



WP

Jan Bartholdy, Dennis Olson & Paula Peare

Conducting event studies on a small stock exchange

Finance
Research Group

Conducting event studies on a small stock exchange

Jan Bartholdy*
jby@asb.dk

Aarhus School of Business
Department of Business Studies
Fuglesangs Allé 4
DK-8210 Aarhus V

Dennis Olson
dolson@usharjah.edu

American University of Sharjah

Paula Peare
ppe@asb.dk

* Corresponding author.

Conducting event studies on a small stock exchange

Abstract

This paper analyses whether it is possible to perform an event study on a small stock exchange with thinly traded stocks. The main conclusion is that event studies can be performed provided that certain adjustments are made. First, a minimum of 25 events appears necessary to obtain acceptable size and power in statistical tests. Second, trade to trade returns should be used. Third, one should not expect to consistently detect abnormal performance of less than about 1% (or perhaps even 2%), unless the sample contains primarily thickly traded stocks. Fourth, nonparametric tests are generally preferable to parametric tests of abnormal performance. Fifth, researchers should present separate results for thickly and thinly traded stock groups. Finally, when nonnormality, event induced variance, unknown event day, and problems of very thin trading are all considered simultaneously, no one test statistic or type of test statistic dominates the others.

Keywords: Event studies, thin trading

1. INTRODUCTION

Following the seminal articles by Ball and Brown (1968) and Fama, Fisher, Jensen and Roll (1969), event studies have become one of the most widely used empirical techniques in finance and accounting.¹ They are normally designed to detect abnormal price changes in financial assets in the time period around various events. Crucial to the process is the ability for the researcher to accurately determine what constitutes abnormal performance, regardless of the institutional setting. Although event study methods are well developed and often used to test financial theories for the US and other well-established stock exchanges, there is some concern regarding efficiency when applied to small stock exchanges dominated by thinly traded stocks (i.e. stocks that do not trade every day).

Heinkel and Kraus (1988), Campbell and Wasley (1993) and Cowan and Sergeant (1996) have discussed various modifications in event study techniques to adjust for thin trading.

¹Although event studies date back to the 1930s, the papers by Ball and Brown (1968) and Fama et. al. (1969) introduced the methods used today. MacKinlay (1997) contains an excellent description of the history and implementation of event studies.

However, these studies are based upon US stock return data derived from the CRSP tapes—the primary source of data for most event studies. The institutional setting in other countries differs, and on some stock exchanges the problem is not just thin trading, but one of very thinly traded stocks. In such situations, questions arise as to whether event studies can be reliably conducted with daily data, or whether researchers need to use weekly or monthly returns data. Maynes and Rumsey's (1993) study of thin, moderate, and thickly traded Canadian stocks provides a good framework for conducting event studies on a small stock exchange. Using trade to trade returns to adjust for thin trading and nonparametric test statistics to deal with nonnormality, they were able to successfully detect daily abnormal returns in all stocks for all three trading frequencies. Their findings are encouraging, but it is not known if their results generalise to event studies outside of North America. A particular concern is that Maynes and Rumsey's (1993) sample of thinly traded stocks averaged 3.67 days between trades. Even the average stock on some small exchanges may trade less frequently than this. For example, the average number of days between trades for a stock on the Copenhagen Stock Exchange (CSE) was 6.86 days in 1990 and 4.17 days in 2001. During various years from 1990 to 2001, the trading frequency for thinly traded CSE stocks ranged from 11 to 23 days.

The purpose of this paper is to provide further evidence about the efficiency of event study techniques when applied to thinly traded stocks by examining data from the Copenhagen Stock Exchange (CSE). It is a small stock exchange where thin (or very thin) trading is extensive. The study applies Brown and Warner's (1980, 1985) well-event study methods to three samples of Danish stocks differentiated by their relative trading frequencies (thick, medium, and thin). Following Maynes and Rumsey (1993), we use trade to trade returns to deal with thin trading and apply a battery of parametric and nonparametric tests, as discussed by Brown and Warner (1985) to detect abnormal performance. In addition, we apply the procedures of Boehmer, Musumeci, and Poulsen (1991) to examine event induced variance and consider the case where the event day is uncertain.

Specifically, we simulate an event study by randomly selecting event days for various stocks. On the event day, daily returns are increased by increments ranging from zero to two percent to simulate the impact of new information. This procedure is repeated 1000 times for portfolios of 10, 25, and 50 stocks to determine the power and size of various parametric and nonparametric test statistics. Abnormal performance in both frequently and infrequently traded Danish stocks was successfully detected—providing evidence that event studies probably can be successfully conducted on other small stock exchanges.

2. CALCULATION OF RETURN UNDER THIN TRADING

The impact of new information on the value of a given stock is measured by the difference between the actual return at time t , r_t , and the expected return $E(r_t)$. This difference is called the abnormal return, A_t , and is given by:

$$A_t = r_t - E[r_t]$$

Expected return can be obtained from estimation of the market model so

$$E[r_t] = \hat{\alpha} + \hat{\beta} r_{mt}$$

where r_{mt} is the market return at time t and $\hat{\alpha}$ and $\hat{\beta}$ are obtained from OLS estimation of the regression

$$r_t = \alpha + \beta r_{mt} + \varepsilon_t$$

during the estimation period.² The standard estimation period is between 200 and 250 observations; i.e. about a year of trading prior to a three day event window. Estimation of abnormal returns for frequently traded stocks is therefore relatively straightforward.³

Nontrading, and the subsequent problem of missing return observations, is not encountered in most event studies. Most empirical evidence for event studies in finance comes from a single data source—the CRSP data files for US stocks. Daily price and returns data are readily available for large stocks and for most small stocks. If a stock does not trade on a given day, price is typically recorded on the CRSP tapes as the average of the bid-ask quotes. Provided these quotes are realistic, a daily return series with no missing observations can be calculated. In contrast, there may be no meaningful bid or ask prices in the order book on any given day for thinly traded stocks on other stock exchanges. There is no standard method to calculate daily

² Brown and Warner (1980, 1985) used two additional methods for measuring expected return: mean adjusted return and market adjusted return. For mean adjusted return, expected return is set equal to the mean of the return over the estimation period (equivalent to setting $\beta = 0$ in the market model). For market adjusted return, the mean of the market return over the estimation period is used for expected return (equivalent to setting $\alpha = 0$ and $\beta = 1$ in the market model). These methods are simpler than the market model and significantly reduce computing requirements—a major concern in the 1980s. Since the market model includes both of these methods as special cases, and computing requirements are no longer a concern, only the market model is considered here.

³ As pointed out by Scholes and Williams (1977), OLS beta estimates are biased downward for securities more thinly traded than the market index. However, OLS probably remains the best technique for estimating abnormal returns in the presence of thin trading. Cowan and Sergeant (1996) report that the Scholes-Williams correction for nonsynchronous trading provides no significant benefit in coping with the problems caused by thin trading. In fact, the Scholes-Williams (1977) and Dimson (1979) corrections for nonsynchronicity can even make matters worse. For example, in a study of the biases induced by thin trading in New Zealand, Bartholdy and Riding (1994) found that OLS beta estimates were less biased, more efficient, and at least as consistent as estimates made using the Dimson or Scholes-Williams corrections.

stock returns on the days when a stock does not trade, but four possible techniques could be employed to adjust for nontrading.

Probably the least satisfactory means of dealing with nontrading is to calculate simple returns for each stock only for days for which consecutive prices are available. Then, subtract the market return on these days to obtain daily abnormal return and ignore individual stock and market returns on other days. This technique gives unbiased estimates of abnormal returns on the days calculated, but it fails to use information about returns on other days and estimators based on these returns are therefore not efficient.

The Copenhagen Stock Exchange and other small exchanges generally list the last observed transaction price as a stock's price on nontrading days. Calculating daily returns from the recorded price series therefore gives zero returns for nontrading days and relatively large positive or negative returns on days when a stock trades. Maynes and Rumsey (1993) refer to returns generated in this manner as lumped returns because all of the multi-period return is allocated to the trading day, or the last day of a multi-period interval. The numerous zeroes in the return series lead to underestimates of the variance of returns and bias the test statistics used to judge abnormal performance. In spite of these problems, lumped returns are easy to calculate and it is perhaps the most frequently used method of adjusting for thin trading.

The uniform method is a third technique for handling thin trading. Instead of filling in unknown return days with zeroes, it calculates total return between trades and then allocates the average daily return to each day over the multi-period interval between trades. The same return is recorded for all of the nontrading days, as well as on the day a trade occurs. This technique is more efficient than using simple returns, but the test statistics are potentially biased—just as when calculating lumped returns.⁴ In addition, uniform returns give no particular emphasis to the actual trade day, meaning that some information is ignored when using this method. Maynes and Rumsey (1993) find that the uniform return method performs about the same as lumped returns.

Calculation of trade to trade returns represents a fourth technique for dealing with nontrading. The first step is to calculate an individual stock's return between the days when transactions actually take place. Then, trade-to-trade returns for the market index are calculated over the same period as for the stock. These two sets of trade-to-trade returns are used to

⁴If the difference between the “filled in” value and the underlying unobservable “true” value is white noise then both the lumped and uniform methods provide an unbiased estimate of returns. Also, the bias in the lumped return method may not be too large if volume and returns are positively correlated. If a lack of volume implies small price changes, a zero return on a nontrading day might be a reasonable estimate of the true unobserved return for that day. See Karpoff (1987) for a survey of the relationship between changes in price and volume.

estimate the market model to obtain abnormal returns for the stock over this period. Since empirical estimation of abnormal return by this method is more difficult than for lumped returns, the details are presented in Appendix I.

The trade to trade method uses all available information about total stock and market returns over time and no bias is introduced by attempting to estimate unobserved daily stock returns. Although it ignores information about daily market returns over nontrading periods, the small reduction in efficiency relative to the lumped sum or uniform methods is more than offset by the desirable property of unbiasedness. For our Danish data, we find that trade to trade returns are the best way to adjust for the problem of thin trading. This is consistent with Maynes and Rumsey (1993), who also find that the trade to trade method to out performs the lumped or uniform return techniques. For brevity, in the remainder of the paper, we only report results using trade to trade method of adjusting for thin trading.

3. DATA AND TEST STATISTICS

Price and volume data are from the Copenhagen Stock Exchange (CSE). Daily dividend and split adjusted data were extracted from Børsdatabasen, which was maintained by the Århus School of Business. Børsdatabasen starts in 1985, but data for the main index on the exchange, the KFX (Københavns Fondsbørs Indeksindex), is not available until 1990. Throughout this study the KFX returns series is treated as the market index.⁵ Therefore, the sample starts on the first trading day in 1990 and ends in mid-2001—the last available date for the database.

Please insert Table 1 about here.

3.1 Descriptive statistics

Danish stocks are sorted into one of three groups each year (thick, medium, or thin) based on trading frequency. A stock is defined to trade if volume is positive. A thick-traded stock is defined as one trading on more than 80% of trading days, or an average of more than four days per week. Stocks in the medium trading group show transactions on 40% to 80% of all trading

⁵Real time values for the KFX are provided by several data vendors. For example, it is listed as ^KFX on Yahoo-Finance. The KFX is a value-weighted index of the 20 largest Danish blue chip stocks selected from a list of the 25 most actively traded (liquid) stocks during a 12 month period ending in November of the previous year. Although the index is not intended to reflect all major industry groups, it is a good gauge of the Danish economy. Throughout the sample period the index has generally included firms in food production, banking, insurance, technology, wind energy, transportation, pharmaceutical, and the retail sectors. Like most indexes, it does not included dividends. Since it is composed of the most liquid stocks on CSE, it is unlikely to suffer from thin trading problems. Therefore, it is a better market index than an equal or value-weighted index of all stocks on the CSE.

days, or about two to four days per week. Thinly traded stocks trade on less than 40% of all trading days or on average less than two days per week.

Table 1 illustrates the growth of trading activity on the Copenhagen Stock Exchange over the decade of the 1990s. The total number of stocks listed on the exchange increases from 104 in 1990 to 241 in 2001. Particularly noteworthy is the large increase in the percentage of stocks trading each day. In 1990, about 60% of all Danish stocks were classified as thinly traded (70% in 1991), while by mid 2001 only about 26% of stocks were thinly traded. Similarly, the percentage of all stocks in the thickly traded group rose from about 15% in 1990 to 49% in 2001.

Even with the increase in trading activity on the CSE, thin trading remains an important issue in all years. Restricting a sample to medium and thick traded stocks, as often done in US event studies, may not be feasible for smaller stock exchanges. To obtain enough event dates and large enough samples for statistical inference, a Danish event study would likely need to include the thinly traded stocks which historically comprised 24% to 70% of all CSE stocks. The average number of days between trades for CSE thinly traded stocks has ranged from about 11 to 23 days, while the average across all three trading groups on the CSE has ranged from 4.17 days in 2001 up to 12.71 days in 1992. Note that the average Danish stock trades less frequently than Maynes and Rumsey's (1993) average thinly traded Canadian stock (trading frequency of 3.67 days). Therefore, our tests of the Maynes and Rumsey (1993) approach under conditions of more extreme thin trading conditions may be more representative of the situation in most emerging markets and on other small stock exchanges.

Please insert Table 2 about here.

Table 2 shows the descriptive statistics of stock returns for each of the three trading frequency groups using the trade to trade adjustment for thin trading. Some unusual features of the Danish data are that the average coefficient of skewness is higher for the thickly traded group than for the medium traded group and that thinly traded stocks display less excess kurtosis than their thickly traded counterparts. The data simply show that Danish stock returns, even among thickly traded stocks, deviate considerably from normality.

The average R-squared obtained from estimating the market model (equation 1b in Appendix I) over a 247 day period is quite low--ranging from 1% for the medium traded group to 7% for the thick-traded group. However, there is a relatively large variation over time in these numbers (not reported in the tables). For example, the average R-squared in 1990 for the thick group was around 15%, but it has fallen to 3% to 5% in the latter years of the sample. The

explanation probably lies in the change in the composition of the thick-traded group over time. In the early years of the sample the thick traded group consisted of a few large firms (e.g., 16 firms in 1990) that were generally members of the KFX--the market index. By the end of the sample period, the thick traded group consisted of 118 firms, including many smaller firms not part of the KFX. Since the thick traded group and the market have become less similar over time, the drop in average R-squared values over time seems understandable. Finally, the average Durbin-Watson statistics indicate that autocorrelation is unlikely to pose a significant problem in event studies.

3.2 Test statistics.

Parametric test statistics for abnormal performance on event days are based on a standard t test of the difference between two means. The numerator of the test statistic measures the absolute impact of some event relative to the return expected using some kind of market model. The denominator scales this number by some measure of estimated variance. As discussed in Brown and Warner (1985), parametric test statistics differ from one another primarily in the way they adjust for problems encountered in the data.

Three parametric test statistics frequently used in event studies are calculated for Danish stocks in the simulations in the next section. Details about all test statistics used in this paper, including formulas, are presented in Appendix II. The statistics examined are:

- T₁ T-test with adjusted cross sectional independence
(Brown and Warner, 1985 and Patell, 1976)
- T₂ T-test with standardised abnormal return
(Brown and Warner, 1985)
- T₃ T-test with adjusted standardised abnormal return
(Brown and Warner, 1985 and Patell, 1976)

For the T₁ test statistic, the abnormal return on the event day is assumed to be independent across stocks. (An alternative test statistic can be derived if one believes there is cross sectional dependence in the data.) In its unadjusted form (shown in Appendix II), the variance of the test statistic T₁ is the sum of the variances of abnormal returns of the individual stocks. However, to improve upon the performance of the test statistic, Patell (1976) recommends estimating abnormal returns based on forecasts from the market model, and then

calculating the variance of the test statistic as the sum of the variance of the forecast errors for the individual stocks.⁶

For T_2 , abnormal returns for each stock are scaled by their individual standard deviations and added together to produce the test statistic. It is a t-statistic based on standardised abnormal returns. An adjusted version of this test statistic, T_3 , is calculated using the standard deviation of the forecast errors (rather than actual standard deviations) to scale the abnormal returns.

The second set of tests is nonparametric and not based upon the assumption of normality. Given that the Danish returns data in Table 2 show considerable deviations from normality, these statistics should be more reliable than the parametric measures of abnormal performance. The nonparametric statistics are:

- T_4 Rank test (Corrado, 1989)
- T_5 Sign test (Corrado and Zivney, 1992)
- T_6 Generalised sign test
(Cowan, 1992 and Cowan and Sergeant, 1996)

The rank test, T_4 , is from Corrado (1989) and Corrado and Zivney (1992). The sign test, T_5 , is from Corrado and Zivney (1992), while the generalised sign test, T_6 , is based upon Cowan (1992) and Cowan and Sergeant (1996). The sign test, T_5 , assumes that the probability of observing either a negative or positive abnormal return is 0.5, whereas for T_6 this probability is estimated from actual returns over the estimation period.

A final set of tests is used to detect abnormal performance when there is a change in variance around the event day. Boehmer, Musumeci and Poulsen (1991) developed a parametric test and a simple adjustment to the rank test, T_4 provides a non-parametric alternative.

- T_7 Parametric test with variance adjustment
(Boehmer, Musumeci and Poulsen, 1991)
- T_8 Rank test with variance adjustment
(Corrado and Zivney, 1992 and Maynes and Rumsey, 1993)

For the previous statistics ($T_1 - T_6$) the variance or the variance of the forecast error was estimated over the estimation period-- that is, prior to the event window. In contrast T_7 and T_8 are adjusted for changes in variances over the event window.

⁶ Test statistics for cross sectional dependence and cross sectional independence, as presented in Brown and Warner (1980, 1985) were also calculated. These statistics are related to T_1 , but are known have lower power. Our results confirmed that they did not perform as well as T_1 . Further description of test statistics and results from the lumped return adjustment are available from the authors upon request.

Since the exact event date often is not known, a window is assigned to the event and Cumulative Average Returns (CAR) are calculated and tested over the window. The most common window is plus/minus one day around the expected event day. For the simulations in the next section, an event window of three days is used. The event day is randomly assigned (using a uniform distribution) to one of these days and the three day CAR measure

$$CAR_t = \overline{A_{-1}} + \overline{A_0} + \overline{A_{+1}},$$

is applied to all test statistics to account for event day uncertainty.

3.3 Simulation

The performance of the various test statistics is examined using the simulation approach of Brown and Warner (1980, 1985). A total of 1000 portfolios of 10, 25, and 50 stocks are generated by randomly selecting an event day, and then randomly selecting stocks for inclusion in the portfolio. To be included in the portfolio, the stock must trade every day in the event window and have a minimum of 10 observed prices in the estimation period. The length of the estimation period is 247 days, or approximately a year of trading prior to the three day event window. Specifically, the estimation period is from $t = -249$ to -2 , the event window is from $t = -1$ to $t = +1$, and the actual event is at $t = 0$. To simulate the impact of an event 0, 0.5% and 2% are added to the abnormal return on the event day.

The condition that the stock has to trade every day in the event window replicates the selection criterion typically used in event studies where one wants to ensure that actual prices are used to measure the effect of an event. However, imposing this condition can cause problems in real-world event studies. Stocks affected by an event have a higher than usual probability of trading during the event window, but stocks not influenced by an event are no more likely to trade than on any other day. Stocks for which an event has no effect will be dropped from the sample, thereby biasing the study in favour of finding a significant effect. In terms of a simulation study, a positive correlation of volume and return means that stocks satisfying the trading condition are more likely to have abnormal returns in the event window thereby biasing the test statistics in favour of finding significant an effects.

4. RESULTS

4.1 Properties of test statistics under the null hypothesis

Please insert Table 3 about here.

Under the null hypothesis of zero abnormal return, all test statistics presented in Section 3 are assumed to have a standard normal distribution. Properties of the test statistics, under the null hypothesis of zero abnormal performance, are presented in Table 3. The empirical distributions are based on 1000 values of the test statistics for portfolios of 25 stocks.

Under the null hypothesis, the mean of each test statistic should be zero. Instead, all test statistics at all trading frequencies have a positive mean. This bias will lead to over rejection of the null hypothesis—meaning that we may detect event significance in some instances when there is none. The likely cause of this problem was discussed in the previous section.

All test statistics considered (T_1 to T_8) have lower means for the thickly traded group than for the medium and thinly traded stock groups. As a consequence, researchers are more likely to incorrectly identify significant effects in samples that contain a large proportion of medium and thinly traded stocks. It is therefore important to report results for stocks across different trading frequencies and confirm that the findings are consistent among trading groups. For example, the finding of a significant effect for thinly traded stocks and no effect for thickly traded stocks may suggest problems arising from sampling technique. Although no single test statistic is superior across all trading frequencies, the non-parametric test statistics generally out perform the parametric test statistics.

The standard deviation of each of the test statistics should be one under the null hypothesis. A standard deviation of less than one means the null hypothesis will be rejected too often, whereas a standard deviation greater than one leads to less frequent rejections (failing to find an effect when there is one). For the thickly traded group, except for T_2 and T_3 , the standard deviation for all the test statistics differs from unity by less than 10%. For the medium and thinly traded groups, the standard deviations of the test statistics are generally closer to one for the nonparametric group of statistics. This is consistent with Campbell and Wasley (1993) who find that the rank statistic is better specified than parametric statistics under the null hypothesis. The statistic, T_8 , that accounts for event induced volatility has a standard deviation near one and performs well in our simulations.

Skewness and excess kurtosis of the test statistics should both be zero under the null hypothesis. With nonzero skewness, the rejection frequencies for the null hypothesis differ for positive and negative events. If excess kurtosis is positive, then the tails are too thick leading to incorrect (generally too high) rejection frequencies under the null hypothesis. For the thinly traded group, both skewness and kurtosis are a problem for parametric test statistics; so the researcher should generally rely upon nonparametric test statistics to identify the significance of events on these stocks. This is also the case for the thickly and medium traded stocks, but to a lesser extent.

In general all test statistics are biased since the mean is greater than one and the bias increases with lower trading frequencies. Thus, one is more likely to reject the null hypothesis when it is true for thinly traded stocks than for frequently traded stocks. The non-parametric tests generally appear to be better specified than the parametric tests, but no single tests statistic is superior across all trading frequencies and specification criteria.

4.2 Performance of the test statistics: size and power

Please insert Table 4 about here.

To determine the performance of the different test statistics in terms of size and power, portfolios of stocks with artificial abnormal returns of 0%, 0.5%, and 2% in the event window were randomly generated. The results obtained from these simulations for portfolios of 50 stocks are presented in Table 4.

For the thickly traded group, the rejection rate for a 5% significance level is generally around 5% for all test statistics when no abnormal return is added. All of the test statistics perform reasonably well in terms of size. In terms of power, i.e., detecting an effect when extra return has been added on the event day, the four non-parametric tests have the best performance. For example, they reject the null hypothesis between 78% and 88% of the time when 0.5% has been added on the event day. This is significantly better than the parametric tests for which rejection of the null hypothesis ranges from a low of 41% for T_1 to a high of about 71% for T_2 and T_7 . These results are similar to those obtained by Corrado (1989) for US data. However, unlike Maynes and Rumsey's (1993) results with Canadian data, the rank tests do better than the other parametric or nonparametric test statistics.

For the medium trading group, the non-parametric test statistics perform better, in terms of both size and power than the group of parametric test statistics. For example, when 0.5% is added to the return, the non-parametric test statistics reject the null at least 90% of the time while the parametric statistics reject the null, at most, 71% of the time.⁷ The performance of the parametric tests is not as good as in Cowan and Sergeant (1996) for US data; but the non-parametric tests perform similarly for both US and Danish data.

⁷ Although results for the lumped return adjustment for thin trading are not reported in the paper for sake of brevity, lumped returns actually provide slightly better results in terms of power and size for the medium traded group. Trade to trade returns do better for both thick and thin traded stocks. Nevertheless, since the lumped return adjustment is relatively easy to implement, researchers facing time constraints might consider it.

For the thin trading group, the parametric test statistics perform very badly in terms of size when no abnormal return is added. Rejection rates range from 2.9% for T_1 up to 23% for T_2 when the null hypothesis is true. Power varies considerably across test statistics when a 0.5% abnormal return is added with rejection rates ranging from 10% up to 73%. The situation is slightly better for nonparametric test statistics. The rejection rate when a zero abnormal return is added is between 6.6% and 8.4%. As was the case for the medium trading group, the non-parametric tests are noticeably more powerful than the parametric tests. Rejection rates range from about 77% to 81%. Finally, we note that for an abnormal return of 2% on the event day, all non-parametric tests have a rejection rate of 100%. This is reassuring for detecting abnormal performance in thinly traded stocks. To summarize, the results in Table 4 provide strong support for relying primarily on nonparametric statistics when attempting to detect abnormal performance in the presence of thin trading.

Please insert Table 5 about here.

Please insert Table 6 about here.

Tables 5 and 6 present the same information as Table 4, but for portfolios of 25 and 10 stocks, respectively. With abnormal returns of 0% or 2.0% induced on event day, the test statistics look similar to those in Table 4 in terms of both size and power. However, the power of all of the tests drops dramatically for the case of 0.5% induced abnormal returns. Power of the test was above 80% for all nonparametric test statistics for 50 stock portfolios, but power drops to between 50% and 80% for 25 stock portfolios. Also, notice that the reduced power is observed for all three trading groups—think, medium, and thinly traded stocks. A similar result is observed for 10 stock portfolios, except that the power drops to between 20% and 50%. Again, this is observed for all three groups of stocks. Similar results regarding the dramatic reduction in power of test statistics have been observed in event study simulations with US data.

In general, the power and size of test statistics in Tables 4-6 are not significantly lower than found in US studies--suggesting that researchers may be able to perform event studies on small stock exchanges using daily data. Even in the presence of very thin trading, it seems possible to detect abnormal performance. Some recommendations for other researchers include:

- Try to identify at least 25 occurrences of an event, but preferably find at least 50 events.

This is based on the idea that portfolios of 50 stocks provide good size and power for test

statistics, while portfolios of 25 stocks only provide acceptable size and power in some instances.

- When the number of events is 25 or less, abnormal returns of 2% or greater may be needed for the test statistics to reliably identify abnormal performance. If the number of events is small, researchers can only identify events having major impacts on stock prices.
- Since test statistics for thinly traded stocks are biased under the null hypothesis, the researcher should separate the sample into groups or categories based on trading frequency, and report any differences in results across the frequency groups.
- The trade to trade method of adjusting for thin trading, as done in this study, is preferable to other thin trading adjustment methods, but lumped returns perform nearly as well and could be used if an event study must be done quickly.
- Emphasize nonparametric, rather than parametric test statistics for judging significance. However, since no individual test is superior to the others for all portfolio sizes and trading frequencies, researchers may wish to calculate a battery of test statistics, as done in this study.

5. SOME EXTENSIONS

5.1 Increase in variance around the event day

Both the return and variance of a stock can change as a result of the release of new information on the financial markets. Theories in behavioural finance argue that investors need time to process and price new information, leading to increased returns volatility during this period. Alternatively, new information may lead to an increase in systematic risk which also increases volatility around the event day. For example, different stocks do not respond in the same way to an event such as a positive or negative earnings announcement because the response depends upon the difference between expected and actual earnings. The effect of this difference is twofold: (1) since an additional random amount is added to the abnormal, the variance of the abnormal return for an individual stock increases on the event day; and (2) because the effect is different across companies, the cross-sectional variance is increased around the event day compared to the estimation period. Without adjusting for event induced volatility, the variance estimated over the estimation period is likely to understate the variance in the event window causing the null hypothesis (of no effect) to be rejected too often. Although examination of the causes of increased variance around an event day is itself an interesting topic, the focus of this paper is on the effect of increased variance, regardless of origin, on the various test statistics typically used in event studies.

The procedures from the previous section can be modified to account for event induced volatility. Following Boehmer et. al. (1991), the variance on the event day is assumed to be 50% larger than over the estimation period. In Table 7, abnormal returns of 0%, .5%, and 2% have been added to event day returns for portfolios of 25 stocks.

Please insert Table 7 about here.

Rejection rates can be compared with those in Table 5 for portfolios of 25 stocks without event induced variance. With event induced volatility, both parametric tests and nonparametric rank tests lead to high false rejection rates of the null when there is no abnormal performance (0% abnormal return added). For the thick trading group, the lowest rejection rates using the rank tests are about 12%, versus 5% in Table 5. The two sign tests (T_5 and T_6) have good performance in terms of size. As expected, the T_7 test designed specifically to handle the problem of event induced variance, also performs quite well. For the medium and thinly traded stock groups, the performance of the parametric tests (and to a lesser extent the rank tests) are worse than for thickly traded stocks. For example T_2 , the standardised abnormal return test, has a rejection rate of 37% for thinly traded stocks when there is no abnormal return. Boehmer et. al. (1991) did not consider thinly traded securities, but our results for the thickly traded group are in line with their findings. From the evidence above, large increases in event induced variance appear to lead to a misspecification of parametric test statistics and nonparametric rank statistics to such an extent that they are not reliable in event studies.

Examination of rejection rates for 0% additional abnormal return for the sign tests (T_5 and T_6) and the parametric test designed to deal with increased variance around the event day (T_7) suggests that all three tests are reasonably well specified in terms of size. For an induced abnormal performance of 0.5%, the power of the sign tests are 16% and 20% for the Thickly Traded group and even smaller for the Medium and Thinly Traded groups. For T_7 the power is only marginally better than for the sign tests. For 0.5% abnormal returns, the power of all of the tests is such that one is unlikely to find any effect, even if one exists.

For an induced abnormal performance of 2% in the Thickly Traded group, power is above 90% for the for the sign tests (T_5 and T_6) and above 97% for T_7 and T_8 . The power for all four test statistics drops for the Medium and Thinly Traded groups. The performance of T_7 , in particular, deteriorates for the less frequently traded stocks. However, the power of T_8 does not drop below 89% for any of the trading groups. Overall, T_7 , the variance adjusted standardised abnormal returns test proposed by Boehmer et al (1991), appears to be the best test when there is

an increase in variance around the event day. The sign tests, T_5 and T_6 , also have acceptable size and power and can be used to verify the results obtained using T_7 .

5.2 Unknown event day.

To determine performance of the test statistics for the situation where the event day is not known with certainty, the simulations were repeated for an event window of three days where the event day was randomly assigned (using a uniform distribution) to one of the three days. The most obvious effect of an unknown event day is that 33% of the stocks are expected to experience an event on a given day, so the average abnormal return on the event days is diluted⁸. The question is whether the CAR test can compensate for this dilution by “summing the returns” over the three days. The results from the simulations made for portfolios of 25 stocks are reported in Table 8. The top panel of Table 8 shows the average daily results, while the bottom panel presents rejection rates based upon cumulative average abnormal returns.

Please insert Table 8 about here.

Relative to a known event day, there is a marked drop in rejection rates across all trading frequency groups for .5% and 2% induced additional abnormal returns. For example, for the medium traded group, rejection rates for the best test statistics drop from around 66% in Table 5 to 17% in the top panel of Table 8 for an induced additional abnormal return of 0.5%. For a 2% artificial abnormal return, the drop in rejection rates across all trading frequencies is from around 99% in Table 5 down to 66% (at best). The two standardised abnormal return test statistics, T_2 and T_3 , perform the best overall in terms of power of the test. This is in contrast to the better performance of the rank statistics (T_4 and T_8) in Table 5 and the superior performance of the sign statistics (T_5 and T_6) and the T_7 statistic in Table 7. Thus, the introduction of a random event day not only reduces power across test statistics, it may also influence the researcher’s choice of optimal test statistics.

Examination of the CARs in the bottom panel of Table 8 indicates that the power of tests increases substantially relative to the same statistics in the top panel of the table. Rejection rates for detection of a 2% induced abnormal return increase to about 95% for the T_2 and T_3 statistics

⁸ A normal distribution could be used to assign event days. It assigns more events to the original event day and leads to less dilution of abnormal returns. We chose the uniform distribution method to ensure a more stringent test of the impact of an unknown event day.

across all three trading groups. These parametric test statistics appear to out perform the nonparametric tests in this situation, although the nonparametric test statistics are still preferred on the basis of size. That is, the size of the test statistics for T_2 and T_3 is above 7%, while the size of the rank statistics is below 6%. The CAR method has good power and is able to detect abnormal returns above 2% when the event day is uncertain. However, all test statistics are rather small for induced abnormal returns of 0.5%. The rejection rates are less than 50% across all trading groups. Nevertheless, by using the rank and standardised tests together, CAR analysis seems to have reasonable power and size to identify abnormal returns above 2% for random event days.

5. CONCLUSION

This paper has analysed the efficiency of event study methods in the presence of thin trading using data from the Copenhagen Stock Exchange. The main conclusion is that event studies using daily data can be meaningfully performed on a small exchange provided that certain adjustments are made to account for the problems caused by very thin trading. First, researchers should adopt less restrictive selection criteria than used in US event studies in order to have enough observations or events to reliably detect abnormal performance. For the Danish market, a minimum of 25 events appears necessary to obtain acceptable size and power in statistical tests. If possible, 50 or more events provide even better size and power in the test statistics. Second, an adjustment for thin trading is necessary to obtain daily returns. The trade to trade adjustment method is preferred, but a lumped return adjustment would perform almost as well if event study results were needed quickly. Third, researchers should not expect to be able to consistently detect abnormal performance of less than 1% (or perhaps even less than 2%), unless the sample contains primarily thickly traded stocks. Fourth, due to the nonnormality of Danish stock returns, nonparametric tests are generally preferable to parametric tests of abnormal performance—except in the cases of event induced volatility and unknown event day. Fifth, researchers should present separate results for thickly and thinly traded stock groups. If there are noticeable differences between groups, further research can be undertaken to determine the cause of the inconsistencies. Finally, when nonnormality, event induced variance, unknown event day, and problems of very thin trading are all considered simultaneously, no one test statistic or type of test statistic dominates the others. A researcher can use a variety of parametric and nonparametric tests to detect abnormal performance. If all tests agree, the researcher can be fairly confident of results. When there are differences between tests, researchers can try to determine the source of problems and perhaps refer to results in this paper regarding which tests are likely to be more reliable in the presence of various problems with the data.

REFERENCES

- Ball, R. and Brown, P. (1968) An empirical evaluation of accounting numbers, *Journal of Accounting Research*, **6**(2), 159-178.
- Bartholdy, J. and Riding, A. (1994) Thin trading and the estimation of betas: The efficacy of alternative techniques, *Journal of Financial Research*, **17**(2), 241-254.
- Boehmer, E., Musumeci, J. and Poulsen, A. (1991) Event-study methodology under conditions of event-induced variance, *Journal of Financial Economics*, **30**(2), 253-272.
- Brown, S. J. and Warner, J.B. (1980) Measuring security price performance, *Journal of Financial Economics*, **8**(3), 205-258.
- Brown, S. J. and Warner, J.B. (1985) Using daily stock returns: The case of event studies of event-induced variance, *Journal of Financial Economics*, **14**(1), 3-31.
- Campbell, C. J. and Wasley, C. E. (1993) Measuring security price performance using daily NASDAQ returns, *Journal of Financial Economics*, **33**(1), 73-92.
- Corrado, C. J. (1989) A nonparametric test for abnormal security-price performance in event studies, *Journal of Financial Economics*, **23**(2), 385-395.
- Corrado, C.J. and Zivney, T.L. (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. (1992) Nonparametric event study tests, *Review of Quantitative Finance and Accounting*, **2**(4), 343-358.
- Cowan, A. and Sergeant, A. (1996) Trading frequency and event study test specification, *Journal of Banking & Finance*, **20**(10), 1731-1757
- Dimson, E. (1979) Risk measurement when shares are subject to infrequent trading, *Journal of Financial Economics*, **6**(2), 197-226.
- Fama, E., Fisher, L., Jensen, M. and Roll, R. (1969) The adjustment of stock prices to new information, *International Economic Review*, **10**(1), 1-21.
- Heinkel, R. and Kraus, A. (1988) Measuring event impacts in thinly traded stocks, *Journal of Financial and Quantitative Analysis*, **23**(1), 71-88.
- MacKinlay, G. (1997) Event studies in economics and finance, *Journal of Economic Literature*, **35**(1), 13-39.
- Karpoff, J. M. (1987) The relation between price changes and trading volume: A survey, *Journal of Financial and Quantitative Analysis*, **22**(1), 109-126.
- Lyon, J.D, Barber, B.M. and C. Tsai, C. (1999) Improved methods for tests of long-run abnormal stock returns, *Journal of Finance*, **54**(1), 165-201.

- Maynes, E. and Rumsey, J. (1993) Conducting event studies with thinly traded stocks, *Journal of Banking and Finance*, **17**(1), 145-157.
- Patell, J. M. (1976) Corporate forecasts of earnings per share and stock price behavior: Empirical tests, *Journal of Accounting Research*, **14**(2), 246-276.
- Salinger, M. (1992) Standard Errors in Event Studies, *Journal of Financial and Quantitative Analysis*, **27**(1), 39-53.
- Scholes, M. and Williams, J. (1977) Estimating betas from nonsynchronous data, *Journal of Financial Economics*, **5**(3), 309-327.

APPENDIX I: ESTIMATION OF TRADE TO TRADE ABNORMAL RETURNS

Following Maynes and Rumsey (1993), for a given stock, the observable trade-to-trade return between time $t-n$ and time t (where there are no trades on the $n-1$ days in between) is given by:

$$R_t = \ln \left[\frac{P_t}{P_{t-n}} \right].$$

However, just because a stock does not trade does not mean it does not have a value. Letting \hat{P}_s denote the unobserved (closing) price for the stock on day s where no trade takes place the trade-to-trade return can be written as:

$$R_t = \ln \left[\frac{P_t}{\hat{P}_{t-1}} \times \frac{\hat{P}_{t-1}}{\hat{P}_{t-2}} \times \dots \times \frac{\hat{P}_{t-(n-1)}}{P_{t-n}} \right] = r_t + r_{t-1} + \dots + r_{t-(n-1)}$$

where r_s is the unobserved return for day s . Thus, the trade-to-trade return is the sum of the n unobserved daily returns which occur because there are $n-1$ days with no trades⁹. If the unobserved daily returns are generated by the market model (where the parameters are constant over the n days with no observed daily returns) then we have:

$$r_s = \alpha + \beta r_{ms} + \varepsilon_s \quad s = t - (n - 1), \dots, t.$$

where r_{ms} is the observed return on the market for day s . Thus:

$$R_t = \sum_{s=t-(n-1)}^t r_s = \sum_{s=t-(n-1)}^t (\alpha + \beta r_{ms} + \varepsilon_s) = n\alpha + \beta \sum_{s=t-(n-1)}^t r_{ms} + \sum_{s=t-(n-1)}^t \varepsilon_s.$$

The trade to trade return for the market between day $t-n$ and day t is equal to the sum of the n (in this case observed) daily returns. That is,

$$R_{mt} = \sum_{s=t-(n-1)}^t r_{ms}$$

so we have

$$R_t = n\alpha + \beta R_{mt} + \sum_{s=t-(n-1)}^t \varepsilon_s \quad (1a)$$

By assumption, the market model error terms ε_s are independently and identically distributed.

The error term in (1a), $\sum_{s=t-(n-1)}^t \varepsilon_s$, therefore depends on the number of terms in the sum, namely

⁹If there are no trades on the $n-1$ days between day $t-n$ and day t , then the n daily returns for day's $t-(n-1)$ to day t are unobserved.

n . Aggregation of error terms introduces heteroscedasticity into the model, so for estimation purposes (1a) is divided by \sqrt{n} to remove the induced heteroscedasticity giving:

$$\frac{1}{\sqrt{n}}R_t = \sqrt{n}\alpha + \beta \frac{1}{\sqrt{n}}R_{mt} + \frac{1}{\sqrt{n}} \sum_{s=t-(n-1)}^t \varepsilon_s \quad (1b)$$

From (1a) expected return is given by:

$$E[R_t] = n\hat{\alpha} + \hat{\beta}R_{mt} \quad (2)$$

where $\hat{\alpha}$ and $\hat{\beta}$ are obtained from estimation of (1b) using OLS. Therefore abnormal return is given by:

$$A_t = R_t - E[R_t] = R_t - n\hat{\alpha} + \hat{\beta}R_{mt}. \quad (3a)$$

Dividing by \sqrt{n} again removes the introduced heteroscedasticity and solves the problem for estimation purposes. That is, the following equation is used for estimation:

$$A'_t = \frac{1}{\sqrt{n}}A_t = \frac{1}{\sqrt{n}}R_t - \sqrt{n}\hat{\alpha} + \frac{1}{\sqrt{n}}\hat{\beta}R_{mt}. \quad (3b)$$

Table 2. Descriptive statistics

Our analysis uses an estimation plus event period of 247 days and descriptive statistics for each stock are calculated over twelve 247 day intervals. Within each interval or “year”, a stock is assigned to a trading frequency group. Since stocks move from one trading frequency to another over time, one cannot simply calculate the statistics for each stock over the whole period. The reported statistics are **averages** of the individual statistics for each stock within each group over the 12 intervals. Thin trading is accounted for using the trade to trade adjustment procedure. Beta, R-squared and the Durbin Watson statistics are from equation (1b):

$$\frac{1}{\sqrt{n}} R_t = \sqrt{n} \alpha + \beta \frac{1}{\sqrt{n}} R_{mt} + \frac{1}{\sqrt{n}} \sum_{s=t-(n-1)}^t \varepsilon_s .$$

	Trading frequency		
	Thick	Medium	Thin
Daily return (%)	0.012	0.041	0.021
Standard deviation (%)	2.26	2.96	6.17
Coefficient of Skewness	0.17	0.01	-0.31
Coefficient of Kurtosis	9.66	9.59	5.12
Beta	0.38	0.18	0.19
Standard error of the residuals (%)	2.15	2.44	3.25
R-squared	0.07	0.01	0.03
Durbin Watson statistic	2.02	2.23	2.17

Table 1.**Trading statistics for different trading groups**

Stocks in the sample are divided into trading groups based on their trading frequency. Stocks in the thick trading group trade more than 80% of the time i.e. on average, more than 4 days a week. Stocks in the medium trading group trade between 40% and 80% of the time i.e., on average, between 2 and 4 days a week, and stocks in the thin trading group trade than 40% of the time, i.e. on average, on less than 2 days a week.

Year	Trading Frequency												Total number of stocks
	Thick				Medium				Thin				
	Number of stocks	% of stocks	Average number of days between trades	Average trading Frequency (%)	Number of stocks	% of stocks	Average number of days between trades	Average trading frequency (%)	Number of stocks	% of stocks	Average number of days between trades	Average trading frequency (%)	
1990	16	15,38	1.08	93.26	26	25	1.90	54.73	62	59,62	11.38	17.63	104
1991	14	12,73	1.10	91.80	19	17,27	1.84	56.42	77	70	17.24	18.23	110
1992	19	15,57	1.07	93.56	28	22,95	1.83	56.81	75	61,48	18.84	17.67	122
1993	27	20,93	1.07	94.22	31	24,03	1.72	60.10	71	55,04	15.56	18.93	129
1994	32	22,7	1.06	94.35	46	32,62	1.78	58.51	63	44,68	22.92	18.60	141
1995	37	23,57	1.05	95.39	50	31,85	1.69	61.70	70	44,59	19.68	18.96	157
1996	57	33,14	1.07	94.04	63	36,63	1.76	59.44	52	30,23	12.94	18.58	172
1997	85	44,04	1.06	94.98	60	31,09	1.73	60.38	48	24,87	18.79	18.16	193
1998	83	36,56	1.06	94.45	68	29,96	1.71	61.12	76	33,48	18.76	16.47	227
1999	96	38,55	1.06	94.89	73	29,32	1.64	63.30	80	32,13	20.29	17.87	249
2000	149	54,58	1.05	95.34	58	21,25	1.67	62.28	66	24,18	17.41	17.61	273
2001	118	48,96	1.06	95.02	59	24,48	1.64	63.42	64	26,56	13.36	16.79	241

Note: The number of listed companies in Table 2 is the average number of listed stocks over the year. The “number of stocks” here is the total number of stocks over the year in each group. Thus the sum of the “number of stocks” here correspond s to the maximum number of stocks listed in a given year and is therefore larger than the average number provided in Table 2.

Table 3. Properties of test statistics under the null hypothesis.

Under the null hypothesis of zero abnormal return, the test statistics are assumed to have a standard normal distribution. Empirical distributions are based on 1000 values of the test statistics for portfolios of 25 stocks. The two best performing statistics have been shaded.

Test statistic	Thick trading				Medium trading				Thin trading			
	Mean	Std. Dev	Skew-ness	Kurto-sis	Mean	Std. Dev	Skew-ness	Kurto-sis	Mean	Std. Dev	Skew-ness	Kurto-sis
Parametric tests												
T ₁ – Adj. cross sec. ind.	0.1056	1.0753	0.1188	4.9185	0.2737	0.8511	0.3162	0.9349	0.2997	1.0714	6.9882	127.184
T ₂ – Stand. Ab. Ret.	0.1152	1.4717	2.1051	33.1813	0.3596	1.0393	0.0077	0.2331	0.7698	1.9001	4.0984	41.7675
T ₃ - Adjusted s.a.r.	0.1149	1.4630	2.1201	33.3374	0.3583	1.0283	-0.0069	0.2387	0.7411	1.8467	4.2920	44.9004
Non-parametric tests												
T ₄ - Rank test	0.0931	1.0472	-0.0711	-0.1869	0.3040	0.9833	-0.0937	-0.1336	0.4843	1.0233	-0.0305	0.1046
T ₅ - Sign Test	0.0698	1.0357	0.0036	-0.1878	0.2876	0.9660	-0.1505	-0.0574	0.4873	1.1015	-0.1410	0.1203
T ₆ - Generalised sign test.	0.0653	1.0075	0.0188	-0.2790	0.2821	0.9295	-0.1036	-0.0918	0.5266	1.0388	-0.1452	0.1906
Event induced variance adjusted tests												
T ₇ - Parametric	0.0817	1.056	-0.005	-0.1617	0.3593	1.0173	-0.1385	-0.1650	0.5885	1.0185	-0.2381	0.1061
T ₈ - Rank test	0.0950	1.0437	-0.0703	-0.214	0.3067	0.9797	-0.0938	-0.1279	0.4847	1.0074	-0.0263	0.1181

Table 4. Average rejection rates for portfolios of 50 stocks

Portfolios of 50 stocks were randomly generated and artificial returns of 0, 0.5% and 2% were added during in the event window. The reported numbers are the rejection rates for a 5% significance level. The two best performing test statistics have been shaded.

Test statistics		Thick trading			Medium trading			Thin trading		
		Level of artificial induced return								
		0	0.005	0.02	0	0.005	0.02	0	0.005	0.02
Parametric tests	T ₁ - Adjusted cross sec. ind.	5.6	41.3	99.8	4.6	23.4	99.8	2.9	10.3	77.7
	T ₂ – Stand. abnormal return	7.6	71.4	99.9	10.2	71.4	100.0	22.7	72.5	99.9
	T ₃ - Adjusted s.a.r.	7.1	70.9	99.9	10.0	70.9	100.0	21.6	70.5	99.8
Non-parametric tests	T ₄ - Rank test	5.4	88.0	100.0	7.4	90.1	100.0	8.2	81.4	100.0
	T ₅ - Sign test	3.7	78.8	100.0	7.4	89.0	100.0	6.6	76.8	100.0
	T ₆ - Generalised sign test.	3.7	83.3	100.0	6.6	90.5	100.0	7.6	79.9	100.0
Event induced variance adjusted tests	T ₇ - Parametric test	5.0	71.5	99.5	9.1	71.0	99.7	12.0	60.1	96.3
	T ₈ - Rank test	5.1	87.9	100.0	7.2	90.2	100.0	9.6	81.2	100.0

Table 5. Average rejection rates for portfolios of 25 stocks

Portfolios of 50 stocks were randomly generated and artificial returns of 0, 0.5% and 2% were added during the event window. The reported numbers are the rejection rates for a 5% significance level. The two best performing test statistics have been shaded.

Test statistics		Thick trading			Medium trading			Thin trading		
		Level of artificial induced return								
		0	0.005	0.02	0	0.005	0.02	0	0.005	0.02
Parametric tests	T ₁ - Adjusted cross sec. ind.	5.5	25.6	99.4	3.9	16.3	93.1	3.3	9.8	57.6
	T ₂ - Stand. abnormal return	7.7	47.6	99.9	6.8	46.0	99.8	18.1	51.7	99.6
	T ₃ - Adjusted s.a.r.	7.5	47.1	99.9	6.7	45.5	99.8	15.8	48.7	99.5
Non-parametric tests	T ₄ - Rank test	6.3	60.9	100.0	5.1	66.1	100.0	8.6	57.9	100.0
	T ₅ - Sign Test	4.5	51.1	99.9	4.0	60.0	99.9	6.6	51.7	99.1
	T ₆ - Generalised sign test.	5.4	53.7	99.9	4.1	65.6	99.9	8.1	54.1	98.9
Event induced variance adjusted tests	T ₇ - Parametric test	5.8	49.3	99.2	6.8	49.2	99.3	9.2	43.3	95.7
	T ₈ - Rank test	6.2	60.8	100.0	5.0	65.8	100.0	8.5	57.7	100.0

Table 6. Average rejection rates for portfolios of 10 stocks

Portfolios of 10 stocks were randomly generated. Artificial returns of 0, 0.5% and 2% were added on the event day. The reported numbers are the rejection rates for a 5% significance level.

Test statistics		Thick trading			Medium trading			Thin trading		
		Level of artificial induced return								
		0	0.005	0.02	0.000	0.005	0.020	0.000	0.005	0.020
Parametric tests	T ₁ - Adjusted cross sec. ind.	4.7	13.4	89.6	5.2	10.3	67.9	5.2	8.8	43.5
	T ₂ - Stand. abnormal return	5.9	22.3	98.0	7.2	21.9	95.8	12.2	27.1	90.8
	T ₃ - Adjusted s.a.r.	5.5	22.1	98.0	7.2	21.4	95.7	11.2	24.7	89.8
Non-parametric tests	T ₄ - Rank test	5.3	28.7	99.1	4.2	32.0	97.3	5.2	26.6	92.4
	T ₅ - Sign test	3.4	19.6	85.9	2.3	22.8	79.2	11.7	35.1	87.2
	T ₆ - Generalised sign test.	4.8	26.3	90.8	3.3	31.0	84.7	4.3	23.0	75.5
Event induced variance adjusted tests	T ₇ - Parametric test	7.0	31.1	95.6	7.4	29.5	90.9	7.0	24.3	81.8
	T ₈ - Rank test	5.2	28.6	99.1	4.1	31.7	97.2	4.9	26.0	92.2

Table 7. Average rejection rates for portfolios of 25 stocks with increased variance on the event day

Portfolios of 25 stocks were randomly generated. Abnormal performance added to the event day is a random variable drawn from $N(x, 1.5\sigma^2)$ where x is 0, 0.5% and 2%. σ^2 is the variance of the abnormal return for the security over the event period.

		Thick trading			Medium trading			Thin trading		
		Level of artificial induced return								
		0	0.005	0.02	0	0.005	0.02	0	0.005	0.02
Parametric tests	T ₁ - Adjusted cross sec. ind.	21.4	37.7	96.1	24.5	33.1	82.6	29.4	35.3	60.3
	T ₂ - Stand. abnormal return	23.9	49.3	99.9	29.5	47.5	99.0	37.0	58.4	96.9
	T ₃ - Adjusted s.a.r.	23.7	49.2	99.9	28.9	47.0	99.0	35.9	56.3	96.6
Non-parametric tests	T ₄ - Rank test	12.4	35.4	99.4	15.3	31.4	93.2	16.3	30.8	89.8
	T ₅ - Sign test	4.5	15.7	90.6	4.5	12.9	75.3	5.6	14.1	65.3
	T ₆ - Generalised sign test.	5.5	19.6	92.3	5.8	19.0	81.4	7.0	15.9	69.4
Event induced variance adjusted tests	T ₇ - Parametric	5.8	24.1	97.3	5.9	20.4	89.3	8.9	20.1	80.4
	T ₈ - Rank test	12.4	35.4	99.4	15.2	31.1	93.2	16.3	30.8	89.2

Table 8. Average rejection rates for portfolios of 25 stocks with random event day – CAR analysis.

Portfolios of 25 stocks were randomly generated with a random event day drawn over the interval [-1, 1] assuming a uniform distribution. Artificial returns of 0, 0.5% and 2% were added on the event day. The reported numbers are the rejection rates for a 5% significance level.

Test statistics		Thick trading			Medium trading			Thin trading		
		Level of artificial induced return								
		0	0.005	0.02	0	0.005	0.02	0	0.005	0.02
Event study assuming that the event day is t = 0										
Parametric tests	T ₁ - Adjusted cross sec. ind.	5.5	8.0	43.4	3.9	6.1	26.7	3.3	4.3	14.0
	T ₂ - Stand. abnormal return	7.7	13.5	66.1	6.8	15.2	62.7	18.1	27.2	64.1
	T ₃ - Adjusted s.a.r.	7.5	13.1	65.7	6.7	14.9	61.9	15.8	25.6	61.4
Non-parametric tests	T ₄ - Rank test	6.3	14.3	52.8	5.1	17.2	52.8	8.6	17.5	49.5
	T ₅ - Sign test	4.5	10.8	28.7	4.0	11.9	31.9	6.6	13.0	30.0
	T ₇ - Generalised sign test.	5.4	12.0	31.9	4.1	14.4	35.9	8.1	17.6	36.8
Event induced variance adjusted tests	T ₈ - Parametric	5.8	12.4	51.4	6.8	14.9	50.3	9.2	16.3	47.7
	T ₆ - Rank test.	6.2	14.3	52.8	5.0	17.1	52.6	8.5	17.0	48.6
Cumulative or CAR Analysis										
Parametric tests	T ₁ - Adjusted cross sec. ind.	6.5	13.8	82.4	5.8	12.6	63.8	7.8	12.0	35.6
	T ₂ – stand. abnormal return	7.4	19.9	97.1	10.3	30.4	97.1	25.8	44.6	95.4
	T ₃ - Adjusted s.a.r.	7.2	19.4	96.9	9.9	30.1	97.0	24.8	43.3	94.9
Non-parametric tests	T ₄ - Rank test	5.0	24.7	96.2	7.7	37.8	95.9	12.0	37.3	92.0
	T ₅ - Sign Test	5.5	20.0	68.5	5.8	32.4	77.6	9.1	29.2	70.0
	T ₇ - Generalised sign test.	5.4	22.5	74.5	6.4	36.5	79.0	13.5	38.7	78.9
Event induced variance adjusted tests	T ₈ - Parametric	5.1	19.4	92.4	9.1	32.2	94.1	14.9	32.8	89.9
	T ₆ - Rank test	4.9	24.8	96.2	7.8	37.7	95.8	12.4	37.2	92.4

APPENDIX II—DESCRIPTION OF THE TEST STATISTICS

T₁ -- T-test with adjusted cross sectional independence (Brown and Warner, 1980, 1985)

The cross-sectional average of abnormal returns is given by:

$$\overline{A'_t} = \frac{1}{N_t} \sum_{j=1}^{N_t} A'_{jt} \quad (1)$$

The variance of the average is the average of the individual variances, so the test statistic (as yet unadjusted for forecast error) is:

$$T_1(\text{unadj.}) = \frac{\overline{A'_0}}{S(A')} \sim N(0,1).$$

The degrees of freedom are large, so that the test statistic can be assumed to be unit normal, and the standard deviation is given by:

$$S(A') = \frac{1}{N} \sqrt{\sum_1^N \frac{1}{T_i - 1} \sum_{t=1}^{T_i} [A'_{it}]^2}.$$

The variance of each security is calculated separately. Since returns are assumed to be independently distributed, the standard deviation of the cross-sectional average return on the event day is the average of the individual standard deviations.

Notice that $S(A')$ is an estimate of the standard deviation of abnormal returns over the estimation period, or event window. Since these abnormal returns are the forecasts from the market model, it is necessary to adjust $S(A')$ to account for the variance of the forecast error. The standard deviation of abnormal returns for security i on the event day (including the adjustment for forecast error) is then given by¹⁰:

$$S_i^A = \sqrt{\frac{1}{T_i - 1} \sum_{t=1}^{T_i} [A'_{it}]^2 \left(1 + \frac{1}{T_i} + \frac{(R_{m,0} - \overline{R}_m)^2}{\sum_t (R_{mt} - \overline{R}_m)^2} \right)}.$$

With adjustments for forecast error, the standard deviation used in the test statistic is:

$$S^A(A') = \frac{1}{N} \sqrt{\sum_i^N \frac{1}{T_i - 1} \sum_{t=1}^{T_i} [A'_{it}]^2 \left(1 + \frac{1}{T_j} + \frac{(R_{m,0} - \overline{R}_m)^2}{\sum_t (R_{mt} - \overline{R}_m)^2} \right)}.$$

¹⁰This is often referred to as Patell's (1976) adjustment.

Notice that T_i is the number of observed returns in the estimation window, so $\frac{1}{T_i}$ approaches zero

for large T_i . Also, since the value of the denominator $\sum_t (R_{m,t} - \bar{R}_m)^2$ increases with T_i ,

$\frac{(R_{m,0} - \bar{R}_m)^2}{\sum_t (R_{m,t} - \bar{R}_m)^2}$ approaches zero for large T_i . Given that T_i is typically over 200 for thickly traded

stocks, this adjustment is often ignored. For thinly traded stocks calculated using trade to trade returns, even though the estimation period exceeds 200 days, there may be only 10 to 20 observed returns over the period. Hence, it is not appropriate to ignore the above adjustment.

After adjusting for forecast error, the test statistic becomes:

$$T_1 = \frac{\bar{A}_0}{S^A(A')} \sim N(0,1).$$

The same test statistic applied to the unadjusted cumulative abnormal returns (CARs) is

$$T_1^{CAR}(\text{unadj.}) = \frac{\sum_{t=-1}^1 \bar{A}_t}{\sqrt{3} S(\bar{A}')} \sim N(0,1).$$

When Patell's adjustment is used, the test statistic is

$$T_1^{CAR} = \frac{\sum_{t=-1}^1 \bar{A}_t}{\sqrt{3} S^A(\bar{A}')}.$$

T₂ -- T-test with standardised abnormal return (Brown and Warner, 1985)

Standardised (unit variance) abnormal returns for security i are given by:

$$A_{it}^s = \frac{A_{it}'}{S(A_{it}')},$$

where

$$S(A_{it}') = \sqrt{\frac{1}{T_i - 1} \sum_{t=1}^{T_i} (A_{it}')^2}.$$

The test statistic for the event day is:

$$T_2 = \frac{1}{\sqrt{N_1}} \left(\sum_{t=1}^{N_1} (A_{jt}^s) \right) \sim N(0,1).$$

Assuming unit variance, the test statistic based on a three day event window is:

$$T_2^{CAR} = \frac{\sum_{j=1}^N A_{j-1}^s + \sum_{j=1}^N A_{j0}^s + \sum_{j=1}^N A_{j+1}^s}{\sqrt{3N}} \sim N(0,1).$$

T₃ -- T-test with adjusted standardised abnormal return (Brown and Warner, 1985 and Patell, 1976)

As was the case for T_1 , it is possible to adjust T_2 for the forecast errors, so that

$$A_{it}^{As} = \frac{A_{it}'}{S^A(A_{it}')},$$

where

$$S^A(A_{it}') = \sqrt{\frac{1}{T_i - 1} \sum_{t=1}^{T_i} [A_{it}']^2 \left(1 + \frac{1}{T_i} + \frac{(R_{m,0} - \bar{R}_m)^2}{\sum_t (R_{m,t} - \bar{R}_m)^2} \right)}$$

The adjusted test statistic for the event day is given by:

$$T_3 = \frac{\sum_{j=1}^{N_j} A_{jt}^{As}}{\sqrt{\sum_i \frac{T_j - 2}{T_i - 4}}}$$

The Cumulative Average Residuals for a three day window are

$$CAR_j = \sum_{t=-1}^1 A_{jt}.$$

The CARs are simply the unadjusted abnormal returns. Using the standard errors given by Salinger (1992, p. 41) for a 3-day event window, we construct a test statistic for the adjusted CARs as follows:

$$T_3^{A-CAR} = \sum_j^N \frac{\sqrt{N} \times Car_j}{\sqrt{3} \times \sigma \times \sqrt{1 + \frac{3}{T_j} + \frac{3 \left(\frac{1}{3} \sum_{t=-1}^1 r_{mt} - \sum_{t=-1}^1 r_{mt} \right)^2}{3 \times \sigma_{r_m}^2}}}$$

T₄ -- Rank test (Corrado, 1989 and Corrado and Zivney, 1992)

This test corresponds to T₃ in Corrado and Zivney [1992].

Let K_{it} denote the rank of excess return A_{it} in security i 's estimation and event period of 250 days:

$$K_{it} = \text{rank}(A_{it}), t = -248, \dots, +1.$$

The first 247 observations are used as an estimation period and 1 observation on each side of day "0" are used as the event window. To account for missing observations from thin trading Corrado and Zivney (1992) standardised each rank by the number of non-missing abnormal returns as follows:

$$U_{it} = \frac{K_{it}}{(1 + T_i)},$$

where T_i is the number of non-missing abnormal returns for security i over the entire period. Since the sum of T_i values of K_{it} is $\frac{T_i}{2}(T_i + 1)$, the average value of U_{it} is

$$\bar{U}_i = \frac{\frac{T_i}{2}(T_i + 1)}{T_i} = \frac{1}{2}.$$

The test statistic is given by:

$$T_4 = \frac{\frac{1}{\sqrt{N}} \sum_{j=1}^N \left(U_{jt} - \frac{1}{2} \right)}{S(K)} \quad \text{where} \quad S(K) = \sqrt{\frac{1}{250} \sum_{t=-248}^1 \left(\frac{1}{\sqrt{N_t}} \sum_{i=1}^{N_t} \left(U_{it} - \frac{1}{2} \right) \right)^2}$$

and N_t represents the number of non-missing returns in a cross-section of N -firms at time t . Under the null hypothesis the rank of the abnormal return is drawn from the uniform distribution, so Corrado and Zivney (1992) suggest that the test statistic converges to a standard normal.

The test statistic for CARs using ranks is given by:

$$T_4^{CAR} = \sum_{t=-1}^1 \frac{\frac{1}{\sqrt{N_t}} \sum_{j=1}^N \left(U_{jt} - \frac{1}{2} \right)}{\sqrt{3} S(K)}$$

T₅ -- Sign Test (Corrado and Zivney, 1992)

Let the median abnormal return in security i 's time series of abnormal returns be denoted by median (A_i). For each day in the sample period, the sign of each excess return is calculated as:

$$G_{it} = \text{sign}(A_{it} - \text{median}(A_i)), t = -248, \dots, +1,$$

where $\text{sign}(x)$ is +1, -1 or zero depending on whether x is positive, negative or zero. The expected value of G_{it} under the null hypothesis is zero and the test statistic is given by:

$$T_5 = \frac{\frac{1}{\sqrt{N_0}} \sum_{j=1}^{N_0} G_{j0}}{S(G)}$$

where

$$s(G) = \sqrt{\frac{1}{250} \sum_{t=-244}^5 \left(\frac{1}{\sqrt{N_t}} \sum_{i=1}^{N_t} G_{it} \right)^2}.$$

and N_t is the number of non-missing returns on day t .

The test statistic for CARs using signs is given by:

$$T_5^{CAR} = \sum_{t=-1}^1 \frac{\frac{1}{\sqrt{N_0}} \sum_{j=1}^{N_0} G_{jt}}{\sqrt{3}S(G)}.$$

T₆ -- Generalised sign test (Cowan, 1992 and 1996)

For the traditional sign test, T_5 , the expected number of positive abnormal returns under the null hypothesis is 0.5, whereas for this test the expected number of abnormal returns is estimated from the estimation period across time and stocks. The fraction of positive abnormal returns under the null hypothesis is given by:

$$\hat{p} = \frac{1}{N} \sum_{j=1}^N \frac{1}{T_j} \sum_{t=1}^{T_j} \varphi_{jt},$$

where $\varphi_{it} = 1$ if $A_{it} > 0$, and $\varphi_{it} = 0$ otherwise. The generalised sign test statistic is:

$$T_6 = \frac{w - Nxp}{\sqrt{Nxp(1-\hat{p})}}.$$

Again, the CAR version of the test statistic is the sum of the three test statistics during the event window divided by the square root of three.

T₇ - Variance adjusted standardised abnormal returns (Boehmer, Musumeci and Poulsen, 1991)

Often an event leads to increases in the variance of the returns around the event due to a temporary increase in systematic risk, uncertainty regarding the effects of the event, etc. In this situation, the standard deviation calculated over the usual estimation period under states the standard deviation of returns expected during the three-day event window. To adjust for this problem, the residuals are first standardised and then variance is estimated during the event window. The resulting test is given by:

$$T_9 = \frac{\frac{1}{N} \sum_{i=1}^N A_{it}^{As}}{\sqrt{\frac{1}{N(N-1)} \sum \left[A_{it}^{As} - \frac{1}{N} \sum_{i=1}^N A_{it}^{As} \right]^2}}$$

where

$$A_{it}^{As} = \frac{A'_{it}}{S^A(A_{it})}$$

and

$$S^A(A_{it}) = \sqrt{\frac{1}{T_i - 1} \sum_{t=1}^{T_i} [A'_{it}]^2 \left(1 + \frac{1}{T_i} + \frac{(R_{m,0} - \bar{R}_m)^2}{\sum_t (R_{m,t} - \bar{R}_m)^2} \right)}$$

The CAR version of this test statistic is calculated using Salinger [1992] as for T₃ above.

T₈ -- Rank test of adjusted standardised abnormal returns Corrado and Zivney [1992] and Maynes and Rumsey, (1993.)

This test is equivalent to T₄ except that it makes a cross-sectional variance adjustment and follows Corrado and Zivney [1992]. The standardised returns are given by

$$A_{it}^s = \frac{A_{it}}{S(A_i)}$$

where $S(A_j)$ is given by

$$S(A_i) = \sqrt{\frac{1}{T_i - 1} \sum_{t=1}^{T_i} (A_{it}')^2} \text{ for the estimation period}$$

$$S(A_i) = \sqrt{\frac{1}{T_i - 1} \sum_{t=1}^{T_i} [A_{it}']^2 \left(1 + \frac{1}{T_i} + \frac{(R_{m,0} - \bar{R}_m)^2}{\sum_t (R_{m,t} - \bar{R}_m)^2} \right)} \text{ in the event window}$$

The cross-sectional variance adjustment is then applied:

$$X_{it} = \begin{cases} A_{it}^s, & t \neq 0 \\ \frac{A_{it}^s}{S(A_0^s)}, & t = 0 \end{cases}$$

where

$$S(A_0^s) = \sqrt{\frac{1}{N - 1} \sum_{i=1}^N (A_{i0}^s - \bar{A}_0^s)^2}$$

The test statistic is then derived in the same manner as T_4 .

Working Papers from Finance Research Group

- F-2006-03 Jan Bartholdy, Dennis Olson & Paula Peare: Conducting event studies on a small stock exchange.
- F-2006-02 Jan Bartholdy & Cesário Mateus: Debt and Taxes: Evidence from bank-financed unlisted firms.
- F-2006-01 Esben P. Høg & Per H. Frederiksen: The Fractional Ornstein-Uhlenbeck Process: Term Structure Theory and Application.
- F-2005-05 Charlotte Christiansen & Angelo Ranaldo: Realized bond-stock correlation: macroeconomic announcement effects.
- F-2005-04 Søren Willemann: GSE funding advantages and mortgagor benefits: Answers from asset pricing.
- F-2005-03 Charlotte Christiansen: Level-ARCH short rate models with regime switching: Bivariate modeling of US and European short rates.
- F-2005-02 Charlotte Christiansen, Juanna Schröter Joensen and Jesper Rangvid: Do more economists hold stocks?
- F-2005-01 Michael Christensen: Danish mutual fund performance - selectivity, market timing and persistence.
- F-2004-01 Charlotte Christiansen: Decomposing European bond and equity volatility.



Handelshøjskolen i Århus

Aarhus
School of Business

ISBN 87-7882-129-0

Department of Business Studies

Aarhus School of Business
Fuglesangs Allé 4
DK-8210 Aarhus V - Denmark

Tel. +45 89 48 66 88
Fax +45 86 15 01 88

www.asb.dk