Event Studies and Systems Methods: Some Additional Evidence

Bill McDonald


Stable URL:  
http://links.jstor.org/sici?sici=0022-1090%28198712%2922%3A4%3C495%3AESASMS%3E2.0.CO%3B2-P

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

*The Journal of Financial and Quantitative Analysis* is published by University of Washington School of Business Administration. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/uwash.html.

*The Journal of Financial and Quantitative Analysis*  
©1987 University of Washington School of Business Administration

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office.  
For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR

http://www.jstor.org/  
Mon May 19 14:20:40 2003
Event Studies and Systems Methods: Some Additional Evidence

Bill McDonald*

Abstract

This paper extends a recent study by Malatesta [14] on measuring abnormal performance using joint generalized least squares. For monthly data and a random sample of securities, Malatesta finds that there is little benefit in using more sophisticated econometric techniques to identify abnormal returns. The current study extends these results using a design that is more amenable to the benefits of the generalized methods and is consistent with actual event studies. Most notably, the study uses a sample of securities experiencing an actual event and tests both monthly and daily data. In addition, iterative techniques are compared to the ordinary least squares and estimated generalized least squares methods. The results of this study support the original conclusions of Malatesta, indicating no measurable gain in using any of the systems methods for event study applications.

I. Introduction

A variety of more sophisticated econometric techniques have proven ineffective in increasing the power of tests intended to measure abnormal security returns in event studies. However, systems methods that incorporate information about the interrelationships between groups of securities experiencing the same event have a substantial amount of intuitive appeal as an approach that would improve on methods that ignore this information. Binder [2], Schipper and Thompson [15], and Thompson [19] provide rigorous arguments supporting the use of systems methods. Malatesta [14], in a series of simulations, reports that the methods do not appear to provide any benefit in identifying abnormal returns.

Given the intuitive promise of systems methods in event applications, the purpose of this study is to extend the tests of Malatesta to determine if the disappointing results are simply an artifact of the specified experimental conditions. This is done in a number of ways. First, the tests are extended to include daily data, which have become the more common data format in event studies and, as reported in Brown and Warner [5], provide more power in identifying abnormal returns. Second, one set of samples is based on a group of securities experiencing

* College of Business Administration, University of Notre Dame, Notre Dame, IN 46556. This research was funded in part by the Gallo Foundation Fund. Data for the study were made available through the Jesse H. Jones Research Data Base at the University of Notre Dame. Special thanks to Jim Wruck and his staff for facilitating the computational requirements of this research. In addition, the author thanks an anonymous JFQA referee for help in clarifying the presentation and discussion.
an actual event. The appeal of systems methods is based on their ability to incorporate information about contemporaneous covariance, which may not be appropriately modeled for event study applications by a random sample of securities. In addition to these extensions, a testing format is used that is directly comparable across all methods and the design matrix is specified in a manner more consistent with studies of noncontemporaneous events. Also, the form of the design matrix allows for more heterogeneity across equations, thus taking better advantage of the benefits of systems methods.

II. Previous Research

A. Contemporaneous Events and Systems Methods

Many of the initial applications of systems methods to event studies are based on a single event occurring on one calendar date. This is the special case where the coefficient estimates of generalized least squares (GLS) are identical to those of ordinary least squares (OLS). However, the covariance structure can still be incorporated in the statistical tests to reflect the additional information contained in the cross-sectional correlations.

Gibbons [7] initially suggests the systems method for event-based studies as an artifact of multivariate tests of asset pricing. Collins and Dent [6] provide a modified GLS test that allows for a shift in residual variances. In a series of simulations, the test is compared to various other univariate and multivariate tests. By varying the level of cross-sectional correlation and other relevant factors, they provide explicit measures of the potential gains associated with the GLS approach. Schipper and Thompson [15], Hughes and Ricks [8], and Binder [2] provide applications of the systems method to various regulatory and policy changes. All of these authors subsequently extend their results (see [16], [9], and [3]) to incorporate tests that do not rely on asymptotic distributions. Binder [3] also provides a clarifying summary of the advantages and disadvantages of using multivariate techniques, noting that many test statistics in this case are biased against the null hypothesis.

B. Noncontemporaneous Events and Systems Methods

Malatesta [14] extends previous applications of the systems method to allow for noncontemporaneous events. In addition, a series of Brown and Warner [5] type simulations are used to examine the benefit of the generalized methods over simple univariate tests in actual applications. The subsequent discussion focuses on the Malatesta study, in that the current study is a direct extension of his results.

One of the difficulties of estimating systems methods is the size of the various matrices required for the general equations. In the first segment of his paper, Malatesta provides a derivation of GLS, specifically in the context of event study applications, that utilizes the unique structure of the problem to provide a more compact estimator. The second segment of the paper provides some initial empir-
tical results that suggest the systems method is, in practice, of no benefit. Before the approach is completely discounted, there are a number of extensions to these initial tests that have promise in the context of this application and merit additional study.

One extension that does not appeal so much to the specific method, but more to the general power of the tests, is the use of daily data. In their simulations, Brown and Warner [5] find that the power of tests based on daily versus monthly data, in correctly rejecting the null hypothesis of no abnormal returns, increases by a factor of about three. The use of daily observations also allows for the contemporaneous components of the estimated covariance matrix to be based on data from a period within approximately one year of the event date. Malatesta uses an 80-month estimation period, thus including information from the past seven years.

Another important extension to Malatesta’s empirical tests was noted in the conclusion of his paper where he recognizes that “the potential benefit of using joint GLS in an event study is theoretically greater when disturbance terms are more highly correlated across equations.” Thus, one would expect that the random sample tested by Malatesta would be biased toward accepting the less general model in comparison to tests based on a sample of securities experiencing the same event. The current study derives a sample based on one of the more widely tested events, the stock split.

A series of test statistics is studied by Malatesta. A generalized $F$ statistic is used to identify abnormal performance for the systems method. The most successful test statistic was based on ordinary least squares estimates that were used to develop an average $t$-statistic, labeled $W^{**}$. The current study will report the $W^{**}$ test for comparative purposes. The remaining tests, however, will all be based on a likelihood ratio statistic. Comparing different tests for different methods does not provide for a clear distinction as to the gain or loss associated with a specific method, since the separate test statistics will have differing empirical characteristics.

If, across equations, the covariance is zero or the independent variables are equivalent, the OLS and GLS estimates will be identical. To the extent that these conditions do not hold, the OLS estimate of the regression coefficients is inefficient and the usual OLS estimate of the covariance matrix is biased. Thus, the potential benefit of the systems method over ordinary least squares is a positive function of the absolute level of correlation between the error terms in the equations and a negative function of the homogeneity of the independent variables across equations. In addition, the realized benefit of GLS is related to the ability of the estimation method to accurately characterize the covariance structure. Malatesta examines four sample time intervals, and in each interval the event date is randomly assigned to one of the eighty months. Although this implies that the dummy variables will differ across equations, the independent variable representing the market returns will be identical. The current study places the dummy variable at the end of each estimation interval with the event day being determined either by an actual event or a randomly assigned event within the test interval. In this case, the design matrix more closely parallels the structure used
in event studies and contains a heterogeneous set of observations for the market return variable in each equation.\(^1\)

III. Sample Design

Samples of this study are taken from both the CRSP daily and monthly data files. Due to data availability at the time of the study, the daily samples contain event dates from 1964 to 1984, the monthly samples contain event dates from 1961 to 1985. The estimation period for the monthly tests includes 80 months (as in Malatesta) with the event included in the final month. The estimation period for the daily sample includes 245 days (similar to Brown and Warner) with the event falling on the final day. For each group of data, two predefined sampling populations are developed. The first population consists of all securities with CRSP-documented stock splits greater than or equal to 25 percent. Only cases in which an additional split was not reported during the estimation interval were included (thus creating a smaller group for the monthly observations). This resulted in 2,675 daily events and 840 monthly events. Subsequent tests are based on 250 samples of 30 securities selected from these populations. The daily and monthly sample subsets based on these populations will be labeled Split.

For both the daily and monthly data, an additional sampling population is defined for comparison with the samples based on an actual event. In this case, the event dates are randomly selected from the time series and applied to randomly selected securities satisfying the data requirements. The process is repeated to generate sampling populations of the same size as those based on stock splits. The sample subsets generated from these populations are comparable to those of Malatesta and Brown and Warner, where the securities are not experiencing event commonalities. Samples taken from these populations will be labeled Random. In addition, simulated abnormal returns are introduced in two other sample groups for both the daily and monthly data by adding 0.005 and 0.01 to the event day/month return for the Random populations.\(^2\)

IV. Estimators and Tests

A. The Estimators

The basis for the tests is the standard market model, extended to include an event-day dummy variable,

\[
R_{it} = \alpha_i + \beta_i R_{mt} + \gamma_i \delta_{it} + \mu_{it}, \quad \text{for } i = 1 \text{ to } J \text{ and } t = 1 \text{ to } T,
\]

where \(R_{it}\) is the return on security \(i\) in period \(t\), \(R_{mt}\) is the market return (measured here by the CRSP equal-weighted return series), and \(\delta_{it}\) is a dummy variable assigned a value of one on the event date.\(^3\) \(\gamma_i\) provides a measure of abnormal

\(^1\) This is true to the extent that there is not event day clustering. For the sample of stock split events considered in this study, 68 percent occurred on a unique event date. Approximately 99 percent of the event dates contained four or fewer announcements.

\(^2\) Brown and Warner [5] show that these types of tests are not sensitive to the actual method of introducing abnormal returns.

\(^3\) The method is not developed in detail here. For a more complete derivation, see [14] or [12]. For a discussion of maximum likelihood estimation of the GLS model, see [13] or [11].
returns. By stacking the regression equations for each security, the relationship
can be estimated using systems methods. The estimators tested in this study are
developed using the log-likelihood function for this relationship. Let \( \mu \) be the \( JT \)
vector of disturbance terms, distributed as \( N(0, \Sigma \otimes I_T) \), where \( \Sigma \) is the covariance
matrix of contemporaneous disturbances for the system as defined in Malatesta,
and \( I_T \) is a \( T \times T \) identity matrix. Note that the implied covariance structure in this
case is similar in form to Malatesta’s model; however, the source of covariability in
the errors is very different. In both cases, we assume that the errors are indepen-
dent over time but are contemporaneously correlated. In Malatesta’s case, the
contemporaneous covariance is not so much a result of an event, but is an artifact
of correlated errors in calendar time. In this study, the time series of observations
on \( R_{mt} \) are aligned in event time, although they may be from different calendar
periods. A variety of studies documents preevent patterns in event-time residuals
(e.g., the case of stock splits, dividend changes, or mergers). The covariance
structure assumed in this study implies that the relevant source of contemporane-
ous covariance is in event time.

The likelihood function for the problem is given by
\[
(2) \quad \ln L = -\left(\frac{JT}{2}\right) \ln 2\pi - \left(\frac{1}{2}\right) \ln |\Sigma \otimes I_T| - \left(\frac{1}{2}\right) \left[ \mu' \left( \Sigma^{-1} \otimes I_T \right) \mu \right].
\]

Using an asterisk to indicate parameter estimates, let the estimated covariance
matrix be denoted as \( \Sigma^* = [\sigma_{ij}^*] \) for \( i, j = 1, 2, \ldots, J \), and define \( \mu_{ji}^* = [\mu_i^*,
\mu_j^*, \ldots, \mu_J^*] \). The maximum likelihood estimate of \( \sigma_{ij}^* \) is
\[
(3) \quad \sigma_{ij}^* = T^{-1} \mu_i^{*'} \mu_j^*.
\]

Thus, the concentrated likelihood function from Equation (2) is (see [11], p. 589)
\[
(4) \quad \ln L = -\left(\frac{JT}{2}\right) \ln 2\pi - \left(\frac{1}{2}\right) \ln |\Sigma^* \otimes I_T| - \left(\frac{T}{2}\right).
\]

The function is further simplified by dropping the constant terms and using the
matrix properties,
\[
(5) \quad |\Sigma^* \otimes I_T| = |\Sigma^*|^T |I_T|^T, \quad \text{and}
\]
\[
(6) \quad |I_T| = 1,
\]

providing the condensed likelihood function,
\[
(7) \quad \ln L = -T/2 \ln |\Sigma^*|.
\]

The specific type of estimator being tested is determined by the method used
to generate the estimated covariance matrix, \( \Sigma^* \). If the covariance matrix is re-
stricted to be diagonal, i.e., it is assumed that there is zero covariance between
the disturbances from each equation, the estimates correspond to the OLS esti-
mates. If this restriction is not made, the components of the covariance matrix
can be asymptotically approximated using estimated generalized least squares
(EGLS), where the covariance terms are derived using the first stage OLS esti-
mates. More generally, the covariance matrix can be directly estimated using
maximum likelihood methods. As a limiting approximation to the maximum likelihood estimates, the EGLS approach can be iterated (IGLS). This methodological structure provides a simple format that allows the extension of the diagonal OLS covariance specification to the more general form to be tested, and for the significance of the event parameters to be tested, all using a similar statistical test (the test statistics are described in the next section).

Preliminary tests in this study comparing the maximum likelihood (ML) and IGLS methods indicated that: (1) as expected, the ML method was extraordinarily expensive in terms of computational time, (2) the IGLS method was substantially less expensive and would, on average, converge to within $10^{-5}$ of the ML estimates within 10 iterations (most of the gains occur in the first 3 iterations). Thus, the estimators that will be compared in this study are the OLS, EGLS, and IGLS methods. The extension of EGLS to IGLS provides another source of possible gains over the initial tests in Malatesta.

B. The Tests

Of the various statistical tests in Malatesta, the one with the best performance was based on the OLS estimates. For each system of equations, each estimate of $\gamma$ is divided by its estimated standard deviation. The average of these standardized parameters is asymptotically normal with mean zero and variance $J^{-1}$. The statistic is labeled $W^{**}$, and will be reported for comparative purposes. In addition, the statistic will be estimated from the parameters using each of the three methods. This provides an indirect method of comparing the frequency distributions for the parameter estimates from each period.

The remaining tests will be based on the likelihood ratio (LR) statistic, where twice the difference between the log-likelihood value of Equation (2) for an unrestricted and restricted model is distributed as a $\chi^2(r)$, where $r$ is the number of restrictions. Two general types of comparisons are considered. First, the statistical significance of the event day dummy can be tested for each method by estimating Equation (1) with the restriction that $\gamma_i$ equals zero for all $i$. The likelihood value for the restricted version can be compared to the full version of Equation (1) with the likelihood ratio being distributed as a $\chi^2$ with 30 degrees of freedom. Thus, the ability of the OLS, GLS, and IGLS methods to detect an event will be compared using the same basic statistic.

An additional comparison can be made to determine the significance of extending the estimates of Equation (1) to the more general methods. The likelihood values from Equation (2) for the IGLS and GLS methods can be compared to the OLS value to determine if the more general specification is statistically significant. The LR in this case is distributed as a $\chi^2$ with 435 degrees of freedom (if 30 equations are included in the system, OLS constrains $(30*30 - 30)/2$ unique elements in the covariance matrix to zero).

V. Results

The results for each predefined population are presented in Table 1 and are based on 250 samples of 30 securities. The table reports the proportion of cases in which the null hypothesis ($\gamma_i = 0$, for all $i$) is rejected using the $W^{**}$ and LR
tests. The rejection frequencies are presented for each of the four sampling populations using both monthly and daily data. The “Random” sample corresponds to the case in which the null hypothesis is true.

The W** statistic is computed from the mean and standard deviation of the estimated μs. Thus, the comparison of the statistic across each method reveals the effects of the more general methods on the point estimates of μ. The similarity of the W** statistics in Table 1 indicates that, as in Malatesta, the generalized methods do not provide any notable effects on the point estimates.

<table>
<thead>
<tr>
<th>Monthly</th>
<th>W**</th>
<th>Likelihood Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>EGLS</td>
</tr>
<tr>
<td>Split*</td>
<td>68.0%</td>
<td>67.6%</td>
</tr>
<tr>
<td>Random</td>
<td>6.0</td>
<td>6.4</td>
</tr>
<tr>
<td>Random + 0.005</td>
<td>11.2</td>
<td>12.8</td>
</tr>
<tr>
<td>Random + 0.01</td>
<td>20.8</td>
<td>20.4</td>
</tr>
</tbody>
</table>

| Daily |                |         |       |       |
|-------|----------------|---------|-------|
| Split | 80.0           | 80.0    | 80.4  | 94.8 | 94.4 | 94.4 |
| Random | 4.4            | 4.4     | 4.8   | 14.4 | 20.4 | 20.4 |
| Random + 0.005 | 18.0 | 17.6 | 17.6 | 19.6 | 24.4 | 24.4 |
| Random + 0.01 | 56.4 | 55.2 | 55.2 | 38.0 | 43.6 | 44.0 |

a Samples taken from a predefined population of 840 stock splits for monthly data, and 2,675 stock splits for daily data.
b Samples taken from predefined populations of the same magnitude as the split populations, where in this case the security is randomly selected and randomly assigned an event date within the test interval.

A. Monthly Data

The rejection frequencies for the LR test using monthly data indicate an improvement in power associated with the generalized method for the Split samples and the Random samples with simulated abnormal return. There is some improvement in power associated with IGLS; however, based on confidence limits computed from the binomial distribution, the difference is not statistically significant. More importantly, the rejection frequencies for the random sample using monthly data indicate that the LR test suffers from a high level of Type I error, consistent with the F-test results in Malatesta.

The W** test for the monthly data appears to be well specified. The level of Type I error is not significantly different from the expected level of 5 percent.

---

A 95-percent confidence interval based on the binomial distribution is $1.96(\hat{p})p < p < 1.96(\hat{p})p$, where $p$ is the proportion of rejections and $\hat{p} = \sqrt{\hat{p}(1-p)/250}$. 

The rejection frequencies for the random samples with an induced abnormal return of 1 percent are comparable to those reported by Brown and Warner [4] in their tests of monthly data. For the Split samples, even if the high level of Type I error is ignored for the LR test, the rejection frequencies for the generalized methods are no greater than the OLS-based \( W^* \) statistic.

In general, the monthly results suggest that there is some positive effect on the power of the LR statistic from using the less restrictive methods. However, the gain does not appear to be beneficial in terms of identifying abnormal performance on the event date.

B. Daily Data

The daily results are very similar to those for monthly data, with the conclusions concerning the use of systems methods similarly nonsupportive. The LR tests suffer from Type I errors in this case; however, the false rejections for the LR test are not as extreme as they were for the monthly data. The rejection frequencies for the simulated abnormal returns of 0.005 and 0.01 are not quite as high as those reported in Brown and Warner [5]. (Using the market model method, they report rejection frequencies of approximately 27 percent and 80 percent for simulated abnormal returns of 0.005 and 0.01, respectively.)

Although the rejection frequencies are higher for the LR tests in the Split sample, most of this gain is attributable to higher levels of Type I error. Even if the Type I error differences are ignored, the generalized methods do not have any effect on the power of the tests for the Split sample. The results for the Random populations with simulated abnormal returns indicate some increase in power is attributable to the systems methods; however, confidence intervals based on the binomial distribution indicate that none of the differences is statistically significant.

C. Combined Results

In addition to testing the power of the various methods to identify abnormal returns, the statistical significance of using a generalized covariance structure was also tested. The IGLS method cannot be compared to the EGLS method using the LR test because one is not a constrained form of the other. The comparison of the IGLS and EGLS methods to the OLS approach indicate that the generalization provided statistically significant gains in more than 98 percent of the cases for both the daily and monthly data. Although these tests suggest that the systems methods do provide important econometric benefits in the returns-based models, the critical issue in this case is that the benefits do not provide a more discerning test for identifying abnormal returns.

The high level of Type I error observed for the LR tests in Table 1 for both the monthly and daily data parallels the findings of Stambaugh [18] and Jobson and Korkie [10] in their applications of systems methods to tests of asset pricing. They find that the Type I error increases as the ratio of the number of securities to the number of time-series observations increases (i.e., \( J/T \) increases). The failure of the LR statistic to conform to its limiting distribution is based on the lack of stability in the \( \Sigma^* \) matrix. (Shanken [17], pp. 332–333, provides an insightful
description of the nature of the small-sample bias in the LR statistic.) The results in Table 1 confirm this tendency, with the Type I error decreasing as \( T \) increases from 80 to 245 between the monthly and daily results. If the only limitation to applying systems methods in event studies was the Type I error of the LR test, modified LR statistics or alternative tests could be considered (see [1]). The failure of the systems methods to dominate the simpler OLS-based methods precludes the need for extending the LR test.

VI. Conclusions

After extending the results of Malatesta in a number of ways that would take advantage of the benefits of systems methods, the basic conclusion from this research is a confirmation of Malatesta’s findings. That is, although systems methods have various characteristics that are amenable to event study applications, the promise of these methods is not supported by a variety of empirical tests. The extension of Malatesta’s results to samples containing actual events, daily data, iterative methods, and alternative tests did not change the basic conclusions that the simple OLS market model method is sufficient. In this study, there is limited evidence that the extension of the OLS model to systems methods is in many cases statistically significant; however, these gains do not provide superior power over simpler tests in identifying abnormal returns. Future studies appealing to the characteristics of systems methods in tests of abnormal performance must be careful to show the efficacy of using the more complicated methodology.
References


