Nonparametric Event Study Tests*

Arnold R. Cowan
Department of Finance
College of Business
Iowa State University
Ames, Iowa 50011–2063
515-294-9439/Fax 515-294-3525
arnie@iastate.edu

Review of Quantitative Finance and Accounting 2 (1992):343–358
(December 1992)

This paper provides the first documentation of the power and specification of the generalized sign test, which is based on the percentage of positive abnormal returns in an estimation period. In simulations using daily stock return data, the generalized sign test is well specified with both exchange listed and NASDAQ stocks. A rank test is more powerful under ideal conditions. However, the rank test is more sensitive to increases in the length of the event window, to increases in return variance and to thin trading. The generalized sign test is a viable alternative to the rank test under these conditions.

*Comments by Chris Coyne, Rick Dark, Nandu Nayar, Michael Pescow, Ajai Singh and *RQFA* referees Victor Bernard and Charles Jacklin helped to improve the paper.

1. Introduction

Event studies measure stock price reactions to events. Price reactions are represented by abnormal returns, which are stock returns adjusted for normal daily stock price and market index movements. Researchers examine test statistics to infer whether to attribute observed abnormal returns to chance or to the event under investigation. Many event studies rely on parametric test statistics. However, a disadvantage of parametric statistics is that they embody detailed assumptions about the probability distribution of returns.

In addition to parametric statistics, event studies typically report a nonparametric test. Nonparametric statistics do not require as stringent assumptions about return distributions as parametric tests. The sign test is a nonparametric test often used in event studies.[1] However, the sign test judges the proportion of positive and negative abnormal returns against an assumed 50 percent split under the null hypothesis of no reaction to the event. Brown and Warner (1980, p.220; 1985, p.24) observe that correct specification of the sign test requires equal numbers of positive and negative abnormal returns, absent a reaction to an event. Cowan, Nayar and Singh (1990) and Sanger and Peterson (1990) use a variation of the sign test called the generalized sign test. The generalized sign test compares the proportion of positive abnormal returns around an event to the proportion from a period unaffected by the event. In this way, the generalized sign test takes account of a possible asymmetric return distribution under the null hypothesis. The power and

[1] Some recent examples are Kim and Schatzberg (1987); Doukas and Travlos (1988); Jarrell and Poulsen (1988); Ryngaert (1988); Hite and Vetsuypens (1989); Agrawal and Mandelker (1990); Chan, Martin and Kensinger (1990); Loderer and Martin (1990); and McWilliams (1990).

specification of the generalized sign test have not been documented.

Corrado (1989) reports that another nonparametric test, the rank test, accords more power to detect abnormal stock price changes than standard parametric tests. Like the generalized sign test, the rank test does not require symmetry of the cross-sectional abnormal return distribution. Corrado (1989) reports simulations of the rank test using NYSE and AMEX stock data and single day event windows.

In this paper, I investigate several issues that previous research does not consider. I report the performance of the generalized sign test in extensive simulations with actual stock return data. The simulations provide the first empirical evidence of the power and specification of the generalized sign test. A second issue is the relative power of the generalized sign and rank tests to detect abnormal returns. All simulations include both tests to allow a direct comparison.

Third, the paper examines the power and specification of both tests when the stock return variance increases around the event date. According to Brown and Warner (1985), event-related variance increases cause standard parametric tests to report a price reaction where none actually exists more often than expected. Nonparametric tests, which do not use the return variance, may perform better under variance increases than parametric tests that assume stable variances.

Fourth, I consider the performance of the tests when the sample includes an outlier that has a large abnormal return while the rest of the stocks have no abnormal return. A common application of the traditional sign test in event studies is to verify that the parametric findings do not result from an outlier.[2] The outlier is

[2] For example, Schipper and Smith (1983, p. 446) state, "These results are not driven by a few outliers.... A binomial test of the null hypothesis that half the cumulative prediction errors are positive by chance yields a z-statistic of +3.34, rejecting the null hypothesis at the .01 level."

a special case of an event date variance increase. Thus nonparametric tests again may be less sensitive to this circumstance than parametric tests.

The fifth issue is the extension of the tests to event windows longer than one day. Corrado (1989) reports the rank test for abnormal return on one day, but actual event studies often test windows of two or more days. Finally, the paper includes NASDAQ stocks in addition to exchange listed stocks. Nonparametric tests may be especially appropriate for NASDAQ stocks because thin trading makes violations of parametric test assumptions more likely.

The results show that the rank test is misspecified for NASDAQ stocks while the generalized sign test is correctly specified. Both tests are correct with NYSE and AMEX stocks, but the rank test supplies more power to detect abnormal performance in one and two day event windows. The generalized sign test becomes relatively more powerful as the length of the event window increases.

The generalized sign test is correctly specified when the variance of the stock return increases during the event window. The generalized sign test provides more power than a parametric test based on standard errors from the cross-section of event date abnormal returns. In contrast, the rank test rejects true null hypotheses more often than the nominal significance level in the case of increased variance.

The generalized sign test maintains low rejection rates under the null hypothesis after the introduction of an extreme abnormal return for a single stock. The rank test is more sensitive to the extreme return, rejecting the null marginally more often than the nominal significance level in the single day window.[3]

McWilliams (1990, p. 1632) writes, "Since the t-statistic can be sensitive to extreme values, tests of statistical significance are also performed using a nonparametric sign test."

[3]This paper does not examine the case of clustered event dates. As Bernard (1989) notes,

2. Test Procedures

2..1. Abnormal Returns

Results are reported for abnormal returns based on the market model,

$$R_{jt} = \alpha_j + \beta_j R_{mt} + \epsilon_{jt},$$

where R_{jt} is the rate of return of the j^{th} stock on day t and R_{mt} is the rate of return of a market index on day t. ϵ_{jt} is a random variable that has an expected value of zero, is not correlated with $R_{it,i\neq j}$, is not autocorrelated, and has constant variance.

The abnormal return for the j^{th} common stock on day t is

$$AR_{jt} = R_{jt} - (\widehat{\alpha}_j + \widehat{\beta}_j R_{mt}),$$

where $\widehat{\alpha}_j$ and $\widehat{\beta}_j$ are ordinary least squares estimates of α_j and β_j. The parameter estimation period contains 100 days, E_1 through E_{100}, immediately preceding the first day of the 11-day event period.

The cumulative abnormal return for stock j over an event window, days D_1 through D_d, is

$$CAR_{j,(D_1,D_d)} = \sum_{t=D_1}^{D_d} AR_{jt}.$$

The cumulative average abnormal return for a sample of n stocks over the event window is

$$CAAR_{D_1,D_d} = \frac{1}{n}\sum_{j=1}^{n} CAR_{j,(D_1,D_d)}.$$

---

there are relatively few event studies with clustered event dates. It is unlikely that a researcher would use these tests on clustered data, as they do not account for cross-sectional dependence and therefore would certainly be misspecified.

The null and alternative hypotheses of interest are

$$H_0 : CAAR_{D_1,D_d} \leq 0$$
$$H_A : CAAR_{D_1,D_d} > 0.$$

The generalized sign test is more properly interpreted as a test of the median CAR. In practice, researchers often use sign and related tests to support conclusions about the mean, $CAAR$. The simulations in this paper address the suitability of the generalized sign test for this purpose. I also report results when the hypotheses of interest are

$$H_0 : CAAR_{D_1,D_d} \geq 0$$
$$H_A : CAAR_{D_1,D_d} < 0.$$

2..2. Generalized Sign Test

The generalized sign test examines whether the number of stocks with positive cumulative abnormal returns in the event window exceeds the number expected in the absence of abnormal performance. The number expected is based on the fraction of positive abnormal returns in the 100 day estimation period,

$$\hat{p} = \frac{1}{n}\sum_{j=1}^{n}\frac{1}{100}\sum_{t=E_1}^{E_{100}} S_{jt},$$

where

$$S_{jt} = \begin{cases} 1 & \text{if } AR_{jt} > 0 \\ 0 & \text{otherwise} \end{cases}$$

The test statistic uses the normal approximation to the binomial distribution with parameter $\hat{p}$. Define w as the number of stocks in the event window for which the cumulative abnormal return $CAR_{j,(D_1,D_d)}$ is positive. The generalized sign test statistic is

$$Z_G = \frac{w - n\hat{p}}{\left[n\hat{p}\left(1-\hat{p}\right)\right]^{\frac{1}{2}}}. \tag{1}$$

2..3. Rank Test

Corrado (1989) describes the rank test for a one-day event window. The ranks of the abnormal returns of different days are dependent by construction. However, the effect of ignoring the dependence may be negligible for event windows of a few days. This paper extends the rank test to windows of two to 11 days by assuming that the daily return ranks within the window are independent.

The rank test procedure treats the 100 day estimation period and the 11 day event period as a single time series, and assigns a rank to each daily return for each firm. Following the notation of Corrado (1989), let K_{jt} represent the rank of abnormal return AR_{jt} in the time series of 111 daily abnormal returns of stock j. Rank one signifies the smallest abnormal return; rank 111, the largest. The mean rank is 56. The rank test statistic for the event window composed of days D_1 through D_d is

$$Z_R = d^{\frac{1}{2}} \frac{\overline{K_D} - 56}{\left[\sum_{t=1}^{111}\left(\overline{K_t} - 56\right)^2 /111\right]^{\frac{1}{2}}}. \tag{2}$$

$\overline{K_D}$ is the average rank across the n stocks and d days of the event window ($1 \leq d \leq 11$) and $\overline{K_t}$ is the average rank across n stocks on day t of the 111 day combined estimation and event period.[4]

[4]Note that the expected rank still is 56 for event windows shorter than 11 days, because the full 111 day time series still is used for the assignment of ranks.

2..4. Sample Construction

The simulation samples come from the CRSP 1990 daily NYSE-AMEX and NASDAQ files. The NYSE-AMEX file contains returns for all firms listed on the New York Stock Exchange or the American Stock Exchange from July 1962 through December 1990. The NASDAQ file contains returns from the inception of NASDAQ in December, 1972 through December, 1990. To ensure sufficient data, I restrict the simulations to stocks that have at least 260 trading days of returns on the NYSE-AMEX or NASDAQ file.

I randomly select 50,000 NYSE-AMEX stocks with replacement from the 5,640 that meet the 260 trading day requirement. The probability of selection depends on the amount of data available. The purpose is to make the sample similar in firm longevity and size to samples in actual event studies. [See Kothari and Wasley (1989) for evidence of the effects of firm size on event study methods.] I assign the stocks to quintiles based on the number of trading days reported. Each stock in the quintile with the greatest number of trading days has a five in 16,918 chance of selection on any one drawing. Each in the second greatest quintile has a four in 16,918 chance, and so on. A stock in the middle quintile has a three in 16,918 or one in 5,640 chance of selection, the same probability that any stock would have under uniform random selection. The quintiles affect only the construction of the sample of 50,000, not the subsequent formation of portfolios. A parallel set of 50,000 NASDAQ stocks is formed by the quintile method. The CRSP NASDAQ file contains 8,989 stocks that have at least 260 trading days of data.

For each stock selected, I randomly choose an event date with replacement from the range of dates that returns are available for the stock. To allow for estimation and cumulation, the event date must occur after the first 100 and before the last 11 trading days of data for the stock. Within this constraint, each day has an equal probability of selection. If a stock has more than ten missing returns in the estimation period, or any missing return in the 11 day event period, I eliminate it from the sample. Another randomly selected stock and event date take the place of the eliminated observation. The stocks are grouped into 1000 NYSE-AMEX portfolios and 1000 NASDAQ portfolios. Each portfolio consists of 50 stocks.

The sample selection procedure should produce samples similar to those that would be obtained by the method Brown and Warner (1985) use. They initially allow each stock on the CRSP file an equal probability of selection. However, Brown and Warner do not condition the choice of date on the range of data available for the stock. Stocks that are listed for less than the full period can randomly receive a date outside their trading period and thus be excluded. For example, a stock with returns reported for only one-fifth of the dates on the full CRSP file has only one-fifth the probability of entering the final sample as a stock with data for the full period.

I examine event windows of one, two, five and 11 days. To simulate abnormal performance, I add a constant to the abnormal return of one day during the event window. For the two, five and 11 day windows, the day within the window is randomly selected. The levels of abnormal performance simulated are positive and negative 0.5 percent, one percent and two percent.

3. Properties of Individual Stock Returns

Table 1 reports the properties of daily raw and abnormal returns in the 100 day estimation period for the samples of 50,000 common stocks. Individual stock returns and abnormal returns are highly non-normal for both the NYSE-AMEX and NASDAQ samples. The studentized range statistic is sensitive to departures from normality. The one percent, two tailed critical values of the studentized range from a sample of size 100 are 4.02 and 6.54 [David, Hartley and Pearson (1954)]. A value outside the range between the critical values dictates rejection of the null hypothesis that the sample comes from a normal distribution. In 50,000 independent repetitions of the studentized range test, 500 rejections of the null hypothesis should occur by chance. The NYSE-AMEX sample tests produce 17,373 rejections of the null hypothesis for the raw returns and 16,567 rejections for the abnormal returns. The NASDAQ stock returns deviate more frequently from the normal distribution, with 32,032 and 31,269 rejections for raw and abnormal returns respectively.

The departures from normality consist of both right skewness and leptokurtosis (fat tails). The results are consistent with previous studies of stock return distributions [see Fama (1976) and Brown and Warner (1985)]. However, the NASDAQ stock returns are considerably more leptokurtotic than the returns of NYSE-AMEX stocks. This suggests that parametric tests based on the assumption of normally distributed raw or abnormal returns may be less well specified for NASDAQ stocks than previous studies report for NYSE and AMEX stocks.

The fraction of positive raw returns in the estimation period for NASDAQ stocks is 22.7 percent. The low number of positive returns is due to many zero returns

Table 1
Descriptive statistics of daily return and abnormal return for individual common stocks.[a]

Type of Return	Mean	Standard Deviation	Coefficient of Skewness	Coefficient of Kurtosis	Student-ized Range	% Days Return > 0[b]
			NYSE-AMEX Portfolios			
Raw	0.00068	0.02749	0.49193	3.31626	6.30347	37.2%
Abnormal	−0.00000	0.02591	0.51092	3.22653	6.25709	47.0%
			NASDAQ Portfolios			
Raw	0.00050	0.03469	0.61180	9.58241	7.34247	22.7%
Abnormal	−0.00000	0.03381	0.64159	9.33155	7.29384	46.6%

[a]For each parameter, the table reports the mean of the 50,000 estimates. Each estimate is calculated from the 100 day estimation period with a minimum of 90 non-missing returns. The market index is the equally weighted NYSE-AMEX index. The stocks and event dates are randomly selected from the period 1962–1990 for NYSE-AMEX and 1973–1990 for NASDAQ.
[b]The percent of days with positive returns in the 100 day estimation period.

rather than a high number of negative returns. Raw returns equal to zero account for 53.2 percent of estimation period returns across the 50,000 NASDAQ stocks. Only 23.3 percent of NYSE and AMEX stock returns are zero. The large fraction of zero returns on the NASDAQ file occurs because of infrequent trading. The CRSP files report returns based on the average of bid and ask quotes when no trade price is available. A zero return occurs when there is no trading on two consecutive days and market makers do not adjust their bid and ask quotes. A high fraction of zero returns may increase the power of the tests to detect abnormal returns that are present. The addition of a positive or negative amount to the normal return will be easier to detect when the normal return is zero.

4. Properties of Portfolio Returns with no Abnormal Performance

Table 2 details the cross-sectional properties of the 1000 event date portfolio mean raw returns and mean abnormal returns when no abnormal return is added. Portfolio returns and abnormal returns are right skewed and leptokurtotic relative to the normal distribution, but less seriously than individual stock returns. The studentized range should be within the range of 5.50 to 7.99 (for a sample of 1000) to avoid rejecting the null hypothesis of normally distributed portfolio returns at the one percent level. The studentized range statistics for portfolio returns and abnormal returns are within the critical values. The results indicate that the portfolio returns are close to normally distributed. Brown and Warner (1985) reach a similar conclusion.

5. Properties of Test Statistics with no Abnormal Performance

Table 3 reports the mean, variance and coefficients of skewness and kurtosis of the empirical distributions of the test statistics when no abnormal performance is added, as well as a test of standard normality. The test of standard normality is the Kolmogorov-Smirnov goodness of fit test [see Neave and Worthington (1988), p. 89–90]. The Kolmogorov-Smirnov test compares the empirical cumulative distribution function to the hypothesized distribution.

Table 2
Descriptive statistics of cross-sectional distribution of portfolio mean event date return and abnormal return across 1000 portfolios of 50 stocks each.

Type of Return	Mean	Median	Standard Deviation	Coefficient of Skewness	Coefficient of Kurtosis	Student-ized Range
			NYSE-AMEX Portfolios			
Raw	0.00065	0.00050	0.00447	0.34392[a]	0.81411[a]	7.67187
Abnormal	−0.00002	−0.00014	0.00430	0.36790[a]	0.96973[a]	7.62644
			NASDAQ Portfolios			
Raw	0.00074	0.00047	0.00646	0.37710[a]	1.15540[a]	7.00929
Abnormal	0.00011	−0.00015	0.00641	0.34757[a]	1.12968[a]	6.83038

[a] The difference between the parameter estimate and the expected value under standard normality is significant at the one percent level.

The properties of the generalized sign statistic closely match those of the standard normal distribution. The mean, skewness and kurtosis are within the one percent critical values for the standard normal. The standard deviation is within the one percent limits except in the two day event window for both NYSE-AMEX and NASDAQ portfolios, where it is less than the value of one expected under normality. Even in the two day window, the Kolmogorov-Smirnov test does not reject the null. Thus, the generalized sign test statistic appears to follow the standard normal distribution.

Table 3
Descriptive statistics and a test of standard normality for the generalized sign and rank event study test statistics.

Test	Event Window (Days)	Mean	Standard Deviation	Coefficient of Skewness	Coefficient of Kurtosis	K-S Test[a]
			NYSE-AMEX Portfolios			
Generalized Sign	1	−0.020	0.947	0.107	−0.160	0.029
	2	−0.049	0.933[b]	−0.010	0.139	0.046
	5	0.079	0.997	−0.117	−0.091	0.046
	11	−0.003	0.963	0.096	0.044	0.031
Rank	1	−0.052	0.999	0.029	−0.137	0.029
	2	−0.066	0.948	0.040	0.270	0.046
	5	−0.059	0.910[b]	−0.141	0.059	0.043
	11	−0.053	0.965	−0.038	0.168	0.052[c]
			NASDAQ Portfolios			
Generalized Sign	1	0.020	0.956	0.143	−0.069	0.023
	2	0.060	0.940[b]	0.116	−0.019	0.042
	5	0.042	0.982	−0.025	0.186	0.042
	11	0.039	1.014	0.015	−0.022	0.030
Rank	1	0.001	1.098[b]	0.018	−0.023	0.028
	2	−0.041	1.104[b]	0.080	0.005	0.039
	5	−0.089[b]	1.180[b]	−0.006	0.113	0.048
	11	−0.041	1.154[b]	0.032	−0.137	0.060[c]

[a]The Kolmogorov-Smirnov (K-S) test assesses the goodness of fit of a normal distribution with zero mean and unit variance to the empirical cumulative distribution function.
[b]The difference between the parameter estimate and the expected value under standard normality is significant at the one percent level.
[c]The K-S test rejects the null hypothesis that the event study test statistic follows the standard normal distribution at the one percent level of significance.

For NYSE-AMEX portfolios in the one and two day event windows, the rank test also is close to standard normal. The parameter estimates in the one day window closely resemble those reported by Corrado (1989). The only time the Kolmogorov-Smirnov test indicates that the empirical distribution function differs from the standard normal is in the 11 day event window. With NASDAQ portfolios, the standard deviation of the rank statistic significantly exceeds one for all event windows. This implies that the use of the rank test with NASDAQ stocks may lead to excessive rejection of the null hypothesis. However, the Kolmogorov-Smirnov test detects a departure from standard normality only in the 11 day window.

6. Power and Specification of the Tests in Simulation

6..1. Random Event Dates

Table 4 reports the frequencies with which the generalized sign and rank tests based on a 100 day estimation period reject the null hypothesis at the one percent level of significance when varying amounts of abnormal performance are added to the event window return.

When no abnormal performance is added, both tests reject the null hypothesis close to one percent of the time. I test the difference between the reported rejection rates under the null hypothesis and one percent using the normal approximation to the binomial distribution. The test statistic is

$$Z = \frac{f - .01}{\left[.01\,(.99)\,/1000\right]^{\frac{1}{2}}},$$

where f is the relative rejection frequency in simulation. The rejection rates of the generalized sign and rank tests under the null do not differ significantly from one percent.

Table 4 also shows that the rank test is more powerful than the generalized sign test in the one and two day event windows. However, the power of the rank test drops off more rapidly than that of the sign test as the length of the event window increases. For each stock, the rank of the average day in a long event window does not stand out as much as a single day with abnormal performance. Thus, the mean

Table 4
Rejection frequencies at the 1% nominal significance level of generalized sign and rank tests in simulations using NYSE and AMEX stocks.

Percentage of 1000 portfolios where the null hypothesis is rejected[a]					
Test	Days	Absolute value of abnormal return added			
		0.0%	0.5%	1.0%	2.0%
Generalized Sign	1	0.9%	36.9%	90.5%	100.0%
		(0.5%)	(32.4%)	(81.0%)	(99.8%)
Generalized Sign	2	0.6%	14.6%	56.3%	98.7%
		(0.9%)	(13.2%)	(49.0%)	(96.8%)
Generalized Sign	5	1.0%	5.1%	21.5%	74.8%
		(0.7%)	(4.7%)	(17.2%)	(65.7%)
Generalized Sign	11	0.8%	3.3%	10.0%	38.5%
		(1.1%)	(2.7%)	(8.0%)	(27.3%)
Rank	1	1.1%	43.6%	95.2%	100.0%
		(0.8%)	(46.8%)	(94.7%)	(100.0%)
Rank	2	0.6%	16.6%	62.7%	99.2%
		(0.8%)	(20.7%)	(67.7%)	(99.1%)
Rank	5	0.6%	7.2%	20.7%	56.9%
		(1.0%)	(7.9%)	(21.9%)	(60.9%)
Rank	11	0.5%	2.0%	5.0%	13.5%
		(0.7%)	(3.0%)	(7.7%)	(18.4%)

[a]Lower-tail rejection frequencies when negative abnormal performance is added appear in parentheses.

rank in the event window is diluted. The generalized sign test does not suffer as much dilution because it considers only whether the event window as a whole has a positive (or negative) abnormal return for each stock. Consequently, in the 11 day window, the generalized sign test is more powerful overall than the rank test.[5]

Table 5 reports the rejection frequencies of the generalized sign and rank tests in simulations using portfolios of NASDAQ stocks. With no abnormal return, the generalized sign test does not reject the null hypothesis significantly more or less often than the nominal significance level of one percent. When abnormal perfor-

[5]The conclusions are similar when the nominal significance level of the event study tests is five percent. The details are reported in an appendix available from the author.

Table 5
Rejection frequencies at the 1% nominal significance level of generalized sign and rank tests in simulations using NASDAQ stocks.

Percentage of 1000 portfolios where the null hypothesis is rejected[a]					
Test	Days	Absolute value of abnormal return added			
		0.0%	0.5%	1.0%	2.0%
Generalized Sign	1	0.8%	78.9%	97.7%	100.0%
		(0.2%)	(59.6%)	(92.5%)	(99.4%)
Generalized Sign	2	0.7%	37.4%	75.9%	98.9%
		(0.5%)	(20.5%)	(62.0%)	(93.8%)
Generalized Sign	5	1.1%	8.6%	26.7%	70.6%
		(1.0%)	(5.2%)	(17.9%)	(57.5%)
Generalized Sign	11	1.2%	3.2%	8.1%	27.3%
		(1.0%)	(2.1%)	(5.3%)	(18.9%)
Rank	1	2.1%[b]	91.9%	99.2%	100.0%
		(2.0%)[b]	(91.4%)	(99.3%)	(100.0%)
Rank	2	2.5%[b]	59.4%	86.6%	98.5%
		(1.8%)	(64.3%)	(87.4%)	(99.3%)
Rank	5	2.3%[b]	22.3%	36.1%	58.7%
		(2.5%)[b]	(25.8%)	(40.3%)	(61.3%)
Rank	11	1.9%[b]	9.6%	12.9%	20.1%
		(3.0%)[b]	(12.6%)	(16.5%)	(24.9%)

[a]Lower-tail rejection frequencies when negative abnormal performance is added appear in parentheses.
[b]The rejection rate with no abnormal return added is significantly different from one percent using the normal approximation to the binomial test at the one percent significance level.

mance is added, the power of the generalized sign test is greater with NASDAQ than with NYSE-AMEX portfolios.

The rank test rejects the null significantly more often than one percent in every window. In most cases, the actual rejection frequency is more than twice the nominal significance level of one percent. The rank test therefore is misspecified for NASDAQ samples. A possible reason for the misspecification is the large fraction of zero returns of NASDAQ stocks reported in section 3. Zero returns cause tied ranks and reduce the estimated standard deviation of the ranks for each stock. However, the actual variability of event day ranks across stocks may be greater for NASDAQ because of variability in the number of zero returns. This is consistent with the result from table 3 that the standard deviation of the rank statistic significantly exceeds 1.0 in NASDAQ portfolios. The results suggest that researchers should prefer the generalized sign test over the rank test for use with NASDAQ samples or other samples dominated by thinly traded stocks.[6,7]

A regularity in tables 4 and 5 is that the generalized sign test rejects the null more often when a positive abnormal return is added than when a negative abnormal return is added. The generalized sign test rejects more often in the upper tail because the fraction of positive abnormal returns in the absence of a reaction to an event, $\hat{p}$, is less than 50 percent. When a positive abnormal return is added, assume that all of the stock abnormal returns in a portfolio become positive; when a negative abnormal return is added, assume that all become negative. For $\hat{p} < .5$ the numerator of equation (1) is larger in absolute value in the positive abnormal return case:

$$50 - 50\hat{p} > |0 - 50\hat{p}| \, .$$

Thus the generalized sign test is more powerful in the detection of positive than negative abnormal returns when fewer returns are positive than negative in the estimation period.

[6] The simulations use the equally weighted NYSE-AMEX index with NASDAQ stocks.

[7] The simulations in tables 4 and 5 were repeated with 1000 twenty-stock portfolios. The overall conclusions are unchanged. Both tests are well specified for NYSE-AMEX portfolios, with the rank test more powerful in the shorter windows and the generalized sign test more powerful in the 11 day window. For NASDAQ portfolios, the generalized sign test, but not the rank test, exhibits rejection frequencies under the null hypothesis that are close to the nominal significance level. The power of both tests is less than the power with fifty stock portfolios. The details are reported in an appendix available from the author.

Table 6

Lower-tail rejection frequencies at the 1% nominal significance level of alternative generalized sign test comparing the fraction of negative returns.

Percentage of 1000 portfolios where the null hypothesis is rejected					
Sample	Days	Abnormal return added			
		0.0%	−0.5%	−1.0%	−2.0%
NYSE-AMEX	1	0.5%	31.8%	80.3%	99.6%
NYSE-AMEX	2	0.9%	12.8%	48.5%	96.5%
NYSE-AMEX	5	0.6%	4.6%	17.0%	64.7%
NYSE-AMEX	11	1.0%	2.7%	7.7%	26.7%
NASDAQ	1	0.6%	65.1%	93.5%	99.3%
NASDAQ	2	0.8%	28.2%	68.3%	96.5%
NASDAQ	5	0.7%	6.6%	20.3%	64.4%
NASDAQ	11	0.8%	2.6%	6.3%	24.9%

An alternative lower-tailed test uses the fraction of negative returns in the estimation period. If the fraction of firms with negative returns in the event window exceeds the fraction of negative returns in the estimation period, the null is rejected. Table 6 reports the results for the alternative test. The rejection frequencies with NYSE-AMEX portfolios resemble the rejection frequencies from table 4. The rejection frequencies with NASDAQ portfolios are larger than those in table 5. When the alternative hypothesis is lower-tailed, the comparison of fractions of negative returns rather than positive returns can provide additional power to detect abnormal performance in NASDAQ samples.

The rank test tends to reject the null more often when a negative abnormal return is added, but the pattern is not consistent across all event windows and abnormal return levels.

6..2. Variance Increases during the Event Window

Brown and Warner (1985) document that a variance increase on the event date adversely affects the specification of test statistics based on variance estimates from outside the event window. Tables 7 and 8 report the behavior of the tests using NYSE-AMEX and NASDAQ stocks when I simulate a doubling of the cumulative return variance during the event window. The method of doubling the variance entails adding the abnormal return from the day following the 11 day event period to the event date return. [See Brown and Warner (1985) for a detailed description.] The generalized sign test is correctly specified for all event windows. The rank test exhibits considerable misspecification, rejecting the true null hypothesis too often. The misspecification is especially pronounced with NASDAQ samples.

The poor performance of the rank test results from the fact that ranks by definition consider the magnitude of the abnormal return. When the variance increases, extreme observed abnormal returns become more likely even though there is no change in the mean abnormal performance. The extreme observations receive higher ranks and thus inflate test statistics based on event-window ranks. The generalized sign test does not sustain a similar effect because it does not measure the magnitude of abnormal returns, only their sign. When there has been no shift in the true mean, there should be no change in the observed numbers of positive and negative abnormal returns. The generalized sign test thus appears to be a better choice for both NYSE-AMEX and NASDAQ portfolios when the return variance may increase.

Brown and Warner (1985) report that the use of only the cross-section of event

Table 7
Rejection frequencies at the 1% nominal significance level with increased return variance during the event window in simulations using NYSE and AMEX stocks.

Percentage of 1000 portfolios where the null hypothesis is rejected[a]					
Test	Days	Absolute value of abnormal return added			
		0.0%	0.5%	1.0%	2.0%
Generalized Sign	1	1.2%	13.5%	49.7%	96.4%
		(1.0%)	(9.6%)	(39.2%)	(91.6%)
Generalized Sign	2	0.5%	6.4%	24.3%	78.4%
		(0.6%)	(4.9%)	(18.1%)	(66.3%)
Generalized Sign	5	0.8%	3.4%	9.9%	39.2%
		(0.5%)	(2.7%)	(7.4%)	(29.0%)
Generalized Sign	11	1.3%	2.5%	5.1%	17.5%
		(0.7%)	(1.5%)	(4.3%)	(13.1%)
Rank	1	1.9%[b]	23.2%	69.8%	99.9%
		(4.0%)[b]	(35.1%)	(81.1%)	(99.9%)
Rank	2	0.8%	11.1%	36.6%	89.5%
		(2.3%)[b]	(19.4%)	(47.6%)	(94.0%)
Rank	5	0.8%	4.9%	12.5%	38.8%
		(1.9%)[b]	(6.3%)	(17.3%)	(47.0%)
Rank	11	0.8%	1.7%	3.0%	10.4%
		(1.4%)	(2.5%)	(5.7%)	(13.5%)

[a] Lower-tail rejection frequencies when negative abnormal performance is added appear in parentheses.
[b] Rejection rate with no abnormal return added is significantly different from one percent using the normal approximation to the binomial test at the one percent significance level.

window abnormal returns for variance estimation can provide a better specified parametric test when the variance increases on the event date. However, the cross-sectional procedure lacks power relative to other parametric tests when there is no variance increase. Brown and Warner (table 9, p. 23) report the rejection frequency of the cross-sectional test only at the five percent significance level. With one percent abnormal return added the cross-sectional test rejects the null 80.4 percent of the time when there is no variance increase and 48.8 percent when there is a variance increase. The corresponding results for the generalized sign test, not reported

Table 8
Rejection frequencies at the 1% nominal significance level with increased return variance during the event window in simulations using NASDAQ stocks.

Percentage of 1000 portfolios where the null hypothesis is rejected[a]					
Test	Days	Absolute value of abnormal return added			
		0.0%	0.5%	1.0%	2.0%
Generalized Sign	1	1.7%	40.1%	77.7%	97.8%
		(0.5%)	(19.6%)	(53.2%)	(91.4%)
Generalized Sign	2	1.5%	15.5%	44.7%	86.4%
		(0.4%)	(6.9%)	(24.6%)	(67.6%)
Generalized Sign	5	1.5%	5.5%	15.1%	44.6%
		(0.4%)	(2.3%)	(6.5%)	(25.6%)
Generalized Sign	11	1.3%	3.3%	7.3%	17.2%
		(0.4%)	(1.2%)	(3.7%)	(10.6%)
Rank	1	3.3%[b]	64.5%	90.8%	99.8%
		(5.3%)[b]	(69.8%)	(93.7%)	(99.9%)
Rank	2	1.9%[b]	34.0%	61.8%	93.1%
		(3.2%)[b]	(41.6%)	(72.7%)	(94.5%)
Rank	5	2.0%[b]	13.2%	25.4%	43.7%
		(3.4%)[b]	(17.2%)	(29.9%)	(49.3%)
Rank	11	1.6%	6.3%	10.3%	15.8%
		(3.3%)[b]	(7.8%)	(11.9%)	(19.8%)

[a] Lower-tail rejection frequencies when negative abnormal performance is added appear in parentheses.
[b] Rejection rate with no abnormal return added is significantly different from one percent using the normal approximation to the binomial test at the one percent significance level.

in the tables, are as follows. When the variance does not increase, the rejection frequency is 99.7 percent; when the variance increases, the rejection frequency is 77.0 percent. The results indicate that the generalized sign test can circumvent the trade-off inherent in standard parametric tests between correct specification in the presence of variance increases and power in the absence of increases.

6..3. A Single Outlier

A nonparametric test is sometimes used to check whether parametric test results

Table 9
Percentage of 1000 portfolios with a single outlying stock return where the null hypothesis is rejected at the 1% nominal significance level in simulations using NYSE and AMEX stocks.[a]

Test	Days	Rejection Frequency Upper Tail	Rejection Frequency Lower Tail
Generalized Sign	1	1.3%	0.4%
Generalized Sign	2	0.6%	0.4%
Generalized Sign	5	1.5%	0.3%
Generalized Sign	11	1.0%	0.7%
Rank	1	2.0%[b]	0.3%
Rank	2	0.9%	0.7%
Rank	5	1.0%	0.7%
Rank	11	0.5%	0.5%

[a] An abnormal return of positive 25 percent is added to the event window return of one randomly chosen stock in each 50 stock portfolio. No abnormal return is added to the returns of the other stocks.
[b] Rejection rate with no abnormal return added is significantly different from one percent using the normal approximation to the binomial test at the one percent significance level.

are due to a single outlying stock return. Table 9 provides evidence of the sensitivity of the generalized sign and rank tests to the presence of a single large outlier. Each portfolio contains one stock to which an abnormal return of positive 25 percent is added. No abnormal return is added to the remaining stocks. The generalized sign test continues to be well specified. The rank test rejects the null twice as often as the nominal one percent significance level in the upper tail when the event window is a single day, but otherwise remains well specified. Thus, the generalized sign test is less sensitive than the rank test to the presence of a single outlier, but only in a one day event window, even though the outlying return is extremely large. The empirical results support the theoretical advantage of nonparametric tests over parametric tests in resistance to outliers.

7. Conclusion

The generalized sign test does not require cross-sectional symmetry of the abnormal returns for correct specification. Simulation results for portfolios of actual NYSE-AMEX and NASDAQ stocks show that the generalized sign test is well specified for event windows of one to 11 days. The rank test studied by Corrado (1989) is also correctly specified for NYSE-AMEX stocks, but rejects true null hypotheses more often than the nominal significance level with NASDAQ stocks. The rank test also rejects the null hypothesis too often in the case of an event date variance increase. The generalized sign test is correctly specified for both exchange listed and NASDAQ stocks when the variance increases and is more powerful than the cross-sectional parametric test reported by Brown and Warner (1985). The generalized sign test also appears to be more powerful than the cross-sectional test in the absence of a variance increase.

The generalized sign test continues to be correctly specified under the null hypothesis when a single stock in each portfolio has an extreme positive return. The rank test is more sensitive to an extreme return.

When both tests are correctly specified, the rank test generally provides more power than the sign test to detect an abnormal return that is present in very short event windows. However, the power of the rank test drops off rapidly as the number of days in the event window increases. The generalized sign test thus is better suited to the investigation of cumulative abnormal returns over event windows of several days.

Overall, the rank test performs better than the generalized sign test under ideal conditions. However, when the sample contains thinly traded stocks or the return variance increases on the event date, the generalized sign test will be a better choice.

References

Agrawal, Anup and Gershon N. Mandelker, Large shareholders and the monitoring of managers: The case of antitakeover charter amendments, *Journal of Financial and Quantitative Analysis* 25, 143–161 (1990).

Bernard, Victor L., 1989, Capital markets research in accounting during the 1980s: A critical review, working paper, University of Michigan.

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

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

Chan, Su Han, John D. Martin and John W. Kensinger, Corporate research and development expenditures and share value, *Journal of Financial Economics* 26(2), 255–276 (1990).

Cowan, Arnold R., Nandkumar Nayar and Ajai K. Singh, Stock returns before and after calls of convertible bonds, *Journal of Financial and Quantitative Analysis* 25, 549–554 (1990).

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

David, H.A., H.O. Hartley and E.S. Pearson, The distribution of the ratio, in a single normal sample, of range to standard deviation, *Biometrika* 61, 482–493 (1954).

Doukas, John and Nickolaos G. Travlos, The effect of corporate multinationalism on shareholders' wealth: Evidence from international acquisitions, *Journal of Finance* 43(5), 1161–1175 (1988).

Fama, Eugene F., *Foundations of Finance.* New York: Basic Books, Inc., 1976.

Hite, Gailen L. and Michael R. Vetsuypens, Management buyouts of divisions and shareholder wealth, *Journal of Finance* 44(4), 953–970 (1989).

Jarrell, Gregg A. and Annette B. Poulsen, Dual-class recapitalizations as antitakeover mechanisms: The recent evidence, *Journal of Financial Economics* 20(1/2) (1988), 129–152.

Kim, E. Han and John D. Schatzberg, Voluntary corporate liquidations, *Journal of Financial Economics* 19(2), 311–328 (1987).

Kothari, S.P. and Charles E. Wasley, Measuring security price performance in size-clustered samples, *Accounting Review* 64, 228–249 (1989).

Loderer, Claudio and Kenneth Martin, Corporate acquisitions by listed firms: The experience of a comprehensive sample, *Financial Management* 19(4), 17–33 (1990).

McWilliams, Victoria B., Managerial share ownership and the stock price effects of anti-takeover amendment proposals, *Journal of Finance* 45(5), 1627–1640 (1990).

Neave, Henry R. and P.L. Worthington, *Distribution-free tests.* London: Unwin Hyman, 1988.

Ryngaert, Michael, The effect of poison pill securities on shareholder wealth, *Journal of Financial Economics* 20(1/2), 377–417 (1988).

Sanger, Gary C. and James D. Peterson, An empirical analysis of common stock delistings, *Journal of Financial and Quantitative Analysis* 25(2), 261–272 (1990).

Schipper, Katherine and Abbie Smith, Effects of recontracting on shareholder wealth: The case of voluntary spin-offs, *Journal of Financial Economics* 437–467 (1983).

Appendix A. Rejection frequencies at the .05 significance level

This tables in this appendix report the frequencies at which the generalized sign and rank test reject the null hypothesis of no abnormal performance when the nominal significance level is five percent. Rejection frequencies for the one percent nominal significance level are reported in the text.

Table A.1

Rejection frequencies at the five percent level of generalized sign and rank tests in simulations using NYSE and AMEX stocks.

The generalized sign and rank tests are based on market model abnormal returns. Estimation of the market model parameters over a 100 day period uses the equally weighted index of all NYSE and AMEX common stocks. The stocks and event dates are randomly selected from the period 1963–1990 and grouped into 1000 portfolios of 50 stocks each. Rejection frequencies when negative abnormal performance is added appear in parentheses.

Percentage of 1000 portfolios where the null hypothesis is rejected					
Test	Days	Absolute value of abnormal return added			
		0.0%	0.5%	1.0%	2.0%
Generalized Sign	1	4.7%	66.1%	98.1%	100.0%
		(3.6%)	(59.1%)	(94.3%)	(100.0%)
Generalized Sign	2	3.6%	34.1%	81.2%	99.7%
		(5.3%)	(36.2%)	(75.6%)	(99.2%)
Generalized Sign	5	4.8%	16.5%	44.7%	91.3%
		(4.2%)	(17.6%)	(39.7%)	(85.7%)
Generalized Sign	11	5.5%	14.3%	28.6%	65.3%
		(4.7%)	(11.1%)	(21.4%)	(53.0%)
Rank	1	5.0%	69.9%	99.2%	100.0%
		(5.7%)	(74.5%)	(99.1%)	(100.0%)
Rank	2	3.8%	40.4%	88.1%	99.9%
		(4.9%)	(48.7%)	(89.5%)	(99.9%)
Rank	5	4.5%	21.5%	46.2%	86.3%
		(5.0%)	(22.9%)	(52.7%)	(86.8%)
Rank	11	2.5%[a]	10.0%	22.1%	47.0%
		(4.1%)	(14.1%)	(24.6%)	(47.9%)

[a] The rejection rate with no abnormal return added is significantly different from five percent using the normal approximation to the binomial test at the one percent significance level.

Table A.2

Rejection frequencies at the five percent level of generalized sign and rank tests in simulations using NASDAQ stocks.

The generalized sign and rank tests are based on market model abnormal returns. Estimation of the market model parameters over a 100 day period uses the equally weighted index of all NYSE and AMEX common stocks. The stocks and event dates are randomly selected from the period 1973–1990 and grouped into 1000 portfolios of 50 stocks each. Rejection frequencies when negative abnormal performance is added appear in parentheses.

Percentage of 1000 portfolios where the null hypothesis is rejected					
Test	Days	Absolute value of abnormal return added			
		0.0%	0.5%	1.0%	2.0%
Generalized Sign	1	4.2%	93.9%	99.7%	100.0%
		(4.0%)	(84.1%)	(98.2%)	(100.0%)
Generalized Sign	2	5.1%	62.2%	91.5%	99.9%
		(2.8%)[a]	(49.3%)	(83.5%)	(99.2%)
Generalized Sign	5	6.1%	24.5%	52.3%	88.1%
		(5.0%)	(18.1%)	(41.6%)	(81.7%)
Generalized Sign	11	5.0%	13.0%	23.7%	56.3%
		(4.1%)	(9.5%)	(19.4%)	(46.5%)
Rank	1	5.9%	97.9%	99.9%	100.0%
		(6.8%)[a]	(97.4%)	(99.8%)	(100.0%)
Rank	2	7.0%[a]	82.1%	96.2%	99.9%
		(7.9%)[a]	(83.7%)	(96.9%)	(100.0%)
Rank	5	6.8%[a]	46.6%	64.8%	85.6%
		(9.2%)[a]	(49.5%)	(66.3%)	(86.3%)
Rank	11	7.6%[a]	25.0%	33.6%	45.6%
		(9.4%)[a]	(28.9%)	(39.3%)	(52.8%)

[a] The rejection rate with no abnormal return added is significantly different from five percent using the normal approximation to the binomial test at the one percent significance level.

Table A.3

Rejection frequencies at the five percent level with increased return variance during the event window in simulations using NYSE and AMEX stocks.

The generalized sign and rank tests are based on market model abnormal returns. Estimation of the market model parameters over a 100 day period uses the equally weighted index of all NYSE and AMEX common stocks. The stocks and event dates are randomly selected from the period 1963–1990 and grouped into 1000 portfolios of 50 stocks each. A mean-preserving doubling of the cumulative return variance during the event window is simulated for each stock. Rejection frequencies when negative abnormal performance is added appear in parentheses.

Percentage of 1000 portfolios where the null hypothesis is rejected					
Test	Days	Absolute value of abnormal return added			
		0.0%	0.5%	1.0%	2.0%
Generalized Sign	1	5.1%	33.8%	77.0%	99.3%
		(4.9%)	(26.4%)	(66.9%)	(98.2%)
Generalized Sign	2	4.8%	19.9%	49.0%	93.0%
		(3.7%)	(19.0%)	(42.8%)	(87.0%)
Generalized Sign	5	5.3%	13.1%	29.3%	63.9%
		(4.1%)	(10.5%)	(22.4%)	(56.1%)
Generalized Sign	11	5.1%	9.9%	17.6%	42.6%
		(4.6%)	(8.8%)	(14.4%)	(33.5%)
Rank	1	6.7%	43.6%	85.9%	99.9%
		(12.9%)[a]	(59.0%)	(92.1%)	(100.0%)
Rank	2	4.8%	27.1%	63.1%	98.7%
		(9.1%)[a]	(39.2%)	(73.3%)	(98.8%)
Rank	5	4.8%	14.6%	31.5%	68.6%
		(6.6%)	(20.3%)	(41.0%)	(75.0%)
Rank	11	2.4%[a]	7.8%	15.3%	33.1%
		(5.0%)	(11.0%)	(20.2%)	(38.2%)

[a]Rejection rate with no abnormal return added is significantly different from five percent using the normal approximation to the binomial test at the one percent significance level.

Table A.4

Rejection frequencies at the five percent level with increased return variance during the event window in simulations using NASDAQ stocks.

The generalized sign and rank tests are based on market model abnormal returns. Estimation of the market model parameters over a 100 day period uses the equally weighted index of all NYSE and AMEX common stocks. The stocks and event dates are randomly selected from the period 1973–1990 and grouped into 1000 portfolios of 50 stocks each. A mean-preserving doubling of the cumulative return variance during the event window is simulated for each stock. Rejection frequencies when negative abnormal performance is added appear in parentheses.

Percentage of 1000 portfolios where the null hypothesis is rejected					
Test	Days	Absolute value of abnormal return added			
		0.0%	0.5%	1.0%	2.0%
Generalized Sign	1	4.3%	66.5%	92.5%	99.5%
		(2.8%)[a]	(42.9%)	(77.3%)	(98.7%)
Generalized Sign	2	5.9%	39.2%	69.6%	95.0%
		(2.5%)[a]	(23.1%)	(52.7%)	(87.1%)
Generalized Sign	5	5.4%	17.8%	35.7%	70.8%
		(3.3%)	(10.2%)	(21.9%)	(53.6%)
Generalized Sign	11	5.7%	12.9%	19.8%	41.8%
		(3.5%)	(6.8%)	(12.5%)	(28.9%)
Rank	1	8.9%[a]	82.3%	96.8%	100.0%
		(12.4%)[a]	(85.8%)	(98.6%)	(100.0%)
Rank	2	6.6%	57.8%	84.4%	98.5%
		(9.8%)[a]	(66.2%)	(89.6%)	(98.8%)
Rank	5	6.3%	32.8%	49.2%	72.1%
		(9.8%)[a]	(36.8%)	(54.4%)	(77.5%)
Rank	11	6.7%	18.5%	25.3%	36.6%
		(8.0%)[a]	(23.1%)	(32.2%)	(44.7%)

[a]Rejection rate with no abnormal return added is significantly different from five percent using the normal approximation to the binomial test at the one percent significance level.

Table A.5
Rejection frequencies at the five percent level with clustered event dates in simulations using NYSE and AMEX stocks.

The generalized sign and rank tests are based on market model abnormal returns. Estimation of the market model parameters over a 100 day period uses the equally weighted index of all NYSE and AMEX common stocks. The stocks and event dates are randomly selected from the period 1963–1990 and grouped into 1000 portfolios of 50 stocks each. Within each portfolio all stocks have the same event date. Rejection frequencies when negative abnormal performance is added appear in parentheses.

Percentage of 1000 portfolios where the null hypothesis is rejected					
Test	Days	Absolute value of abnormal return added			
		0.0%	0.5%	1.0%	2.0%
Generalized Sign	1	6.8%[a]	66.4%	95.1%	99.9%
		(7.0%)[a]	(59.9%)	(93.7%)	(99.9%)
Generalized Sign	2	6.6%	41.9%	79.7%	98.0%
		(7.3%)[a]	(38.3%)	(74.7%)	(98.5%)
Generalized Sign	5	6.6%	24.4%	50.8%	87.8%
		(6.8%)[a]	(20.7%)	(44.5%)	(85.6%)
Generalized Sign	11	8.5%[a]	19.3%	33.8%	65.5%
		(5.5%)	(12.6%)	(24.6%)	(55.4%)
Rank	1	5.2%	69.2%	96.8%	100.0%
		(5.7%)[a]	(67.9%)	(97.5%)	(100.0%)
Rank	2	6.3%	44.2%	84.2%	98.3%
		(6.4%)[a]	(43.0%)	(83.5%)	(99.6%)
Rank	5	5.9%	25.0%	48.5%	81.1%
		(8.3%)[a]	(25.6%)	(48.3%)	(82.6%)
Rank	11	4.8%	14.1%	24.8%	47.2%
		(9.9%)[a]	(18.0%)	(28.4%)	(46.2%)

[a] Rejection rate with no abnormal return added is significantly different from five percent using the normal approximation to the binomial test at the one percent significance level.

Table A.6

Rejection frequencies at the five percent level with clustered event dates in simulations using NASDAQ stocks.

The generalized sign and rank tests are based on market model abnormal returns. Estimation of the market model parameters over a 100 day period uses the equally weighted index of all NYSE and AMEX common stocks. The stocks and event dates are randomly selected from the period 1973–1990 and grouped into 1000 portfolios of 50 stocks each. Within each portfolio all stocks have the same event date. Rejection frequencies when negative abnormal performance is added appear in parentheses.

Percentage of 1000 portfolios where the null hypothesis is rejected					
Test	Days	Absolute value of abnormal return added			
		0.0%	0.5%	1.0%	2.0%
Generalized Sign	1	19.0%[a]	88.3%	97.3%	99.6%
		(15.0%)[a]	(78.2%)	(94.6%)	(99.7%)
Generalized Sign	2	12.2%[a]	68.0%	90.2%	97.3%
		(8.5%)[a]	(53.3%)	(82.5%)	(96.7%)
Generalized Sign	5	9.6%[a]	36.1%	61.2%	87.1%
		(6.2%)	(24.1%)	(45.8%)	(77.8%)
Generalized Sign	11	12.0%[a]	24.0%	36.8%	61.4%
		(8.2%)[a]	(15.1%)	(24.8%)	(48.3%)
Rank	1	6.7%	87.7%	97.4%	99.7%
		(5.4%)	(83.9%)	(97.0%)	(100.0%)
Rank	2	7.9%[a]	61.9%	82.0%	94.2%
		(6.5%)	(59.5%)	(81.0%)	(94.0%)
Rank	5	7.4%[a]	32.5%	45.5%	64.1%
		(6.1%)	(33.4%)	(45.2%)	(63.2%)
Rank	11	5.4%	18.8%	24.6%	33.0%
		(6.7%)	(18.4%)	(25.9%)	(36.1%)

[a]Rejection rate with no abnormal return added is significantly different from five percent using the normal approximation to the binomial test at the one percent significance level.

Appendix B. Results for portfolios of twenty stocks each

This tables in this appendix report the frequencies at which the generalized sign and rank test reject the null hypothesis of no abnormal performance at the nominal significance level of one percent using portfolios of twenty stocks. Rejection frequencies using portfolios of fifty stocks are reported in the text.

Table B.1
Rejection frequencies of generalized sign and rank tests in simulations using portfolios of 20 NYSE and AMEX stocks each.

The nominal significance level of the event study tests is one percent. Rejection frequencies when negative abnormal performance is added appear in parentheses.

Percentage of 1000 portfolios where the null hypothesis is rejected					
Test	Days	Absolute value of abnormal return added			
		0.0%	0.5%	1.0%	2.0%
Generalized Sign	1	1.1%	15.3%	46.2%	92.5%
		(0.7%)	(12.1%)	(33.5%)	(78.9%)
Generalized Sign	2	0.4%	6.9%	20.2%	69.6%
		(1.0%)	(5.5%)	(20.4%)	(55.0%)
Generalized Sign	5	1.4%	3.2%	8.0%	32.2%
		(0.7%)	(2.6%)	(7.7%)	(22.2%)
Generalized Sign	11	0.9%	2.2%	4.3%	14.7%
		(0.7%)	(1.9%)	(4.3%)	(10.8%)
Rank	1	1.0%	18.1%	59.4%	98.5%
		(1.4%)	(20.4%)	(58.0%)	(97.8%)
Rank	2	1.2%	7.8%	25.8%	72.6%
		(1.3%)	(9.9%)	(29.4%)	(74.4%)
Rank	5	1.4%	4.0%	8.7%	25.5%
		(1.4%)	(4.7%)	(9.7%)	(25.1%)
Rank	11	0.4%	1.1%	2.2%	7.4%
		(0.7%)	(2.2%)	(4.1%)	(8.3%)

Table B.2

Rejection frequencies of generalized sign and rank tests in simulations using portfolios of 20 NASDAQ stocks each.

The nominal significance level of the event study tests is one percent. Rejection frequencies when negative abnormal performance is added appear in parentheses.

Percentage of 1000 portfolios where the null hypothesis is rejected					
Test	Days	Absolute value of abnormal return added			
		0.0%	0.5%	1.0%	2.0%
Generalized Sign	1	0.4%	38.9%	67.8%	89.7%
		(0.5%)	(25.2%)	(50.9%)	(78.2%)
Generalized Sign	2	0.8%	15.0%	35.9%	67.1%
		(0.9%)	(8.4%)	(24.9%)	(55.9%)
Generalized Sign	5	1.0%	3.3%	9.8%	29.6%
		(0.9%)	(2.6%)	(8.2%)	(22.8%)
Generalized Sign	11	0.5%	0.2%	0.4%	10.8%
		(0.7%)	(1.5%)	(2.9%)	(8.1%)
Rank	1	1.6%	58.5%	81.6%	97.4%
		(1.1%)	(56.8%)	(81.5%)	(98.5%)
Rank	2	2.2%[a]	27.9%	46.7%	74.9%
		(1.5%)	(28.2%)	(48.7%)	(76.3%)
Rank	5	1.6%	11.0%	16.4%	27.1%
		(1.6%)	(11.6%)	(18.4%)	(30.9%)
Rank	11	1.4%	0.5%	6.7%	8.7%
		(2.2%)[a]	(6.5%)	(8.5%)	(11.7%)

[a]The rejection rate with no abnormal return added is significantly different from one percent using the normal approximation to the binomial test at the one percent significance level.

Table C.1
Rejection frequencies with clustered event dates in simulations using NYSE and AMEX stocks.

The generalized sign and rank tests are based on market model abnormal returns. Estimation of the market model parameters over a 100 day period uses the equally weighted index of all NYSE and AMEX common stocks. The stocks and event dates are randomly selected from the period 1963–1990 and grouped into 1000 portfolios of 50 stocks each. Within each portfolio all stocks have the same event date. The nominal significance level of the event study tests is one percent. Rejection frequencies when negative abnormal performance is added appear in parentheses.

Percentage of 1000 portfolios where the null hypothesis is rejected					
Test	Days	Absolute value of abnormal return added			
		0.0%	0.5%	1.0%	2.0%
Generalized Sign	1	1.8%	41.1%	87.1%	99.4%
		(2.1%)[a]	(31.8%)	(81.7%)	(99.5%)
Generalized Sign	2	1.5%	19.4%	58.3%	95.6%
		(1.9%)[a]	(18.0%)	(53.5%)	(94.5%)
Generalized Sign	5	1.9%[a]	8.3%	28.5%	72.0%
		(1.7%)	(6.8%)	(22.1%)	(65.9%)
Generalized Sign	11	2.2%[a]	6.5%	14.1%	45.0%
		(1.9%)[a]	(4.0%)	(9.1%)	(32.9%)
Rank	1	1.0%	41.3%	91.5%	100.0%
		(1.6%)	(43.0%)	(91.9%)	(99.9%)
Rank	2	1.1%	20.9%	61.7%	95.7%
		(2.6%)[a]	(19.6%)	(61.4%)	(97.1%)
Rank	5	1.5%	8.7%	23.2%	54.1%
		(3.1%)[a]	(10.6%)	(24.7%)	(55.4%)
Rank	11	1.0%	3.4%	7.9%	17.9%
		(4.0%)[a]	(8.4%)	(11.9%)	(21.2%)

[a]Rejection rate with no abnormal return added is significantly different from one percent using the normal approximation to the binomial test at the one percent significance level.

Appendix C. Clustered event dates

Tables C.1 (NYSE-AMEX stocks) and C.2 (NASDAQ stocks) report the rejection frequencies when every stock in a single event study portfolio has the same event date. Brown and Warner (1985) refer to the coincidence of event dates as clustering.

In a majority of cases, the rejection rates under the null hypothesis with event date clustering exceed the corresponding rejection rates with random event dates reported in

Tables 4 and 5. Neither test is correctly specified in every window. However, the rank test is better specified overall than the generalized sign test. The rank test does not excessively reject true null hypotheses in the upper tail for any window. The rejection rate is excessive in the lower tail for the two, five and 11 day windows using NYSE-AMEX portfolios. With NASDAQ portfolios, the rejection frequency of the rank test under the null hypothesis does not significantly exceed one percent except in the lower tail for the 11 day event window. In contrast, the rejection frequency of the generalized sign test using NASDAQ portfolios ranges from 2.5 to 9.6 times the nominal significance level in seven of eight cases.

The results indicate that the generalized sign test is misspecified when event dates are clustered, especially when NASDAQ stocks are involved. Consistent with results reported by Corrado (1989), the rank test may be a reasonable choice when the event window is a single day. The rank test does not handle clustered event dates well when the alternative hypothesis is lower tailed and the event window spans two or more days. As Bernard (1989) notes, however, there are relatively few event studies with clustered event dates. Thus, poor specification in this case is not a major deficiency of the nonparametric tests.

Table C.2
Rejection frequencies with clustered event dates in simulations using NASDAQ stocks.

The generalized sign and rank tests are based on market model abnormal returns. Estimation of the market model parameters over a 100 day period uses the equally weighted index of all NYSE and AMEX common stocks. The stocks and event dates are randomly selected from the period 1973–1990 and grouped into 1000 portfolios of 50 stocks each. Within each portfolio all stocks have the same event date. The nominal significance level of the event study tests is one percent. Rejection frequencies when negative abnormal performance is added appear in parentheses.

Percentage of 1000 portfolios where the null hypothesis is rejected					
Test	Days	Absolute value of abnormal return added			
		0.0%	0.5%	1.0%	2.0%
Generalized Sign	1	9.6%[a]	77.6%	93.5%	98.8%
		(6.0%)[a]	(62.0%)	(88.2%)	(98.4%)
Generalized Sign	2	5.6%[a]	49.3%	79.1%	95.1%
		(2.5%)[a]	(31.8%)	(64.2%)	(91.7%)
Generalized Sign	5	3.0%[a]	17.7%	41.1%	76.1%
		(1.7%)	(8.5%)	(25.7%)	(59.2%)
Generalized Sign	11	4.4%[a]	10.2%	20.0%	42.1%
		(3.1%)[a]	(5.9%)	(11.2%)	(25.9%)
Rank	1	1.1%	66.8%	91.2%	98.9%
		(0.5%)	(64.5%)	(88.9%)	(99.1%)
Rank	2	1.7%	34.9%	57.3%	81.9%
		(1.7%)	(34.2%)	(56.5%)	(82.0%)
Rank	5	1.2%	13.3%	21.8%	36.4%
		(1.7%)	(15.4%)	(22.3%)	(36.3%)
Rank	11	1.3%	5.5%	7.6%	11.9%
		(2.1%)[a]	(6.2%)	(8.8%)	(13.3%)

[a]Rejection rate is significantly different from one percent using the normal approximation to the binomial test at the one percent significance level.