power of calendar time portfolio approach decreases as the holding period increases. Mitchell and Stafford (2000), on the other hand, find no evidence that supports the concern raised by Loughran and Ritter (2000). This disagreement gives us the motivation to investigate whether the CTP approach really lacks power in detecting the long-run anomalies.

The empirical findings reveal that the mean monthly calendar time abnormal return method has low power than the buy-and-hold abnormal return approach. Although each of these methodologies is well-specified in random samples and in samples with small firms and with calendar clustering, the CTAR approach lacks power in detecting the long-term anomalies. The study, therefor, proposes to refine the mean monthly calendar time abnormal return methodology by considering the standardized abnormal returns of the event firms forming the monthly portfolios. In our modified approach, the monthly portfolios are also weighted in such a way that periods of heavy event activity receives more loadings than the periods of low event activity. Our analysis further shows that the refined calendar time approach produces well-specified test statistics andimproves the power in all the

sampling schemes under consideration.However, one potential limitation with our proposed approach is that it is, like other existing long-run event study methodologies, is not well specified in all types of nonrandom samples.

The remainder of the paper is organized as follows. Section 2 reviews the existing literature. Section 3 outlines the data and methodology. Results are discussed in Section 4, and Section 5 concludes the paper.

2. Literature Review

Previous research documents that the buy-and-hold abnormal return methodology and the calendar time portfolio approach are commonly employed for examining the long-term abnormal stock returns. Barber and Lyon (1997) and Lyon et al. (1999) claim that the BHARs most accurately capture investor experience. Fama (1998), however, argues against the BHAR methodology as it experiences the bad model problems. Fama and later Mitchell and Stafford (2000)strongly recommend the use of calendar time methodology to deal with the bad model problems.

Barber and Lyon (1997) and Lyon et al. (1999) also identify new listing, re-balancing, and skewness biases with inference in long-run event studies using the BHAR. They use simulations to investigate the impact of these biases on inference when BHAR is exercised to measure the abnormal performance and standard tests are applied. However, in case of using a reference portfolio to capture normal or expected return, the new listing and rebalancing biases can be addressed in a relatively simple way by careful construction of the reference portfolio. Unfortunately, considering reference portfolio also gives rise to the skewness bias.

To avoid thisskewness bias, a control firm rather than a reference portfolio can be used as the long-run return benchmark. The BHAR is then measured as the difference between the long-run holding-period returns of the event firm's equity and that of a control firm. Although the distribution of each asset's holding-period return is highly skewed, the distribution of their difference is not. As a result, standard statistical tests based on the control firm approach have the right size in random samples.

Barber and Lyon (1997), however, report that standard tests based on the control firm approach are not as powerful as those based on the reference portfolio approach. This is because of the fact that the use of a control firm is a noisier way to control for expected returns than is the use of a reference portfolio and this added noise reduces the power of the test. The variance of the difference between the returns on two individual assets is generally much higher than the variance of the difference between the return of an asset and that of a portfolio, even when the control firm is chosen carefully. Powerful tests thus require very large samples when control firm approach is applied.

To resolve these problems that the BHAR methodology encounters, the calendar time portfolio approach is considered as a possible alternative. Unfortunately, Loughran and Ritter (2000) argue that the CTP approach lacks power while identifying the abnormal performance. Ang and Zhang (2004) also report the same in their simulation study. In this paper, we, therefore, make an attempt to modify the conventional calendar time approach such that its power improves. In the following section, we discuss our proposed methodology. Reviewing the conventional approaches will follow.

3. Data and Methodology

The data employed in this paper comprise NYSE, Amex, and Nasdaq stocks, and our sample period ranges from July 1980 to December 2012. We obtain monthly stock prices, market value (MV) or size and book-to-market (BM) value data from DataStream.

In this study, we consider a size-BM-matched control firm to calculate the abnormal returns. Identifying such a 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. We do not use reference portfolios as test statistics based on buy-and-hold abnormal return calculated employing a reference portfolio approach are generally misspecified.

3.1 Mean Monthly Calendar Time Abnormal Return (CTAR)

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

$$mean\ monthly\ CTAR = \frac{1}{T}\sum_{t=1}^{T} CTAR_t, \tag{1}$$

where

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

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 a control firm and T is the total number of months in the

sample period. To test the null hypothesis that there is no abnormal performance, the *t*-statistic of the mean monthly CTAR is calculated using the intertemporal standard deviation of the monthly CTARs defined in equation (2).

3.2 Standardized Calendar Time Approach (SCTA)

In our proposed approach, the monthly portfolios are constructed in an alternative way. We consider standardized abnormal returns to compute the monthly CTARs. Using standardized returns is advantageous in the sense that many firms, especially small firms, often produce volatile returns and consequently, the distributions of long-run returns tend to have fat tails. One possible solution to this problem is standardizing the abnormal returns by their volatility measures. However, to reduce the skewness problem, we consider using log returns.

Now the formation of the monthly portfolios involves 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 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 on a control firm and H is the holding period which equals 12, 36 or 60 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:

$$CTAR_t = \frac{1}{N_t}\sum_{t=1}^{N_t} z_{it} \tag{3}$$

We also propose to weight each of the monthly CTARs by $\sqrt{N_t}$. This weighting scheme is lucrative as it gives more loadings to periods of heavy event activity than the periods of low event activity. However, Loughran and Ritter (2000) argue that when a small number of firms include a large proportion of a value-weighted portfolio, unsystematic risk is not diversified away. In this paper, we, therefore, consider only equally-weighted portfolios to estimate the abnormal returns. The grand mean monthly abnormal return, denoted by $\overline{CTAR}$, is then calculated as:

$$\overline{CTAR} = \frac{1}{T}\sum_{1}^{T} CTAR_t \tag{4}$$

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).

3.3 Buy-and-Hold Abnormal Return (BHAR)

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

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

where 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:

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

where $\overline{BHAR_H}$ 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.

Table 1 summarizes the test statistics of different methods used in our study to investigate the long-term abnormal performance. We use size-BM matched control firm to measure the anomalies. The standard error shown in the numerator is the traditional standard error computed using the CTARs defined in equation (2) and equation (3).

Table 1. Summary of alternative methodologies

Method Description	Test Statistics
Standardized Calendar Time Approach (SCTA)	$t = \dfrac{\overline{CTAR}}{Standard\ Error}$
Buy-and-Hold Abnormal Return (BHAR)	$t_{BHAR} = \dfrac{\overline{BHAR_H}}{\sigma(BHAR_H)/\sqrt{n}}$
Mean Monthly Calendar Time Abnormal Return (CTAR)	$t = \dfrac{Mean\ Monthly\ \overline{CTAR}}{Standard\ Error}$

Note. This table summarizes the test statistics of different empirical procedures employed in this study. The standard error is the conventional standard error computed using the CTARs defined in equation (2) and equation (3).

3.4 Simulation Method

To assess the specification of the employed methodologies, we randomly select 1000 samples of 200 event months without replacement. 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. Empirical Results

This section reports the specification and power of various methodologies used in our study. We first discuss the specification issue. Later, we focus on the power issue. In addition to random samples, two types of nonrandom samples based on small firms and calendar clustering are also considered in our analysis.

4.1 Specification of Tests

Table 2 indicates the rejection rates in 1000 simulations with a sample of 200 firms. Panels A, B and C present the size of alternative tests for random samples and samples with small firms and with calendar clustering respectively. The simulated results reveal that all the empirical methods considered in our analysis are well-specified in random samples as well as in nonrandom samples. For example, Panel A suggests that when the holding period is one year, the rejection rates at the 5% level of significance are 2.4% and 2.8% for SCTA, 2.8% and 0.4% for BHAR, and 2.7% and 1.3% for CTAR. Panel B, on the other hand, indicates that with a five-year investment horizon, the rejection rates at the 5% level of significance are 3.5% and 1.6% for SCTA, 3.6% and 0.7% for BHAR, and 2.4% and 1.2% for CTAR. Panel C reveals that with a three-year holding period the rejection rates at the 5% level of significance are 3.6% and 2.2% for SCTA, 2.7% and 2.1% for BHAR, and 2.1% and 2.8% for CTAR.

Table 2. Specification of tests in random samples and samples with small firms and with calendar clustering

	Holding Period					
	12 Months		36 Months		60 Months	
	Theoretical Cumulative Density Function (%)					
Method	2.5	97.5	2.5	97.5	2.5	97.5
Panel A: Random Samples						
SCTA	2.4	2.8	1.6	2.8	2.0	3.6
BHAR	2.8	0.4	3.2	1.2	2.6	2.3
CTAR	2.7	1.3	2.4	2.4	1.1	1.6
Panel B: Samples with Small Firms						
SCTA	3.2	0.9	2.6	2.1	3.5	1.6
BHAR	2.4	3.6	3.2	2.8	3.6	0.7
CTAR	1.1	0.3	2.9	1.2	2.4	1.2
Panel C: Samples with Calendar Clustering						
SCTA	1.9	2.7	3.6	2.2	2.0	3.1
BHAR	3.4	0.8	2.7	2.1	3.6	1.1
CTAR	3.1	0.9	2.1	2.8	3.2	2.0

Note. This table presents the percentages of 1000 samples of 200 firms that reject the null hypothesis of no abnormal returns over one-year, three-year, and five-year holding periods. Panel A shows the results for random samples, while Panel B and Panel C indicate the findings for samples with small firms and with calendar clustering.

4.2 Power

Tables 3–5 show the power of alternative methodologies in random samples and in samples with small firms and with calendar clustering respectively. Since the methods we employ in this paper are not, in general, well-specified in samples based on book-to-market-ratio or large firms, we do not report the power of different approaches using these nonrandom samples. To examine the power of the employed methods, we introduce a constant level of abnormal return ranging from -20% to 20% at an interval of 5% to event firms. Tables 3–5 indicate the percentages of 1000 samples of 200 firms that reject the null hypothesis of zero abnormal returns over one, three and five-year holding periods.

It is evident from Table 3 that in case of detecting the anomalies, the calendar time methodologies have more power than the BHAR approach does. For example, with +10% (-10%) per year abnormal returns and with a three-year holding period, the rejection rate is 25% (16%) for SCTA, 23% (16%) for BHAR, and 17% (13%) for CTAR. Table 4, however, indicates that with +15% (-15%) per year abnormal returns and with a five-year investment horizon, the rejection rate is 30% (26%) for SCTA, 29% (25%) for BHAR, and 21% (17%) for CTAR. Finally, Table 5 reveals that with +20% (-20%) per year abnormal returns and with a three-year holding period, the rejection rate is 49% (41%) for SCTA, 43% (39%) for BHAR, and 41% (36%) for CTAR.

Table 3. Power of Alternative methodologies in random samples

Methods	Induced Level of Abnormal Returns (%)								
	-20	-15	-10	-5	0	5	10	15	20
Panel A: One Year Holding Period									
SCTA	0.98	0.81	0.51	0.19	0.05	0.27	0.62	0.87	1.0
BHAR	0.96	0.79	0.46	0.17	0.03	0.21	0.55	0.83	0.99
CTAR	0.81	0.67	0.43	0.12	0.04	0.18	0.52	0.74	0.91
Panel B: Three Years Holding Period									
SCTA	0.54	0.35	0.16	0.08	0.04	0.12	0.25	0.44	0.63
BHAR	0.51	0.34	0.16	0.08	0.04	0.10	0.23	0.39	0.69
CTAR	0.47	0.31	0.13	0.07	0.05	0.07	0.17	0.37	0.63
Panel C: Five Years Holding Period									
SCTA	0.36	0.19	0.08	0.06	0.05	0.07	0.12	0.26	0.41
BHAR	0.33	0.17	0.09	0.06	0.05	0.08	0.12	0.24	0.40
CTAR	0.32	0.14	0.06	0.04	0.02	0.06	0.09	0.21	0.36

Note. This table documents the percentages of 1000 random samples of 200 firms that reject the null hypothesis of no abnormal performance over one-year (Panel A), three-year (Panel B), and five-year (Panel C) holding periods. We add the levels of annual abnormal return indicated in the column heading.

Table 4. Power of alternative methodologies in samples with small firms

Induced Level of Abnormal Returns (%)								
Methods	-20	-15	-10	-5	0	5	10	15	20
Panel A: One Year Holding Period								
SCTA	0.96	0.77	0.56	0.19	0.04	0.21	0.63	0.69	0.88
BHAR	0.91	0.74	0.54	0.18	0.06	0.15	0.59	0.66	0.82
CTAR	0.63	0.45	0.31	0.12	0.02	0.06	0.26	0.38	0.58
Panel B: Three Years Holding Period								
SCTA	0.61	0.40	0.22	0.11	0.06	0.14	0.29	0.49	0.71
BHAR	0.57	0.38	0.19	0.10	0.05	0.13	0.31	0.47	0.68
CTAR	0.42	0.29	0.12	0.07	0.05	0.08	0.21	0.34	0.50
Panel C: Five Years Holding Period								
SCTA	0.52	0.26	0.09	0.07	0.06	0.08	0.11	0.30	0.56
BHAR	0.48	0.25	0.10	0.08	0.07	0.07	0.11	0.29	0.54
CTAR	0.29	0.17	0.07	0.05	0.04	0.08	0.09	0.21	0.38

Note. This table shows the percentages of 1000 samples of 200 small firms that reject the null hypothesis of no abnormal performance over one-year (Panel A), three-year (Panel B), and five-year (Panel C) holding periods. We add the levels of annual abnormal return indicated in the column heading.

Table 5. Power of alternative methodologies in samples with calendar clustering

Induced Level of Abnormal Returns (%)								
Methods	-20	-15	-10	-5	0	5	10	15	20
Panel A: One Year Holding Period								
SCTA	0.73	0.52	0.31	0.12	0.05	0.17	0.39	0.61	0.82
BHAR	0.68	0.47	0.29	0.11	0.04	0.14	0.38	0.58	0.82
CTAR	0.61	0.41	0.25	0.08	0.04	0.11	0.31	0.49	0.68
Panel B: Three Years Holding Period								
SCTA	0.67	0.43	0.24	0.09	0.06	0.11	0.31	0.51	0.76
BHAR	0.59	0.36	0.21	0.06	0.04	0.10	0.32	0.49	0.72
CTAR	0.54	0.32	0.18	0.06	0.05	0.08	0.19	0.37	0.52
Panel C: Five Years Holding Period								
SCTA	0.41	0.28	0.10	0.06	0.05	0.08	0.14	0.32	0.49
BHAR	0.39	0.26	0.10	0.06	0.05	0.07	0.11	0.30	0.43
CTAR	0.36	0.24	0.08	0.06	0.04	0.07	0.11	0.31	0.41

Note. This table documents the percentages of 1000 samples of 200 firms with calendar clustering that reject the null hypothesis of no abnormal performance over one-year (Panel A), three-year (Panel B), and five-year (Panel C) holding periods. We add the levels of annual abnormal return indicated in the column heading.

These findings suggest three important implications. First, with the increase in the investment period, the power starts decreasing for all the methods employed in this study. Ang and Zhang (2004), however, also report the same. Second, the mean monthly calendar time abnormal return has low power to detect the long-term anomalies. Third, our proposed standardized calendar time approach improves the power in all the sampling schemes under study.

5. Conclusion

Although long-run event studies have an extensive history, the power issue generally receives less attention than the specification issue. To conceal this research gap, the present study aims to compare the power of alternative methodologies under different sampling schemes. In doing so, this paper extends the prior literature in two aspects. First, the study inspects whether the calendar time methodology really lacks power in detecting the long-term abnormal performance of the firms following major corporate events. To serve this purpose, we compare the results of buy-and-hold abnormal return approach and the mean monthly calendar time abnormal return methodology. The empirical analysis indicates that although each of these methodologies is well-specified in random samples as well as in nonrandom samples, the mean monthly calendar time abnormal return method lacks power in detecting the anomalies. Second, we propose to refine the mean monthly calendar time approach by forming the monthly portfolios in a variant way and further analysis shows that our proposed calendar time approach, with only a few exceptions, improves the power in random samples and in samples with small firms and with calendar clustering.

Acknowledgement

References

Ang, S., & Zhang, S. (2004). An evaluation of testing procedures for long horizon event studies. *Review of Quantitative Finance and Accounting*, *23*(3), 251–274. http://dx.doi.org/10.1023/B:REQU.0000042344.27369.0d

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. http://dx.doi.org/10.1016/S0304-405X(96)00890-2

Fama, E. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, *49*(3), 283–306. http://dx.doi.org/10.1016/S0304-405X(98)00026-9

Jegadeesh, N., & Karceski, J. (2009). Long-run performance evaluation: correlation and hetero skedasticity-consistent tests. *Journal of Empirical Finance*, 16(1), 101–111. http://dx.doi.org/10.1016/j.jempfin.2008.06.002

Kothari, S. P., & Warner, J. B. (2007). Econometrics of event studies. In Eckbo, E. (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance* (vol. 1, pp. 3–36). Elsevier/North-Holland, Amsterdam.

Kothari, S., & Warner, J. B. (1997). Measuring long-horizon security price performance. *Journal of Financial Economics, 43*(3), 301–339. http://dx.doi.org/10.1016/S0304-405X(96)00899-9

Loughran, T., & Ritter, J. (2000). Market efficiency, uniformly least powerful tests of market efficiency. *Journal of Financial Economics*, *55*(3), 361–389. http://dx.doi.org/10.1016/S0304-405X(99)00054-9

Lyon, J., Barber, B., & Tsai, C. (1999). Improved methods of tests of long-horizon abnormal stock returns. *Journal of Finance*, *54*(1), 165–201. http://dx.doi.org/10.1111/0022-1082.00101

Mitchell, M., & Stafford, E. (2000). Managerial decisions and long-term stock price performance. *Journal of Business*, *73*(3), 287–329. http://dx.doi.org/10.1086/209645

Ritter, J. (1991). The long-term performance of initial public offerings. *Journal of Finance*, *46*(1), 3–27. http://dx.doi.org/10.1111/j.1540-6261.1991.tb03743.x

Copyrights