USING DAILY STOCK RETURNS:
THE CASE OF EVENT STUDIES

by
Jerold B. Warner*

and

Stephen J. Brown**

Working Paper Series Number
MERC 84-05

February 1983
Revised March 1984

University of Rochester
Graduate School of Management
Rochester, New York 14627

* Graduate School of Management, University of Rochester, Rochester, NY.
** Bell Communications Research, Murray Hill, NJ and Yale University, New Haven, CT.
ABSTRACT

This paper examines properties of daily stock returns and how the particular characteristics of these data can affect event study methodologies. We find no evidence that either nonnormality in the time-series of daily excess returns or bias in OLS estimates of Market Model parameters affect the specification or power of tests for abnormal performance. However, under plausible conditions, both autocorrelation in excess returns and changes in the variance of daily returns conditional on an event can affect the tests; simple procedures to deal with these issues are sometimes quite useful. We also show that taking into account dependence in the cross-section of the daily excess returns can be harmful, resulting in tests with relatively low power and which are no better specified than those which assume independence.
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), hereinafter 'BW') in which we investigated 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., BW (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 nonsynchronous trading. This complication is described as "especially severe" by Scholes and Williams (1977, p. 324).

This 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 that which we developed previously is applied to observed daily returns. Various event study
methodologies are simulated by repeated application of each methodology to samples which 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. Initial results are presented in Section 4, and extended in Sections 5 and 6. In Section 7, we 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.

Nonnormality

The daily stock return for an individual security exhibits substantial departures from normality which are not observed with monthly data; the evidence generally suggests that distributions of daily returns are fat-tailed relative to a normal (Fama (1976,
p. 21). As shown later, the same holds true for the daily excess return. 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 standard Central Limit Theorem are violated with these data.

Nonsynchronous Trading and Market Model Parameter Estimation

When the return on a security and the return on the market index are measured over different intervals, 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.
Variance Estimation

With both daily and monthly data, a topic of interest is estimation of the variance of the sample mean excess return in the context of tests for statistical significance. There are several variance estimation issues which this paper investigates.

The first issue is the time-series properties of daily data. As a consequence of nonsynchronous 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 will be 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., BW, 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) and Patell and Wolfson (1979)). How that possibility affects event study procedures will be illustrated.

1Other time-series properties, such as day of the week or weekend effects (e.g., French (1980), Gibbons and Hess (1981)), are not explicitly studied.
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 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 nonsynchronous 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 use data presumably from the true return-generating process and are a direct way to summarize how, in practice, 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.

3.2 Excess Return Measures

Define $A_{i,t}$ as the excess return for security $i$ at day $t$. For every security, excess returns for each day in the event period are estimated using a number of different procedures:

**Raw Returns:**

(1) $A_{i,t} = R_{i,t}$

where $R_{i,t}$ is the observed arithmetic return for security $i$ at day $t$.

**Mean Adjusted Returns:**

(2) $A_{i,t} = R_{i,t} - \bar{R}_i$

(3) $\bar{R}_i = \frac{1}{239} \sum_{t=-244}^{-6} R_{i,t}$

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.
R_i is the simple average of security i's daily returns in the (-244, -6) estimation period.¹

**Market Adjusted Returns:**

(4) \[ A_{i,t} = R_{i,t} - R_{m,t} \]

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

**OLS Market Model**

(5) \[ A_{i,t} = R_{i,t} - \hat{\alpha}_i - \hat{\beta}_i R_{m,t} \]

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 BW (pp. 207-209).² The Scholes/Williams and Dimson procedures for estimating market model parameters will be considered later.

¹Equation (3) 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.

²Each of these methods is consistent with the following model of the return-generating process:

\[ \tilde{R}_{it} = \alpha_{it} + \beta_{it} R_{mt} + \delta_{it} P_{it} + \epsilon_{it} \]

where \( \tilde{R}_{it} \) is the return on security i at day t, \( R_{mt} \) is the return on a market index, \( \delta_{it} \) is a dummy variable equal to one if for firm i an event occurs at time t and equal to zero otherwise, \( P_{it} \) is the impact of the event on the security's return, \( \epsilon_{it} \) is a disturbance term with unconditional mean of zero, and \( \alpha_{it} \) and \( \beta_{it} \) are constants (see Schipper and Thompson (1983), Stillman (1983), and Thompson (1983). With Raw Returns, \( \alpha = \beta = 0 \). With Mean Adjusted Returns, \( \alpha \) is an estimate of the security's mean return and \( \beta = 0 \). With Market Adjusted Returns, \( \alpha = 0 \) and \( \beta = 1 \). With the Market Model, \( \alpha \) and \( \beta \) are OLS estimates.
3.3 Test Statistics Under the Null Hypothesis

Given the excess returns based on each method, the statistical significance of the event period excess return 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.\(^1\) This hypothesis is important in many event studies, where the main concern is the average affect of an event on returns to shareholders. Other hypotheses, for example concerning changes in the variance of the excess returns, are sometimes important as well. The effect of such alternative experimental situations is discussed in Section 6.3.

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:

\[
\frac{\bar{A}_t}{S(\bar{A}_t)}
\]

\[
\bar{A}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} A_{i,t}
\]

where

\[
\hat{S}(\bar{A}_t) = \sqrt{\frac{\sum_{t=-6}^{t=-244} (A_t - \bar{A})^2}{238}}
\]

\[
\bar{A} = \frac{1}{239} \sum_{t=-244}^{t=-6} \bar{A}_t
\]

\(^1\)Later, tests over the entire (-5, +5) interval are examined. Construction of the tests is discussed in the Appendix.
$N_t$ is the number of sample securities whose excess returns are available at day $t$.\footnote{For Mean Adjusted Returns and the Market Model, the denominator of equation (6) 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.}

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 BW. A constant is added to the observed day 0 return for each security. For example, to simulate 1% abnormal performance, .01 is added. In the initial simulations, the level of abnormal performance is the same for all sample securities.

With constant samplewide abnormal performance, the procedure for introducing abnormal performance is equivalent to taking the test statistic in (6) which obtains 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 from 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 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 excess returns.

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 .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 .99 fractile of the respective distribution which would obtain under normality. These various figures do not change markedly for the measures of excess returns. For example, with the market model,

1Autocorrelations for excess returns are examined in Section 6.1.
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 .0267; the mean standard deviation of excess returns from the Market Model is .0258, only slightly lower, and the mean Market Model $R^2$ is only .10. These figures suggest that the measures of excess returns will exhibit similar ability to detect abnormal performance when it is present.

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 figures applied when the properties of the mean excess returns in the estimation period for each sample were 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 .08 for Raw Returns to .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 samples smaller than 50. 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 will be 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 -.0001, with a t-statistic of -0.46. For Raw Returns, however, the mean performance measure has a value of .0006, with a t-statistic of 2.33. This merely indicates that average returns are positive. As one would expect and as we shall see, the Raw Returns methodology will tend to indicate positive abnormal performance when none is present; the misspecification is more readily apparent the longer the event period over which excess returns are examined.

4.2 Properties of the Test Statistics

Figure 1 compares the cumulative distribution of the 250 test statistics when no abnormal performance is introduced with cumulative values from a unit normal; the sample size is 50, using the Market Model. 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-squared 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, as well as leptokurtic; for all methodologies, the studentized range of the test statistics is in excess of six. In addition, for Raw Returns the average test statistic is significantly positive.

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 .05 significance level. For illustrative purposes, rejection frequencies for this significance level will be used throughout the paper.\(^1\)

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 (BW, p. 216); although the highest rejection rate is for Raw Returns, based on these figures the bias in that methodology towards too many rejections is not large. 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 Raw Returns and the Market Model. These rejection frequencies indicate little

\(^1\)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.
difference in the power of alternative procedures. Moreover, these rejection frequencies are roughly three times those reported by BW (Table 3) for monthly data. Thus, as was emphasized in BW, there are 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, Figure 2 compares the empirical and theoretical power for the Market Model.\textsuperscript{1} For a variety of abnormal performance levels, rejection frequencies are similar, and this also applied to test at the .1 and .01 significance levels. Since the test statistics in the experimental situation examined in Table 3 were approximately unit normal under the null hypothesis, such results should not be 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) and Dyckman, Philbrick, and Stephen (1983)), and is not highly sensitive to several changes in the experimental procedure.

\textsuperscript{1}For computational simplicity, the derivation of each power function in Figure 2 assumes that the estimated standard deviation of the mean excess return is the same for all samples and equal to the value of .0038 reported in Panel B of Table 1. The empirical power in Table 3 differs slightly from that in the figure because estimates of the standard deviation can differ across samples.
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.

Longer Event Periods

The test statistics also continue to be generally well-specified when the event period is longer than one day. Table 4 shows results for the case where, in samples of size 50,

1. the null hypothesis 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 interval (-5, +5), with each day having an equal probability of being selected.

Although Raw Returns rejects the null hypothesis in 10.0% of the samples, rejection rates under the null hypothesis for the other methodologies range from 2.8% to 4.0%, and various goodness of fit tests for the other methodologies failed to find marked evidence of test statistic misspecification; 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. For example, with 1% abnormal performance, the rejection frequency for Market Adjusted Returns is 13.6%, compared to the earlier figure of 79.6% reported in Table 3.
Clustering

In addition, the results on specification are not radically altered in experiments where there is clustering in event dates and hence nonindependence 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.\(^1\) 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, the 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. As in the nonclustering case, for Raw Returns there are too many rejections.

However, 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 nonclustering 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 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 methodologies are either poorly specified or inefficient.

\(^1\)The day is randomly selected for a given sample and selection of the day is carried out without replacement.
5. NONSYNCHRONOUS TRADING: ALTERNATIVE PROCEDURES FOR $\beta$ ESTIMATION

Nonsynchronous trading introduces potentially serious difficulties into empirical studies using daily stock returns (Scholes and Williams (1977) and Dimson (1979)). One reason is that in the presence of nonsynchronous trading, OLS estimates of market model $\beta$ are biased and inconsistent; the evidence is that shares traded relatively infrequently have downward biased $\beta$ estimates, while those traded relatively frequently have upward biased $\beta$ estimates. However, the results presented thus far in this paper indicate that the failure to take into account nonsynchronous trading in estimating Market Model coefficients does not result in misspecification of event study methodologies using the OLS market model.

Such results are consistent with the evidence presented by Scholes and Williams and Dimson that OLS estimates of $\beta$ are biased; even when such biases exist they do not necessarily imply misspecification when using OLS 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 $\beta$ is compensated for by a bias in $\alpha$. With stationarity, the event period excess returns for an individual security can be shown to have mean zero unconditional on the market return. Although the excess return conditional on the market is biased for an individual security, 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 nonclustering of event dates. Nevertheless, even when OLS is well specified, its use could still result in less precise (as opposed to biased) estimates of excess returns, thus yielding tests with relatively low power. Thus, it is of interest to investigate alternative procedures for market model parameter estimation.

**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.\(^1\) A representative set of results, for a sample size of 20, is presented in Table 6.\(^2\)

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 were similar for the three

\(^1\)These procedures are discussed in the Appendix. For a recent discussion of the Dimson procedure, see Fowler and Rorke (1983).

\(^2\)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.
procedures and reasonably close to unit normal under the null hypothesis.

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 do not have 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.¹ 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 \( \beta \) are systematically higher than the estimates using the Scholes/Williams or Dimson procedure; the average values are .96, .94, and .91, respectively. For AMEX stocks, the relative magnitudes of the \( \beta \) estimates are reversed, with the average estimates of \( \beta \) ranging from 1.05 with the OLS Market Model to 1.11 with the Dimson procedure. These figures indicate detectable biases in OLS estimates of \( \beta \) similar to those discussed by Scholes and Williams, and by Dimson; they also indicate a relation between true \( \beta \) and trading volume (Scholes and Williams (1977, p. 320).

¹Average 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).
However, there is no evidence that procedures other than OLS improve either the specification or the power of the tests. In the 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 .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.

6. ESTIMATING THE VARIANCE OF THE MEAN EXCESS RETURN

6.1 Time-Series Dependence and Nonsynchronous Trading

Although the simulations in Table 6 indicate that biases in $\beta$ due to nonsynchronous trading did not affect tests for abnormal performance, nonsynchronous 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. However, 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 would appear that autocorrelation plays a minor role.

1Similar results using an alternative proxy for trading volume are reported by Dyckman, Philbrick, and Stephen (1983).
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 -.101, -.037, and -.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.

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

\(^1\)For a sample where these values represent the true parameters, ignoring autocorrelation results in an 18% overestimate of the standard deviation of the 11-day cumulative mean excess return; see equations (A.11) and (A.12) of the Appendix.
of average Mean Adjusted Returns; the 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 instead 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, thus taking into account any cross-sectional dependence in the security-specific excess returns (see BW (1980), Beaver (1981), and Dent and Collins (1981)). However, in cases where the degree of cross-sectional dependence is small, this widely used procedure is unnecessary to assure reasonable test statistic specification. Furthermore, dependence adjustment
can result in tests with relatively low power compared to those where the variance estimator for the mean excess return explicitly assumes cross-sectional independence. If the independence assumption is correct, use of such information increases the efficiency of the variance estimator, and more precise estimation of the variance should make it easier to detect abnormal performance when it is present.

Table 8 reports simulation results when the variance estimator assumes cross-sectional independence. As discussed in the Appendix, the test statistic is assumed unit normal under the null hypothesis.\(^1\) To save space, results are reported only for Mean Adjusted Returns and the Market Model.

**Comparison of Alternative Estimators: Nonclustering and Clustering Results**

Panel A indicates that when there is no clustering of event dates, the gains from assuming independence are substantial.\(^2\) With no abnormal performance, rejection rates are approximately equal to the significance level of the tests. With .5% abnormal performance, the rejection rate with the Market Model assuming independence is 53.2%, almost double the figure of 27.2% reported

---

\(^1\)A similar test statistic is used in some daily data event studies (e.g., Dodd and Warner (1983) and Larcker (1983)). 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.

\(^2\)Increases in power from assuming independence were not detected with monthly data (BW, 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 estimators.
earlier with dependence adjustment, and similar results apply for Mean Adjusted Returns.

From Panel B, the gains from 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; the rejection frequencies with .5% abnormal performance are 61.2% and 39.2%, respectively. However, while extraction of the market factor via the Market Model appears to be a sufficient adjustment for dependence, this result is for randomly selected securities; if instead the securities came from the same industry group, with clustering there could be a higher degree of cross-sectional dependence in Market Model excess returns, and measurable misspecification (Dent and Collins (1981)). The consequences of 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;¹ as with monthly data (BW, Table 6), the Mean Adjusted Returns methodology with dependence

¹Misspecification when independence is assumed with clustering was also found for procedures similar to those of Schipper and Thompson (1981). Using a statistic which constrained the abnormal performance to be the same across sample securities (Schipper and Thompson (1981, pp. 17-18, equations 4 and 5)), the rejection frequency in the absence of abnormal performance was 11.2%. The corresponding figure for nonclustering (and assuming independence) was 6.8%; for nonclustering the power of the tests was similar to that reported in Panel A for the Market Model assuming independence.
adjustment, while well-specified, is not very powerful with clustering.

6.3 Variance Increases During the Event Period

There is evidence that the variance of a security's return increases for the days around some types of events (Beaver (1968), Patell and Wolfson (1979), and Kalay and Lowenstein (1983)). Christie (1983) argues that the variance of daily excess returns in some event studies increases by a factor in excess of four. 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.

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. Each security's day 0 return, $R_{i,0}^*$, is transformed to double the variance but leave the expected return unchanged:

$$R_{i,0}^* = R_{i,0} + (R_{i,-6} - \bar{R}_i),$$

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.
where $R_{i0}'$ is the transformed return, $R_{i,-6}$ is the security's return at day -6, and $\bar{R}_i$ 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 12.0%, almost three times the figure of 4.4% obtained with no variance increase.

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)). The 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 (Mikkelson, 1981, p. 257); other procedures for variance shift detection can also be 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.

---

1 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.
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 the conclusions about abnormal performance. ¹ 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 thus 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.

**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 standard time-series estimation procedures to test the significance of subsample mean excess returns. Although we leave

¹Similarly, the nonparametric 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, the test is 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 (BW, p. 222).
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 examined how the particular characteristics of daily stock return data affect event study methodologies. Using simulation procedures with actual daily data, the paper investigated the impact of a number of potential problems. These include (1) nonnormality of returns and excess returns, (2) bias in OLS estimates of Market Model parameters in the presence of nonsynchronous trading, and (3) estimation of the variance to be used in hypothesis tests, 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, was also investigated.

The results from simulations with daily data generally reinforce the conclusions we drew in previous work with monthly data: methodologies based on the OLS Market Model and using parametric tests are typically well-specified under a variety of conditions. However, explicit recognition of the characteristics of daily data can sometimes be advantageous, for example, in cases involving autocorrelation and variance increases. The major findings are detailed below.

**Nonnormality and the Properties of Tests**

The nonnormality of daily returns had no obvious impact on event study methodologies. Although daily excess returns are
highly nonnormal, there was evidence that the mean excess return in a cross-section of securities converged to normality as the number of sample securities increased. Standard parametric tests for significance of the mean excess return at day 0 were well-specified. In samples of only 5 securities, and even when event days were clustered, the tests typically had the appropriate probability of Type I error.

As in the case of monthly data, the conclusion that the methodologies are well-specified applied 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 outperformed simpler procedures: the Raw Returns methodology was misspecified, and Mean Adjusted Returns had low power in cases involving event-date clustering. In addition, exchange-listing was an important correlate of the power of the various tests, with samples of NYSE securities exhibiting dramatically higher power than AMEX securities.

Alternative Procedures for Market Model Parameter Estimation

Procedures other than OLS for estimating the Market Model in the presence of nonsynchronous trading conveyed no clear-cut benefit in detecting abnormal performance. Methodologies based on the procedures suggested by Scholes and Williams and of Dimson did reduce biases in OLS estimates of $\beta$. However, the specification and power of the actual tests for abnormal performance was similar to that obtained with the OLS Market
Model, and this conclusion applied to samples having trading frequencies systematically different from average.

**Variance Estimation: Autocorrelation, Cross-Sectional Dependence, and Variance Increases**

While nonnormality and biases in estimating the Market Model were unimportant in tests for abnormal performance, the choice of variance estimator to be used in hypothesis tests was of some concern, affecting both the specification and power of the tests. For hypothesis tests involving periods as short as about ten days, there was evidence that the specification of the test statistic was 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 were small, and only applied in special cases, for example, event studies concentrating on AMEX firms. Nonsynchronous 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 were studied, only in special cases was such adjustment necessary to prevent misspecification. Moreover, there was a potentially large cost. In results reported in the paper, tests which assume non-zero cross-sectional dependence were only about half as powerful and usually no better specified than those which assume 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.
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 based on its estimation period excess returns are made. For each parameter, the table reports the mean of the 12,500 estimates.

Maximum Number of Observations Per Security = 239
CRSP Equal Weighted Index

<table>
<thead>
<tr>
<th>Performance Measure</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Studentized Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw Returns</td>
<td>.0006</td>
<td>.0267</td>
<td>.99</td>
<td>6.87</td>
<td>7.59</td>
</tr>
<tr>
<td>Mean Adjusted Returns</td>
<td>.0000</td>
<td>.0267</td>
<td>.99</td>
<td>6.87</td>
<td>7.59</td>
</tr>
<tr>
<td>Market Adjusted Returns</td>
<td>.0000</td>
<td>.0258</td>
<td>.97</td>
<td>6.66</td>
<td>7.49</td>
</tr>
<tr>
<td>Market Model</td>
<td>.0000</td>
<td>.0253</td>
<td>1.01</td>
<td>6.80</td>
<td>7.54</td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Method</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Studentized Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Raw Returns</td>
<td>0.0005</td>
<td>0.0127</td>
<td>0.39</td>
<td>4.68</td>
<td>7.13</td>
</tr>
<tr>
<td></td>
<td>Mean Adjusted Returns</td>
<td>-0.0001</td>
<td>0.0127</td>
<td>0.42</td>
<td>4.71</td>
<td>7.05</td>
</tr>
<tr>
<td></td>
<td>Market Adjusted Returns</td>
<td>0.0001</td>
<td>0.0122</td>
<td>0.40</td>
<td>4.91</td>
<td>7.42</td>
</tr>
<tr>
<td></td>
<td>Market Model</td>
<td>0.0001</td>
<td>0.0121</td>
<td>0.42</td>
<td>5.04</td>
<td>7.64</td>
</tr>
<tr>
<td>20</td>
<td>Raw Returns</td>
<td>0.0007</td>
<td>0.0062</td>
<td>0.00</td>
<td>3.30</td>
<td>6.42</td>
</tr>
<tr>
<td></td>
<td>Mean Adjusted Returns</td>
<td>0.0001</td>
<td>0.0062</td>
<td>0.00</td>
<td>3.33</td>
<td>6.41</td>
</tr>
<tr>
<td></td>
<td>Market Adjusted Returns</td>
<td>0.0001</td>
<td>0.0060</td>
<td>0.02</td>
<td>3.28</td>
<td>6.40</td>
</tr>
<tr>
<td></td>
<td>Market Model</td>
<td>0.0001</td>
<td>0.0059</td>
<td>0.03</td>
<td>3.43</td>
<td>6.47</td>
</tr>
<tr>
<td>50</td>
<td>Raw Returns</td>
<td>0.0006</td>
<td>0.0040</td>
<td>0.08</td>
<td>3.17</td>
<td>5.69</td>
</tr>
<tr>
<td></td>
<td>Mean Adjusted Returns</td>
<td>0.0000</td>
<td>0.0040</td>
<td>0.08</td>
<td>3.19</td>
<td>5.73</td>
</tr>
<tr>
<td></td>
<td>Market Adjusted Returns</td>
<td>-0.0001</td>
<td>0.0039</td>
<td>0.10</td>
<td>3.23</td>
<td>5.81</td>
</tr>
<tr>
<td></td>
<td>Market Model</td>
<td>-0.0001</td>
<td>0.0038</td>
<td>0.10</td>
<td>3.10</td>
<td>5.59</td>
</tr>
</tbody>
</table>

Upper Percentage Points, Samples of 250
Drawn From a Normal Population

<table>
<thead>
<tr>
<th>Variable</th>
<th>.95</th>
<th>.99</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>.063</td>
<td>.129</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.52</td>
<td>3.87</td>
</tr>
<tr>
<td>Studentized Range (N = 200)</td>
<td>6.15</td>
<td>6.85</td>
</tr>
</tbody>
</table>
### 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

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean</th>
<th>Standard deviation for Mean</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Studentized Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw Returns</td>
<td>0.133</td>
<td>0.92</td>
<td>2.26</td>
<td>0.091</td>
<td>3.55</td>
</tr>
<tr>
<td>Mean Adjusted Returns</td>
<td>-0.014</td>
<td>0.92</td>
<td>-0.24</td>
<td>0.094</td>
<td>3.61</td>
</tr>
<tr>
<td>Market Adjusted</td>
<td>-0.019</td>
<td>0.91</td>
<td>-0.33</td>
<td>0.073</td>
<td>3.52</td>
</tr>
<tr>
<td>Returns</td>
<td>-0.028</td>
<td>0.91</td>
<td>-0.49</td>
<td>0.084</td>
<td>3.42</td>
</tr>
</tbody>
</table>

**X^2 Statistic**  
(9 intervals concentrating in the tail regions)  
(20 equally spaced intervals)

<table>
<thead>
<tr>
<th>X^2 Statistic</th>
<th>X^2 Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>(9 intervals)</td>
<td>(20 equally</td>
</tr>
<tr>
<td></td>
<td>spaced intervals)</td>
</tr>
<tr>
<td><strong>Upper Percentage</strong></td>
<td><strong>Upper</strong></td>
</tr>
<tr>
<td><strong>Points</strong></td>
<td><strong>Percentage</strong></td>
</tr>
<tr>
<td><strong>X^2 (8)</strong></td>
<td>15.5</td>
</tr>
<tr>
<td><strong>X^2 (19)</strong></td>
<td>30.1</td>
</tr>
</tbody>
</table>

---

1. The intervals are 0 - .01, .01 - .02, .02 - .05, .05 - .1, .1 - .9, .9 - .95, .95 - .98, .98 - .99, .99 - 1.0.

---

**Upper Percentage Points**

\[ X^2 \text{ (8)} \quad \frac{.95}{15.5} \quad \frac{.99}{20.1} \]

\[ X^2 \text{ (19)} \quad 30.1 \quad 36.2 \]
Table 3

A COMPARISON OF ALTERNATIVE PERFORMANCE MEASURES
Percentage of 250 samples where the null hypothesis is rejected

H₀: mean abnormal performance at day "0" = 0.

Actual level of Abnormal Performance at day "0"

<table>
<thead>
<tr>
<th>Method</th>
<th>0</th>
<th>.005</th>
<th>.01</th>
<th>.02</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw Returns</td>
<td>6.4</td>
<td>30.8</td>
<td>80.4</td>
<td>99.6</td>
</tr>
<tr>
<td>Mean Adjusted Returns</td>
<td>6.4</td>
<td>25.2</td>
<td>75.6</td>
<td>99.6</td>
</tr>
<tr>
<td>Market Adjusted Returns</td>
<td>4.8</td>
<td>26.0</td>
<td>79.6</td>
<td>99.6</td>
</tr>
<tr>
<td>Market Model</td>
<td>4.4</td>
<td>27.2</td>
<td>80.4</td>
<td>99.6</td>
</tr>
</tbody>
</table>

Sample Size = 50 Securities
One-tailed test, α = .05
Randomly Selected Securities and Event Dates, 1962-1979
Table 4

ALTERNATIVE PERFORMANCE MEASURES 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 an 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.

<table>
<thead>
<tr>
<th>Method</th>
<th>Days in Event Period</th>
<th>0</th>
<th>.01</th>
<th>.02</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw Returns</td>
<td>11</td>
<td>10.0</td>
<td>28.8</td>
<td>54.4</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>6.4</td>
<td>80.4</td>
<td>99.6</td>
</tr>
<tr>
<td>Mean Adjusted Returns</td>
<td>11</td>
<td>4.0</td>
<td>13.6</td>
<td>37.6</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>6.4</td>
<td>75.6</td>
<td>99.6</td>
</tr>
<tr>
<td>Market Adjusted Returns</td>
<td>11</td>
<td>4.0</td>
<td>13.2</td>
<td>52.0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>4.8</td>
<td>79.6</td>
<td>99.6</td>
</tr>
<tr>
<td>Market Model</td>
<td>11</td>
<td>2.8</td>
<td>13.2</td>
<td>37.2</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>4.4</td>
<td>80.4</td>
<td>99.6</td>
</tr>
</tbody>
</table>

Sample Size = 50
One-tailed test, α = .05
Randomly Selected Securities and Event Dates, 1962–1979
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 nonclustering case reported in Table 4 are also shown.

H₀: Cumulative Mean Daily Abnormal Performance in the Interval (-5, +5) = 0.

<table>
<thead>
<tr>
<th>Method</th>
<th>Actual Level of Abnormal Performance in Interval (-5, +5)¹</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Raw Returns</td>
<td>Clustering</td>
</tr>
<tr>
<td></td>
<td>Nonclustering</td>
</tr>
<tr>
<td>Mean Adjusted Returns</td>
<td>Clustering</td>
</tr>
<tr>
<td></td>
<td>Nonclustering</td>
</tr>
<tr>
<td>Market Adjusted Returns</td>
<td>Clustering</td>
</tr>
<tr>
<td></td>
<td>Nonclustering</td>
</tr>
<tr>
<td>Market Model</td>
<td>Clustering</td>
</tr>
<tr>
<td></td>
<td>Nonclustering</td>
</tr>
</tbody>
</table>

¹For each security, abnormal performance is introduced for one day in the interval, with each day having an equal probability of selection.

Sample Size = 50
One-tailed test, α = .05
Randomly Selected Securities and Event Dates, 1962-1979
### Table 6
ALTERNATIVE PROCEDURES FOR MARKET MODEL ESTIMATION

Percentage of 250 Samples Where the Null Hypothesis is Rejected

\[ H_0: \text{Mean Abnormal Performance at Day "0" = 0.} \]

<table>
<thead>
<tr>
<th>Actual Level of Abnormal Performance at Day 0</th>
<th>Mean Value in Estimation Period (N = 5000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.01</td>
</tr>
</tbody>
</table>

#### Panel A:
Randomly Selected Securities

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean Abnormal Performance at Day 0</th>
<th>Mean Value in Estimation Period (N = 5000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS Market Model</td>
<td>2.8</td>
<td>46.8</td>
</tr>
<tr>
<td>Scholes/Williams Procedure</td>
<td>2.8</td>
<td>46.8</td>
</tr>
<tr>
<td>Dimson Aggregated Coefficients Method</td>
<td>2.4</td>
<td>47.2</td>
</tr>
</tbody>
</table>

#### Panel B:
Sample Formation by Exchange Listing

<table>
<thead>
<tr>
<th>Exchange Samples</th>
<th>Method</th>
<th>Mean Abnormal Performance at Day 0</th>
<th>Mean Value in Estimation Period (N = 5000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NYSE SAMPLES</td>
<td>OLS Market Model</td>
<td>5.2</td>
<td>64.4</td>
</tr>
<tr>
<td></td>
<td>Scholes/Williams Procedure</td>
<td>4.4</td>
<td>64.8</td>
</tr>
<tr>
<td></td>
<td>Dimson Aggregated Coefficients Method</td>
<td>4.4</td>
<td>65.2</td>
</tr>
<tr>
<td>AMEX SAMPLES</td>
<td>OLS Market Model</td>
<td>5.6</td>
<td>28.8</td>
</tr>
<tr>
<td></td>
<td>Scholes/Williams Procedure</td>
<td>5.2</td>
<td>31.2</td>
</tr>
<tr>
<td></td>
<td>Dimson Aggregated Coefficients Method</td>
<td>5.2</td>
<td>30.0</td>
</tr>
</tbody>
</table>

Sample Size = 20 Securities
One-tailed test, \( \alpha = .05 \)
Time Period: 1962-1979
Table 7

TIME-SERIES PROPERTIES OF ESTIMATION PERIOD EXCESS RETURNS AND MEAN EXCESS RETURNS
250 Samples of 50 Securities
Time Period: 1962-1979
Nonclustering Case

Each number is the mean value of the estimated autocorrelation coefficient. The top number is based on for 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.

<table>
<thead>
<tr>
<th>Method</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Adjusted Returns</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual Securities</td>
<td>-.006</td>
<td>-.014</td>
<td>-.014</td>
</tr>
<tr>
<td>Samples of 50</td>
<td>-.060</td>
<td>-.027</td>
<td>-.021</td>
</tr>
<tr>
<td>Market Adjusted Returns</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual Securities</td>
<td>-.032</td>
<td>-.025</td>
<td>-.024</td>
</tr>
<tr>
<td>Samples of 50</td>
<td>-.093</td>
<td>-.036</td>
<td>-.031</td>
</tr>
<tr>
<td>OLS Market Model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual Securities</td>
<td>-.044</td>
<td>-.025</td>
<td>-.026</td>
</tr>
<tr>
<td>Samples of 50</td>
<td>-.101</td>
<td>-.037</td>
<td>-.030</td>
</tr>
</tbody>
</table>

Mean Autocorrelations, OLS Market Model Residuals for Individual Securities, by Exchange-Listing

<table>
<thead>
<tr>
<th>Exchange</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NYSE</td>
<td>-.027</td>
<td>-.020</td>
<td>-.025</td>
</tr>
<tr>
<td>AMEX</td>
<td>-.071</td>
<td>-.031</td>
<td>-.027</td>
</tr>
</tbody>
</table>
Table 8
THE EFFECT OF ASSUMING CROSS-SECTIONAL INDEPENDENCE
Percentage of 250 samples where the null hypothesis is rejected.

<table>
<thead>
<tr>
<th>Actual Level of Abnormal Performance at Day &quot;0&quot;</th>
<th>Mean Test Statistic 0% Abnormal Performance</th>
<th>Mean Test Statistic .5% Abnormal Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-0.01</td>
<td>1.13</td>
</tr>
<tr>
<td>.005</td>
<td>-0.00</td>
<td>1.69</td>
</tr>
<tr>
<td>.01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel A: Nonclustering

MEAN ADJUSTED RETURNS

<table>
<thead>
<tr>
<th>Dependence Adjustment</th>
<th>6.4</th>
<th>25.2</th>
<th>75.6</th>
<th>-0.01</th>
<th>1.13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assume Independence</td>
<td>6.0</td>
<td>48.0</td>
<td>96.0</td>
<td>-0.00</td>
<td>1.69</td>
</tr>
</tbody>
</table>

MARKET MODEL

<table>
<thead>
<tr>
<th>Dependence Adjustment</th>
<th>4.4</th>
<th>27.2</th>
<th>80.4</th>
<th>-0.02</th>
<th>1.17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assume Independence</td>
<td>6.4</td>
<td>53.2</td>
<td>97.6</td>
<td>0.00</td>
<td>1.79</td>
</tr>
</tbody>
</table>

Panel B: Clustering

MEAN ADJUSTED RETURNS

<table>
<thead>
<tr>
<th>Dependence Adjustment</th>
<th>7.2</th>
<th>19.6</th>
<th>39.6</th>
<th>0.17</th>
<th>0.82</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assume Independence</td>
<td>25.2</td>
<td>56.4</td>
<td>84.4</td>
<td>0.35</td>
<td>2.05</td>
</tr>
</tbody>
</table>

MARKET MODEL

<table>
<thead>
<tr>
<th>Dependence Adjustment</th>
<th>8.0</th>
<th>39.2</th>
<th>84.4</th>
<th>0.18</th>
<th>1.43</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assume Independence</td>
<td>8.0</td>
<td>61.2</td>
<td>96.0</td>
<td>0.20</td>
<td>1.98</td>
</tr>
</tbody>
</table>

Sample Size = 50
One-tailed test, α = .05
Randomly selected securities and event dates 1962-1979
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 Residuals

<table>
<thead>
<tr>
<th>Variance Estimation Method:</th>
<th>Actual Level of Abnormal Performance at Day 0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Use time-series data from estimation period</td>
<td>Variance Increase: 12.0</td>
</tr>
<tr>
<td></td>
<td>No Increase: 4.4</td>
</tr>
<tr>
<td>Use cross-sectional data from event period</td>
<td>Variance Increase: 4.0</td>
</tr>
<tr>
<td></td>
<td>No Increase: 3.6</td>
</tr>
</tbody>
</table>

Sample size = 50
One-tailed test, $\alpha = .05$
Randomly Selected Securities and Event Dates 1962-1979
Figure 1

THEORETICAL AND EMPIRICAL CUMULATIVE DISTRIBUTION FUNCTIONS

VALUES OF TEST STATISTICS UNDER NULL HYPOTHESIS
Figure 2

THEORETICAL AND EMPIRICAL POWER FUNCTIONS

THEORETICAL POWER \((\alpha = 5\%)\)

EMPIRICAL POWER \((\alpha = 5\%)\)

BASED ON MARKET MODEL TEST STATISTICS
Figure 1
THEORETICAL AND EMPIRICAL CUMULATIVE DISTRIBUTION FUNCTIONS

VALUES OF TEST STATISTICS UNDER NULL HYPOTHESIS

CUMULATIVE PROBABILITY DISTRIBUTION

- NORMAL DISTRIBUTION
- EMPIRICAL DISTRIBUTION OF MARKET MODEL TEST STATISTICS
Figure 2

THEORETICAL AND EMPIRICAL POWER FUNCTIONS

- THEORETICAL POWER ($\alpha = 5\%$)
- EMPIRICAL POWER ($\alpha = 5\%$)

BASED ON MARKET MODEL TEST STATISTICS

FREQUENCY OF REJECTING NULL HYPOTHESIS

0.0001 0.0005 0.001 0.005 0.01 0.05

ABNORMAL PERFORMANCE
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

Scholes/Williams Procedure

\begin{align}
A_{i,t} &= R_{i,t} - \alpha_i - \beta_i R_{m,t} \\
\alpha_i &= \frac{1}{237} \sum_{t=-7}^{t=-243} R_{i,t} - \beta_i \frac{1}{237} \sum_{t=-7}^{t=-243} R_{m,t} \\
\beta_i' &= \frac{\beta_i^- + \beta_i^+ + \beta_i^+}{1 + 2\hat{\rho}_m}
\end{align}

\(\beta_i^-\) and \(\beta_i^+\) are the OLS estimation period values of

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

\(\hat{\rho}_m\) is the estimation period value for the first order autocorrelation coefficient of the CRSP Equal Weighted Market Index (see Scholes and Williams (1977, p. 317, equations 13-15, 19, and 20)).
Dimson Aggregated Coefficients Method

(A.4) \[ A_i,t = R_i,t - \hat{\alpha} - \hat{\beta}^* R_m,t, \]

where

(A.5) \[ \hat{\alpha}^* = \frac{1}{233} \sum_{t=-241}^{-9} R_i,t - \hat{\beta}_i^* \frac{1}{233} \sum_{t=-241}^{-9} R_m,t, \]

(A.6) \[ \hat{\beta}_i^* = \sum_{k=-3}^{+3} \hat{\beta}_{ik}, \]

\( \hat{\beta}_{ik} \) is the OLS estimate of

\[ \frac{\text{cov}(R_i,t, R_m,t-k)}{\sigma(R_m,t)\sigma(R_m,t-k)} \]

from the estimation period (see Dimson (1979, p. 204, equation 12)).

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.7) \[ A_i,t' = \frac{A_i,t}{S(A_i,t)} \]

---

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 were made to reflect this.
where

\[(A.8) \quad S(A_i, t) = \sqrt{\frac{\sum_{t=-6}^{t=0} (A_i, t - A_i^*)^2}{238}} \quad \text{and} \]

\[(A.9) \quad A_i^* = \frac{1}{239} \sum_{t=-244}^{t=0} A_i, t \]

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

\[(A.10) \quad (\sum_{i=1}^{N_t} A_i, t) \cdot (\mu_t)^{-\frac{1}{2}} . \]

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 \( N_t \).\(^1\)

A.3 Tests Over Multi-day Intervals\(^2\)

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 by:

\(^1\)As in the case of dependence adjustment, adjustments to equation A.8 to reflect the predictive nature of the excess returns are omitted. These adjustments had no detectable impact on the paper's tables.

\(^2\)We only discuss test statistics and autocorrelation adjustments for the case of dependence adjustment; the derivation under cross sectional independence is similar. Multi-day simulations under independence did not change the paper's conclusions.
where the terms in the denominator are from equation (8) in the text. The test statistic is assumed unit normal in the absence of abnormal performance.

**Autocorrelation Adjustments**

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

\[
(A.12) \quad \Sigma = \begin{bmatrix}
\sigma^2(\bar{\alpha}_5) & \text{cov}(\bar{\alpha}_5, \bar{\alpha}_4) - & \text{cov}(\bar{\alpha}_5, \bar{\alpha}_5) \\
\text{cov}(\bar{\alpha}_4, \bar{\alpha}_4) & \sigma^2(\bar{\alpha}_4) & \text{cov}(\bar{\alpha}_4, \bar{\alpha}_5) \\
\text{cov}(\bar{\alpha}_5, \bar{\alpha}_5) & \text{cov}(\bar{\alpha}_5, \bar{\alpha}_4) & \sigma^2(\bar{\alpha}_5)
\end{bmatrix}
\]

**Parameter Estimation**

The estimated standard deviations of the mean daily excess return \( \bar{\alpha}_t \) is given in equation (8) in the text. Estimates of the diagonal elements of (A.12) follow directly. Noting that each off-diagonal element (cov \( \bar{\alpha}_t, \bar{\alpha}_{t-\ell} \)) = \( \rho(\bar{\alpha}_t, \bar{\alpha}_{t-\ell}) \sigma(\bar{\alpha}_t) \sigma(\bar{\alpha}_{t-\ell}) \),
References


sample estimates of each off-diagonal element use estimates of the autocorrelation coefficients of mean excess returns.

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

$$|\rho_\ell| > \frac{2.0}{\sqrt{239-\ell}}$$

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 ARIMA modelling of time-series properties would lead to improved specification or power.

**Test Statistics**

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

¹Once the estimation process is completed, we also check that estimates of $\rho_\ell$ imply that $\Sigma$ is positive definite; if not, all values of $\rho_\ell$ are also constrained to zero. See Box and Jenkins (1976, p. 29) for a discussion.


