# Event Studies: A Methodology Review

**Charles J. Corrado**
**Deakin University**
**Melbourne, Australia**
**corrado@deakin.edu.au**

**26 July 2010**

**Abstract**

Originally developed as a statistical tool for empirical research in accounting and finance, event studies have since migrated to other disciplines as well, including economics, history, law, management, marketing, and political science. Despite the elegant simplicity of a standard event study, variations in methodology and their relative merits continue to attract attention in the literature. This paper reviews some of the fundamental topics in short-term event study methodology, with an attempt to add new perspectives to some pressing topics.

## 1. Introduction

Originally introduced to a broad audience of accounting and financial economists in two landmark papers by Ball and Brown (1968) and Fama, Fisher, Jensen, and Roll (1969), event studies have since become ubiquitous in capital markets research. There have been many advances in event study methodology over the years, but the core elements of a typical event study can be found in these early papers. However, Ball-Brown and Fama-Fisher-Jensen-Roll were not the first event studies. MacKinlay (1997) reports an early event study by Dolley (1933) examining stock price reaction to stock splits and refers to several other published papers indicating that by the 1960s event studies had made their way into leading business economics journals (Myers and Bakay, 1948; Barker, 1956, 1957, 1958; Ashley, 1962).

It is common in science for the pieces of an idea to float around for a time and then coalesce in an inspired form at the right time. Factors contributing to the success of the Ball-Brown and Fama-Fisher-Jensen-Roll event studies are their use of the 'market model' patterned after the then recently developed capital asset pricing model (CAPM) of Sharpe (1964), their use of data from the newly established Center for Research in Security Prices (CRSP) at the University of Chicago, which quickly became a standard data source in capital markets research, and the rapidly expanding access to computer systems equipped with sophisticated statistical software. Also contributing in no small measure to the success of financial event studies was the paradigm-shifting revolution in corporate finance initiated in the classic papers of Modigliani and Miller (1958), Miller and Modigliani (1961, 1963) that brought capital structure issues to the forefront of financial research. Event studies became a key empirical tool in studies devoted to these issues.

No one really knows how many event studies have been published. Kothari and Warner (2005) report that over the period 1974-2000 five major finance journals published 565 articles containing event study results. This is clearly a very conservative number as it does not include the many event studies published in accounting journals and other finance journals. Moreover, event studies are now regularly seen outside the realm of mainstream accounting and finance journals. Schwert (1981) is an early example. A parsimonious, recent sampling of papers in various areas where event studies are now applied might include Ghosh et al. (1995), Meznar et al. (1998), Lamdin (2001), McKenzie and Thomsen (2001), Bhagat and Romano (2002), Chatterjee et al. (2002), Nicolau (2002), Rose (2003), Thomsen and McKenzie (2001), Drakos (2004), Nicolau and Sellers (2004), Wulf (2004), Tuck (2005), Cichello and Lamdin (2006), Fornell et al. (2006), Calvo-Gonzalez (2007), Johnston (2007), Darkow (2008), Gong (2009), Misra and Rao (2009), and Roztocki and Weistroffer (2009). There are, of course, many others too numerous to mention here. What is clear is that event studies continue to be popular and appear destined to be a part of capital markets research for some time.

This paper represents a modest attempt to review event study methodology and its applications. The topic is so broad that failure to cover all aspects, or even all important aspects, must be admitted at the outset. A complete bibliography has yet to be compiled – it would be quite large. A number of previous reviews of event study methodology assisted in writing the current review, including Peterson (1989), Henderson (1990), Armitage (1995), Thompson (1995), MacKinlay (1997), McWilliams and Siegel (1997), Binder (1998), McWilliams and McWilliams (2000), Lamdin (2001), Serra (2002), Kothari and Warner (2005), Cichello and Lamdin (2006), and Johnston (2007). Some of these are now dated, but were useful nonetheless in understanding the evolution of financial event study methodology. I have tried to avoid recycling topics competently reviewed in previous surveys. Instead, I have restricted coverage to what I believe are important topics in short-term event studies, with attempts to add some new material and viewpoints to these topics.

The structure of this review proceeds along the following sequence: Section 2 outlines the econometric skeleton of an event study, which is followed in Section 3 by a survey of results obtained from studies of event study methodology. Section 4 discusses the important topic of event-induced variance and attempts to offer some new insights into the scope of the problem. Event-induced returns are discussed in Section 5. Section 6 summarises and concludes the review.

## 2. The econometric skeleton of an event study

The skeletal econometric structure of an event study is well illustrated by the case of a single security-event date combination. While such studies rarely find their way into a research journal, they are common in certain legal proceedings. For example, Mitchell and Netter (1994) discuss their application in securities' law actions such as disgorgement proceedings related to insider-trading cases. A typical penalty is disgorgement of any gain attributed to the insider trading activity, plus a fine of some multiple of that gain. As a hypothetical example, suppose that shortly before a merger announcement a corporate insider with access to private information buys a large block of the company's shares in anticipation of a stock price appreciation on the date of the merger announcement. In a disgorgement proceeding, it must be shown that the insider materially benefited from access to the private information. This might be demonstrated by showing that on the merger announcement date the stock price enjoyed a significant increase in value. Event study procedures are commonly accepted as evidence in deciding whether insiders benefited from their use of private information and in determining the magnitude of their gain.

> "An event study is a statistical technique that estimates the stock price impact of occurrences such as mergers, earnings announcements, and so forth. The basic notion is to disentangle the effects of two types of information on stock prices – information that is specific to the firm under question (e.g., dividend announcement) and information that is likely to affect stock prices marketwide (e.g., change in interest rates." (Mitchell and Netter, 1994)

To illustrate how event study evidence might be presented in an insider-trading case, let ‘day-0’ identify the merger announcement date under scrutiny and let days $t = \ldots, -3, -2, -1$ represent trading days leading up to the event. A naïve event study might compare the event date return $R_0$ with returns observed during a control period before the event. If the event date return $R_0$ was statistically large compared to pre-event control period returns, we might conclude that the event had a significant stock price impact. However, this naïve comparison of the event date return $R_0$ with pre-event date returns does not disentangle the effects of firm-specific information from market-wide information affecting the stock price.

A more sophisticated procedure uses the market model to adjust the event date return to remove the influence of the overall market. Mathematically, the market model used in event studies is specified in equation (1), where $R_t$ and $RM_t$ denote the return on day $t$ for the stock and the overall market, respectively, $e_t$ represents a firm-specific return, and the parameters $a, b$ specify the linear structure of the market model.

$$R_t = a + b \times RM_t + e_t \tag{1}$$

By assumptions inherent in the structure of the market model, the firm-specific return $e_t$ is unrelated to the overall market and has an expected value of zero. Hence, the expected event date return conditional on the event date market return is,

$$E\left(R_0 \middle| RM_0\right) = a + b \times RM_0 \tag{2}$$

Adjusting the observed event date return $R_0$ by subtracting the conditional expected return yields an abnormal return specified in equation (3). Notice that the abnormal return $A_0$ is simply the day-zero firm-specific return $e_0$ identified by the market model in equation (1).

$$A_0 = R_0 - E\left(R_0 \middle| RM_0\right) = R_0 - a - b \times RM_0 \tag{3}$$

For comparison and to determine statistical significance, a series of abnormal returns are obtained from a control period before the event date. By convention, this is normally a period ending a few days before the event date.

$$\begin{aligned} A_t &= R_t - E\left(R_t \middle| RM_t\right) = R_t - a - b \times RM_t \\ t &= -n-5, -n-4, \ldots, -7, -6 \end{aligned} \tag{4}$$

The control period indicated above in equation (4) contains $n$ days beginning $n+5$ days and ending 6 days before the event date. Typically, a value of $n = 250$ days is chosen to correspond approximately to the number of trading days in a calendar year. The market model parameters $a, b$ are usually obtained by an ordinary least-squares regression of firm returns $R_t$ on market returns $RM_t$ over the control period. The event date abnormal return $A_0$ is then assessed for statistical significance relative to the distribution of abnormal returns $A_t$ in the control period. A common assumption used to formulate tests of statistical significance is that abnormal returns are normally distributed.

Under a null hypothesis of no abnormal performance the event date abnormal return $A_0$ will have an expected value of zero and a variance given by least-squares regression theory as stated in equation (5), where $\hat{\sigma}_e$ denotes the standard error of the regression used to obtain the market model parameters $a, b$ (Greene, 2006).

$$Var\left(A_0\right) = \hat{\sigma}_e^2 \left( 1 + \frac{1}{n} + \frac{\left(RM_0 - \overline{RM}\right)^2}{\sum_{h=-(n+5)}^{-6} \left(RM_h - \overline{RM}\right)^2} \right) \tag{5}$$

Assuming returns are normally distributed, the test statistic in equation (6) below is distributed as Student-*t* with $n-2$ degrees of freedom ($df$). A Student-*t* variable has a variance and kurtosis of $df/\left(df-2\right)$ and $\left(3df-6\right)/\left(df-4\right)$, respectively. For large values of $df = n-2$, these are close to the standard normal distribution variance and kurtosis of one and three, respectively.

$$\frac{A_0}{\sqrt{Var(A_0)}} \sim T_{n-2} \tag{6}$$

Selecting a critical value $C_\alpha$ such that $\Pr(T_{n-2} > C_\alpha) = \alpha$, the null hypothesis of no positive abnormal performance is rejected with a confidence level of $1-\alpha$ if the test statistic is greater than the critical value, that is,

$$\frac{A_0}{\sqrt{Var(A_0)}} > C_\alpha \tag{7}$$

A test statistic larger than the upper-tail critical value $C_\alpha$ provides statistical evidence that the merger announcement had a significant positive impact on the stock price on the event date. Similarly, a test statistic less than the lower-tail critical value $-C_\alpha$ would provide evidence that the announcement had a significant negative impact.

However, a defence attorney armed with a knowledge of statistics might question this result and argue that the validity of the test depends critically on the assumption that stock returns are normally distributed. This argument might have more force were it shown that the sample skewness and/or kurtosis of returns were far from those expected from normally distributed data. Fortunately, there is a simple, robust procedure to assess statistical significance that does not depend on an assumption of normally distributed returns. This procedure merely involves counting the number of returns from the control period that are larger or smaller than the event date return $A_0$. It would be especially convincing were the abnormal return $A_0$ larger than all abnormal returns from the control period, since this would only occur with a probability of $1/(n+1)$ under a true null hypothesis. For $n = 250$ this probability is less than 0.00398. This can be established by a simple thought experiment. Suppose balls numbered zero through $n$ are randomly scattered along a line from left to right. There are $n+1$ equally likely positions for the ball marked zero and the probability that this ball occupies the rightmost position is $1/(n+1)$. In general, the probability that the ball marked zero occupies one of the $k$ rightmost positions is $k/(n+1)$. Choosing a value for $k$ such that $k/(n+1) \le \alpha$ yields a test of significance at level $\alpha$. The validity of this simple test does not depend on the assumption that stock returns follow a normal distribution.

Most event studies contain a sample with multiple security-event date combinations, where it is desirable to aggregate results into a single hypothesis test. For example, an insider-trading case might involve multiple security-event date combinations, as in the classic case of R. Foster Winans, the *Wall Street Journal* reporter charged and convicted of insider trading for tipping off a stock broker about items that would appear in his influential *Heard On The Street* column (Winans, 1984).

A simple method of aggregation across security-event date combinations is to sum the individual $T$-statistics in equation (6) and divide this sum by the square root of the sample size, as shown in equation (8) for a sample of size $m$. This classic test statistic was proposed in Patell (1976) and Dodd and Warner (1983) and is commonly referred to as the Patell *T*-test. Assuming security returns are normally distributed, this test statistic, here denoted by $T_P$, is distributed as a convolution of Student-*t* variables.

$$T_P = \frac{1}{\sqrt{m}} \sum_{j=1}^{m} \frac{A_{j,0}}{\sqrt{Var\left(A_{j,0}\right)}} \qquad (8)$$

The Patell *T*-test continues to be popular among capital market researchers, in no small measure because of its intuitively appealing characteristic of placing equal statistical weight on each security-event date combination.

If we wish to avoid any assumption of normality in the underlying returns data, we can construct a simple test statistic along the lines of the thought experiment described earlier. Let $h$ denote the position of the ball marked zero among $n+1$ randomly scattered balls along a line from left to right. Since $h$ can vary with equal probability from one to $n+1$, the statistic $h/(n+2)$ follows a discrete uniform distribution, which for large $n$ is nearly continuous. Indexing each security-event date combination with the subscript $j$, subtracting one-half, then summing across a sample of size $m$ and multiplying the sum by $\sqrt{12/m}$ yields the statistic $Z_U$ shown in equation (9).

$$Z_U = \sqrt{\frac{12}{m}} \sum_{j=1}^{m} \left( \frac{h_j}{n_j + 2} - \frac{1}{2} \right) \tag{9}$$

In the current context, $h_j$ represents the position of the abnormal return $A_{j,0}$ among the sorted vector of $n_j + 1$ abnormal returns for the $j^{\text{th}}$ security-event date combination in the sample. Maynes and Rumsey (1993) point out that the statistic $Z_U$ has approximately a unit variance and a kurtosis of $3 - 1.2/m$, which converges quickly to the standard normal kurtosis of three as the sample size $m$ increases.

If the sample size $m$ is small, precision might be gained by transforming the statistics $h_j / \left(n_j + 2\right)$ using Tukey's lambda distribution as shown in equation (10).

$$Z_L = \frac{4.91}{\sqrt{m}} \sum_{j=1}^{m} \left( u_j^{0.14} - \left(1 - u_j\right)^{0.14} \right) \qquad u_j = \frac{h_j}{n_j + 2} \tag{10}$$

The statistic $Z_L$ is close to exactly distributed as standard normal even for very small values of $m$ (Joiner and Rosenblatt, 1971).

### 3. 'Brown and Warner' methodology studies

With the emergence of event studies in capital markets research, there soon followed a stream of studies examining event study methodology. Early methodology studies examined test statistic performance with monthly security returns data, e.g., Brown and Warner, 1980; Larcker, 1980; Shevlin, 1981. These were quickly superseded by studies based on daily data as they became available from the Center for Research in Security Prices (CRSP). The best known among these is the study by Brown and Warner (1985), which has since come to eponymously define the genre. Brown and Warner-type studies are categorically distinguished by their use of computer simulation experiments based on random drawings from a large population of actual security returns data. Other early studies in this genre include Collins and Dent (1984), Dyckman, Philbrick, and Stephan (1984), Jain (1986), Bernard (1987), and Heinkel and Krause (1988). These early studies generally concluded that standard parametric event study tests are well-specified with good test power when used with returns data from the New York stock exchange. These classic studies can still be profitably read by the student of capital markets research.

A subsequent branch of Brown and Warner studies involved testing the performance of robust tests that did not rely on an assumption of normally distributed returns for correct specification. The most successful among these tests were the non-parametric sign and rank tests advanced in McConnell and Muscarella (1985), Corrado (1989), Lummer and McConnell (1989), Zivney and Thompson (1989), and Corrado and Zivney (1992). Well-known studies of this type are Chandra, Rohrbach, and Willinger (1992, 1995), Cowan (1992, 1993), Campbell and Wasley (1993), Campbell and Wasley (1996), Cowan and Sergeant (1996), Giaccotto and Sfiridis (1996), Affleck-Graves, Callahan, and Ramanan (2000), Dombrow, Rodriguez, and Sirmans (2000), Seiler (2000), Bartholdy and Peare (2007), and Corrado and Truong (2008). These studies generally concluded that sign and rank tests are well-specified and provide an improvement in test power compared to standard parametric tests.

Nonparametric tests are motivated by concerns that non-normally distributed security returns data may cause parametric tests derived under an assumption of normality to be poorly specified and so yield imprecise inferences. As shown in most of the studies cited above, this does not appear to be an urgent concern with returns data from the New York stock exchange. However, several studies find that non-normality is an issue of potentially serious concern with returns data from other exchanges. Campbell and Wasley (1993) provided an early alert to the severity of this problem with their finding that standard parametric event study tests were poorly specified with Nasdaq returns data, but the nonparametric rank test advanced in Corrado (1989) maintained good specification with these data. Similarly, Maynes and Rumsey (1993) found that a rank test outperformed parametric tests with data from the Toronto stock exchange. More recently, Bartholdy, Olson, and Peare (2007) find that rank and sign tests outperform parametric tests with security returns data from the Copenhagen stock exchange. In a comprehensive study of event study test statistic performance with data from eleven Asia-Pacific stock exchanges, Corrado and Truong (2008) find that standard parametric event study tests can be poorly specified with these returns data, but nonparametric rank and sign tests are both well specified.

Subsequent advances in parametric test methodology have also been evaluated using Brown and Warner-type simulation experiments. Worthy of particular mention is a variation of the Patell *T*-test that adds a correction for a potential cross-sectional increase in returns variance on the event date. This test was independently developed in Boehmer, Musumeci, and Poulsen (1991) and Sanders and Robins (1991), but is usually referred to as the Boehmer-Musumeci-Poulsen *T*-test. In the same notation used earlier, the Boehmer-Musumeci-Poulsen *T*-test statistic, here denoted by $T_{BMP}$, is constructed as shown in equation (11).

$$T_{BMP} = \frac{\sum_{j=1}^{m} \frac{AR_{j,0}}{\sqrt{Var\left(A_{j,0}\right)}}}{\sqrt{Var\left(\sum_{j=1}^{m} \sqrt{Var\left(A_{j,0}\right)}\right)}} \qquad (11)$$

$$Var\left(\sum_{j=1}^{m} \frac{AR_{j,0}}{\sqrt{Var\left(A_{j,0}\right)}}\right) = \sum_{j=1}^{m}\left(\frac{AR_{j,0}}{\sqrt{Var\left(A_{j,0}\right)}} - \frac{1}{N}\sum_{i=1}^{m} \frac{AR_{i,0}}{\sqrt{Var\left(A_{i,0}\right)}}\right)^2$$

In simulation experiments based on New York stock exchange data, Boehmer, Musumeci and Poulsen (1991) and Corrado and Truong (2008) report that this test statistic is well-specified with good power properties.

The Boehmer-Musumeci-Poulsen *T*-test is a candidate example of the pieces of an idea floating around for a time before coalescing in a form attracting widespread interest. Similar cross-sectional variance adjustments can be found in several early studies, for example, Charest (1978) and Penman (1982). The success of the $T_{BMP}$ statistic might be attributed to its use of standardized abnormal returns adopted from the Patell *T*-test and so involves only a simple adjustment to the popular Patell *T*-test to implement. However, like the Patell *T*-test, the Boehmer-Musumeci-Poulsen *T*-test may not perform well when the distribution of security returns data departs markedly from a normal distribution. Corrado and Truong (2008) find that the BMP *T*-test tends to reject a true null too often when used with data from the American stock exchange, Nasdaq, and stock exchanges in the Asia-Pacific region. Bartholdy, Olson, and Peare (2007) report a similar finding with security returns data from the Copenhagen stock exchange.

Corrado and Zivney (1992) show that the rank test introduced in Corrado (1989) can benefit from the same adjustment underlying the Boehmer-Musumeci-Poulsen modification of the Patell *T*-test, while maintaining robustness against non-normality in returns data. This adjustment is implemented as follows: let $SAR_{j,t} = AR_{j,t} / \sqrt{Var\left(AR_{j,t}\right)}$ denote the standardized abnormal return series for the $j^{\text{th}}$ security over the control period and the event date, i.e., $t = -n-5,\ldots,-6,0$. The cross-sectional variance adjustment is applied only to event date standardized abnormal returns, so each series is then specified as shown in equation (12).

$$SAR_{j,t} = \begin{cases} AR_{j,t} / \sqrt{Var\left(AR_{j,t}\right)} & t = -n-5,\ldots,-6 \\ \dfrac{AR_{j,t} / \sqrt{Var\left(AR_{j,t}\right)}}{\sqrt{Var\left(AR_{.,t} / \sqrt{Var\left(A_{.,t}\right)}\right)}} & t = 0 \end{cases} \tag{12}$$

Let $r\left(SAR_{j,0}\right)$ denote the rank of the event date standardized abnormal return $SAR_{j,0}$ within the vector of $n+1$ standardized abnormal returns for the $j^{\text{th}}$ security. These ranks are then used to compute the Corrado-Zivney rank test statistic $T_{CZ}$ as shown in equation (13).

$$T_{CZ} = \frac{1}{\sqrt{m}} \sum_{j=1}^{m} \frac{r\left(SAR_{j,0}\right) - \left(\frac{n+1}{2}\right)}{\sqrt{n\left(n+1\right)/12}} \tag{13}$$

In comparisons with a battery of popular alternative parametric and nonparametric tests, Corrado and Truong (2008) find that this refined rank test has the best specification overall with United States and Asia-Pacific security market returns data.

Brown and Warner-type studies have also found that the issue of non-normally distributed data can be severe in event studies based on data types other than daily returns. Campbell and Wasley (1996) find that parametric tests can perform quite poorly in event studies examining abnormal trading volume, but the Corrado (1989) rank test is well-specified with these data. Affleck-Graves, Callahan, and Ramanan (2000) compare several event study methods with bid-ask spread data. They find that these data are generally non-normally distributed, but not severely so. Consequently, standard parametric tests performed well and the Corrado rank test provided only slightly more test power with bid-ask spread data. Mucklow (1994) finds the rank test to be well-specified with good test power in detecting event effects using intraday returns data.

Several alternative methods for dealing with non-normally distributed returns data in event studies have been proposed. Among these, perhaps the most interesting are those based on the bootstrap, which have been examined in a number of studies, including Kramer (2001), Hamill et al. (2002), Chou (2004), Hein and Westfall (2004), Jackson, Kline, and Skinner (2006), Baixauli (2007), Ford and Kline (2007), and Corrado and Truong (2008). So far, evidence on bootstrap performance has been mixed and appears to vary with the type of problem to which it is applied. Seemingly, for this reason the bootstrap has not become popular in event studies, but further study may alter this result.

Besides providing a method to compare the performance of various event study test statistics, Brown and Warner-type studies have also provided other interesting revelations regarding event study methodology. For example, the choice of market index employed with the market model can be an important factor affecting test specification. Brown and Warner (1980) found that event study tests based on a market model using a value weight index were poorly specified. Campbell and Wasley (1993) found that the choice between an equal weight or value weight market index was important in event studies using Nasdaq data and strongly recommended using the Nasdaq equal weight market index. Similarly, Corrado and Truong (2008) find that tests based on an equal weight index provide noticeably improved performance over tests based on a value weight index with both United States and

Asia-Pacific data. Corrado and Truong (2008) also find that tests based on logarithmic returns generally produce better test specification than tests based on arithmetic returns. Ahern (2009) provides interesting evidence demonstrating how severe biases can be introduced into an event study when sample selection includes securities grouped by certain common characteristics such as market capitalisation, prior returns, book-to-market ratio, and earnings-price ratio. Samples concentrated within any of these characteristics can easily lead to imprecise and potentially erroneous inferences. Anderson (2009) examines the effectiveness of friction modelling in dealing with biases induced by thin trading.

**4. Event-induced volatility**

Event-induced volatility has been a source of concern in event studies for some time. In an early discussion, Brown and Warner (1985, pp.22-23) articulate the crux of the issue:

"There is evidence of substantial increases in the variance of a security's return for the days around some types of events [Beaver (1968) Patell and Wolfson (1979) and Kalay and Lowenstein (1983)]. Christie (1983, table 4) suggests that the variance in some event studies could increase by a factor of almost two. … The most obvious implication of a variance increase is that standard procedures using a time-series of non-event period data to estimate the variance of the mean excess return will result in too many rejections of the null hypothesis that the mean excess return is equal to zero."

The issue continues to be a source of concern among capital market researchers. In particular, Aktas, de Bodt and Cousin (2007), Harrington and Shrider (2008) and Higgins and Peterson (1998) keenly argue that all events induce an increase in cross-sectional variance that must be estimated and adjustments embodied in all tests used to assess the statistical significance of event date abnormal returns.

The argument is easily overstated. Recall the defence attorney in our hypothetical insider trading case who might prepare an all-events-induce-variance argument that the event-date abnormal return measuring the insider's gain is not significant at a conventional level once it is scaled with a returns variance estimated across a collection of merger announcements. The argument raises an important issue: should an event date abnormal return $A_0$ that is significantly large relative to abnormal returns in the pre-event control period be accepted as *prima facie* statistical evidence? Or should the result be rejected when the abnormal return $A_0$ is not significantly large relative to abnormal returns across a collection of similar events? As shall be discussed immediately below, the issue hinges essentially on whether interest is focused on the immediate sample or on an extraneous population of similar events.

Consider the issue within the framework suggested in Harrington and Shrider (2008), where an event date abnormal return is a sum comprising two parts: 1) an event-induced return $\delta$ related only to the event and unrelated to the market, and 2) a firm-specific return $e_0$ unrelated to the event and the market.

$$A_0 = e_0 + \delta \qquad (14)$$

The event-induced return $\delta$ can be treated as a constant, or it can be treated as a drawing from a larger actual or hypothetical population of similar events. Treatment as a constant implies a conditional variance:

$$Var\left(A_0 \left|\delta\right.\right) = Var\left(e_0\right) \qquad (15)$$

Treatment as a drawing from a population of similar events implies an unconditional variance:

$$Var\left(A_0\right) = Var\left(e_0\right) + Var\left(\delta\right) \qquad (16)$$

When interest is focused solely on the sample, we treat $\delta$ as a constant. In our hypothetical insider trading case, interest is focused on the single event-induced return $\delta$ and not on a population from which it was allegedly drawn. Event-induced returns across an extraneous population of events have no import. To see this, consider a thought experiment where event-induced returns can be observed directly and so we know the sign and magnitude of $\delta$ exactly. In our hypothetical insider trading case, $\delta$ is the only datum of interest. If observed directly, no statistical test is required. However, the addition of a firm-specific return $e_0$ qua background noise obscures our view and limits us to probabilistic inference, specifically, that the event-induced return $\delta$ is positive, that is, $\Pr(\delta > 0) = \Pr(e_0 < A_0)$. This probability makes no reference to returns from an extraneous population of similar events.

The discussion above extends *mutatis mutandis* to samples with multiple event-induced returns, where $\overline{A}_0$, $\overline{\delta}$, and $\overline{e}_0$ would represent sample means as shown in equation (17).

$$\overline{A}_0 = \overline{e}_0 + \overline{\delta} \qquad (17)$$

When interest is concentrated on the sample mean event-induced return $\overline{\delta}$ but our view is obscured by background noise then we are limited to a probabilistic inference, that is, $\Pr(\overline{\delta} > 0) = \Pr(\overline{e}_0 < \overline{A}_0)$. This probability makes no reference to returns from an extraneous population of similar events.

Beyond applications in securities' law and litigation actions, there are many event-study settings where interest naturally concentrates on the sample, in particular the mean event-induced return $\overline{\delta}$. For example, Fields et al. (1998) examine intra-industry effects on publicly traded US insurance companies stemming from reports of Lloyd's financial distress appearing in the *New York Times* on April 27, 1993 and the *Wall Street Journal* three days later; Meznar et al. (1998) investigate the effects of corporate announcements to withdraw from South Africa during the apartheid era; Drakos (2004) investigates the effects of the September 11 terror attacks on a set of airline stocks listed on various international stock exchanges; and Calvo-Gonzalez (2007) examines how the Spanish-American agreements in the Pact of Madrid signed on 26 September 1953 affected asset prices in Franco-era Spain. In each of these studies, interest is focused on inferences regarding the mean event-induced return $\overline{\delta}$. Each of these studies examine historical events for which, in essence, the sample is the population.

At the other extreme of the sample-population spectrum, a sample represents a drawing from an infinite, hypothetical population of event-induced returns and $\overline{\delta}$ is merely the mean of the sample. In turn, this sample mean event-induced return has an expected value of $\mu$, the mean of the hypothetical population. If interest is focused on this population, then we might wish to project an inference beyond the sample mean $\bar{\delta}$ to the population mean $\mu$. In this case, a population variance estimate is required to form a probabilistic inference, specifically, that the population mean $\mu$ is positive, that is, $\Pr(\mu > 0)$.

A perusal of the finance literature suggests that many event studies limit themselves to statistical inferences about the mean event-induced return within the sample, i.e., $\bar{\delta}$, without projecting inferences onto the mean return of a parent population, i.e., $\mu$. This may suggest a conservative bias, though this bias is diminished by the inevitable follow-up studies with extended data sets that form ongoing streams of research into interesting and important topics. Nevertheless, projecting inferences onto a population larger than the sample can often be instructive. In forming these inferences, cross-sectional variance adjustment procedures advanced in Boehmer, Musumeci and Poulsen (1991), Sanders and Robins (1991), or Corrado and Zivney (1992) are aptly recommended.

However, cross-sectional variances adjustments are impractical in event studies with a small sample size and may be difficult even with medium-sized samples. Hilliard and Savickas (2002) suggest that estimating and making inferences regarding event-induced volatilities can be unreliable even when using advanced econometric techniques. As a potential detour, resort might be made to the rank test proposed in Corrado (1989), which has a high degree of robustness against event-induced variance. This has been demonstrated in several Brown and Warner-type studies cited earlier in Section 3. A theoretical basis for this robustness can be found in Pratt (1964), which is adapted immediately below to a brief mathematical demonstration of this robustness property.

Let $a_0$ denote an event date standardised abnormal return and let $a_t$ represent standardised abnormal returns within an $n$-day control period. To facilitate algebraic analysis, we assume that $a_0$ and the set of $a_t$ are distributed as standard normal, where $N(a)$ denotes the standard normal distribution function. The rank of the event day return $a_0$ among the $n+1$ returns from the event day plus control period, scaled as shown immediately below, has asymptotically the same distribution as $N(a_0)$.

$$\lim_{n\to\infty}\frac{Rank(a_0)}{n+2} \sim N(a_0)$$

Given that $a_0$ is distributed as standard normal, the function $N(a_0)$ is distributed as uniform with the variance $Var(N(a_0)) = 1/12$. However, if the variance of $a_0$ is greater than one, then $N(a_0)$ has a variance greater than 1/12 and is no longer distributed as uniform. To measure the impact of event-induced volatility on the rank test, we assess the variance of $N(a_0)$ when the variance of $a_0$ is $\lambda^2 \geq 1$. The expression for the asymptotic variance $Var(N(a_0))$ as a function of $\lambda$ is derived in Appendix A. In that expression, with $\lambda = 1$ we obtain the variance of a uniform distribution, i.e., $Var(N(a_0)) = 1/12$. However, $\lambda > 1$ yields $Var(N(a_0)) > 1/12$ causing an actual test level to deviate from its nominal level.

Table 1 below compares test levels for the rank test $R$ and the $t$-test $T$ when the nominal test level is $\alpha = 0.05$, where the event day abnormal return variance $\lambda^2 = Var(a_0)$ varies from 1 to 8 and where both test statistics are distributed as standard normal when $\lambda = 1$.

Table 1: Comparison of rank test and $t$-test specification with event-induced volatility, where $\lambda^2 = Var\left(a_0\right)$ measures the event day abnormal return variance. The rank test $R$ and $t$-test $T$ are both asymptotically distributed as standard normal when $\lambda = 1$.

$$R = \sqrt{\frac{12}{m}} \sum_{j=1}^{m} \left( N\left(a_{j,0}\right) - 1/2 \right) \qquad T = \frac{1}{\sqrt{m}} \sum_{j=1}^{m} a_{j,0}$$

| $\lambda^2$ | $\sqrt{12Var\left(N\left(a_0\right)\right)}$ | Rank test level (%) | Misspec-ification | T-test level (%) | Misspec-ification |
|---|---|---|---|---|---|
| 1 | 1.000 | 5.0 | 0.0 | 5.0 | 0.0 |
| 2 | 1.180 | 8.2 | 3.2 | 12.2 | 7.2 |
| 4 | 1.331 | 10.8 | 5.8 | 20.5 | 15.5 |
| 8 | 1.449 | 12.7 | 7.7 | 28.0 | 23.0 |

With a doubling of the event date abnormal return variance, i.e., $\lambda^2 = 2$, the actual Type-I error rate for the $t$-test is 12.2%, but just 8.2% for the rank test. Notice that in all cases where $\lambda > 1$, the degree of misspecification (as measured by deviations of actual Type-I error rates from the nominal 5% level) for the $t$-test is more than double the misspecification for the rank test. Thus, while affected by event-induced volatility, the rank test provides notable robustness against it.

**5. Event-induced returns**

The basic purpose of a financial event study is to detect the presence of event-induced returns within an event period – whether a single day or a longer period. This is normally done using a classical statistical approach specifying null and alternative hypotheses. Under a true null hypothesis, event-induced returns are not present within the event period, but they are present under a true alternative hypothesis. The power of a test determines its ability to detect a true alternative when it is present.

A parametric *t*-test statistic based on a weighted average of event date abnormal returns can be expressed in the general form shown in equation (18), where $A_{k,0}$ represents a day zero abnormal return for the $k^{\text{th}}$ security, $\sigma_k^2$ is its variance and $w_k$ is the weight attached to $A_{k,0}$.

$$T_P = \frac{\sum_{k=1}^{N} w_k A_{k,0}}{\sqrt{\sum_{k=1}^{N} w_k^2 \sigma_k^2}} \tag{18}$$

Let $\delta_k$ indicate an event-induced return for the $k^{\text{th}}$ firm in a sample of size *N*. Since the expected abnormal return is $E\left(A_{k,0}\right) = \delta_k$, the expected value of the *t*-test statistic is given as shown in equation (19).

$$E\left(T_P\right) = \frac{\sum_{k=1}^{N} w_k \delta_k}{\sqrt{\sum_{k=1}^{N} w_k^2 \sigma_k^2}} \tag{19}$$

Under a null hypothesis, $\delta_k = 0$ for all *k* and the expected value of the *t*-test statistic is zero. Under various alternative hypotheses $\delta_k \neq 0$ and the expected value of the *t*-test statistic departs from zero. An optimal test statistic must take into account the specific alternative being considered. The most popular alternative hypotheses in the event-studies literature are: 1) event-induced returns constant across the sample, that is, $\delta_k = \delta$ for all *k*, and 2) event-induced returns proportionate to abnormal return standard deviations, that is, $\delta_k = \delta \times \sigma_k$. The first alternative $\delta_k = \delta$ is motivated by simplicity and tradition. The second alternative $\delta_k = \delta \times \sigma_k$ is motivated by consideration of leverage effects, where changes in equity values are leveraged from underlying changes in asset values. It is not obvious which alternative $\delta_k = \delta$ or $\delta_k = \delta \times \sigma_k$ is more realistic across all, or even most applications. Conceivably, we may wish to consider testing for both alternatives as a robustness check.

For each of the two alternatives $\delta_k = \delta$ and $\delta_k = \delta \times \sigma_k$ there are optimal weights that maximise the theoretical power of a test to detect them. Seemingly, the choice of weights should be consistent with the choice of alternative hypothesis deemed most applicable. However, it has become an entrenched anomaly in the event studies literature that the most popular tests used in event studies, including the Patell *T*-test

(Patell, 1976), the Boehmer-Musumeci-Poulsen *T*-test (Boehmer, Musumeci and Poulsen, 1993) and the Corrado rank test (Corrado, 1989), are constructed to be optimal for the second alternative, i.e., $\delta_k = \delta \times \sigma_k$, yet almost all Brown and Warner-type simulation studies use the first alternative, i.e., $\delta_k = \delta$, as a basis for power comparisons among these tests. In what follows immediately below, we consider the potential consequences of mismatching weights optimal for one alternative but applied to detecting the other alternative.

Optimal weights for the *t*-test represented in equation (19) are $w_k \sim 1/\sigma_k^2$ under the alternative $\delta_k = \delta$ and $w_k \sim 1/\sigma_k$ under the alternative $\delta_k = \delta \times \sigma_k$. Test statistic expected values under the alternatives $\delta_k = \delta$ and $\delta_k = \delta \times \sigma_k$ with the optimal weights $w_k \sim 1/\sigma_k^2$ and $w_k \sim 1/\sigma_k$, respectively, are stated in equations (20) and (21).

$$E\left(T_P\right) = \delta \times \sqrt{\sum_{k=1}^{N} 1/\sigma_k^2} \tag{20}$$

$$E\left(T_P\right) = \delta \times \sqrt{N} \tag{21}$$

It is interesting to see by how much test power is affected when weights that are optimal for one alternative are mismatched to an alternative for which they are not optimal. For example, test statistic expected values under the alternatives $\delta_k = \delta$ and $\delta_k = \delta \times \sigma_k$ with the mismatched weights $w_k \sim 1/\sigma_k$ and $w_k \sim 1/\sigma_k^2$, respectively, are stated in equations (22) and (23).

$$E\left(T_P\right) = \delta \times \frac{\sum_{k=1}^{N} 1/\sigma_k}{\sqrt{N}} \tag{22}$$

$$E\left(T_P\right) = \delta \times \frac{\sum_{k=1}^{N} 1/\sigma_k}{\sqrt{\sum_{k=1}^{N} 1/\sigma_k^2}} \tag{23}$$

The ratio of equation (22) divided by equation (20) or equation (23) divided by equation (21) both yield the same ratio of test statistic expected values with sub-optimal weights to optimal weights. This ratio has the form of a Cauchy-Schwartz inequality as shown in equation (24), where a ratio value less than one measures the efficiency loss due to the use of non-optimal weights.

$$\frac{\sum_{k=1}^{N} 1/\sigma_k}{\sqrt{N \times \sum_{k=1}^{N} 1/\sigma_k^2}} \le 1 \tag{24}$$

The magnitude of the efficiency loss measured by equation (24) depends on the distribution of return standard deviations $\sigma_k$. If all standard deviations are equal, the ratio in equation (24) is equal to one and no efficiency loss occurs. More realistically, there may be considerable variation among the standard deviations. For example, suppose in a sample of size $N = 50$ the standard deviations $\sigma_k$ are uniformly distributed from $\sigma_1 = 0.50\%$ to $\sigma_{50} = 4.0\%$. In this case, the calculation shown immediately below yields a ratio value indicating a considerable loss of test power from the use of mismatched weights. Both more and less extreme examples are plausible.

$$\frac{\sum_{k=1}^{N} 1/\sigma_k}{\sqrt{N \sum_{k=1}^{N} 1/\sigma_k^2}} = \frac{\sum_{k=1}^{50} 1 \Big/ \left(0.5 + \frac{3.5 \times (k-1)}{49}\right)}{\sqrt{50 \times \sum_{k=1}^{50} 1 \Big/ \left(0.5 + \frac{3.5 \times (k-1)}{49}\right)^2}} = 0.829$$

As a consequence, Brown and Warner-type studies examining the most popular tests used in event studies have consistently underestimated their power in detecting the alternative for which they are optimal, i.e., $\delta_k = \delta \times \sigma_k$. However, this would have little effect on the legitimacy of studies where the mismatch of weights and alternatives has been consistent, specifically, where all tests optimal for the alternative

$\delta_k = \delta \times \sigma_k$ are compared using the alternative $\delta_k = \delta$. Relative test power would still be accurately assessed, though absolute test power would be understated.

Nonetheless, confusion arises when a newly proposed test constructed to be optimal for the alternative $\delta_k = \delta$ is compared to traditionally popular tests that are optimal for the alternative $\delta_k = \delta \times \sigma_k$, while using the alternative $\delta_k = \delta$ to make the comparison. Predictably, the proposed tests outperform the traditional tests in these studies. For example, Savickas (2003), Baixauli (2007), and Bremer and Zhang (2007) propose tests constructed to be optimal for the alternative $\delta_k = \delta$ and find that their proposed tests are more powerful than the popular Patell *T*-test and Boehmer-Musumeci-Poulsen *T*-test. A proper comparison would re-weight the traditional tests to be optimal for the alternative $\delta_k = \delta$. The optimal weights $w_k \sim 1/\sigma_k^2$ for the alternative $\delta_k = \delta$ were used to obtain equations (20) and (21). Re-weighting the Patell *T*-test using the weights $w_k \sim 1/\sigma_k^2$ yields the test statistic proposed in Collins and Dent (1984). Re-weighting the Boehmer-Musumeci-Poulsen *T*-test using the same weights yields the test statistic proposed in Sanders and Robbins (1993). These Collins-Dent and Sanders-Robbins *T*-tests are optimally weighted for the alternative $\delta_k = \delta$. Optimal weights for the rank test are considered next.

Appendix B shows that a nonparametric rank test statistic based on a weighted average of event date abnormal return ranks has the form specified in equation (25), where $F\left(A_{k,0}/\sigma_k\right)$ is the cumulative distribution function of standardized abnormal returns.

$$T_R = \frac{\sqrt{12}\sum_{k=1}^{N} w_k \left(F\left(A_{k,0}/\sigma_k\right) - 1/2\right)}{\sqrt{\sum_{k=1}^{N} w_k^2}} \qquad (25)$$

With $E\left(A_{k,0}\right)=\delta_k$, a first-order expansion of the expected value of the rank test statistic is stated in equation (27) below, where $\int f^2\left(a\right)da$ represents the expected value of the density function of standardised abnormal returns.

$$E\left(T_R\right)=\sqrt{12}\int f^2\left(a\right)da\times\frac{\sum_{k=1}^{N}w_k\delta_k/\sigma_k}{\sqrt{\sum_{k=1}^{N}w_k^2}} \tag{26}$$

Optimal weights for the non-parametric rank test are $w_k \sim 1/\sigma_k$ under the alternative $\delta_k=\delta$ and $w_k \sim 1$ under the alternative $\delta_k=\delta\times\sigma_k$. Rank test statistic expected values under the alternative $\delta_k=\delta$ with optimal weights $w_k \sim 1/\sigma_k$ and under the alternative $\delta_k=\delta\times\sigma_k$ with optimal weights $w_k \sim 1$ are stated in equations (27) and (28), respectively.

$$E\left(T_R\right)=\delta\sqrt{12}\int f^2\left(a\right)da\times\sqrt{\sum_{k=1}^{N}1/\sigma_k^2} \tag{27}$$

$$E\left(T_R\right)=\delta\sqrt{12N}\int f^2\left(a\right)da \tag{28}$$

Rank test statistic expected values under the alternatives $\delta_k=\delta$ and $\delta_k=\delta\times\sigma_k$ with the mismatched weights $w_k \sim 1$ and $w_k \sim 1/\sigma_k$, respectively, are stated in equations (29) and (30).

$$E\left(T_R\right)=\delta\sqrt{12}\int f^2\left(a\right)da\times\frac{\sum_{k=1}^{N}1/\sigma_k}{\sqrt{N}} \tag{29}$$

$$E\left(T_R\right)=\delta\sqrt{12}\int f^2\left(a\right)da\times\frac{\sum_{k=1}^{N}1/\sigma_k}{\sqrt{\sum_{k=1}^{N}1/\sigma_k^2}} \tag{30}$$

Dividing equation (29) by equation (27) or dividing equation (30) by equation (28) both yield the same Cauchy-Schwartz inequality stated in equation (24). Thus, the loss of efficiency for the nonparametric rank test due to mismatched weights is essentially the same proportion as for the parametric *t*-tests.

Appropriate comparisons of equations (20) through (23) with equations (27) through (30) indicate that the theoretical power of the rank test relative to the that of the *t*-test is proportionate to the value of $\sqrt{12}\int f^2(a)\,da$, which in turn depends on the shape of the density function $f(a)$. When $f(a)$ represents a standard normal density we have $\sqrt{12}\int f^2(a)\,da = \sqrt{3/\pi} = 0.977$, which accords with the well-known result that a nonparametric rank test has near-equivalent power compared to a parametric *t*-test when the underlying data is normally distributed. It is also a well-known result that the rank test will typically offer greater test power with non-normally distributed data, in particular with fat-tailed distributions (Hollander and Wolfe, 1973; Hettmansperger, 1984; Lehman, 1986; Manoukian, 1986). Campbell and Wasley (1996) provide a simple illustration of this point using the Edgeworth density specified in equation (31), where $\mu_3, \mu_4$ denote, respectively, the skewness and kurtosis of the distribution.

$$f(a) = \frac{e^{-a^2/2}}{\sqrt{2\pi}}\left(1 + \frac{\mu_3}{6}\left(a^3 - 3a\right) + \frac{\mu_4 - 3}{24}\left(a^4 - 6a^2 + 3\right)\right) \tag{31}$$

Integrating the squared density $f^2(a)$ yields,

$$\sqrt{12}\int f^2(a)\,da = \sqrt{\frac{3}{\pi}}\left(1 + \frac{5}{96}\mu_3^2 + \frac{105}{96^2}\left(\mu_4 - 3\right)^2 + \frac{\mu_4 - 3}{16}\right) \tag{32}$$

In the case of a normal density, skewness is zero and kurtosis is three and the right side of equation (33) collapses to $\sqrt{3/\pi} = 0.977$. However, with zero skewness $\mu_3 = 0$ and only modest kurtosis of, say, $\mu_4 = 5$, equation (32) yields a value of $\sqrt{12}\int f^2(a)\,da = 1.171$, indicating that the rank test is 17.1% more efficient than the *t*-test. This can be a notable advantage in event studies with a limited sample size, even when the underlying returns data only modestly departs from normality.

## Summary and conclusions

Over the last four decades, event studies have made an enormous contribution to capital markets research. Their contribution continues to expand, not only in accounting and finance, but in many other disciplines as well. The basic structure of an event study in which abnormal returns are measured by deviations from market model predictions is largely the same as when first introduced in Fama, Fisher, Jensen, and Roll (1969). However, the methods by which statistical inferences are obtained from abnormal returns have been and continue to be a popular area of investigation with constant attempts at innovation and refinement. Since event studies are applied in so many different settings, no single method dominates or even applies to all possible settings.

This paper has reviewed short-term event study methodology, with a particular focus on a select few important topics. First, considering the relative merits of parametric versus nonparametric methods it was concluded that parametric tests are suitable in event studies utilizing returns data from the New York stock exchange, but may yield inaccurate inferences with returns data from other markets where non-normality in the data can be severe. Nonparametric sign or rank tests are recommended for applications where robustness against non-normally distributed data is desirable. Second, the issue of event-induced variance was considered and it was concluded that event-induced variance was relevant in event studies attempting to extend inferences to a population larger than the specific sample examined. However, for many event studies where the sample constitutes the entire population of interest cross-sectional variance adjustments to compensate for event-induced variance may result in incorrect inferences. Finally, it was concluded that an optimal scheme for cross-sectional weighting of abnormal returns depends on the specific alternative hypothesis being tested. Comparing tests optimally weighted for one alternative with other tests not optimally weighted for the same alternative biases any conclusion related to the relative power of the tests. A proper comparison would use a weighting scheme optimal for the alternative considered across all test comparisons.

**Appendix A**

The asymptotic variance of $N\left(a_0\right)$ is mathematically determined as shown below, where $N_2\left(x,y;\rho\right)$ denotes the bivariate normal distribution function.

$$Var\left(N\left(a_0\right)\right)=\int_{-\infty}^{\infty}N^2\left(a_0\right)n\left(\lambda a_0\right)da_0-\left(\int_{-\infty}^{\infty}N\left(a_0\right)n\left(\lambda a_0\right)da_0\right)^2$$
$$=\frac{\lambda}{2\pi}\left(2\arctan\left(1/\lambda\right)-\arccos\left(\frac{\lambda^2}{1+\lambda^2}\right)\right)+2N_2\left(0,0;\frac{\lambda}{\sqrt{1+\lambda^2}}\right)-\frac{3}{4}$$

The error function is defined by,

$$erf\left(x\right)=\frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt=2\int_0^{\sqrt{2}x}\frac{e^{-z^2/2}}{\sqrt{2\pi}}dz=2N\left(\sqrt{2}x\right)-1\quad x\geq 0$$

From Gradshteyn and Ryzhik (1994, p.941), we have this equality:

$$\int_0^{\infty}\left(1-erf\left(x\right)\right)^2 e^{-\beta x^2}dx=\frac{1}{\sqrt{\beta\pi}}\left(2\text{arctg}\left(\sqrt{\beta}\right)-\arccos\left(\frac{1}{1+\beta}\right)\right)=\frac{A\left(\beta\right)}{\sqrt{\beta\pi}}$$

We wish to obtain the following integral:

$$\int_{-\infty}^{\infty}N^2\left(x\right)\frac{e^{-x^2/2\lambda^2}}{\sqrt{2\pi\lambda^2}}dx=\int_{-\infty}^{\infty}N^2\left(\lambda z\right)\frac{e^{-z^2/2}}{\sqrt{2\pi}}dz=\int_{-\infty}^{\infty}N^2\left(\lambda z\right)n\left(z\right)dz$$

The first part of the derivation proceeds as follows:

$$\int_{-\infty}^{\infty}N^2\left(\lambda z\right)n\left(z\right)dz=\int_0^{\infty}N^2\left(\lambda z\right)n\left(z\right)dz+\int_{-\infty}^{0}N^2\left(\lambda z\right)n\left(z\right)dz$$
$$=\int_0^{\infty}N^2\left(\lambda z\right)n\left(z\right)dz+\int_0^{\infty}N^2\left(-\lambda z\right)n\left(z\right)dz$$
$$=\int_0^{\infty}N^2\left(\lambda z\right)n\left(z\right)dz+\int_0^{\infty}\left(1-N\left(\lambda z\right)\right)^2 n\left(z\right)dz$$
$$=2\int_0^{\infty}N^2\left(\lambda z\right)n\left(z\right)dz+\frac{1}{2}-2\int_0^{\infty}N\left(\lambda z\right)n\left(z\right)dz$$
$$=2\int_0^{\infty}N^2\left(\lambda z\right)n\left(z\right)dz+\frac{1}{2}-2\left(\int_{-\infty}^{\infty}N\left(\lambda z\right)n\left(z\right)dz-\int_{-\infty}^{0}N\left(\lambda z\right)n\left(z\right)dz\right)$$
$$=2\int_0^{\infty}N^2\left(\lambda z\right)n\left(z\right)dz+\frac{1}{2}-2\left(N\left(0\right)-N_2\left(0,0;\frac{-\lambda}{\sqrt{1+\lambda^2}}\right)\right)$$
$$=2\int_0^{\infty}N^2\left(\lambda z\right)n\left(z\right)dz+\frac{1}{2}-2N_2\left(0,0;\frac{\lambda}{\sqrt{1+\lambda^2}}\right)$$

Using the change of variable $\lambda z = \sqrt{2}x$ and the error function $erf\left(x\right) = 2N\left(\sqrt{2}x\right) - 1$ defined above yields,

$$\int_0^\infty N^2\left(\lambda z\right) n\left(z\right) dz = \frac{1}{\lambda\sqrt{\pi}} \int_0^\infty N^2\left(\sqrt{2}x\right) e^{-x^2/\lambda^2} dx$$

$$= \frac{1}{4\lambda\sqrt{\pi}} \int_0^\infty \left(1 + erf\left(x\right)\right)^2 e^{-x^2/\lambda^2} dx$$

$$= \frac{1}{4\lambda\sqrt{\pi}} \int_0^\infty \left(2 - \left(1 - erf\left(x\right)\right)\right)^2 e^{-x^2/\lambda^2} dx$$

$$= \frac{1}{4\lambda\sqrt{\pi}} \int_0^\infty \left(4 + \left(1 - erf\left(x\right)\right)^2 - 4\left(1 - erf\left(x\right)\right)\right) e^{-x^2/\lambda^2} dx$$

$$= \frac{1}{\lambda\sqrt{\pi}} \int_0^\infty erf\left(x\right) e^{-x^2/\lambda^2} dx + \frac{1}{4\lambda\sqrt{\pi}} \int_0^\infty \left(1 - erf\left(x\right)\right)^2 e^{-x^2/\lambda^2} dx$$

$$= \frac{1}{\lambda\sqrt{\pi}} \int_0^\infty \left(2N\left(\sqrt{2}x\right) - 1\right) e^{-x^2/\lambda^2} dx + \frac{\lambda A\left(1/\lambda^2\right)}{4\pi}$$

$$= \int_0^\infty \left(2N\left(\lambda z\right) - 1\right) n\left(z\right) dz + \frac{\lambda A\left(1/\lambda^2\right)}{4\pi}$$

$$= 2\left(N\left(0\right) - N_2\left(0, 0; \frac{-\lambda}{\sqrt{1+\lambda^2}}\right)\right) - \frac{1}{2} + \frac{\lambda A\left(1/\lambda^2\right)}{4\pi}$$

$$= 2N_2\left(0, 0; \frac{\lambda}{\sqrt{1+\lambda^2}}\right) - \frac{1}{2} + \frac{\lambda A\left(1/\lambda^2\right)}{4\pi}$$

Finally, putting it all together yields,

$$\int_{-\infty}^\infty N^2\left(x\right) \frac{e^{-x^2/2\lambda^2}}{\sqrt{2\pi\lambda^2}} dx = \int_{-\infty}^\infty N^2\left(\lambda z\right) n\left(z\right) dz =$$

$$\frac{\lambda}{2\pi}\left(2\operatorname{arctg}\left(1/\lambda\right) - \arccos\left(\frac{\lambda^2}{1+\lambda^2}\right)\right) + 2N_2\left(0, 0; \frac{\lambda}{\sqrt{1+\lambda^2}}\right) - \frac{1}{2}$$

Subtracting $\left(\int_{-\infty}^\infty N\left(\lambda z\right) n\left(z\right) dz\right)^2 = \frac{1}{4}$ yields $Var\left(N\left(a_0\right)\right)$ in equation (17).

## Appendix B

Scaling the abnormal return ranks $Rank\left(A_{k,0}\right)$ by the number of returns in the control period $n_k$ for the $k^{\text{th}}$ firm in the sample yields an empirical distribution function, which asymptotically converges to the true distribution under the null hypothesis.

$$\lim_{n_k \to \infty} \frac{Rank\left(A_{k,0}\right)}{n_k + 2} = F_k\left(A_{k,0}\right)$$

For convenience, we shall assume that the distribution of the abnormal returns $A_{k,0}$ and $A_{j,0}$ differ only in scale and so the standardised abnormal returns $A_{k,0}/\sigma_k$ and $A_{j,0}/\sigma_j$ are identically distributed. The rank test in a standardised asymptotic form is,

$$T_R = \frac{\sqrt{12}\sum_{k=1}^{N} w_k\left(F\left(A_{k,0}/\sigma_k\right) - 1/2\right)}{\sqrt{\sum_{k=1}^{N} w_k^2}}$$

Under a null hypothesis, the rank test form above has a mean of zero and a variance of one. Under the general alternatives $\delta_k$, the rank test has the following expected value:

$$E\left(T_R\right) = \frac{\sqrt{12}\sum_{k=1}^{N} w_k E\left(F\left(\left(A_{0,k} + \delta_k\right)/\sigma_k\right) - 1/2\right)}{\sqrt{\sum_{k=1}^{N} w_k^2}}$$

$$= \frac{\sqrt{12}\sum_{k=1}^{N} w_k Ef\left(A_{0,k}/\sigma_k\right)\delta_k/\sigma_k}{\sqrt{\sum_{k=1}^{N} w_k^2}}$$

$$= \sqrt{12}\int f^2\left(a\right)da \times \frac{\sum_{k=1}^{N} w_k \delta_k/\sigma_k}{\sqrt{\sum_{k=1}^{N} w_k^2}}$$

**References**

Affleck-Graves, J., Callahan, C.M., and R. Ramanan, 2000, Detecting abnormal bid-ask spread: A comparison of event study methods, *Review of Quantitative Finance and Accounting* 14, 45-65.

Ahern, K.R., 2009, Sample selection and event study estimation, *Journal of Empirical Finance* 16, 466-482.

Aktas, N., de Bodt, E., and J.G. Cousin, 2007, Event studies with a contaminated estimation period, *Journal of Corporate Finance* 13(1), 129-145.

Anderson, W., 2009, Alternative event study methodology for detecting dividend signals in the context of joint dividend and earnings announcements, *Accounting and Finance* 49, 247-265.

Armitage, Seth, 1995, Event study methods and evidence on their performance, *Journal of Economic Surveys* 8(4), 25-52.

Ashley, J., 1962, Stock prices and changes in earnings and dividends: Some empirical results, *Journal of Political Economy* 70(1), 82-85.

Baixauli, J.S., 2007, Abnormal performance in small portfolios with event-induced volatility: The case of stock splits, *Journal of Financial Research* 30(1), 35–52.

Ball, R. and P. Brown, 1968, An empirical evaluation of accounting income numbers, *Journal of Accounting Research* 6, 159-178.

Barker, C.A., 1956, Effective stock splits, *Harvard Business Review* 34(1), 101-106.

Barker, C.A., 1957, Stock splits in a bull market, *Harvard Business Review* 35(3), 72-79.

Barker, C.A., 1958, Evaluation of stock dividends, *Harvard Business Review* 36(4), 99-114.

Bartholdy, J., Olson, D., and P. Peare, 2007, Conducting event studies on a small stock exchange, *European Journal of Finance* 13(3), 227-252.

Beaver, W.H., 1968, The information content of annual earnings announcements, *Empirical research in accounting: Selected studies, Supplement to the Journal of Accounting Research*, 67-92.

Bernard, V., 1987, Cross-sectional dependence and problems in inference in market-based accounting research, *Journal of Accounting Research* 25(1), 1-48.

Bhagat, S. and R. Romano, 2002, Event studies and the law: part I. Technique and corporate litigation, *American Law and Economics Review* 4(1), 141-167.

Binder, J.J., 1998, The event study methodology since 1969, *Review of Quantitative Finance and Accounting* 11, 111-137.

Boehmer, E., Musumeci, J., and A.B. Poulsen, 1991, Event-study methodology under conditions of event-induced variance. *Journal of Financial Economics* 30(2), 253-272.

Bremer, R. and Z.H. Zhang, 2007, Extensions of the standardized cross-sectional approach to short-horizon event studies, *Journal of Financial Research.* 30(4), 495-513.

Brown, S.J. and J.B. Warner, 1980, Measuring security price performance, *Journal of Financial Economics* 8(3), 205-258.

Brown, S.J. and J.B. Warner, 1985, Using daily stock returns: The case of event studies, *Journal of Financial Economics* 14(1), 3-31.

Calvo-Gonzalez, O., 2007, American military interests and economic confidence in Spain under the Franco dictatorship, *Journal of Economic History* 67(3), 740-767.

Campbell, C.J. and C.E. Wasley, 1993, Measuring security price performance using daily NASDAQ returns, *Journal of Financial Economics* 33(1), 73-92.

Campbell, C.J. and C.E. Wasley, 1996, Measuring abnormal daily trading volume for samples of NYSE/ASE and NASDAQ securities using parametric and nonparametric test statistics, *Review of Quantitative Finance and Accounting* 6(3), 309-326.

Chandra, R., Rohrbach, K., and G.L. Willinger, 1992, Rank tests for detecting location shift in the distribution of abnormal returns: An extension, *Contemporary Accounting Research* 9, 296-305.

Chandra, R., Rohrbach, K., and G.L. Willinger, 1995, A comparison of the power of parametric and nonparametric tests of location shift for event studies, *The Financial Review* 30(4), 685-710.

Charest, G., 1978, Dividend information, stock returns, and market efficiency – II. *Journal of Financial Economics* 6, 297-330.

Chatterjee, C., Pacini, C., and V. Sambamurthy, 2002, The shareholder-wealth and trading-volume effects of information-technology infrastructure investments. *Journal of Management Information Systems* 19(2), 7-42.

Christie, A., 1983, On information arrival and hypothesis testing in event studies, Working paper, University of Rochester.

Chou, P.H., 2004, Bootstrap tests for multivariate event studies, *Review of Quantitative Finance and Accounting* 23, 275-290.

Cichello, M. and D.J. Lamdin, 2006, Event studies and the analysis of antitrust, *International Journal of the Economics of Business* 13(2), 229-245.

Collins, D.W. and W.T. Dent, 1984, A comparison of alternative testing methodologies used in capital markets research, *Journal of Accounting Research* 22(1), 48-84.

Corrado, C.J., 1989, A nonparametric test for abnormal security-price performance in event studies, *Journal of Financial Economics* 23, 385-395.

Corrado, C.J. and C. Truong, 2008, Conducting event studies with Asia-Pacific security market data, *Pacific-Basin Finance Journal* 16, 493-521.

Corrado, C.J. and T.L. Zivney, 1992, The specification and power of the sign test in event study hypothesis tests using daily stock returns, *Journal of Financial and Quantitative Analysis* 27(3), 465-478.

Cowan, A.R., 1992, Nonparametric event study tests, *Review of Quantitative Finance and Accounting* 2, 343-358.

Cowan, A.R., 1993, Tests for cumulative abnormal returns over long periods: Simulation evidence, *International Review of Financial Analysis* 2(1), 51-68.

Cowan, A.R. and A.M.A. Sergeant, 1996, Trading frequency and event study test specification, *Journal of Banking & Finance* 20(10), 1731-1757.

Darkow, I.L., Kaup, C., and D. Schiereck, 2008, Determinants of merger & acquisition success in global logistics, *International Journal of Logistics Research and Applications* 11(5), 333-345.

Dodd, P. and J.B. Warner, 1983, On corporate governance: A study of proxy contests, *Journal of Financial Economics* 11(1), 401-438.

Dolley, J.C., 1933, Characteristics and procedure of common stock split-ups, *Harvard Business Review* 11, 316-326.

Dombrow, J. Rodriguez, M., and C.F. Sirmans, 2000, A complete nonparametric event study approach, *Review of Quantitative Finance and Accounting* 14, 361-380.

Drakos, K., 2004, Terrorism-induced structural shifts in financial risk: airline stocks in the aftermath of the September 11th terror attacks, *European Journal of Political Economy* 20, 435-446.

Dyckman, T., Philbrick, D., and J. Stephan, 1984, A comparison of event study methodologies using daily stock returns: A simulation approach, *Journal of Accounting Research* 22, 1-30.

Fama, E.F., Fisher, L., Jensen, M.C., and R. Roll, 1969, The adjustments of stock prices to new information, *International Economic Review* 10(1), 1-21.

Fields, J.A., Klein, L.S., and E.G. Myskowski, 1998, Lloyd's financial distress and contagion within the US property and liability insurance industry, *Journal of Risk and Uncertainty* 16(2), 173-185.

Ford, G.S. and A.D. Kline, 2007, Event studies for merger analysis: An evaluation of the effects of non-normality on hypothesis testing, available at SSRN: http://ssrn.com/abstract=925953

Fornell, C., Mithas, S., and F.V. Morgeson, 2006, Customer satisfaction and stock prices: High returns, low risk, *Journal of Marketing* 70(1), 3-14.

Ghosh, C., Rodriguez, M., and C.F. Sirmans, 1995, Gains from corporate-headquarters relocations – Evidence from the stock market, *Journal of Urban Economics* 38(3), 291-311.

Giaccotto, C. and J.M. Sfiridis, 1996, Hypothesis testing in event studies: The case of variance changes, *Journal of Economics and Business* 48(4), 349-370.

Gong, S.X.H., 2009, Event study in transport research: Methodology and applications, *Transport Reviews* 29(2), 207-222.

Gradshteyn, I.S. and I.M. Ryzhik, 1994, *Table of Integrals, Series, and Products* (5th ed.), Academic Press: New York.

Greene, W. H., 2003, *Econometric Analysis*, Upper Saddle River, NJ: Prentice-Hall.

Hamill, P.A., Opong, K.K., and P. McGregor, 2002, Equity option listing in the UK: a comparison of market-based research methodologies, *Journal of Empirical Finance* 9, 91-108.

Harrington, S.E. and D.G. Shrider, 2007, All events induce variance: Analyzing abnormal returns when effects vary across firms, *Journal of Financial and Quantitative Analysis* 42(1), 229-256.

Hein, S.E. and P. Westfall, 2004, Improving tests of abnormal returns by bootstrapping the multivariate regression model with event parameters, *Journal of Financial Econometrics* 2(3), 2004, 451-471.

Heinkel, R. and A. Kraus, 1988, Measuring event impacts in thinly traded stocks, *Journal of Financial and Quantitative Analysis* 23(1), 71-88.

Henderson, G.V., Jr., 1990, Problems and solutions in conducting event studies, *Journal of Risk and Insurance* 57, 282-306.

Hettmansperger, T.P., 1984, *Statistical Inference Based on Ranks*, New York: John Wiley & Sons.

Higgins, E.J. and D.R. Peterson, 1998, The power of one and two sample t-statistics given event-induced variance increases and nonnormal stock returns: A comparative study, *Quarterly Journal of Business and Economics* 37(1), 27-50.

Hilliard, J.E. and R. Savickas, 2002, On the statistical significance of event effects on unsystematic volatility, *Journal of Financial Research* 25(4), 447-463.

Hollander, M. and D.A. Wolfe, 1973, *Nonparametric Statistical Methods*, New York: John Wiley & Sons.

Jackson, J.D., Kline, A.D., and S.J. Skinner, 2006, The impact of non-normality and misspecification on merger event studies, *International Journal of the Economics of Business* 13(2), 247-264.

Jain, P., 1986, Analyses of the distribution of market model prediction errors for daily returns data, *Journal of Accounting Research* 24(1), 76-96.

Johnston, M.A., 2007, A review of the application of event studies in marketing, *Academy of Marketing Science Review* 11(4), 1-31.

Joiner, B.L. and J.R. Rosenblatt, 1971, Some properties of the range in samples from Tukey's symmetric lambda distribution, *Journal of the American Statistical Association* 66, 394-399.

Kalay, A. and U. Lowenstein, 1983, Are 'excess returns' necessarily economic profits? An analysis of changes in risk and their pricing effects in 'event time', Working paper, New York University.

Kothari, S.P. and Warner, J.B. Econometrics of event studies, 2005, Chapter 1 in B. Espen Eckbo (ed.), *Handbook of Corporate Finance: Empirical Corporate Finance* (Handbooks in Finance Series, Elsevier/North-Holland).

Kramer, L., 2001, Alternative methods for robust analysis in event study applications, *Advances in Investment Analysis and Portfolio Management* 8, 1-26.

Lamdin, D.J., 2001, Implementing and interpreting event studies of regulatory changes, *Journal of Economics and Business* 53, 171–83.

Larcker, D.F. et al., 1980, Testing for market efficiency: a comparison of the cumulative average residual methodology and intervention analysis, *Journal of Financial and Quantitative Analysis* 2, 267-287.

Lehman, E.L., 1986, *Testing Statistical Hypotheses*, New York: John Wiley & Sons.

Lummer, S. and J. McConnell, 1989, Further evidence on the bank lending process and the capital market response to bank loan agreements, *Journal of Financial Economics* 25(1), 88-122.

MacKinlay, A.C., 1997, Event studies in economics and finance, *Journal of Economic Literature* 35(1), 13-39.

,
Manoukian, E.B., 1986, *Mathematical Nonparametric Statistics*, New York: Gordan & Breach Science Publishers.

Maynes, E. and J. Rumsey, 1993, Conducting event studies with thinly traded stocks. *Journal of Banking & Finance* 17(1), 145-157.

McConnell, J. and C. Muscarella, 1985, Corporate capital expenditure decisions and the market value of the firm, *Journal of Financial Economics* 14(3), 399-422.

McKenzie, A.M. and M. Thomsen, 2001, The effect of E-coli O157: H7 on beef prices. *Journal of Agricultural Resource Economics* 26(2), 431-444.

McKenzie, A.W., Thomsen, M., and B.L. Dixon, 2004, The performance of event study approaches using daily commodity futures returns, *Journal of Futures Markets* 24(6), 533-555.

McWilliams, A. and D. Siegel, 1997, Event studies in management research: Theoretical and empirical issues, *Academy of Management Journal* 40, 626-657.

McWilliams, T.P. and V.B. McWilliams, 2000, Another look at theoretical and empirical issues in event study methodology, *Journal of Applied Business Research* 16(3), 1-11.

Meznar, M.B., Nigh, D., and C.C.Y. Kwok, 1998, Announcements of withdrawal from South Africa revisited: Making sense of contradictory event study findings. *Academy of Management Journal* 41(6), 715-730.

Misra, L. and V.S. Rao, 2009, The launching of transactional web sites: Market response to announcements by incumbent B2C companies, *Journal of Organizational Computing and Electronic Commerce* 19(1), 50-82.

Miller, M. H. and F. Modigliani, F., 1961, Dividend policy, growth and the valuation of shares, *Journal of Business* 34, 411-433.

Miller, M.H. and F. Modigliani, 1963, Corporate income taxes and the cost of capital: a correction, *American Economic Review* 53(3), 433-443.

Mitchell, M. L. and J.M. Netter, 1994, The role of financial economics in securities fraud cases: applications at the Securities and Exchange Commission, *Business Lawyer* 49(2), 545-590.

Modigliani, F. and M.H. Miller, 1958, The cost of capital, corporation finance and the theory of investment, *American Economic Review* 48(3), 261-297.

Mucklow, B., 1994, Market microstructure: An examination of the effects on intraday event studies, *Contemporary Accounting Research* 10(2), 355-382.

Myers, J.A. and A. Bakay, 1948, Influence of stock split-ups on market price, *Harvard Business Review* 26, 251-255.

Nicolau, J.L., 2002, Assessing new hotel openings through an event study. *Tourism Management* 23(1), 47-54.

Nicolau, J.L. and R. Sellers, 2002, The stock market's reaction to quality certification: Empirical evidence from Spain, *European Journal of Operational Research* 142(3), 632-641.

Patell, J.M., 1976, Corporate forecasts of earnings per share and stock price behavior: Empirical tests, *Journal of Accounting Research* 14, 246-276.

Patell, J. and M. Wolfson, 1979, Anticipated information releases reflected in call option prices, *Journal of Accounting and Economics* 1, 117-140.

Penman, S. 1982, Insider trading and the dissemination of firms' forecast information, *Journal of Business* 55, 479-503.

Peterson, P.P. 1989, Event studies: A review of issues of methodology, *Quarterly Journal of Business and Economics* 28, 36-66.

Pratt, J.W., 1964, Robustness of some procedures for the two-sample location problem, *Journal of the American Statistical Association* 59(307), 665-680.

Rose, C., 2003, Impact of investor meetings/presentations on share prices, insider trading and securities regulation. *International Review of Law and Economics* 23(3), 227-236.

Roztocki, N. and H.R. Weistroffer, 2009, Event studies in information systems research: An updated review, *Proceedings of the Fifteenth Americas Conference on Information Systems*, Available at: http://ssrn.com/abstract=1457432.

Sanders, R.W. Jr. and R.P. Robins, 1991, Discriminating between wealth and information effects in event studies in accounting and finance research, *Review of Quantitative Finance and Accounting* 1(3), 307-330.

Savickas, R., 2003, Event induced volatility and tests for abnormal performance. *Journal of Financial Research* 26(2), 165-178.

Schwert, G. W., 1981, Measuring the effects of regulation: evidence from the capital markets, *Journal of Law and Economics* 24, 121-145.

Seiler, M.J., 2000, The efficacy of event-study methodologies: Measuring EREIT abnormal performance under conditions of induced variance, *Journal of Financial and Strategic Decisions* 13(1) 101-112.

Serra, A.P., 2002, Event study tests: A brief survey, *Working Papers da FEP* 117.

Sharpe, W.F., 1964, Capital asset prices: A theory of equilibrium under conditions of risk, *Journal of Finance* 19, 425-442.

Shevlin, T.J. 1981, Measuring abnormal performance on the Australian securities market, *Australian Journal of Management* 6(1), 67-107.

Thompson, R., 1995, Empirical methods of event studies in corporate finance, Ch. 29 of *Finance* (eds., R. Jarrow, V. Maksimovic and W. Ziemba, Vol. 9 of *Handbooks in Operations Research and Management Science*). Amsterdam: North Holland.

Thomsen, M. and A.M. McKenzie, 2001, Market incentives for safe foods: An examination of shareholder losses from meat and poultry recalls, *American Journal of Agricultural Economics* 82(3), 2003, 526-538.

Tuck, C.E., 2005, A quality award and stock market reaction: Evidence from the European Union, *Total Quality Management* 16(8-9), 979-986.

Winans, R.F., 1984, *Trading Secrets: Seduction and Scandal at the Wall Street Journal*, New York: St. Martin's Press.

Wulf, J., 2004, Do CEOs in mergers trade power for premium? Evidence from "mergers of equals", *Journal of Law, Economics, and Organization* 20(1), 60-101.

Zivney, T.L. and D.J. Thompson II, 1989, The specification and power of the sign test in measuring security price performance: Comments and analysis, *The Financial Review* 24, 581-588.