Journal of Financial Economics 30 (1991) 253–272. North-Holland

Event-study methodology under conditions of event-induced variance*

Ekkehart Boehmer
Louisiana State University, Baton Rouge, LA 70803, USA

Jim Musumeci
American University, Washington, DC 20016, USA

Annette B. Poulsen
University of Georgia, Athens, GA 30602, USA

Received September 1990, final version received October 1991

Many authors have identified the hazards of ignoring event-induced variance in event studies. To determine the practical extent of the problem, we simulate an event with stochastic effects. We find that when an event causes even minor increases in variance, the most commonly-used methods reject the null hypothesis of zero average abnormal return too frequently when it is true, although they are reasonably powerful when it is false. We demonstrate that a simple adjustment to the cross-sectional techniques produces appropriate rejection rates when the null is true and equally powerful tests when it is false.

1. Introduction

Since Fama, Fisher, Jensen, and Roll's 1969 study of stock splits, event studies have become the predominant methodology for determining the effects of an event on the distribution of security returns. In this paper, we investigate an often-ignored aspect of event studies: whether event-induced increases in the variance of returns affect the ability of event-study methods to detect whether the event's average effect on stock returns is zero. Brown and Warner (1980, 1985) verify that event studies work well when an event

* We thank Stephen Brown (the referee) and John B. Long (the editor) for their many helpful and thoughtful comments. We also thank Mary Dehner, David Jobson, Scott Lee, Rob Nash, Jeffry Netter, Ralph Sanders, Joseph Sinkey, Stephen D. Smith, and Jerold Warner for their valuable suggestions.

has an identical effect on all firms, but they also warn that when an event has differing effects on firms, the variance of returns will increase and common methods may fail (1985, pp. 22–25). In fact, Brown, Harlow, and Tinic (1988, 1989) show that many events cause changes in both risk and return for individual securities, as indicated by a temporary increase in the variance of abnormal returns accompanying the mean shift. They argue that an increase in variance accompanying an event is due to a temporary change in the firm's systematic risk. While we do not discuss the cause of event-induced variance in this paper, we show that it is necessary to control for variance changes to obtain appropriate tests of the null hypothesis that the average abnormal return is zero.

To determine the ability of commonly-used methods to identify abnormal returns in the presence of event-induced variance, we simulate the occurrence of an event with stochastic effects on stock returns for 250 samples of 50 securities each. We compare the results of several tests and find that when an event causes even minor increases in variance, the most commonly-used methods frequently cause the null hypothesis of zero average abnormal returns to be rejected when it is, in fact, true. We show, however, that a simple adjustment to the cross-sectional method results in equally-powerful tests when the null is false and appropriate rejection rates when it is true. Both the size and the power of the adjusted test are unaffected when applied to portfolios subject to event-date clustering.

The remainder of the paper is organized as follows. In section 2, we briefly define the problem of event-induced variance in event-study methodology and summarize solutions proposed by Christie (1983), Collins and Dent (1984), and Ball and Torous (1988). Our experimental design and the test statistics are described in section 3. In section 4, we present simulation results comparing the ability of six test statistics to detect abnormal returns when there are event-induced changes in variance. Our conclusions are in section 5.

2. Dealing with event-induced variance

It is not uncommon for an event to be accompanied by increases in the cross-sectional dispersion of stock returns. More than 20 years ago, Beaver (1968) concluded that an increase in the cross-sectional dispersion of abnormal returns at the time of earnings announcements implied that the announcement conveyed information. Certainly, if a researcher fails to appropriately control for factors that lead to varying announcement effects across firms, he or she will generally measure a dispersion increase on the event day. For example, we would expect a different abnormal stock return associated with the adoption of a poison pill by a firm in the midst of a takeover contest as compared to a firm with 75% insider ownership. Similarly, the varying levels of announced earnings across firms will lead to varying abnormal stock returns. Such varying announcement effects lead to an increase in measured cross-sectional dispersion that really reflects the failure to control for all factors affecting returns.

Brown and Warner (1980, 1985) suggest that, in general, event-study tests are well-specified and reasonably powerful. However, they identify potential testing problems created by an event-induced increase in variances (1985, pp. 22–25). They note that if the variance is underestimated, the test statistic will lead to rejection of the null hypothesis more frequently than it should, even when the average abnormal performance is zero. Although Brown and Warner conduct several simulations, they point out the need for further research. Several other authors [e.g., Beaver (1968), Christie (1983), Dann (1981), Kalay and Lowenstein (1985), Patell and Wolfson (1979), and Rosenstein and Wyatt (1990)] find that the variance of returns increases significantly when certain events occur. Dann (in his table 3), for example, shows the event-period standard deviation to be more than three and a half times as great as the estimation period in his study of stock repurchases.

One remedy, of course, is to ignore the estimation-period residual variance and to use instead the cross-sectional variance in the event period itself to form the test statistic. At least eight event studies have used this cross-sectional technique and reported both the estimation-period and event-period cross-sectional standard deviations. [See Charest (1978), Dann (1981), Mikkelson (1981), Penman (1982), and Rosenstein and Wyatt (1990); several of the papers contained more than one event study.] In each study, the event-period standard deviation was larger than the estimation-period standard deviation.

The event-study literature contains various other proposals for coping with the problem of event-induced variance. Christie (1983) suggests that event-induced variance may be estimated if multiple events are observed for each firm. Mitchell and Netter (1989) take several approaches to dealing with the large increase in variance surrounding the stock market crash of 1987, including tests doubling the variance based on estimation-period variance (pre- and post-event), cross-sectional tests, and nonparametric tests. Collins and Dent (1984) simulate stock returns and an event with stochastic effects. They demonstrate that with event-date clustering, a generalized least squares (GLS) technique works better than the traditional ordinary least squares (OLS) methods. The Collins and Dent approach has the additional advantage of being able to accommodate cross-sectional correlation of returns. Nonetheless, researchers generally do not use any of these suggestions for dealing with event-induced variance because of data limitations and difficulty of implementation.

More recently, Ball and Torous (1988) simulate an event that increases the mean and variance of stock returns and apply a maximum likelihood estimation (MLE) technique to stock return data, simultaneously estimating event-period returns, the variance of these returns, and the probability of the event's occurrence for any given day in the event window. When abnormal performance is present, their simulations show that the MLE method rejects the null hypothesis more frequently than does the traditional Brown and Warner method, although the null is not rejected too frequently when it is true. Finally, Corrado (1989) proposes a nonparametric rank test that recognizes the asymmetry in cross-sectional excess returns. He shows that the rank test performs better than the parametric t-tests used by Brown and Warner if there is an event-induced increase in the variance of the abnormal returns.

We propose another test that works well in the presence of event-induced variance. This 'standardized cross-sectional' test is easy to implement and is a hybrid of Patell's (1976) standardized-residual methodology and the ordinary cross-sectional approach suggested, for example, by Charest (1978) and Penman (1982). The test incorporates variance information from both the estimation and the event periods and is closely related to the special case of Ball and Torous' (1988) estimator in which there is no event-day uncertainty. To determine the robustness of our method relative to several commonly-used methods that ignore changes in variance, we simulate the occurrence of an event with stochastic effects on stock returns. Our test yields appropriate rejection rates when the null is true and yet is equally powerful when the null is false.

3. Experimental design

3.1. Sample selection

We construct 250 samples of 50 securities each. The securities in each sample are randomly selected from all securities included in the 1987 CRSP Daily Returns File and are assigned randomly-selected event dates. Although the 1987 CRSP Daily Returns File includes security returns from July 1962 through December 1987, we exclude 1987 event dates because of the volatility in stock returns in the latter part of the year. (We do not believe that the exclusion of this period affects the generality of our results.) All securities and event dates are sampled with replacement such that each security/date combination has an equal chance of being chosen at each selection. We also check each security/date combination for sufficient return data in the CRSP file. To remain in a sample, a security must have at least 50 daily returns in the estimation period (-249 through -11) and no missing returns in the 30 days surrounding the event date (-19 through $+10$). Each of the 250 portfolios is drawn independently of the others.

3.2. Simulating abnormal performance with event-induced variance

Our simulations focus on detection of abnormal performance using daily returns. The level of abnormal performance on day 0 for each security in each portfolio is drawn from a normal distribution with a mean of 0%, 1%, or 2% and a nonzero variance. The nonzero variance implies that, unlike most previous simulation studies, each security will not have an identical level of abnormal performance.

Our simulated increase in the variance of abnormal performance (i.e., the event-induced variance) for each security i takes the general form $k_i\sigma^2$, where σ^2 represents either the variance of security i's estimation-period residuals or that of the average security's estimation-period residuals, and k_i is a constant equal to 0, 0.5, 1, or 2. Charest (1978), Mikkelson (1981), Penman (1982), and Rosenstein and Wyatt (1989) generally found the event-period standard deviation to be about 1.2 to 1.5 times the estimation-period standard deviation, corresponding to values of k_i ranging from about 0.44 to 1.25. Beaver (1968) used a different method to conclude that variance increased by an amount corresponding to a value of $k_i = 0.67$. Dann's (1981) study of stock repurchases reported the largest increase in cross-sectional standard deviation (by a factor of about 3.625, corresponding to a value for k_i of about 11). The case of $k_i = 0$, indicating event-induced variance is zero, is simply a replication of earlier studies that add a constant abnormal return to stock residuals.

When event-induced variance is proportional to the average firm variance, the abnormal performance is independent and identically distributed for each security in each portfolio. When event-induced variance is proportional to the individual security variance, returns are not identically distributed since the variance differs for each security but the abnormal returns are still drawn independently from normal distributions.

For each of the 250 portfolios, we estimate excess returns using the market-model method (with the CRSP equally-weighted index) described in Brown and Warner (1980, 1985). We then use several different test statistics to test for abnormal returns. Each test statistic is described below and formally defined in the appendix. Our results are reported in section 4.[1]

3.3. *Test statistics*

3.3.1. *The traditional test*

Brown and Warner's 'no dependence adjustment' method (1980, app. A.3), henceforth the 'traditional method', implicitly assumes that security residuals

[1]While not reported in this paper, we also simulate a stochastic variance increase where k_i is a constant multiplied by an independent drawing from a χ^2 distribution with one degree of freedom. The expected (or specified) value of k_i is either 0, 0.5, 1, or 2. Since a chi-squared distribution with one degree of freedom has a mean of 1, we multiplied by the appropriate constant (0.5, 1, or 2) to get the desired mean value of k_i in the cases where k_i is a random variable. The simulation results using a stochastic variance increase are essentially identical to those obtained for deterministic variance increases.

are uncorrelated and that event-induced variance is insignificant.[2] The test statistic equals the sum of the event-period abnormal returns divided by the square root of the sum of all securities' estimation-period residual variances.

3.3.2. Standardized-residual test

Like the traditional method, Patell's (1976) method (henceforth the 'standardized-residual method') assumes that security residuals are uncorrelated and that event-induced variance is insignificant. However, the residuals are standardized before forming portfolios. This standardization serves two purposes. First, it adjusts for the fact that the event-period residual is an out-of-sample prediction and hence it will have a higher standard deviation than estimation-period residuals [see, for example, Judge, Hill, Griffiths, Lutkepohl, and Lee (1988, p. 170)]. Second, standardizing the event-period residuals before forming portfolios allows for heteroskedastic event-day residuals, and prevents securities with large variances from dominating the test. The standardized residual equals the event-period residual divided by the standard deviation of the estimation-period residuals, adjusted to reflect the forecast error. This standardized residual is approximately unit normal, so that the appropriate t-statistic is the sum of the standardized residuals divided by (approximately) the square root of the number of sample firms.[3] The standardized-residual method is used by Brown and Warner (1985).

3.3.3. Sign test

The traditional and standardized-residual tests are often conducted in conjunction with a sign test to help verify that a few firms are not driving the results. The test statistic for the sign method is the observed proportion of positive returns minus 0.5, divided by the standard deviation of a binomial distribution. A problem with this approach is that it assumes that 50% of security returns are negative, while returns are, in fact, skewed to the right [see, for example, Fama (1976) or Brown and Warner (1980)]. Several researchers have adjusted for the skewness by comparing the event-period

[2]This method is not appropriate if securities' residuals are cross-sectionally correlated, which may be the case if the securities have a common event date (also known as event-date clustering). To solve this problem, Brown and Warner's 'crude dependence adjustment' uses the variance of portfolio residuals (rather than the sum of the variances of residuals for individual securities) from the estimation period. The t-statistic equals the portfolio abnormal return divided by the portfolio residual's standard deviation from the estimation period. If clustering is present, this portfolio approach will impound any residual cross-sectional correlation in its estimate of portfolio residual's standard deviation. We included the crude dependence adjustment in our simulations (though not reported here) and, like Brown and Warner, found it to be less powerful than the traditional test.

[3]The actual denominator is $\sqrt{\sum_{i=1}^{N}(T_i-2)/(T_i-4)}$, where T_i is the number of days in security i's estimation period and N is the number of firms in the sample. If for most firms there are a large number of days in the estimation period, $\sum_{i=1}^{N}(T_i-2)/(T_i-4) \approx N$. Brown and Warner (1985) used the approximate test statistic in their hypothesis tests assuming cross-sectional independence.

proportion of positive returns to the proportion in the estimation period [see, for example, McConnell and Muscarella (1985) and Lummer and McConnell (1989)].

3.3.4. Cross-sectional test

The ordinary cross-sectional method conducts a t-test by dividing the average event-period residual by its contemporaneous cross-sectional standard error. As do the preceding methods, the ordinary cross-sectional method requires security residuals to be uncorrelated across firms. It does not require event-induced variance to be insignificant, although if the event-period residuals for different firms are drawn from different distributions, the ordinary cross-sectional test will be misspecified.

3.3.5. Method-of-moments estimation

Froot's (1989) method-of-moments estimator allows residuals to be contemporaneously correlated and heteroskedastic. The procedure requires, however, that we identify certain groups of firms for which residuals across groups are independent, though residuals within groups may be correlated. We form suitable groups within each of our 250 samples by classifying firms according to industry (one-digit SIC code), consistent with Froot's proposed classification scheme except that we allow the number of firms per industry to vary. (Froot's requirement of the same number of firms per industry is too restrictive for practical purposes.)

To obtain the test statistic, we first calculate the average residual for each industry. This residual is then divided by its standard error to form a standardized industry residual (SIR). Since by assumption the SIRs are independent, the resulting test statistic is the sum of the SIRs divided by the square root of the number of industries.[4]

3.3.6. Standardized cross-sectional test

Our proposed procedure addresses the misspecification problem of the ordinary cross-sectional technique. By combining the standardized-residual and the ordinary cross-sectional approaches, we form a hybrid which we call the standardized cross-sectional test. First, the residuals are standardized by the estimation-period standard deviation (adjusted for forecast error) to eliminate the misspecification problem of the ordinary cross-sectional test. The ordinary cross-sectional technique is then applied to the standardized residuals; the test statistic is found by dividing the average event-period standardized residual by its contemporaneous cross-sectional standard error. Like the ordinary cross-sectional method, this test allows event-induced

[4]The test statistic used here corresponds to a special case of Froot's estimator. As he suggests, we use the estimation-period variance of the mean industry residuals as an instrumental variable. In our simulation there is only one time-series observation (the event day), with an average of seven industries for each 50-security portfolio.

variance changes. It also incorporates information from the estimation period, which may enhance its efficiency and power. This method also requires that security residuals be cross-sectionally uncorrelated.[5]

4. Results

Our simulation results are summarized in tables 1–4. Each table presents the rejection frequencies for the six tests of zero average abnormal performance. Table 1 reports the results of simulations in which event-induced variance is zero. Our simulations are conducted under the same conditions as Brown and Warner's (1985) simulations, and for the two cases which are directly comparable, we obtain similar results. Specifically, for a one-tailed test with $\alpha = 0.05$, their simulations (using the standardized-residual method) reject the null hypothesis 6.4% of the time, while our rejection rate for the standardized-residual method is 7.2%. When there is abnormal performance of 1%, their simulations reject the null 97.6% of the time, while ours reject it 96% of the time.

When there is no abnormal performance, all test statistics for average abnormal performance reject the null at about the correct frequency. When there is abnormal performance, all of the tests reject the null at reasonable rates except for the method-of-moments estimator, which is distinctly less powerful as measured by its relatively low rejection frequency.

In tables 2, 3, and 4, panel A reports rejection rates from tests that assume that event-induced variance is some multiple of the individual security's estimation-period residual variance, while panel B reports rejection rates from tests that assume that event-induced variance is some multiple of the average of the portfolio's estimation-period residual variances. All tests reported in this paper assume that the ratio of event-induced variance to estimation-period variance is a constant.

4.1. Event-induced variance proportional to individual-security residual variance (panel A of tables 2, 3, and 4)

In panel A of table 2, we report our results for the case in which there is zero abnormal performance but individual-security variance increases by a constant proportion, k_i. The problem with the traditional and the standardized-residual methods is clearly demonstrated. When there is zero abnormal

[5]Our standardized cross-sectional test is similar to the test statistic derived by Ball and Torous (1988), though we consider several different cases of event-induced variance and compare our estimator to all of the standard methodologies. The Ball–Torous model is very general in that it allows for alternative specifications of the return-generating process, as well as event-date uncertainty. If the market model generates returns such that the event occurs with probability one on a certain day, and any event-induced variance is proportional to the individual security's estimation-period variance, the standardized returns in Ball and Torous are normal, independent, and identically distributed. In this case, the MLE coincides with OLS.

Table 1

Average rejection rates for various test statistics for 250 portfolios of 50 securities (from 1962 to 1986) each at several significance levels. The test statistics test the null hypothesis that the average abnormal return is zero. Abnormal performance is equal to μ (no event-induced variance increase).

	One-tailed tests		Two-tailed tests		
	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
	$\mu = 0\%$				
Traditional[a]	0.072	0.032	0.128	0.076	0.024
Standardized-residual[b]	0.072	0.020	0.124	0.076	0.032
Sign[c]	0.016	0.004	0.140	0.088	0.020
Ordinary cross-sectional[d]	0.060	0.016	0.132	0.076	0.028
Method-of-moments[e]	0.060	0.040	0.092	0.072	0.048
Standardized cross-sectional[f]	0.048	0.016	0.116	0.072	0.032
	$\mu = 1\%$				
Traditional	0.792	0.532		0.664	0.468
Standardized-residual	0.960	0.872		0.936	0.816
Sign	0.952	0.756		0.892	0.680
Ordinary cross-sectional	0.812	0.572		0.720	0.504
Method-of-moments	0.332	0.208		0.248	0.158
Standardized cross-sectional	0.960	0.836		0.928	0.752
	$\mu = 2\%$				
Traditional	1.000	0.996		1.000	0.996
Standardized-residual	1.000	1.000		1.000	1.000
Sign	1.000	1.000		1.000	1.000
Ordinary cross-sectional	1.000	0.988		0.992	0.988
Method-of-moments	0.760	0.544		0.632	0.480
Standardized cross-sectional	1.000	1.000		1.000	1.000

[a]The traditional test statistic equals the sum of the event-period abnormal returns divided by the square root of the sum of all securities' estimation-period residual variances.

[b]The standardized-residual test statistic equals the sum of the residuals standardized by their standard deviations divided by the (approximate) square root of the number of sample firms.

[c]The sign test statistic is the observed proportion of positive returns minus 0.50 divided by the standard deviation of a binomial distribution.

[d]The ordinary cross-sectional test statistic divides the average event-period residual by its contemporaneous cross-sectional standard error.

[e]For the method-of-moments test statistic, the average industry residual is calculated. Then the residual is standardized by its standard error and divided by the square root of the number of industries.

[f]For the standardized cross-sectional test statistic, the residuals are standardized by their standard deviations. Then the average event-period standardized residual is divided by its contemporaneous cross-sectional standard error.

performance, the underestimation of event-period variance causes the null hypothesis to be rejected too frequently. For example, even when k_i is only 0.5 (i.e., the incremental or event-induced variance is only half of the estimation-period variance), the traditional and standardized-residual methods reject the null hypothesis anywhere from 2.2 to 6.4 times as often as they

Table 2

Average rejection rates for various test statistics for 250 portfolios of 50 securities (from 1962 to 1986) each at several significance levels. The test statistics test the null hypothesis that the average abnormal return is zero. Abnormal performance is a random variable drawn from $N(0, k\sigma^2)$, where k is a constant and σ^2 is either the individual-security or the average estimation-period variance.

	Panel A Event-induced variance proportional to individual-security estimation-period variance				*Panel B* Event-induced variance proportional to average estimation-period variance			
	One-tailed tests		Two-tailed tests		One-tailed tests		Two-tailed tests	
	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
				$k = 0.5$				
Traditional[a]	0.128	0.052	0.152	0.064	0.124	0.064	0.136	0.056
Standardized-residual[b]	0.148	0.052	0.152	0.052	0.188	0.100	0.236	0.128
Sign[c]	0.052	0.012	0.088	0.016	0.068	0.008	0.080	0.000
Ordinary cross-sectional[d]	0.092	0.004	0.080	0.012	0.088	0.020	0.072	0.032
Method-of-moments[e]	0.068	0.036	0.100	0.036	0.100	0.028	0.100	0.044
Standardized cross-sectional[f]	0.084	0.020	0.068	0.028	0.092	0.028	0.088	0.028
				$k = 1$				
Traditional	0.140	0.080	0.208	0.084	0.164	0.076	0.192	0.088
Standardized-residual	0.180	0.080	0.188	0.084	0.208	0.148	0.320	0.204
Sign	0.068	0.008	0.068	0.004	0.056	0.012	0.076	0.016
Ordinary cross-sectional	0.072	0.008	0.076	0.024	0.064	0.020	0.084	0.028
Method-of-moments	0.100	0.064	0.112	0.068	0.112	0.056	0.140	0.076
Standardized cross-sectional	0.060	0.012	0.060	0.012	0.080	0.020	0.068	0.020
				$k = 2$				
Traditional	0.208	0.112	0.284	0.156	0.200	0.108	0.272	0.164
Standardized-residual	0.224	0.120	0.284	0.164	0.308	0.196	0.408	0.268
Sign	0.060	0.004	0.076	0.000	0.056	0.004	0.076	0.000
Ordinary cross-sectional	0.068	0.016	0.072	0.012	0.064	0.012	0.064	0.020
Method-of-moments	0.140	0.052	0.160	0.092	0.144	0.080	0.208	0.124
Standardized cross-sectional	0.088	0.028	0.084	0.028	0.068	0.012	0.064	0.016

[a]The traditional test statistic equals the sum of the event-period abnormal returns divided by the square root of the sum of all securities' estimation-period residual variances.

[b]The standardized-residual test statistic equals the sum of the residuals standardized by their standard deviations divided by the (approximate) square root of the number of sample firms.

[c]The sign test statistic is the observed proportion of positive returns minus 0.50 divided by the standard deviation of a binomial distribution.

[d]The ordinary cross-sectional test statistic divides the average event-period residual by its contemporaneous cross-sectional standard error.

[e]For the method-of-moments test statistic, the average industry residual is calculated. Then the residual is standardized by its standard error and divided by the square root of the number of industries.

[f]For the standardized cross-sectional test statistic, the residuals are standardized by their standard deviations. Then the average event-period standardized residual is divided by its contemporaneous cross-sectional standard error.

should. When k_t is 1 (i.e., event-induced variance is equal to ordinary-residual variance), they reject the null 2.8 to 8.4 times as often as they should. To demonstrate the seriousness of this problem, the 0.208 rejection rate of the traditional method in panel A of table 2 ($\mu = 0$, $k_t = 1$, $\alpha = 0.05$, two-tailed test) corresponds to rejecting the null hypothesis at the 5% significance level when $|t| > 1.26$ (instead of $|t| > 1.96$). In the same manner, the 0.084 rejection rate of both the traditional and standardized-residual methods in panel A of table 2 ($\mu = 0$, $k_t = 1$, $\alpha = 0.01$, two-tailed test) corresponds to rejecting the null hypothesis at the 1% significance level when $|t| > 1.73$ (instead of $|t| > 2.58$).

Both the ordinary and standardized cross-sectional tests reject the null at about the appropriate significance level. The method-of-moments test rejects with approximately the correct frequency when the event-induced variance is small ($k = 0.5$). When k increases, however, the method-of-moments test rejects too frequently. The sign test rejects the null at about the appropriate significance level, which suggests that use of a sign test in conjunction with either the traditional, standardized-residual, or method-of-moments tests may mitigate the problem of event-induced variance.

In panel A of table 3, we report our results when average abnormal performance equals 1% and the variance increase is proportional to the individual-security variance. We find a moderate reduction of power in all tests when $\alpha = 0.05$ and a severe reduction when $\alpha = 0.01$, as compared to the case of no variance increase reported in table 1. In most cases, the ordinary cross-sectional test, the method-of-moments test, and the sign test are the weakest. In contrast, the standardized cross-sectional technique is not only more powerful than the ordinary cross-sectional test and sign test, but it is often more powerful than the traditional test, particularly at lower levels of event-induced variance. For example, panel A of table 3 reports that in a two-tailed test ($\alpha = 0.05$) the standardized cross-sectional test rejects the null 81.6% of the time when $k_t = 0.5$ and 70.1% of the time when $k_t = 1$. The corresponding traditional method's rejection frequencies are only 69.6% and 61.6%.

When abnormal performance averages 2% and event-induced variance is proportional to the individual-security variance (panel A of table 4), the traditional, standardized-residual, and standardized cross-sectional methods rejected virtually 100% of the time when $\alpha = 0.05$ for a two-tailed test. The sign test and ordinary cross-sectional test are not quite as powerful, and the method-of-moments test is generally the least powerful of all.

4.2. *Event-induced variance proportional to average residual variance (panel B of tables 2, 3, and 4)*

The traditional, standardized-residual, and method-of-moments tests also perform poorly when event-induced variance is proportional to average residual variance and $\mu = 0$ (zero average abnormal performance). For

Table 3

Average rejection rates for various test statistics for 250 portfolios of 50 securities (from 1962 to 1986) each at several significance levels. The test statistics test the null hypothesis that the average abnormal return is zero. Abnormal performance is a random variable drawn from $N(0.01, k\sigma^2)$, where k is a constant and σ^2 is either the individual-security or the average estimation-period variance.

	Panel A Event-induced variance proportional to individual-security estimation-period variance				*Panel B* Event-induced variance proportional to average estimation-period variance			
	One-tailed tests		Two-tailed tests		One-tailed tests		Two-tailed tests	
	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
				$k = 0.5$				
Traditional[a]	0.780	0.572	0.696	0.492	0.748	0.516	0.668	0.448
Standardized-residual[b]	0.920	0.840	0.896	0.788	0.876	0.764	0.820	0.724
Sign[c]	0.780	0.464	0.688	0.344	0.576	0.248	0.448	0.176
Ordinary cross-sectional[d]	0.692	0.440	0.584	0.356	0.636	0.360	0.500	0.284
Method-of-moments[e]	0.360	0.200	0.260	0.156	0.360	0.216	0.268	0.160
Standardized cross-sectional[f]	0.880	0.708	0.816	0.628	0.744	0.472	0.604	0.376
				$k = 1$				
Traditional	0.700	0.528	0.616	0.452	0.740	0.556	0.652	0.480
Standardized-residual	0.900	0.800	0.876	0.760	0.840	0.736	0.824	0.704
Sign	0.648	0.276	0.492	0.168	0.464	0.196	0.352	0.144
Ordinary cross-sectional	0.520	0.280	0.396	0.232	0.560	0.336	0.464	0.252
Method-of-moments	0.328	0.212	0.268	0.176	0.356	0.256	0.320	0.224
Standardized cross-sectional	0.804	0.516	0.704	0.416	0.628	0.356	0.488	0.288
				$k = 2$				
Traditional	0.668	0.516	0.592	0.468	0.660	0.492	0.580	0.456
Standardized-residual	0.844	0.752	0.812	0.696	0.768	0.704	0.748	0.684
Sign	0.548	0.240	0.452	0.172	0.356	0.116	0.276	0.080
Ordinary cross-sectional	0.408	0.224	0.324	0.180	0.428	0.212	0.324	0.152
Method-of-moments	0.424	0.272	0.368	0.224	0.396	0.232	0.328	0.216
Standardized cross-sectional	0.624	0.392	0.504	0.316	0.420	0.220	0.328	0.148

[a]The traditional test statistic equals the sum of the event-period abnormal returns divided by the square root of the sum of all securities' estimation-period residual variances.

[b]The standardized-residual test statistic equals the sum of the residuals standardized by their standard deviations divided by the (approximate) square root of the number of sample firms.

[c]The sign test statistic is the observed proportion of positive returns minus 0.50 divided by the standard deviation of a binomial distribution.

[d]The ordinary cross-sectional test statistic divides the average event-period residual by its contemporaneous cross-sectional standard error.

[e]For the method-of-moments test statistic, the average industry residual is calculated. Then the residual is standardized by its standard error and divided by the square root of the number of industries.

[f]For the standardized cross-sectional test statistic, the residuals are standardized by their standard deviations. Then the average event-period standardized residual is divided by its contemporaneous cross-sectional standard error.

Table 4

Average rejection rates for various test statistics for 250 portfolios of 50 securities (from 1962 to 1986) each at several significance levels. The test statistics test the null hypothesis that the average abnormal return is zero. Abnormal performance is a random variable drawn from $N(0.02, k\sigma^2)$, where k is a constant and σ^2 is either the individual-security or the average estimation-period variance.

	Panel A Event-induced variance proportional to individual-security estimation-period variance				*Panel B* Event-induced variance proportional to average estimation-period variance			
	One-tailed tests		Two-tailed tests		One-tailed tests		Two-tailed tests	
	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
				$k = 0.05$				
Traditional[a]	0.996	0.976	0.992	0.972	1.000	0.984	0.992	0.972
Standardized-residual[b]	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Sign[c]	1.000	0.968	0.996	0.936	0.992	0.876	0.964	0.832
Ordinary cross-sectional[d]	0.992	0.940	0.964	0.916	0.992	0.952	0.988	0.920
Method-of-moments[e]	0.724	0.556	0.644	0.480	0.732	0.520	0.660	0.464
Standardized cross-sectional[f]	1.000	1.000	1.000	0.996	1.000	0.988	1.000	0.972
				$k = 1$				
Traditional	0.980	0.952	0.972	0.944	0.984	0.968	0.972	0.952
Standardized-residual	1.000	1.000	1.000	0.996	0.996	0.996	0.996	0.992
Sign	0.988	0.908	0.968	0.840	0.916	0.720	0.860	0.612
Ordinary cross-sectional	0.960	0.872	0.916	0.816	0.956	0.888	0.932	0.824
Method-of-moments	0.736	0.560	0.640	0.480	0.740	0.508	0.628	0.476
Standardized cross-sectional	1.000	0.984	0.996	0.976	0.980	0.904	0.956	0.844
				$k = 2$				
Traditional	0.964	0.924	0.952	0.900	0.968	0.944	0.960	0.936
Standardized-residual	1.000	1.000	1.000	1.000	0.988	0.980	0.988	0.968
Sign	0.920	0.724	0.868	0.644	0.824	0.484	0.740	0.364
Ordinary cross-sectional	0.888	0.720	0.836	0.672	0.892	0.700	0.828	0.636
Method-of-moments	0.716	0.496	0.620	0.456	0.652	0.500	0.572	0.432
Standardized cross-sectional	0.992	0.928	0.972	0.900	0.896	0.704	0.840	0.616

[a] The traditional test statistic equals the sum of the event-period abnormal returns divided by the square root of the sum of all securities' estimation-period residual variances.

[b] The standardized-residual test statistic equals the sum of the residuals standardized by their standard deviations divided by the (approximate) square root of the number of sample firms.

[c] The sign test statistic is the observed proportion of positive returns minus 0.50 divided by the standard deviation of a binomial distribution.

[d] The ordinary cross-sectional test statistic divides the average event-period residual by its contemporaneous cross-sectional standard error.

[e] For the method-of-moments test statistic, the average industry residual is calculated. Then the residual is standardized by its standard error and divided by the square root of the number of industries.

[f] For the standardized cross-sectional test statistic, the residuals are standardized by their standard deviations. Then the average event-period standardized residual is divided by its contemporaneous cross-sectional standard error.

example, panel B of table 2 reports that when $\mu = 0$ and $k_t = 1$ (event-induced or incremental variance is equal to average residual variance), the traditional, standardized-residual, and method-of-moments tests reject the null hypothesis of zero abnormal performance 19.2%, 32%, and 14% of the time, respectively, for the two-tailed test at the 5% significance level. The problem is more severe at the 1% significance level, where the rejection rates are 8.8%, 20.4%, and 7.6%, respectively, for the two-tailed test. In contrast, the sign, ordinary cross-sectional, and standardized cross-sectional tests reject the null hypothesis at about the appropriate significance level.

When abnormal performance equals 1% and 2% (panel B of tables 3 and 4, respectively), the standardized cross-sectional tests are still more powerful than the sign test, but lose their advantage over the ordinary cross-sectional tests, especially as event-induced variance increases. This is not surprising since the cross-sectional test is well-specified when event-induced variance is large and proportional to average security residual variance.

4.3. *Event-date clustering*

We also simulated the effect of event-date clustering on our results. The simulation design is identical to that described in section 3.1, except that all securities within each portfolio have the same randomly-selected event date (selected without replacement across portfolios). Securities are selected without replacement within each portfolio. Overall, we find little differences in the rejection rates from those obtained without clustering.

In panel A of table 5, we report results for the case of zero abnormal performance with no change in variance ($k_t = 0$) and with event-induced variance equal to twice the variance of each security's estimation-period variance ($k_t = 2$). When there is no event-induced variance, all tests still reject at about the correct level, with the exception of the sign test. In particular, the standardized cross-sectional test is not affected by event-date clustering. As in the case with no event-date clustering, the traditional, standardized-residual, and method-of-moments tests reject substantially more often than expected in the presence of event-induced variance. The sign, ordinary cross-sectional, and standardized cross-sectional tests reject at close to the appropriate rates.

In panel B, we report rejection rates when abnormal performance averages 1% and the variance conditions are the same as in panel A. The results with event-date clustering are again very similar to those without clustering. All tests lose power in the presence of event-induced variance, but the standardized cross-sectional test performs better than the sign, ordinary cross-sectional, and method-of-moments tests.[6]

[6] Note that event-date clustering simulations using randomly-generated samples may not be directly applicable to 'real-world' examples of event-date clustering. In those cases, the fact that some event occurs to all firms on a specific day may imply other commonalities (and thus cross-sectional correlation), such as all firms being in the same industry. Our randomly-generated samples would not necessarily have such correlations.

Table 5

Event-date clustering

Average rejection rates for various test statistics for 250 portfolios of 50 securities (from 1962 to 1986) each at several significance levels. Each security within each portfolio has the same event date. The test statistics test the null hypothesis that the average abnormal return is zero. Abnormal performance is a random variable drawn from $N(\mu, k\sigma^2)$, where μ equals 0% or 1%, k is a constant, and σ^2 is the individual-security estimation-period variance

Panel A: **Abnormal performance $\mu = 0\%$**

	One-tailed tests		Two-tailed tests	
	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
	$k = 0$			
Traditional[a]	0.064	0.016	0.048	0.012
Standardized-residual[b]	0.068	0.016	0.060	0.024
Sign[c]	0.044	0.012	0.148	0.024
Ordinary cross-sectional[d]	0.020	0.004	0.040	0.012
Method-of-moments[e]	0.068	0.008	0.076	0.016
Standardized cross-sectional[f]	0.032	0.008	0.036	0.020
	$k = 2$			
Traditional	0.196	0.116	0.248	0.132
Standardized-residual	0.216	0.128	0.320	0.188
Sign	0.064	0.016	0.060	0.012
Ordinary cross-sectional	0.064	0.020	0.052	0.004
Method-of-moments	0.056	0.008	0.232	0.140
Standardized cross-sectional	0.080	0.032	0.088	0.016

Panel B: Abnormal performance $\mu = 1\%$

	One-tailed tests		Two-tailed tests	
	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
	$k = 0$			
Traditional	0.836	0.656	0.768	0.560
Standardized-residual	0.960	0.860	0.916	0.820
Sign	0.904	0.720	0.860	0.636
Ordinary cross-sectional	0.840	0.664	0.784	0.564
Method-of-moments	0.560	0.316	0.420	0.212
Standardized cross-sectional	0.928	0.840	0.904	0.792
	$k = 2$			
Traditional	0.704	0.600	0.652	0.568
Standardized-residual	0.836	0.752	0.816	0.704
Sign	0.500	0.212	0.404	0.120
Ordinary cross-sectional	0.508	0.244	0.364	0.168
Method-of-moments	0.492	0.356	0.464	0.300
Standardized cross-sectional	0.640	0.436	0.552	0.328

[a]The traditional test statistic equals the sum of the event-period abnormal returns divided by the square root of the sum of all securities' estimation-period residual variances.

[b]The standardized-residual test statistic equals the sum of the residuals standardized by their standard deviations divided by the (approximate) square root of the number of sample firms.

[c]The sign test statistic is the observed proportion of positive returns minus 0.50 divided by the standard deviation of a binomial distribution.

[d]The ordinary cross-sectional test statistic divides the average event-period residual by its contemporaneous cross-sectional standard error.

[e]For the method-of-moments test statistic, the average industry residual is calculated. Then the residual is standardized by its standard error and divided by the square root of the number of industries.

[f]For the standardized cross-sectional test statistic, the residuals are standardized by their standard deviations. Then the average event-period standardized residual is divided by its contemporaneous cross-sectional standard error.

4.4. Summary of simulation results

Overall, our suggested standardized cross-sectional procedure works well. Our simulations lead to the following conclusions:

- In the presence of event-induced variance but zero average abnormal performance, our suggested standardized cross-sectional test rejects the null hypothesis at about the appropriate significance level, as do the sign and the ordinary cross-sectional tests (i.e., they have the right size). The traditional and standardized-residual tests result in too-frequent rejection of the null, while the method-of-moments test performs well at low levels of event-induced variance but poorly as event-induced variance increases.
- When average abnormal performance is 1%, the presence of event-induced variance lowers the power of all of the tests to identify abnormal performance. The standardized-residual test is the most powerful in rejecting the null but the standardized cross-sectional test generally rejects the null more frequently than the other tests.
- When average abnormal performance averages 2%, all of the tests except the method-of-moments method reject the null hypothesis close to 100% of the time.
- The results are essentially unaffected by the presence of event-date clustering.

5. Conclusions

We have demonstrated that traditional event-study methods too frequently reject a null hypothesis of zero abnormal performance if the event itself causes additional variance of event-period returns. The two most common remedies to this problem are to use an ordinary cross-sectional test or to use a sign test in conjunction with a parametric test. We suggest an alternative, easy-to-use test. We find that if event-period returns are normalized and a cross-sectional test is then applied to these standardized residuals, the results are better than with the two common approaches: too frequent rejections of true nulls are avoided without significantly reducing the test's power. In addition, we show that event-date clustering does not affect our results.

Appendix

Throughout this appendix we use the following notation:

N = number of firms in the sample,
A_{iE} = security i's abnormal return on the event day,
A_{it} = security i's abnormal return on day t,
T_i = number of days in security i's estimation period (we omit the subscript i when there is no possible confusion),
R_{mt} = market return on day t,
$\bar{R}_m$ = average market return during the estimation period,
$\hat{s}_i$ = security i's estimated standard deviation of abnormal returns during the estimation period,
SR_{iE} = security i's standardized residual on the event day

$$= A_{iE} \Bigg/ \hat{s}_i \sqrt{1 + \frac{1}{T_i} + \frac{(R_{mE} - \bar{R}_m)^2}{\sum_{t=1}^{T_i} (R_{mt} - \bar{R}_m)^2}}\,, \tag{1}$$

J = number of industries per portfolio,
N_j = number of firms in industry j,
$\bar{A}$ = average portfolio abnormal return.

All of the test statistics described below assume that the null distribution is normal with mean zero and variance equal to one.

The *traditional method* [Brown and Warner (1980, app. A.3)] implicitly assumes that security residuals are uncorrelated and that event-induced variance is insignificant. Its test statistic equals

$$\frac{1}{N}\sum_{i=1}^{N} A_{iE} \Bigg/ \frac{1}{N}\sqrt{\sum_{i=1}^{N} \frac{1}{T-1} \sum_{t=1}^{T} \left(A_{it} - \sum_{t=1}^{T} \frac{A_{it}}{T}\right)^2}\,. \tag{2}$$

The *standardized-residual method* [Patell (1976)] normalizes the residuals before forming portfolios. Its test statistic equals

$$\sum_{i=1}^{N} SR_{iE} \Bigg/ \sqrt{\sum_{i=1}^{N} \frac{T_i - 2}{T_i - 4}}\,. \tag{3}$$

The *sign test* is a simple binomial test of whether the frequency of positive residuals equals one half. For the two-sided test, its test statistic equals

$$|P-\tfrac{1}{2}|\left[\frac{(\frac{1}{2})^2}{N}\right]^{-1/2}, \tag{4}$$

where P is the frequency of positive residuals. For the one-sided test, the test statistic is

$$(P-\tfrac{1}{2})\left[\frac{(\frac{1}{2})^2}{N}\right]^{-1/2}. \tag{5}$$

The *ordinary cross-sectional method* ignores estimation-period estimates of variance and uses the event-day cross-sectional standard deviation for its t-test. The resulting t-statistic is

$$\frac{1}{N}\sum_{i=1}^{N}A_{iE}\Bigg/\sqrt{\frac{1}{N(N-1)}\sum_{i=1}^{N}\left(A_{iE}-\sum_{i=1}^{N}\frac{A_{iE}}{N}\right)^2}. \tag{6}$$

Our hybrid of the standardized-residual and the ordinary cross-sectional approach, the *standardized cross-sectional method*, first finds standardized residuals as Patell did, then applies the ordinary cross-sectional technique just described. Our test statistic is

$$\frac{1}{N}\sum_{i=1}^{N}SR_{iE}\Bigg/\sqrt{\frac{1}{N(N-1)}\sum_{i=1}^{N}\left(SR_{iE}-\sum_{i=1}^{N}\frac{SR_{iE}}{N}\right)^2}. \tag{7}$$

Froot's (1989) *method-of-moments estimator* requires residuals across industries to be independent. Our test statistic differs from Froot's in that it does not restrict the number of firms per industry to be the same across industries. It is given by

$$\frac{1}{\sqrt{J}}\sum_{j=1}^{J}\frac{1}{N_j}\sum_{i=1}^{N_j}A_{ijE}\Bigg/\sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(\frac{1}{N_j}\sum_{i=1}^{N_j}A_{ijt}-\sum_{t=1}^{T}\sum_{i=1}^{N_j}\frac{A_{ijt}}{TN_j}\right)^2}. \tag{8}$$

Ball and Torous (1988) derive a *maximum-likelihood estimator* (MLE) that requires residuals to be normally distributed. For the special case of no event-date uncertainty, the marginal distribution of the standardized event-day abnormal return is

$$f\left(\frac{A_{iE}}{\hat{s}_i}\right) = \frac{1}{\sqrt{2\pi\delta^2}} \exp\left[-\left(\frac{A_{iE}}{\hat{s}_i} - \bar{A}\right)^2 \Big/ 2\delta^2\right], \tag{9}$$

where $\bar{A}$ is a measure of the standardized abnormal return on the event day and δ^2 is the variance change on the event day. $\bar{A}$ and δ^2 are the parameters to be estimated. Since all standardized event-day abnormal returns are independent and identically distributed as a normal random variable $[N(\bar{A}, \delta^2)]$, the estimator for $\bar{A}$ is

$$\bar{A} = \frac{1}{N} \sum_{i=1}^{N} \frac{\bar{A}_{iE}}{\hat{s}_i}. \tag{10}$$

The test statistic for $\bar{A}$ is identical to that for the standardized cross-sectional test without adjustment for the out-of-sample forecast error.

References

Ball, Clifford and Walter Torous, 1988, Investigating security price performance in the presence of event-date uncertainty, Journal of Financial Economics 22, 123–154.

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

Brown, Keith, W.V. Harlow, and Seha M. Tinic, 1988, Risk aversion, uncertain information, and market efficiency, Journal of Financial Economics 22, 355–385.

Brown, Keith, W.V. Harlow, and Seha M. Tinic, 1989, The effect of unanticipated events on the risk and return of common stock, Working paper (University of Texas, Austin, TX).

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

Brown, Stephen and Jerold Warner, 1985, Using daily stock returns: The case of event studies, Journal of Financial Economics 14, 3–32.

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

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

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

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

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

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

Fama, Eugene F., Lawrence Fisher, Michael Jensen, and Richard Roll, 1969, The adjustment of stock prices to new information, International Economic Review 10, 1–21.

Froot, Kenneth, 1989, Consistent covariance matrix estimation with cross-sectional dependence and heteroskedasticity in financial data, Journal of Financial and Quantitative Analysis 24, 333–355.

Judge, George G., R. Carter Hill, William E. Griffiths, Helmut Lutkepohl, and Tsoung-Chao Lee, 1988, Introduction to the theory and practice of econometrics (Wiley, New York, NY).

Kalay, Avner and Uri Lowenstein, 1985, Predictable events and excess returns: The case of dividend announcements, Journal of Financial Economics 14, 423–450.

Lummer, Scott and John McConnell, 1989, Further evidence on the bank lending process and the capital market response to bank loan agreements, Journal of Financial Economics 25, 99–122.

McConnell, John and Chris Muscarella, 1985, Corporate capital expenditure decisions and the market value of the firm, Journal of Financial Economics 14, 399–422.

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

Mitchell, Mark and Jeffry Netter, 1989, Triggering the 1987 stock market crash: Antitakeover provisions in the proposed House Ways and Means tax bill?, Journal of Financial Economics 24, 37–68.

Patell, James, 1976, Corporate forecasts of earnings per share and stock price behavior: Empirical tests, Journal of Accounting Research, 246–276.

Patell, James and Mark Wolfson, 1979, Anticipated information releases reflected in call option prices, Journal of Accounting and Economics 1, 117–140.

Penman, Stephen, 1982, Insider trading and the dissemination of firms' forecast information, Journal of Business 55, 479–503.

Rosenstein, Jeffrey and Stuart Wyatt, 1990, Outside directors, board independence and shareholder wealth, Journal of Financial Economics 26, 175–192.