Using Daily Stock Returns: The Case of Event Studies

Article *in* Journal of Financial Economics · February 1985

Source: RePEc

2 authors, including:

Stephen J Brown
Monash University (Australia)

128 PUBLICATIONS   **12,594** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Hedge Funds View project

USING DAILY STOCK RETURNS
The Case of Event Studies*

Stephen J. BROWN

Yale University, New Haven, CT 06520, USA

Jerold B. WARNER

University of Rochester, Rochester, NY 14627, USA

This paper examines properties of daily stock returns and how the particular characteristics of these data affect event study methodologies. Daily data generally present few difficulties for event studies. Standard procedures are typically well-specified even when special daily data characteristics are ignored. However, recognition of autocorrelation in daily excess returns and changes in their variance conditional on an event can sometimes be advantageous. In addition, tests ignoring cross-sectional dependence can be well-specified and have higher power than tests which account for potential dependence.

1. Introduction

This paper examines properties of daily stock returns and how the particular characteristics of these data affect event study methodologies for assessing the share price impact of firm-specific events. The paper extends earlier work [Brown and Warner (1980)] in which we investigate event study methodologies used with monthly returns. In our previous work, we conclude that a simple methodology based on the market model is both well-specified and relatively powerful under a wide variety of conditions, and in special cases even simpler methods also perform well. However, the applicability of these conclusions to event studies using daily data is an open question [e.g., Brown and Warner (1980, p. 21), Masulis (1980, p. 157), Dann (1981, p. 123), DeAngelo and Rice (1983, p. 348), McNichols and Manegold (1983, p. 58)]. Daily and monthly data differ in potentially important respects. For example, daily stock returns

depart more from normality than do monthly returns [Fama (1976, ch. 1)]. In addition, the estimation of parameters from daily data is complicated by non-synchronous trading, a complication described as 'especially severe' by Scholes and Williams (1977, p. 324).

The paper first studies the statistical properties of both observed daily stock returns and of daily excess returns, given a variety of alternative models for measuring excess returns. To examine the implications of these properties for event studies, a procedure similar to one we developed previously is applied to observed daily returns. Various event study methodologies are simulated by repeated application of each methodology to samples that have been constructed by random selection of securities and random assignment of an 'event-date' to each security. With randomly selected securities and event dates there should be no abnormal performance on average if performance is measured correctly. We study the probability of rejecting the null hypothesis of no average abnormal performance when it is true. We also evaluate the power of the tests, that is, their probability of detecting a given level of abnormal performance.

Section 2 outlines important methodological issues associated with using daily data. Section 3 discusses the experimental design for examining those issues. The paper's results are presented in sections 4 through 6. In section 7, we summarize the results and present our conclusions. An appendix gives additional details of the methodologies examined in the paper.

2. Using daily data: The issues

The use of daily data in event studies involves a number of potentially important problems. These can be summarized as follows.

2.1. Non-normality

The daily stock return for an individual security exhibits substantial departures from normality that are not observed with monthly data. The evidence generally suggests that distributions of daily returns are fat-tailed relative to a normal distribution [Fama (1976, p. 21)]. As shown later, the same holds true for daily excess returns. Since event studies generally focus on the cross-sectional sample mean of security excess returns, this paper will examine the small sample properties of the mean excess return. The Central Limit Theorem [see Billingsley (1979, pp. 308–319)] guarantees that if the excess returns in the cross-section of securities are independent and identically distributed drawings from finite variance distributions, the distribution of the sample mean excess return converges to normality as the number of securities increases. There is some evidence that the distribution of the cross-sectional daily mean return converges to a normal [Blattberg and Gonedes (1974), Hagerman (1978)]. A

chief concern here is whether and for what sample size this result applies to the excess returns, even though the assumptions of the Central Limit Theorem (at least in its standard version) are violated with these data.

2.2. Non-synchronous trading and market model parameter estimation

When the return on a security and the return on the market index are each measured over a different trading interval, ordinary least squares (OLS) estimates of market model parameters are biased and inconsistent. With daily data, the bias can be severe [Scholes and Williams (1977, p. 324), Dimson (1979, p. 197)]. Concerned with that problem, authors of event studies with daily data have used a variety of alternative techniques for parameter estimation [e.g., Gheyara and Boatsman (1980, p. 111), Holthausen (1981, p. 88)]. This paper investigates the use of both OLS and other procedures.

2.3. Variance estimation

With both daily and monthly data, estimation of the variance of the sample mean excess return is important for tests of statistical significance. This paper investigates several variance estimation issues.

The first issue is the time-series properties of daily data. As a consequence of non-synchronous trading, daily excess returns can exhibit serial dependence. Attempts to incorporate such dependence into variance estimates have appeared in the event study literature [e.g., Ruback (1982)]. Serial dependence in excess returns and its implications for event studies are examined.[1]

The second issue is cross-sectional dependence of the security-specific excess returns. Advantages of incorporating cross-sectional dependence into the variance estimator for the mean excess return are well-known [e.g., Brown and Warner (1980), Beaver (1981), Dent and Collins (1981)], and are not limited to daily data. In contrast to the existing literature, this paper focuses on the potential costs of dependence adjustment.

The third issue is stationarity of daily variances. There is evidence that the variance of stock returns increases for the days immediately around events such as earnings announcements [e.g., Beaver (1968), Patell and Wolfson (1979)]. We illustrate how this possibility affects event study procedures.

2.4. Important properties captured by simulation

The basis for inference in event studies is a test statistic, typically the ratio of the mean excess return to its estimated standard deviation. Studying the

---

[1] Other time-series properties, such as day of the week or weekend effects [e.g., French (1980), Gibbons and Hess (1981)], are not explicitly studied.

properties of the test statistic analytically requires detailed knowledge of the distributional properties of excess returns both in the time-series and cross-section. A complication is that variables such as the degree of non-synchronous trading can simultaneously affect both mean and variance estimators [Scholes and Williams (1977, pp. 313–314)]. This paper employs simulation procedures using actual stock return data to investigate the distribution of excess returns and, in particular, the empirical properties of the test statistics. The procedures thus use data from the true return-generating process and are a direct way to summarize how the various problems of daily data jointly affect the event study methodologies.

3. Experimental design

3.1. Sample construction

Two hundred and fifty samples of 50 securities are constructed. The securities are selected at random and with replacement from the population of all securities for which daily return data are available on the files of the Center for Research in Security Prices at the University of Chicago (CRSP). Each time a security is selected, a hypothetical event day is generated. Events are selected with replacement, and are assumed to occur with equal probability on each trading day from July 2, 1962, through December 31, 1979.

Define day '0' as the day of a hypothetical event for a given security. For each security we use a maximum of 250 daily return observations for the period around its respective event, starting at day -244 and ending at day $+5$ relative to the event. The first 239 days in this period (-244 through -6) is designated the 'estimation period', and the following 11 days (-5 through $+5$) is designated the 'event period'. For a security to be included in a sample, it must have at least 30 daily returns in the entire 250 day period, and no missing return data in the last 20 days.

3.2. Excess return measures

Let $R_{i,t}$ designate the observed arithmetic return for security i at day t. Define $A_{i,t}$ as the excess return for security i at day t. For every security, the excess return for each day in the event period is estimated using the following procedures:

Mean adjusted returns

$$A_{i,t} = R_{i,t} - \overline{R}_i, \tag{1}$$

$$\overline{R}_i = \frac{1}{239} \sum_{t=-244}^{-6} R_{i,t}, \tag{2}$$

where $\overline{R}_i$ is the simple average of security i's daily returns in the $(-244, -6)$ estimation period.[2]

Market adjusted returns

$$A_{i,t} = R_{i,t} - R_{m,t}, \tag{3}$$

where $R_{m,t}$ is the return on the CRSP equally weighted index for day t.

OLS market model

$$A_{i,t} = R_{i,t} - \hat{\alpha}_i - \hat{\beta}_i R_{m,t}, \tag{4}$$

where $\hat{\alpha}_i$ and $\hat{\beta}_i$ are OLS values from the estimation period.

Most of these methods and models of the return-generating process on which they are based are discussed in Brown and Warner (1980, pp. 207–209). The Scholes–Williams and Dimson procedures for estimating market model parameters are considered later.

3.3. Test statistics under the null hypothesis

Given the excess returns based on each method, the statistical significance of the event period excess returns is assessed for each sample. The null hypothesis to be tested is that the mean day '0' excess return (e.g., the simple average of market model excess returns) is equal to zero,[3] and thus concerns the average affect of an event on returns to shareholders. The test statistic is the ratio of the day '0' mean excess return to its estimated standard deviation; the standard deviation is estimated from the time-series of mean excess returns. The test statistic for any event day t (in this case $t = 0$) is

$$\overline{A}_t / \hat{S}(\overline{A}_t), \tag{5}$$

where

$$\overline{A}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} A_{i,t}, \tag{6}$$

$$\hat{S}(\overline{A}_t) = \sqrt{\left(\sum_{t=-244}^{t=-6} (\overline{A}_t - \overline{\overline{A}})^2 \right) \Big/ 238}, \tag{7}$$

$$\overline{\overline{A}} = \frac{1}{239} \sum_{t=-244}^{t=-6} \overline{A}_t, \tag{8}$$

[2] Eq. (2) and subsequent equations apply if there are no missing returns in the estimation period. With missing returns, parameter estimation excludes both the day of the missing return and the return for the subsequent day.

[3] Later, tests over longer intervals are examined.

and where N_t is the number of sample securities whose excess returns are available at day t.[4]

A statistic of this form is widely used in event studies [e.g., Masulis (1980), Dann (1981), Holthausen (1981), Leftwich (1981)]. If the $\bar{A}_t$ are independent, identically distributed, and normal, the test statistic is distributed Student-t under the null hypothesis. Since the degrees of freedom exceeds 200, the test statistic is assumed unit normal. Note that by using a time-series of average excess returns (i.e., 'portfolio' excess returns), the test statistic takes into account cross-sectional dependence in the security-specific excess returns. However, the test statistic we initially use ignores any time-series dependence in excess returns.

3.4. Test statistics with abnormal performance

Procedures for introducing a given level of abnormal performance are similar to those in Brown and Warner (1980). A constant is added to the observed day 0 return for each security. For example, to simulate 1% abnormal performance, 0.01 is added. In the initial simulations, the level of abnormal performance is the same for all sample securities.

With constant sample-wide abnormal performance, the procedure for introducing abnormal performance is equivalent to taking the test statistic in (5) under the null hypothesis and adding to it the level of abnormal performance divided by the estimated standard deviation of the mean excess return. Thus, it is computationally easy to study the test statistics for different abnormal performance levels, and hence to estimate the power function by modifying the empirical distribution of test statistics under the null hypothesis.

4. Initial results

4.1. Properties of daily excess returns

Table 1 shows the properties of the various event study performance measures when no abnormal performance is introduced. Panel A of the table shows the properties of the daily returns and excess returns based on time-series data in the estimation period for each security. Panel B details the cross-sectional properties of the 250 day 0 sample-wide mean returns and mean excess returns.

---

[4] For Mean Adjusted Returns and the market model, the denominator of eq. (5) should, in principle, be adjusted because the excess returns are prediction errors. All of the paper's simulations were repeated with the appropriate variance adjustments, but there was no detectable impact. Details of the adjustment procedures were discussed in a previous version of the paper and will be furnished on request.

Table 1

Properties of daily data event study performance measures when no abnormal performance is introduced. Randomly selected securities and event dates; 250 samples of 50 securities; time period: 1962-1979.

Panel A

Properties of daily performance measures for individual common stocks. For each security, parameter estimates are based on estimation period excess returns. For each parameter, the table reports the mean of the 12,500 estimates. Maximum number of observations per security = 239; equal weighted index used as market portfolio; mean of 12,500 values.

Performance measure	Mean	Standard deviation	Skewness	Kurtosis	Studentized range
Returns	0.0006	0.0267	0.99	6.87	7.59
Mean adjusted returns	0.0000	0.0267	0.99	6.87	7.59
Market adjusted returns	0.0000	0.0258	0.97	6.66	7.49
Market model	0.0000	0.0253	1.01	6.80	7.54

Panel B

Cross-sectional properties of sample-wide mean performance measures at day '0'. Each number reported in the table is based on 250 values of the mean performance measure, one for each sample. For a given sample, the mean performance measure is the simple average of the performance measures for the individual securities in the sample.

Sample size	Method	Mean	Standard deviation	Skewness	Kurtosis	Studentized range
5	Returns	0.0005	0.0127	0.39	4.68	7.13
	Mean adjusted returns	− 0.0001	0.0127	0.42	4.71	7.05
	Market adjusted returns	0.0001	0.0122	0.40	4.91	7.42
	Market model	0.0001	0.0121	0.42	5.04	7.64
20	Returns	0.0007	0.0062	0.00	3.30	6.42
	Mean adjusted returns	0.0001	0.0062	0.00	3.33	6.41
	Market adjusted returns	0.0001	0.0060	0.02	3.28	6.40
	Market model	0.0001	0.0059	0.03	3.43	6.47
50	Returns	0.0006	0.0040	0.08	3.17	5.69
	Mean adjusted returns	0.0000	0.0040	0.08	3.19	5.73
	Market adjusted returns	− 0.0001	0.0039	0.10	3.23	5.81
	Market model	− 0.0001	0.0038	0.10	3.10	5.59

Upper percentage points; samples of 250 drawn from a normal population.

Variable	0.95	0.99
Skewness	0.063	0.129
Kurtosis	3.52	3.87
Studentized range ($N = 200$)	6.15	6.85

4.1.1. Results for individual securities

From panel A of table 1, it appears that daily returns and daily excess returns are highly non-normal. The mean studentized range of the returns is 7.59, compared to a value of 6.85 for the 0.99 fractile of the studentized range of samples drawn from a normal population of size 200. Mean values of skewness and kurtosis coefficients for the returns exceed the value of the 0.99 fractile of the respective distribution under normality. These results do not change markedly for the measures of excess returns. For example, with the market model, the mean studentized range is 7.54. Although not reported in table 1, use of continuously compounded returns or of the value weighted index as the market portfolio also yields similar results.

One additional point about panel A is that the various performance measures have similar standard deviations. The mean standard deviation of the returns is 0.0267; the mean standard deviation of excess returns from the market model is 0.0258, and the mean market model R^2 is only 0.10. These results suggest that the alternative measures of excess returns will exhibit similar ability to detect abnormal performance when it is present.

4.1.2. Results for mean excess returns

Panel B of table 1 indicates that departures from normality are less pronounced for cross-sectional mean excess returns than for individual security excess returns, as would be expected under the Central Limit Theorem. For samples of size 50, the mean excess return seems close to normal. The studentized range of the 250 day 0 mean performance measures ranges from 5.59 to 5.81, and similar average values apply when the properties of the mean excess returns in the estimation period for each sample are studied. However, while values of the studentized range in panel B are consistent with normality of the mean performance measure in samples of 50, there is still more skewness than would be expected under normality, with skewness coefficient values ranging from 0.08 for the returns to 0.10 for Market Adjusted Returns and the market model. Furthermore, the conclusion that the mean performance measure is somewhat close to normal does not apply to smaller samples. In samples of five, the studentized range of the mean performance measure is typically in excess of seven, and in samples of 20 it is still on the order of 6.5. These differing results for different sample sizes raise the possibility that the degree of misspecification in the event study methodologies is sensitive to sample size. This topic is investigated in section 4.3.

One final point about panel B is that the various performance measures typically have an expected value of approximately zero in the event period. For example, for samples of size 50, the average of the 250 mean excess returns from the market model is -0.0001, with a t-statistic of -0.46. The only
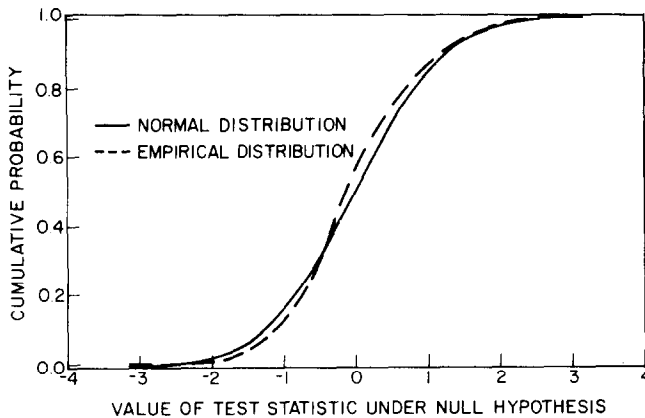
Fig. 1. Theoretical and empirical cumulative distribution of the test statistics using the market model. The solid figure is the theoretical distribution assumed under the null hypothesis, a Student-t with large degrees of freedom, approximated with a unit normal. The dashes indicate the empirical distribution of the 250 test statistics under the null hypothesis, where each test statistic is for a sample of 50 randomly selected securities with 50 randomly selected event dates from the 1962–1979 period.

exception is for the mean returns, where the average value is 0.0006, with a t-statistic of 2.33. This merely indicates that average returns in the cross-section are positive.[5]

4.2. Properties of the test statistics

For the market model and a sample size of 50, fig. 1 compares the cumulative distribution of the 250 test statistics when no abnormal performance is introduced with cumulative values from a unit normal. The similarity between the empirical and theoretical distributions is striking, and this conclusion also applies when the corresponding figure for this experimental situation is examined for the other methodologies.

From table 2, the formal tests also indicate that the empirical distributions of the various test statistics are close to unit normal. The chi-square tests for goodness of fit typically fail to find misspecification, even when the goodness of fit tests concentrate in the tail regions. However, there is some evidence that the test statistics are slightly skewed and leptokurtic. For all methodologies, the studentized range of the test statistics is in excess of six.

---

[5]An event study methodology which ignores the positive average return will tend to find positive abnormal performance when none is present. Such a 'Raw Returns' methodology [i.e., Mean Adjusted Returns where R_t in eq. (1) is constrained to zero] was examined in an earlier version of the paper.

Table 2

Summary measures for the actual frequency distribution of each test statistic, based on 250 values, one for each sample; sample size = 50. Randomly selected securities and event dates; no abnormal performance introduced; time period: 1962–1979.

Method	Mean	Standard deviation	t-statistic for mean	Skewness	Kurtosis	Studentized range	χ^2 statistic[a]	χ^2 statistic[b]
Mean adjusted returns	−0.014	0.92	−0.24	0.094	3.61	6.68	8.7	26.4
Market adjusted returns	−0.019	0.91	−0.33	0.073	3.52	6.60	5.3	22.3
Market model	−0.028	0.91	−0.49	0.084	3.42	6.50	7.0	19.6

	Upper percentage points	
	0.95	0.99
χ^2 (8)	15.5	20.1
χ^2 (19)	30.1	36.2

[a] Nine intervals concentrating in the tail regions. The intervals are 0 − 0.01, 0.01 − 0.02, 0.02 − 0.05, 0.05 − 0.1, 0.1 − 0.9, 0.9 − 0.95, 0.95 − 0.98, 0.98 − 0.99, 0.99 − 1.0.
[b] Twenty equally spaced intervals.

4.2.1. The power of the tests

Table 3 shows rejection frequencies for various levels of abnormal performance ranging from 0 to 2%, using one-tailed tests at the 0.05 significance level. For illustrative purposes, rejection frequencies for this significance level will be used throughout the paper.[6]

From table 3, with no abnormal performance rejection rates range from 4.4% to 6.4%, well within the interval of 2% to 8% for a 95% confidence band under correct specification [Brown and Warner (1980, p. 216)]. With 1% abnormal performance at day 0, the frequency of detecting abnormal performance ranges from 75.6% for Mean Adjusted Returns to 80.4% with the market model. These rejection frequencies indicate little difference in the power of alternative procedures. Moreover, these rejection frequencies are roughly three times those reported for monthly data [Brown and Warner (1980, table 3)], thus highlighting the substantial gains to more precise pinpointing of an event.

Furthermore, the power of the methodologies in table 3 is similar to the theoretical power derived assuming unit normality of the test statistics under the null hypothesis. For example, fig. 2 compares the empirical and theoretical power for the market model. For a variety of abnormal performance levels,

---

[6] For every experiment in the paper, the entire distribution of test statistics under the null hypothesis was also examined, as in table 2. Details for each experiment will be summarized in the text.

Table 3

A comparison of different procedures for detecting abnormal performance; percentage of 250 samples where the null hypothesis is rejected. H_0: mean abnormal performance at day '0' = 0.[a]

Method	Actual level of abnormal performance at day '0'			
	0	0.005	0.01	0.02
Mean adjusted returns	6.4%	25.2%	75.6%	99.6%
Market adjusted returns	4.8	26.0	79.6	99.6
Market model	4.4	27.2	80.4	99.6

[a] Sample size = 50 securities; one-tailed test, α = 0.05; randomly selected securities and event dates, 1962–1979.
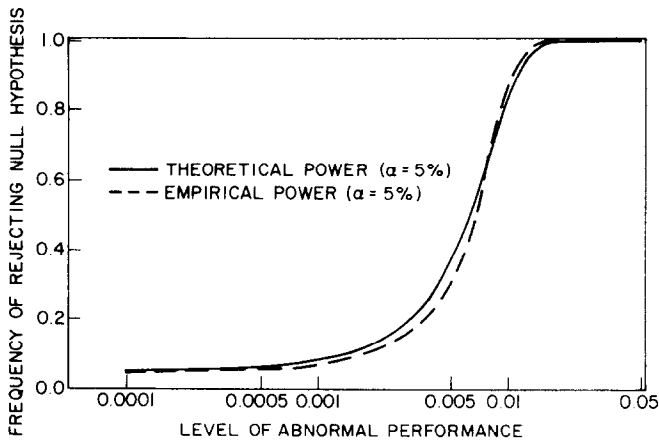

Fig. 2. Theoretical and empirical power using the market model. The theoretical power is the rejection frequency assuming unit normality of the test statistics under the null hypothesis. The empirical power is the rejection frequency using the empirical distribution of the 250 test statistics under the null hypothesis. For the empirical power, each test statistic is for a sample of 50 randomly selected securities with 50 randomly selected event dates from the 1962–1979 period.

For computational simplicity, the derivation of each power function in this figure assumes that the standard deviation of the mean excess return is the same for all samples, and equal to the value of 0.0038 reported in panel B of table 1. Thus, for each power function, the rejection frequency is the proportion of test statistics whose value exceeds [1.645 − (abnormal performance/0.0038)], where 1.645 is the 0.95 fractile of a unit normal. The empirical power in table 3 differs from that in the figure because the standard deviation of the mean excess return can differ across samples.

rejection frequencies are similar, and this also applied to tests at the 0.1 and 0.01 significance levels. Since the test statistics in the experimental situation examined in table 3 were well-specified and approximately unit normal under the null hypothesis, such results are not surprising.

4.3. Sensitivity analysis

The conclusion from the baseline simulations that the test statistics for most methods are reasonably well-specified is consistent with simulation results reported elsewhere [Dodd and Warner (1983), Dyckman, Philbrick and Stephan (1984)], and is not highly sensitive to several changes in the experimental procedure.

4.3.1. Smaller samples

For samples of either five or 20 securities, the specification of the test statistics is not dramatically altered. The goodness of fit tests do not indicate misspecification. However, the degree of skewness and kurtosis in the test statistics is higher for samples of size five and 20 than for samples of 50. For example, kurtosis is typically in excess of four for test statistics from samples of size five. Thus, stated significance levels should not be taken literally.

4.3.2. Longer event periods

Event study test statistics also continue to be generally well-specified when the event period is longer than one day. Table 4 shows results for samples of

Table 4

Comparison of different methods when the event period is longer than one day. The table shows the percentage of 250 samples where the null hypothesis is rejected. For each security, abnormal performance is introduced for one day in the 11-day interval (− 5, + 5), with each day having an equal probability of selection. The null hypothesis is that the cumulative mean excess return in the interval (− 5, + 5) = 0. Rejection rates for a one-day event period from table 3 are also shown: for each security abnormal performance is introduced at day '0', and the null hypothesis in that the mean abnormal performance at day '0' is equal to 0.[a]

Method	Days in event period	Actual level of abnormal performance		
		0	0.01	0.02
Mean adjusted	11	4.0%	13.6%	37.6%
returns	1	6.4	75.6	99.6
Market adjusted	11	4.0	13.2	32.0
returns	1	4.8	79.6	99.6
Market model	11	2.8	13.2	37.2
	1	4.4	80.4	99.6

[a] Sample size = 50; one-tailed test, $\alpha = 0.05$; randomly selected securities and event dates, 1962–1979.

size 50 where:

(1) the null hypothesis (using a test statistic discussed in the appendix) is that the cumulative mean daily excess return over the $(-5, +5)$ interval is equal to zero, and

(2) for each security, abnormal performance is introduced for one day in the interval $(-5, +5)$, with each day having an equal probability of being selected.

Rejection rates under the null hypothesis range from 2.8% to 4.0%. Various goodness of fit tests failed to find marked evidence of test statistic misspecification, despite the failure of the test statistic to explicitly incorporate autocorrelation in the mean daily excess returns.[7] These findings also apply in samples of sizes five and 20. As expected, the power of the tests decreases when the abnormal performance occurs over the $(-5, +5)$ interval rather than at day 0.[8] For example, with 1% abnormal performance, the rejection frequency for Market Adjusted Returns is 13.2%, compared to the earlier figure of 79.6% reported in table 3.

4.3.3. Clustering

The results on specification are not radically altered in experiments where there is clustering in event dates and hence dependence of the excess return measures. Table 5 shows rejection frequencies for the case where day 0 is restricted to be a particular calendar day which is common to all securities in a given sample.[9] As in table 4, the null hypothesis is that the cumulative mean daily excess return over the interval $(-5, +5)$ is equal to zero. From table 5, Market Adjusted Returns and the market model have rejection rates with no abnormal performance of 4.0% and 3.2%, respectively, and the goodness of fit tests typically fail to detect misspecification.

A striking result from table 5 is that for Mean Adjusted Returns, the rejection rate with no abnormal performance is 13.6%. This apparent misspecification was not observed for Mean Adjusted Returns in the non-clustering case, nor is it observed for the clustering case when the hypothesis test is conducted at day 0. As we shall see later, the misspecification is at least partly

---

[7]Similar results also apply for event periods of other lengths; see Dyckman, Philbrick and Stephan (1984).

[8]As in fig. 2, we also compared theoretical and empirical power functions for an 11-day event period. The striking similarity in theoretical and empirical power was still present, even in samples of 5 securities.

[9]The day is randomly selected for a given sample and selection of the day is carried out without replacement.

Table 5

The effect of event-day clustering. The table shows the percentage of 250 samples where the null hypothesis is rejected. For a given sample, day '0' is the same calendar date for each security. The calendar day differs from sample to sample. Corresponding results for non-clustering case reported in table 4 are also shown. H_0: cumulative mean daily abnormal performance in the interval $(-5, +5) = 0.$[a]

Method		Actual level of abnormal performance in interval $(-5, +5)$[b]		
		0	0.01	0.02
Mean adjusted	Clustering	13.6%	21.2%	29.6%
returns	Non-clustering	4.0	13.6	37.6
Market adjusted	Clustering	4.0	14.4	46.0
returns	Non-clustering	4.0	13.2	32.0
Market model.	Clustering	3.2	15.6	46.0
	Non-clustering	2.8	13.2	37.2

[a] Sample size = 50; one-tailed test, $\alpha = 0.05$; randomly selected securities and event dates, 1962–1979.
[b] For each security, abnormal performance is introduced for one day in the interval, with each day having an equal probability of selection.

related to autocorrelation in the time-series of average mean adjusted returns. Furthermore, misspecification with daily data event study methodologies is not limited to cases involving autocorrelation. The remainder of the paper studies a number of experimental situations and identifies several where various method-ologies are either poorly specified or inefficient.

5. Non-synchronous trading: Alternative procedures for β estimation

Non-synchronous trading introduces potentially serious difficulties into em-pirical studies using daily stock returns [Scholes and Williams (1977) and Dimson (1979)]. One reason is that in the presence of non-synchronous trading, OLS estimates of market model β are biased and inconsistent. The evidence is that shares traded relatively infrequently have downward biased β estimates, while those traded relatively frequently have upward biased β estimates. However, the results presented thus far in this paper indicate that the failure to take into account non-synchronous trading in estimating market model coefficients does not result in misspecification of event study methodolo-gies using the OLS market model.

Correct specification of event study methodologies using the OLS market model is consistent with the evidence presented by Scholes and Williams and Dimson that OLS estimates of β are biased. Even when biases in β exist they do not necessarily imply misspecification in an event study. By construction, OLS residuals for a security sum to zero in the estimation period so that a bias in the estimate of β is compensated for by a bias in α. With stationarity, the event period excess returns for an individual security can be shown to have a mean equal to zero unconditional on the market return. The excess return

Table 6

Alternative procedures for market-model estimation. Percentage of 250 samples where the null hypothesis is rejected. H_0: mean abnormal performance at day '0' = 0.[a]

	Actual level of abnormal performance at day 0		Mean value in estimation period ($N = 5000$)		
	0	0.01	$\overset{\bullet}{\sigma}$	$\hat{\alpha}$	$\hat{\beta}$
Panel A: Randomly selected securities					
OLS market model	2.8%	46.8%	0.028	0.0	1.00
Scholes–Williams procedure	2.8	46.8	0.028	0.0	1.00
Dimson aggregated coefficients method	2.4	47.2	0.030	0.0	1.00
Panel B: Sample formation by exchange listing					
New York Stock Exchange samples					
OLS market model	5.2	64.4	0.022	− 0.0002	0.96
Scholes–Williams procedure	4.4	64.8	0.022	− 0.0001	0.94
Dimson aggregated coefficients method	4.4	65.2	0.022	− 0.0001	0.91
American Stock Exchange samples					
OLS market model	5.6	28.8	0.037	0.0002	1.05
Scholes–Williams procedure	5.2	31.2	0.037	0.0001	1.07
Dimson aggregated coefficients method	5.2	30.0	0.037	0.0001	1.11

[a] Sample size = 20 securities; one-tailed test, $\alpha = 0.05$; time period: 1962–1979.

conditional on the market is biased for an individual security. However, no misspecification in an event study is implied if the average bias in the conditional excess returns of the sample securities is zero. This can occur if the securities are drawn from a representative range of trading frequencies or if there is non-clustering of event dates. Nevertheless, even when OLS is well-specified, its use could result in imprecise (as opposed to biased) estimates of excess returns, yielding tests with relatively low power. Thus, it is of interest to investigate alternative procedures for market model parameter estimation.

5.1. Specification and power of Scholes–Williams- and Dimson-based procedures

All of the paper's experiments were repeated using both (1) the Scholes–Williams procedures and (2) the Dimson aggregated coefficients method, with three leads and three lags.[10] A representative set of results, for a sample size of 20, is presented in table 6.[11]

[10] These procedures are discussed in the appendix. For a recent discussion of the Dimson procedure, see Fowler and Rorke (1983).

[11] As we will discuss, the experiments in panel B involve American Stock Exchange securities; given our data base and the requirement that there be 250 samples, 20 is roughly the maximum sample size which can be specified there.

From panel A of table 6, it appears that these alternative methodologies convey no clear-cut benefit in an event study. For the various levels of abnormal performance, rejection rates using both the Scholes–Williams and Dimson procedures are similar to the results obtained with OLS. For example, with 1% abnormal performance, rejection frequencies range from 46.8% for both the Scholes–Williams procedure and the OLS market model to 47.2% for the Dimson method. Although not reported in table 6, the properties of the test statistics are similar for the three procedures and reasonably close to unit normal under the null hypothesis.

5.2. Sample formation by trading frequency: NYSE versus AMEX securities

Panel B reports results when the sample securities tend to trade with frequencies systematically different from average. We have no data on trading frequency. The proxy we use is exchange-listing. There is reason to believe that New York Stock Exchange (NYSE) stocks tend to trade more frequently than average and that American Stock Exchange (AMEX) stocks trade less frequently than average.[12] In panel B of table 6, simulation results for the different methodologies are compared for both NYSE and AMEX stocks.

For NYSE stocks, OLS estimates of β are on average higher than the estimates using the Scholes–Williams or Dimson procedure; the average values are 0.96, 0.94, and 0.91, respectively. For AMEX stocks, the relative magnitudes of the average β estimates are reversed, with the average estimates ranging from 1.05 with the OLS market model to 1.11 with the Dimson procedure. These figures suggest biases in OLS estimates of β similar to those discussed by Scholes and Williams and by Dimson. The figures also indicate a relation between true β and trading volume [Scholes and Williams (1977, p. 320)].

However, there is no evidence that procedures other than OLS improve either the specification or the power of the tests.[13] In absence of abnormal performance, rejection rates for all methodologies are approximately equal to the significance level of the test, ranging from 4.4% to 5.6%. Furthermore, when abnormal performance is present, rejection rates are similar for OLS, the Scholes–Williams, and the Dimson-based procedures. Rejection frequencies with 0.01 abnormal performance range from 64.4% to 65.2% for the NYSE samples and from 28.8% to 31.2% for the AMEX samples. Rejection frequencies are higher for NYSE stocks because residual standard deviations are lower, averaging about 60% those of AMEX stocks.

---

[12]Average annual NYSE volume per security over the sample period was several times greater than for AMEX stocks; for a discussion of the relationship between volume and trading frequency, see Scholes and Williams (1977, p. 319).

[13]Similar results using an alternative proxy for trading volume are reported by Dyckman, Philbrick and Stephan (1984).

6. Estimating the variance of the mean excess return

6.1. Time-series dependence and non-synchronous trading

Although the simulations in table 6 indicate that biases in β due to non-synchronous trading did not affect tests for abnormal performance, non-synchronous trading could also induce serial correlation in the excess return measures. For hypothesis tests over intervals of more than a day, the failure to take into account autocorrelation in estimating the variance of the cumulative mean excess return could result in misspecification. Since the test statistics reported thus far did not account for autocorrelation yet seemed generally well-specified for the 11-day period examined in tables 4 and 5, it appears that autocorrelation plays a minor role.

Table 7 shows the time-series properties of the various excess return measures. While the magnitudes in the table are not obviously large, they are highly statistically significant. For example, the first three estimated autocorrelations for mean market model excess returns in samples of size 50 are -0.101, -0.037, and -0.030, with t-values (assuming independence across the 250 samples) of -20.1, -8.5, and -7.2, respectively. Furthermore, the Scholes–Williams and Dimson procedures do not purge autocorrelation from the excess return measures. Although not reported in the table, results almost identical to these for the market model were found for such procedures.

Table 7

Time-series properties of estimation period excess returns and mean excess returns. 250 samples of 50 securities; time period: 1962–1979; non-clustering case. Each number is the mean value of the estimated autocorrelation coefficient. The top number is based on the 12,500 individual time-series of security excess returns; the bottom number is for the 250 time-series of mean excess returns, where for each sample the mean excess return is the simple average of the excess returns for the 50 individual securities.

Method	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\rho}_3$
Mean adjusted returns			
Individual securities	-0.006	-0.014	-0.014
Samples of 50	-0.060	-0.027	-0.021
Market adjusted returns			
Individual securities	-0.032	-0.025	-0.024
Samples of 50	-0.093	-0.036	-0.031
OLS market model			
Individual securities	-0.044	-0.025	-0.026
Samples of 50	-0.101	-0.037	-0.030
Mean autocorrelations, OLS market model residuals for individual securities, by exchange-listing			
New York Stock Exchange	-0.027	-0.020	-0.025
American Stock Exchange	-0.071	-0.031	-0.027

6.1.1. Cases where autocorrelation adjustments are appropriate

To examine conditions where explicit recognition of autocorrelation is useful, all of the paper's previous experiments were repeated with a simple autocorrelation adjustment discussed in the appendix. No dramatic changes occurred. Thus, the benefits from autocorrelation adjustments appear to be limited. However, there were two instances of small but notable improvement in test statistic specification, and no cases where specification was adversely affected.

First, recall from table 5 that with clustering, under the null hypothesis the Mean Adjusted Returns method had a rejection rate of 13.6% in tests over the $(-5, +5)$ period. However, clustering induces positive serial correlation in the time-series of average Mean Adjusted Returns; the average estimated first-order autocorrelation coefficient based on samples of 50 is 0.25. With autocorrelation adjustments, the rejection rate under the null hypothesis was reduced to 8.8%; although unit normality of the test statistics was generally still rejected, the standard deviation of the test statistics was reduced from 1.54 to 1.29.

Second, note from table 7 that AMEX stocks have an average first-order autocorrelation for market model excess returns of -0.071, compared to -0.027 for NYSE stocks. For AMEX stocks, the day 0 tests reported in table 6 for the market model were carried out over the $(-5, +5)$ interval. Without adjustment for autocorrelation, the test statistics had a standard deviation of 0.76 and the goodness of fit tests rejected unit normality; with an adjustment, the standard deviation of the test statistics rose to 0.86 and departures from normality based on the goodness of fit tests disappeared. Similar results applied to the Scholes–Williams and Dimson procedures. However, no such improvements could be found for NYSE stocks.

6.2. Accounting for cross-sectional dependence: Advantages and disadvantages

The simulations thus far have estimated the variance of the mean excess return from the time-series of estimation period mean excess returns. This procedure takes into account any cross-sectional dependence in the security-specific excess returns. When there is positive cross-sectional dependence, failure to make such an adjustment results in a systematic underestimation of the variance of the mean excess return, implying too many rejections of the null hypothesis, both when it is true and when abnormal performance is present. [See Brown and Warner (1980), Beaver (1981), and Dent and Collins (1981).]

However, adjustment for cross-sectional dependence is not always necessary for reasonable test statistic specification. If the degree of dependence is small, as in studies where event dates are not clustered, ignoring the dependence induces little bias in variance estimates. Furthermore, dependence adjustment

Table 8

The effect of procedures assuming cross-sectional independence. Percentage of 250 samples where the null hypothesis is rejected.[a]

	Actual level of abnormal performance at day '0'			Mean test statistic	
	0	0.005	0.01	0% abnormal performance	0.5% abnormal performance
Panel A: Non-clustering					
Mean adjusted returns					
Dependence adjustment	6.4%	25.2%	75.6%	−0.01	1.13
Assume independence	6.0	48.0	96.0	−0.00	1.69
Market model					
Dependence adjustment	4.4	27.2	80.4	−0.02	1.17
Assume independence	6.4	53.2	97.6	0.00	1.79
Panel B: Clustering					
Mean adjusted returns					
Dependence adjustment	7.2	19.6	39.6	0.17	0.82
Assume independence	25.2	56.4	84.4	0.35	2.05
Market model					
Dependence adjustment	8.0	39.2	84.4	0.18	1.43
Assume independence	8.0	61.2	96.0	0.20	1.98

[a]Sample size = 50; one-tailed test, α = 0.05; randomly selected securities and event dates 1962–1979.

can actually be harmful compared to procedures which assume independence. Even if the independence assumption is only an approximation, explicit use of this information can increase the efficiency of the variance estimator. By permitting more precise estimation of the variance used in the test statistic, the independence assumption can make it easier to detect abnormal performance when it is present.

Table 8 examines the specification and power of event study procedures assuming cross-sectional independence. As discussed in the appendix, the test statistic is assumed unit normal under the null hypothesis.[14] To save space, results are reported only for Mean Adjusted Returns and the market model.

6.2.1. Comparison of alternative estimators: Non-clustering and clustering results

Panel A of table 8 indicates that when there is no clustering of event dates, the gains from procedures assuming independence are substantial.[15] With no

---

[14]A similar test statistic is used in some daily data event studies [e.g., Dodd and Warner (1983) and Larcker (1983)]. Following the typical procedure, each excess return is standardized by its estimated standard deviation. This explicitly accounts for heteroscedasticity in the excess returns, and can also increase the power of the tests.

[15]Increases in power from procedures assuming independence were not detected with monthly data [Brown and Warner (1980, table 6)]. However, in our simulations with daily data, the number of observations in the estimation period is roughly three times greater; there are also differences in the specific variance estimators.

abnormal performance, rejection rates are approximately equal to the signifi-
cance level of the tests, suggesting no detectable bias in variance estimation.
With 0.5% abnormal performance, the rejection rate with the market model
assuming independence is 53.2%, almost double the figure of 27.2% reported
earlier with dependence adjustment, and similar results apply for Mean
Adjusted Returns.

From panel B, the gains from procedures assuming independence can apply
even when there is clustering, and all securities of a given sample have the same
event date. For the market model, the rejection frequency with clustering and
no abnormal performance is 8.0% with both independence and dependence
adjustment. Furthermore, the distribution of market model test statistics for
both independence and dependence adjustment are similar and approximately
unit normal under the null hypothesis; the standard deviations of the 250 test
statistics are 1.02 and 1.01, respectively. The rejection frequencies with 0.5%
abnormal performance are 61.2% under independence and 39.2% with depen-
dence adjustment, indicating substantial power increases from procedures
assuming independence.

However, while extraction of the market factor via the market model
appears to be a sufficient adjustment for dependence, this result is for ran-
domly selected securities. If instead the securities came from the same industry
group, with clustering there could be a higher degree of cross-sectional depen-
dence in market model excess returns, and measurable misspecification [Dent
and Collins (1981)]. The consequences of ignoring extreme cross-sectional
dependence are indicated in panel B, where, for Mean Adjusted Returns, the
rejection rate with no abnormal performance is 25.2% assuming independence,
compared to 7.2% with dependence adjustment;[16] as with monthly data
[Brown and Warner (1980, table 6)], the Mean Adjusted Returns methodology
with dependence adjustment, while well-specified, is not very powerful with
clustering.

6.3. Variance increases during the event period

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

---

[16] Misspecification when independence is assumed with clustering was also found under proce-
dures similar to those of Schipper and Thompson (1983). Using a statistic which constrained the
abnormal performance to be the same across sample securities [Schipper and Thompson (1983, pp.
198–199, equations 4 and 5)], the rejection frequency in the absence of abnormal performance and
assuming independence was 11.2%. The corresponding figure for non-clustering (and assuming
independence) was 6.8%; for non-clustering the power of the tests was similar to that reported in
panel A of table 8 for the market model assuming independence. For related evidence, see Collins
and Dent (1984) and Malatesta (1984).

Table 9

The effect of variance increases during the event period. For each security, the day '0' variance is doubled (see text). Percentage of 250 samples where the null hypothesis is rejected; market model. H_0: mean abnormal performance at day '0' = 0.[a]

Variance estimation method		Actual level of abnormal performance at day 0	
		0	0.01
Use time-series	Variance increase	12.0%	70.4%
data from	No increase	4.4	80.4
estimation period			
Use cross-sectional	Variance increase	4.0	48.8
data from event	No increase	3.6	80.4
period			

[a] Sample size = 50; one-tailed test, $\alpha = 0.05$; randomly selected securities and event dates 1962–1979.

almost two. Although detailed study of variance increases is beyond the scope of this paper, it is useful to briefly outline several implications for event study methodologies.

6.3.1. Misspecification using time-series procedures

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. Table 9 uses simulation procedures to illustrate the magnitude of the misspecification.[17] Each security's day 0 return, $R_{i,0}$, is transformed to double the variance but leave the expected return unchanged,

$$R'_{i,0} = R_{i,0} + (R_{i,-6} - \overline{R}_i), \tag{9}$$

where $R'_{i,0}$ is the transformed return, $R_{i,-6}$ is the security's return at day -6, and $\overline{R}_i$ is the security's average daily return in the estimation period; the use of the return from day -6 is arbitrary. This procedure is equivalent to simulating a situation where the abnormal performance differs across sample securities but is, on average, zero; based on one cross-section of day 0 returns, such a situation cannot be distinguished from a variance increase. From table 9, doubling the variance results in a rejection rate under the null hypothesis of

---

[17] To save space, results are reported only for the market model. Note that results similar to those reported here can be obtained analytically for time-series procedures. Use of simulation procedures with actual return data will allow examination of the specification of the cross-sectional estimators discussed later.

12.0%, almost three times the figure of 4.4% obtained with no variance increase.

6.3.2. Cross-sectional procedures

The variance of the mean excess return is sometimes estimated using only the cross-section of event period excess returns. Typically, the day 0 excess returns are assumed independent and identically distributed, and the variance of the mean excess return is estimated by the cross-sectional variance [e.g., Penman (1982, p. 482)].[18] Cross-sectional estimates also permit construction of a time-series of estimated variances, one for each day around the event, thus allowing detection of event period variance increases [e.g., Mikkelson (1981, p. 257)]. Other procedures for variance shift detection are also used [e.g., Beaver (1968)].

From table 9, a cross-sectional variance estimate can lead to well-specified tests, both with and without a doubling of the variance; the rejection rates under the null hypothesis are 4.0% and 3.6%, respectively. Thus, such procedures can provide a useful check on the robustness of conclusions about abnormal performance.[19] However, cross-sectional procedures have limitations. For example, if the variance shift *differs* across sample securities, the test statistic is likely to be misspecified because the assumption of identically distributed excess returns is violated. In addition, if there is no variance increase, the cross-sectional procedures will not be very powerful because they ignore estimation period data; with 1% abnormal performance at day 0 and no variance increase, the rejection rate using the cross-sectional procedure is only 80.4%, compared to 97.6% in table 8 with time-series procedures assuming independence.

6.3.3. Sample partitioning

A direct way of addressing variance increases is to partition the sample based on an economic model of the effects of the event, such as whether the event is 'good news' or 'bad news' [e.g., Ball and Brown (1968)]. Such procedures can reduce the (conditional) return variances of securities in each subsample. Thus, they can reduce the degree of misspecification in using

---

[18]The test statistic is the ratio of the mean excess return to the cross-sectional standard error, and is approximately unit normal under the null hypothesis.

[19]Similarly, the non-parametric sign and Wilcoxon signed-rank tests were also found to be unaffected by variance shifts. However, because daily excess returns are skewed to the right, these tests are badly specified if the expected proportion of positive excess returns under the null hypothesis is assumed to be 0.5. The degree of misspecification is similar to that with monthly data; some reduction in misspecification was obtained with median regression based procedures [see Brown and Warner (1980, p. 222)].

standard time-series estimation procedures to test the significance of subsample mean excess returns. Although we leave to future work an examination of such procedures, they also seem valuable because they can increase the power of event study methodologies by reducing the unexplained component of returns.

7. Summary and conclusions

This paper examines how the particular characteristics of daily stock return data affect event study methodologies. Using simulation procedures with actual daily data, the paper investigates the impact of a number of potential problems of concern in the literature. These include (1) non-normality of returns and excess returns, (2) bias in OLS estimates of market model parameters in the presence of non-synchronous trading, and (3) estimation of the variance to be used in hypothesis tests concerning the mean excess return, and specifically the issues of autocorrelation in daily excess returns and of variance increases on the days around an event. In addition, the effect of cross-sectional dependence of excess returns on variance estimation, which is an issue even with monthly data, is also investigated.

The results from simulations with daily data generally reinforce the conclusions of our previous work with monthly data: methodologies based on the OLS market model and using standard parametric tests are well-specified under a variety of conditions. Although explicit recognition of the characteristics of daily data can sometimes be advantageous, for example in cases involving variance increases or unusually high autocorrelation, the characteristics of daily data generally present few difficulties in the context of event study methodologies. Furthermore, some of the paper's results indicate a striking similarity between the empirical power of the event study procedures and the theoretical power implied by a few simple assumptions and 'back of the envelope' calculations. This reinforces the view that the use of daily data is straightforward. The paper's findings are discussed in more detail below.

7.1. Non-normality and the properties of tests

The non-normality of daily returns has no obvious impact on event study methodologies. Although daily excess returns are also highly non-normal, there is evidence that the mean excess return in a cross-section of securities converges to normality as the number of sample securities increases. Standard parametric tests for significance of the mean excess return are well-specified. In samples of only 5 securities, and even when event days are clustered, the tests typically have the appropriate probability of Type I error.

As in the case of monthly data, the conclusion that the methodologies are well-specified applies to excess returns measured in a variety of ways, including Market Adjusted Returns and the OLS market model. With daily data, these

two methodologies have similar power, and, as expected, the power of each is much greater with daily than with monthly data. Market Adjusted Returns and the OLS market model also outperform a simpler Mean Adjusted Returns procedure, which has low power in cases involving event-date clustering. In addition, exchange-listing is an important correlate of the power of the various tests, with samples of NYSE securities exhibiting dramatically higher power than AMEX securities.

7.2. Alternative procedures for market model parameter estimation

Procedures other than OLS for estimating the market model in the presence of non-synchronous trading convey no clear-cut benefit in detecting abnormal performance. Methodologies based on the procedures suggested by Scholes and Williams and of Dimson do seem to reduce biases in OLS estimates of β. However, the specification and power of the actual tests for abnormal performance is similar to that obtained with the OLS market model, and this conclusion applies to samples having trading frequencies systematically different from average.

7.3. Variance estimation: Autocorrelation, cross-sectional dependence, and variance increases

While non-normality and biases in estimating the market model are unimportant in tests for abnormal performance, the choice of variance estimator to be used in hypothesis tests is of some concern, affecting both the specification and power of the tests. For hypothesis tests over multi-day intervals, there is evidence that the specification of the test statistic is improved by using simple procedures to adjust the estimated variance to reflect autocorrelation in the time-series of mean daily excess returns. However, the improvements are small, and only apply in special cases, for example event studies concentrating on AMEX firms. Non-synchronous trading, which can induce the autocorrelation, appears to have a detectable but limited impact on the choice of appropriate methodology.

When the implications of adjusting variance estimates to account for dependence in the cross-section of excess returns are studied, only in special cases is such adjustment necessary to prevent misspecification. Moreover, there is a potentially large cost. In results reported in the paper, tests which assume non-zero cross-sectional dependence are only about half as powerful and usually no better specified than those employed assuming independence.

Finally, we illustrated how variance increases can cause hypothesis tests using standard event study procedures to become misspecified. Several procedures to deal with the possibility of variance increases were outlined. However,

further research is necessary to fully understand the properties of alternative procedures for measuring abnormal performance in such situations.

Appendix: Hypothesis tests using daily data

This appendix gives additional details of the procedures used in simulating event study test methodologies.

A.1. Excess return measures

A.1.1. Scholes–Williams procedure

$$A_{i,t} = R_{i,t} - \hat{\alpha}'_i - \hat{\beta}'_i R_{m,t},$$ (A.1)

where

$$\hat{\alpha}'_i = \frac{1}{237} \sum_{t=-243}^{t=-7} R_{i,t} - \hat{\beta}'_i \frac{1}{237} \sum_{t=-243}^{t=-7} R_{m,t},$$ (A.2)

$$\hat{\beta}'_i = (\hat{\beta}_i^- + \hat{\beta}_i + \hat{\beta}_1^+)/(1 + 2\hat{\rho}_m);$$ (A.3)

$\hat{\beta}_i^-$ and $\hat{\beta}_i^+$ are the OLS estimation period values of

$$\frac{\mathrm{cov}(R_{i,t}, R_{m,t-1})}{\sigma(R_{m,t})\sigma(R_{m,t-1})} \quad \text{and} \quad \frac{\mathrm{cov}(R_{i,t}, R_{m,t+1})}{\sigma(R_{m,t})\sigma(R_{m,t+1})},$$

respectively. $\hat{\rho}_m$ is the estimation period value for the first-order autocorrelation coefficient of the Equally Weighted Market Index [see Scholes and Williams (1977, p. 317, eqs. 13–15, 19, 20)].

A.1.2. Dimson aggregated coefficients method

$$A_{i,t} = R_{i,t} - \hat{\alpha}'' - \hat{\beta}'' R_{m,t},$$ (A.4)

where

$$\hat{\alpha}''_i = \frac{1}{233} \sum_{t=-241}^{t=-9} R_{i,t} - \hat{\beta}''_i \frac{1}{233} \sum_{t=-241}^{t=-9} R_{m,t},$$ (A.5)

$$\hat{\beta}''_i = \sum_{k=-3}^{+3} \hat{\beta}_{ik};$$ (A.6)

$\hat{\beta}_{ik}$ are multiple regression estimates from the estimation period [see Dimson (1979, p. 203, eq. 8)].[20]

A.2. Hypothesis tests assuming cross-sectional independence

Each excess return $A_{i,t}$ is first divided by its estimated standard deviation to yield a standardized excess return, $A'_{i,t}$:

$$A'_{i,t} = A_{i,t}/\hat{S}(A_{i,t}),\tag{A.7}$$

where

$$\hat{S}(A_{i,t}) = \sqrt{\left(\sum_{t=-244}^{t=-6}\left(A_{i,t} - A_i^*\right)^2\right)\Bigg/238},\tag{A.8}$$

$$A_i^* = \frac{1}{239}\sum_{t=-244}^{t=-6}A_{i,t}.\tag{A.9}$$

The test statistic for any given day (in this case $t = 0$) is given by

$$\left(\sum_{i=1}^{N_t}A'_{i,t}\right)\cdot(N_t)^{-\frac{1}{2}},\tag{A.10}$$

where N_t is the number of sample securities at day t. If the standardized excess returns are independent and identically distributed with finite variance, in the absence of abnormal performance the test statistic will be distributed unit normal for large N_t.[21]

A.3. Tests over multi-day intervals [22]

For tests over the $(-5, +5)$ interval, the test statistic is the ratio of the cumulative mean excess return to its estimated standard deviation, and is given

---

[20] For the Scholes–Williams and Dimson procedures the maximum number of estimation period excess returns is reduced from 239 to 237 and 233, respectively. Adjustments to all variance estimators in the paper are made to reflect this.

[21] As in the case of dependence adjustment, adjustments to eq. (A.8) to reflect the predictive nature of the excess returns are omitted. These adjustments had no detectable impact.

[22] For multi-day intervals we only discuss test statistics and autocorrelation adjustments for the case of cross-sectional dependence adjustment; the derivation under cross-sectional independence is similar. Multi-day simulations under cross-sectional independence did not change the paper's conclusions.

by

$$\sum_{t=-5}^{+5} \overline{A}_t \bigg/ \left(\sum_{t=-5}^{+5} \hat{S}^2(\overline{A}_t) \right)^{\frac{1}{2}}, \tag{A.11}$$

where the terms in the denominator are from eq. (7) in the text. The test statistic is assumed unit normal in the absence of abnormal performance.

A.3.1. Autocorrelation adjustments

Autocorrelation adjustments use alternative estimates of the standard deviation in the denominator of eq. (A.11). This involves estimation of Σ, the autocovariance matrix of the event period mean excess returns,

$$\Sigma = \begin{bmatrix} \sigma^2(\overline{A}_{-5}) & \mathrm{cov}(\overline{A}_{-5}, \overline{A}_{-4}) & \cdots & \mathrm{cov}(\overline{A}_{-5}, \overline{A}_{+5}) \\ \mathrm{cov}(\overline{A}_{-4}, \overline{A}_{-5}) & \sigma^2(\overline{A}_{-4}) & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathrm{cov}(\overline{A}_{+5}, \overline{A}_{-5}) & & & \sigma^2(\overline{A}_{+5}) \end{bmatrix}. \tag{A.12}$$

A.3.2. Parameter estimation

The estimated standard deviation of the mean daily excess return $\overline{A}_t$ is given in eq. (7) in the text. Estimates of the diagonal elements of (A.12) follow directly. Noting that each off-diagonal element $\mathrm{cov}(\overline{A}_t, \overline{A}_{t-l}) = \rho(\overline{A}_t, \overline{A}_{t-l})\sigma(\overline{A}_t)\sigma(\overline{A}_{t-l})$, sample estimates of each off-diagonal element in addition use estimates of the autocorrelation coefficients of mean excess returns.

Autocorrelation coefficients ρ_l are estimated from the estimation period $\overline{A}_t$. We estimate ρ_l for values of l up to three; for $l > 3$, we arbitrarily set ρ_l equal to zero. The estimation process begins with $l = 1$. The process continues if and only if ρ_l is statistically significant; if not, its value and ρ_l for higher values of l are set equal to zero.[23] The criteria for statistical significance is that

$$|\rho_l| > 2.0/\sqrt{239 - l}. \tag{A.13}$$

---

[23] Once the estimation process is completed, we also check that estimates of ρ_l imply that Σ is positive definite; if not, all values of ρ_l are also constrained to zero. See Box and Jenkins (1976, p. 29) for a discussion.

Truncation of the estimation procedure by lag three and use of a cut-off for statistical significance helps assure simple representation of the autocorrelation structure of excess returns. However, our investigation of autocorrelation is exploratory. Although several minor variations of our procedure yielded similar results to those reported in the paper, it is an open question whether explicit ARIMA modelling of time-series properties would lead to improved specification or power.

A.3.3. Test statistics

The sum of the estimated elements of Σ is the estimated variance of the cumulative mean excess return; its square root becomes the denominator of the test statistic in (A.11).

References

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

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.

Beaver, W.H., 1981, Econometric properties of alternative security return methods, Journal of Accounting Research 19, 163–184.

Billingsley, P., 1979, Probability and measure (Wiley, New York).

Blattberg, R. and N. Gonedes, 1974, A comparison of the stable and student distributions as statistical models for stock prices, Journal of Business 47, 244–280.

Box, G. and G.M. Jenkins, 1976, Time series analysis: Forecasting and control (Holden-Day, San Francisco, CA).

Brown, S. and J. Warner, 1980, Measuring security price performance, Journal of Financial Economics 8, 205–258.

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

Collins, D.W. and W.T. Dent, 1984, A comparison of alternative testing models used in capital market research, Journal of Accounting Research 22, 48–84.

Dann, L., 1981, Common stock repurchases: An analysis of returns to bondholders and stockholders, Journal of Financial Economics 9, 113–138.

DeAngelo, H. and E. Rice, 1983, Antitakeover amendments and stockholder wealth, Journal of Financial Economics 11, 329–359.

Dent, W.T. and D.W. Collins, 1981, Econometric testing procedures in market-based accounting and finance research, Working paper, Nov. (University of Iowa, Iowa City, IA).

Dimson, E., 1979, Risk measurement when shares are subject to infrequent trading, Journal of Financial Economics 7, 197–226.

Dodd, P. and J. Warner, 1983, On corporate governance: A study of proxy contests, Journal of Financial Economics 11, 401–438.

Dyckman, T., D. Philbrick and J. Stephan, 1984, A comparison of event study methodologies using daily stock returns: A simulation approach, Working paper, May (Cornell University, New York).

Fama, E.F., 1976, Foundations of finance (Basic Books, New York).

French, K., 1980, Stock returns and the weekend effect, Journal of Financial Economics 8, 59–69.

Fowler, D. and C.H. Rorke, 1983, Risk measurement when shares are subject to infrequent trading: Comment, Journal of Financial Economics 12, 279–283.

Gheyara, K. and J. Boatsman, 1980, Market reaction to the 1976 replacement cost disclosures, Journal of Accounting and Economics 2, 107–125.

Gibbons, M. and P. Hess, 1981, Day of the week effects and asset returns, Journal of Business 54, 579–596.

Hagerman, R., 1978, More evidence on the distribution of security returns, Journal of Finance 33, 1213–1221.

Holthausen, R., 1981, Evidence on the effect of bond covenants and management compensation contracts on the choice of accounting techniques: The case of the depreciation switchback, Journal of Accounting and Economics 3, 73–109.

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, June (New York University, New York).

Larcker, D., 1983, The association between performance plan adoption and corporate capital investment, Journal of Accounting and Economics 5, 3–30.

Leftwich, R., 1981, Evidence on the impact of mandatory changes in accounting principles on corporate loan agreements, Journal of Accounting and Economics 3, 3–36.

Malatesta, P., 1984, Measuring abnormal performance: The event parameter approach using joint generalized least squares, Working paper, May (University of Washington, Seattle, WA).

Masulis, R., 1980, The effects of capital structure change on security prices: A study of exchange offers, Journal of Financial Economics 8, 139–177.

McNichols, M. and J. Manegold, 1983, The effect of the information environment on the relationship between financial disclosure and security price variability, Journal of Accounting and Economics 5, 49–74.

Mikkelson, W., 1981, Convertible calls and security returns, Journal of Financial Economics 9, 237–264.

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.

Ruback, R., 1982, The effect of discretionary price control decisions on equity values, Journal of Financial Economics 10, 83–105.

Schipper, K. and R. Thompson, 1983, The impact of merger related regulations on the shareholders of acquiring firms, Journal of Accounting Research 21, 184–221.

Scholes, M. and J. Williams, 1977, Estimating betas from nonsynchronous data, Journal of Financial Economics 5, 309–328.