Cuadernos de Economía, Vol. 42 (Noviembre), pp. 307-328, 2005

Measuring Security Price Performance Using Chilean Daily Stock Returns: The Event Study Method*

Rodrigo Saens**
Universidad de Talca
Eduardo Sandoval
Instituto Tecnológico de Monterrey

Following the Brown-Warner simulation approach and using Chilean daily security returns data, we examine the specification and power of three parametric t-tests commonly used in event studies: the standardized, the cross-sectional and the portfolio t-test. Our findings show that even though symptoms of nonnormality in security returns and security abnormal returns persist even at the portfolio level, methods based on the use of parametric tests for samples of ten or more securities are well specified, at least at a significance level of 5%. In terms of power, our simulation results show the standardized t-test is always more effective in detecting the presence of an abnormal return than its two competitors: the cross-sectional and the portfolio t-test. We also find, however, that the power of the three t-tests is very sensitive to both the sample size and the length of the event period. In particular, conclusions from event studies conducted in the Latin American equity market involving multiday event periods have to be taken with caution.

JEL: G14
Keywords: Event Studies Method, Specification Tests

1. Introduction

Over the last twenty years, the performance of the events-study methodology has been the subject of a number of studies. The main concern of this research is to analyze the power and the degree of specification of test statistics

*We are extremely grateful to Claudio Jorquera for his computational assistance in this research project.
**E-mails: rsaens@utalca.cl, esandov@itesm.mx

used in short-run and long-run event studies. Brown and Warner (1985), Dyckman, Philbrick and Stephan (1984), Campbell and Wasley (1993) and Cowan and Sergeant (1996) analyze how the particular properties of daily stock returns affect the performance of several test statistics used in short-run event studies. On the other hand, Barber and Lyon (1997), Kothari and Warner (1997), Brav (2000) and Jegadeesh and Karceski (2004) examine the performance of alternative test statistics used in long-run event studies. An important focus of all this research is the effect of nonnormality in daily return data on parametric test performance.

As McWilliams and Siegel (1997) point out, normality of abnormal returns is a key assumption when parametric test statistics are used in the event-study method. Brown and Warner (1985) and Dyckman *et al.* (1984) study the effect of nonnormality in daily return data on tests performance using samples of randomly selected New York Stock Exchange (NYSE) and American Stock Exchange (AMEX) securities. They report that the nonnormality problem does not have a serious impact on the power of the short-run events study method and that the common parametric t-test used in these studies is well specified under the null hypothesis. They also indicate that the parametric t-test has a statistical power comparable to the theoretical power obtained under the normality assumption. However, Campbell and Wasley (1993) and Maynes and Rumsey (1993) find that with thinly traded samples, the conventional standardized and portfolio t-tests are poorly specified. They report that these parametric tests reject a true null hypothesis too often with NASDAQ and Toronto Stock Exchange samples, respectively. Moreover, Cowan and Sergeant (1996) also report a similar misspecification of the portfolio t-test for thinly traded samples using NYSE-AMEX and NASDAQ daily stock returns files.

Unfortunately, extension of this research to Latin American security returns data is not a clear-cut. Given that the usefulness of the event study methodology is directly related to the market's ability to quickly reflect new information, the thinner trading of the Latin American stock market has a substantial impact on this method. In fact, although empirical research applying this procedure is increasingly incorporating Latin American securities, there are virtually no studies analyzing the performance of the event study methodology using security return data drawn from the Latin American stock market.

This paper examines statistical properties of daily stock returns and how the particular characteristics of these data affect the empirical performance of the short-run events-study methodology when security returns data are drawn from the Chilean stock market. Our study has two main objectives: First, we examine normality in the distributions of daily security returns and daily abnormal security returns data drawn from the Chilean stock market. Given the thinner trading we observe in Latin American stock markets, it is reasonable to expect a more severe degree of nonnormality in the distribution of these security abnormal returns than those found by previous authors in NYSE-AMEX daily abnormal returns. Second, we analyze the performance of the events-study method conducted in the Chilean equity market. We address three issues that determine the ability of an event study to detect abnormal returns: the portfolio size, the magnitude of an eventual abnormal performance and the event date uncertainty. We examine the interaction over ranges of all these three variables simultaneously to determine their effect on the researcher's capacity to identify abnormal performance when event studies are conducted in thinly traded markets.

Our findings show that (1) while symptoms of nonnormality in security returns and security abnormal returns persist even at the portfolio level, methods based on the use of parametric tests for samples of 10 or more securities are well specified, at least for a significance level of 5%; (2) in terms of power, the standardized t-test is always more effective in detecting the presence of an abnormal return that its two competitors: the cross-sectional and the portfolio t-test; and (3) the power of the three parametric t-tests is very sensitive to both the sample size and the length of the event period. Although the evidence presented in this study provides an initial basis for selecting between alternative procedures in event studies conducted in the Latin American stock market, our simulation results also suggest that conclusions from event studies involving multiday event periods must be taken with some caution.

The examination is carried out using a simulation approach analogous to that introduced by Brown and Warner (1980). Unlike a Monte Carlo simulation, where the researcher samples artificially generates values from a specified theoretical probability distribution, the Brown-Warner approach randomly selects event dates and stocks to simulate event studies without assuming a particular distribution of stock returns. Our contribution attempts to help test statistic selection, reducing the probability of misspecification when studies involve Latin American equity market securities. Although this technique overcomes the theoretical question, it allows us to examine the statistical validation of different alternative methods. After all, as Brown and Warner (1980, p.210) indicate, "…the performance of alternative models (in event studies) is an empirical question."

2. USING DAILY DATA: THE NONNORMALITY PROBLEM

An important assumption underlying the use of parametric t-tests in the events study methodology is normality of abnormal returns. Fama (1976, p.21), conversely, documents evidence that the distributions of daily returns exhibit substantial departures from normality, suggesting that they are *fat-tailed* relative to a normal distribution. Brown and Warner (1985) support the same result for the case of NYSE-AMEX daily abnormal returns. They document that daily returns depart considerably from normality in term of skewness and kurtosis. Additionally, Cowan (1992), Campbell and Wasley (1993), and Cowan and Sergeant (1996) show that this is also the case for NASDAQ daily abnormal returns. Even though these findings are not consistent with the normality assumption in excess returns, Dyckman *et al.* (1984) and Brown and Warner (1985) report that the degree of nonnormality in daily NYSE security abnormal returns does not represent a serious problem for a correct test specification. They also show that the portfolio and the standardized t-tests have an empirical power comparable to the theoretical power obtained under the normality assumption. As Brown and Warner (1985) indicate, this result is based on the Central Limit Theorem that guarantees that if the abnormal returns in the cross-section of securities are independent and identically distributed, then the distribution of the sample mean abnormal return converges to a normal distribution.

Extension of these findings to a Latin American stock market, however, is not clear. According to Urrutia (1995) and Rouwenhorst (1999), Latin American stock markets have higher average ex-post returns but, at the same time, their market capitalization, amounts traded, and level of integration are relatively small. It is reasonable to expect thus a more severe degree of infrequent trading and, therefore, nonnormality in the distribution of Latin American security returns. As Campbell and Wasley (1993) and Cowan and Sergeant (1996) show, in markets with thinner trading there is a significant degree of nonnormality in daily return securities that persists even at the portfolio level. As a result, the t-statistics used in event studies depart from their theoretical unit normal distribution under the null hypothesis. Using daily Chilean stock returns, this is one of the issues we examine in this paper.

3. Experimental Design

As Brown and Warner (1980) and Dyckman *et al.* (1984) argue, given the problems of using an analytical approach to compare different properties of alternative return-generating models (RGM), the simulation approach provides a useful method for dealing with conditions where either the analytical approach becomes extraordinarily difficult or the same approach yields results suggesting just directions but not magnitudes. In this paper, we resemble the positive approach of Brown and Warner (1980, 1985) with emphasis on the specification and statistical power of three different return-generating models (RGM) when event studies are conducted using samples from the Chilean stock market.

3.1 Abnormal Returns

An event study attempts to measure the effect of an observed event on the firm's market value. In general, the main purpose of any event study is to find empirical evidence that shows whether a security performance is statistically different from what would be expected under the assumptions of one specific RGM. As MacKinlay (1997, p.13) indicates, "the usefulness of such a study comes from the fact that, assuming rationality in the market place, the effect of an event will be reflected immediately in assets prices." If the event conveys new –relevant– information to the stock market, the mean or the variance of the security abnormal returns must reflect the new economic conditions. Thus, for firm *i* and event date *t* the conditional *abnormal* return is given by:

(1) $AR_{it} = R_{it} - E(R_{it} / \Omega_{t-1})$

where AR_{it}, R_{it} and $E(R_{it} / \Omega_{t-1})$ are the abnormal, actual and normal (expected) return for time *t*, respectively. Notice that Ω_t is the conditional information set in period *t* and that the approach followed for the event study methodology assumes that securities returns are generated by some RGM. Then, it is necessary to specify a model that generates *normal* returns before abnormal returns can be measured. This model can be based on simple statistical relationships such as the OLS Market Model or on more theoretical economic models such as the Capital Asset Pricing Model (CAPM) and the Arbitrage Pricing Model (APT).[1]

We report abnormal performance measures based on the three RGM

A. OLS Market Model

(2) $AR_{it} = R_{it} - \hat{\alpha} - \hat{\beta}_i R_{mt}$

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

B. Market-Adjusted Returns Model

(3) $AR_{it} = R_{it} - R_{mt}$

where R_{mt} is the market index return for day *t*.

C. Mean-Adjusted Returns Model

(4) $AR_{it} = R_{it} - \overline{\overline{R_i}}$

where $\overline{R_i}$ the simple average of security i's daily returns in the estimation period.

These three RGM are discussed in Brown and Warner (1980) and MacKinlay (1997).

[1]Unfortunately, obtaining a more accurate return-generating model (RGM) is not a sufficient condition to generate a well-specified and powerful test of abnormal return. First, as Brown and Warner (1980) indicate, there is a measurement error in each of the variables on which returns depend in the model. For example, in the case of the CAPM, as Roll (1977) argues, it is not possible to observe directly the market portfolio. Second, the efficiency of using either a statistical or an economic model depends critically on the additional statistical assumptions about ε_t , the error term. If the assumed sampling distribution under the null hypothesis is incorrect then we can obtain false inferences.

3.2 Test Statistics

The test statistics for day 0 analyzes whether or not the portfolio mean abnormal return in day 0 is equal to zero. We study the specification and power of three parametric t-tests: The standardized, the cross-sectional and the portfolio t-test.

A. The Standardized t-test (θ_1)

Following Patell (1976) and Dodd and Warner (1983) many authors have used a standardized abnormal return (SAR), where each abnormal security return is normalized by its estimation period standard deviation:

(5) $$SAR_{it} = \frac{AR_{it}}{SD(AR_{it})}$$

The standard deviation $SD(AR_{it})$ of each abnormal return is given by:

(6) $$SD(AR_{it}) = \sqrt{\frac{1}{T_0 - 1}\sum_{t=1}^{T_0} AR_{it}^2}$$

where T_0 is the number of days in the estimation period. Then, the day 0 of the standardized t-test is:

(7) $$\theta_1 = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} SAR_{i0}$$

The standardized t-test assumes that the individual abnormal returns are cross-sectionally independent and identically distributed. By the Central Limit Theorem, the standardized t-test converges to unit normal under the null hypothesis of no abnormal return. Brown and Warner (1985) report that this test is well specified under the null hypothesis for NYSE-AMEX daily security returns data. However, Campbell and Wasley (1993) and Cowan and Sergeant (1996) document that the standardized test is misspecified for NASDAQ samples.

B. The Cross-Sectional t test (θ_2)

As the standardized t-test, this method also assumes that the day 0 abnormal returns are independent and identically distributed. The t-statistic is estimated dividing the average event-period abnormal return ($\overline{AR}_0$) by its contemporaneous cross-sectional standard deviation.

(8) $$\theta_2 = \frac{\overline{AR}_0}{SD(\overline{AR}_0)}$$

The cross-sectional test ignores the estimation period variance and the standard deviation $SD(\overline{AR}_0)$ is given by:

(9) $$SD(\overline{AR}_0) = \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^{N} (AR_{i0} - \overline{AR}_0)^2}$$

This procedure, however, has some limitations. If the variance differs across sample securities or security abnormal performances are correlated across firms the test statistic is likely to be misspecified.

C. The Portfolio t-test (θ_3)

For each day *t*, the cross-sectional average excess return of *N* securities is computed. The portfolio *t-test* is the ratio of the mean excess return in $t = 0$ to its estimated standard deviation:

(10) $$\theta_3 = \frac{\overline{AR}_0}{SD(\overline{AR}_t)}$$

Where for each day the cross-section average excess return of *N* securities is obtained as:

(11) $$\overline{AR}_t = \frac{1}{N} \sum_{i=1}^{N} AR_{it}$$

And the standard deviation is computed over an estimation period of T_0 days. Thus,

(12) $$SD(\overline{AR}_t) = \sqrt{\frac{1}{T_0 - 1} \sum_{t=1}^{-T_0} (\overline{AR}_t - \overline{\overline{AR}})^2}$$

(13) $$\overline{\overline{AR}} = \frac{1}{T_0} \sum_{t=1}^{-T_0} \overline{AR}_t$$

If $\overline{AR}_t$ are independent, identically, and normally distributed, the test statistic is distributed t-student with (T_0-1) degrees of freedom and is asymptotically unit normal under the null hypothesis. Brown and Warner (1980) call this method "Crude Dependence Adjustment" because when using this test, the standard deviation of the day 0 average excess return is estimated using the cross-sectional mean abnormal returns from the estimation period. Thus, the portfolio t-test explicitly takes into account any potential cross-sectional dependence in the security specific abnormal returns. However, Campbell and Wasley (1993) find that, although less pervasive than in the standardized t-test, misspecification is also present in the portfolio t-test when event studies include thinly traded samples.

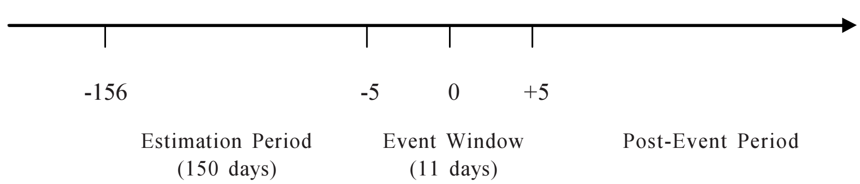
3.3 Data and Sampling Procedure

The data comes from daily closing prices series for stocks traded in the Santiago Stock Exchange from January 1985 to July 2003. As market *proxy* we use the domestic index of stock prices IPSA, which includes the forty most traded stocks in the Chilean market.

Series of 161 observations –trading and non trading days– are drawn randomly with replacement to conform portfolios of 10, 20, 30, 40 and 50 series each. Event dates are assumed to take place with equal probability on each trading day from 01/02/1985 to 07/30/2003. Each stock series is selected by generating two random numbers. The first number identifies a row –a date– and the second number a column –a stock– over the feasible database. With these directions we obtain the initial observation of a 161 daily observations series. This process is repeated over and over again until a portfolio is conformed.

The period -156 through -6 of each series (150 days) is the estimation period, in which the parameters of expected return models are estimated (see Figure 1).[2] The period -5 through +5 (11 days) and the day 0 are the event period and the event date. In order to include a security in a sample it must have at least 40 daily returns in the entire 161 days period, and no missing return data in the last 10 days.

FIGURE 1



3.4 Simulating Abnormal Performance

In order to artificially introduce a given level of abnormal return we follow the Brown-Warner procedure. A constant sample-wide abnormal performance –from 0% to 2.5%– is added to the actual day 0 return for each security. For

[2]As Kothari (2001) and Kothari and Warner (2005) indicate, the length of the estimation period is arbitrary. It has to be long enough to contain a «reasonable» number of observations to estimate the parameters of the model and short enough to avoid an eventual instability of the parameters. In general, the literature uses a length between 120 (Dyckman *et al.*, 1984) and 250 days (MacKinlay, 1997).

example, to simulate 1.5% abnormal return, 0.015 is added. This procedure allows us to analyze the power of test statistics for different abnormal returns level.

4. RESULTS

4.1 Time Series Properties for Individual Chilean Securities

Table 1 reports some statistical properties of daily returns and daily abnormal returns for individual securities selected with replacement from the Chilean stock market. Parameter estimates are computed based on 500 samples of 50 securities, randomly selected. Using the time series of estimation period data, we calculate the mean, standard deviation, skewness and kurtosis coefficients, and the studentized range. Each value on Table 1 represents the mean of 25,000 estimates. In the case of daily abnormal returns, they are based on the three different RGM we examine in this paper.

Results on Table 1 indicate that for the case of the Chilean stock market, daily returns and daily abnormal returns of individual securities depart significantly from the theoretical normal distribution. For example, mean values for the skewness and kurtosis coefficients equal or exceed 0.30 and 6.38, respectively. Additionally, the studentized ranges are 6.93 or greater. All the skewness and kurtosis coefficients and the studentized range for the daily returns and daily abnormal returns exceed the 99th percentile of the respective distribution under the normality hypothesis. Departures from normality are comparable to those documented in Brown and Warner (1985, Table 1) but less severe than those showed by Campbell and Wasley (1993, Table 1) for individual NYSE and NASDAQ daily security returns, respectively[3].

Table 1 also shows that the results are not sensitive to different performances measures. For example, the mean abnormal return using the three RGM is 0.0 % with a very similar standard deviation around 2.1 %. These findings are also consistent with those in Brown and Warner (1980, 1985), which suggest that simple statistical models such as the mean-adjusted returns model often produce comparable results to those of more sophisticated models.

[3]As Brown and Warner (1984) indicate, for the cases of skewness and kurtosis coefficients, and the studentized range, the 95th and 99th percentiles for a normal population are:

Variable	0.95	0.99
Skewness	0.06	0.13
Kurtosis	3.52	3.87
Studentized Range	6.15	6.85

TABLE 1

Properties of daily returns and daily excess returns for individual Chilean securities when no abnormal performance is introduced. For each security, parameters estimates are based on time-series data in the estimation period. Each number in the table shows the mean of 25.000 estimates. Securities and event dates are randomly selected (with replacement) from 02/01/1985 through 07/31/2003.

Performance measure	Mean	Standard deviation	Skewness	Kurtosis	Studentized range
Returns	0.0014	0.0225	0.40	6.92	7.08
OLS market model	-0.0001	0.0206	0.32	6.55	6.98
Market-adjusted	0.0003	0.0198	0.30	6.38	6.93
Mean- adjusted	0.0000	0.0225	0.40	6.92	7.08

4.2 Properties of Sample Mean Abnormal Returns

Table 2 shows cross-sectional properties of the sample mean of daily excess return at day zero. Similarly to Table 1, the different measures of abnormal performance are based on three different RGM. Parameter estimates are computed based on 500 samples of 50, 40, 30, 20 and 10 securities. For each sample, the mean sample estimate is the simple average of the abnormal performance measures for the individual securities in the sample. Mean, standard deviation, skewness and kurtosis coefficients, and the studentized range are computed based on 500 values of the sample mean estimate, one for each sample. As we should expect under the Central Limit Theorem, results of Table 2 show that departures from normality are less severe for cross-sectional mean abnormal returns than for individual abnormal returns. For samples of 50 securities and using the OLS market model, the cross-sectional distribution of the sample mean excess returns seems close to normal. Results in Table 2 also indicate that departures from normality still persist for portfolios less or equal than 40 securities. For example, the cross-sectional distributions of day-0 mean abnormal performance for portfolios of 30 securities exhibit skewness above 0.23. These departures, however, are less pronounced than those documented in Campbell and Wasley (1993, Table 1) using NASDAQ daily security returns.

TABLE 2

Cross-sectional properties of sample-wide mean abnormal performance measures on day 0 using three different return-generating models (RGM) when no abnormal performance is introduced. Each number in the table is based on 500 estimates of the mean, one for each sample. Securities and event dates are randomly selected (with replacement) from 02/01/1985 through 07/31/2003.

Size	Performance measure	Mean	Standard deviation	Skewness	Kurtosis	Studentized range	Jarque-Bera test
50	Market model	0,0002	0,0031	0,04	3,03	6,06	0,17
	Market-adjusted	0,0005	0,0031	0,06	3,18	5,78	0,92
	Mean-adjusted	0,0002	0,0033	0,09	3,05	6,17	0,72
40	Market model	0,0000	0,0036	0,09	3,59	6,99	7,88
	Market-adjusted	0,0004	0,0034	0,03	3,59	7,12	7,26
	Mean-adjusted	0,0000	0,0038	0,17	3,93	7,72	20,69
30	Market model	0,0001	0,0040	0,24	3,38	6,56	7,96
	Market-adjusted	0,0005	0,0039	0,25	3,36	6,02	7,84
	Mean-adjusted	0,0002	0,0043	0,23	3,76	6,68	16,41
20	Market model	0,0000	0,0051	0,15	3,12	5,85	2,11
	Market-adjusted	0,0005	0,0049	0,23	3,57	7,35	11,43
	Mean-adjusted	0,0002	0,0054	0,36	3,27	6,18	12,42
10	Market model	-0,0002	0,0065	0,18	3,70	7,21	12,96
	Market-adjusted	0,0001	0,0063	0,20	3,93	7,52	21,20
	Mean-djusted	-0,0002	0,0074	0,05	3,97	6,40	19,86

4.3 Properties of the Test Statistics

Using the OLS market model as a RGM, Table 3 summarizes the empirical distributions of each test statistic based on 500 portfolios when no abnormal performance is introduced. Under the null hypothesis of no abnormal performance, the distribution of each test statistic should be unit normal. For a portfolio size equal to or above 30 securities, the empirical distribution of the standardized statistic shows small departures from its theoretical distribution. However, Table 3 indicates that if the portfolio size decreases to 10 securities, the degree of nonnormality of the standardized test strongly increases. Our findings in Table 3 also show that the cross-sectional test presents a highly negative skewness.

4.4 Power of the Test

We also examine how the test statistics perform when the null hypothesis is false. To simulate an abnormal performance, a particular abnormal return is introduced into the mean abnormal returns of the sample. Then, the hypothesis of no abnormal performance is tested again. Thus, failing to reject the null hypothesis of no abnormal return when it is false constitutes a Type II error.

Table 4 shows, for three tests and three RGM, the frequency with which the hypothesis of no abnormal performance in day 0 is rejected. For example, for a significance level of α =5% and using the OLS market model, when we introduce a 0.5% level of abnormal performance, the rejection rate for the standardized test is 59% compared to 44% and 38% for the cross-sectional and the portfolio test, respectively. Moreover, the higher power of the standardized test does not depend on the level of significance. For a test of α =1%, and also with 0.5% of abnormal performance, the rejection rate of the standardized test ranges from 36% using the mean-adjusted model to a comparative 50% when the market-adjusted model is examined. Thus, our findings indicate that using Chilean daily security return data the standardized t-test is more likely to detect abnormal return than both the cross-sectional and the portfolio t- test.

TABLE 3

Summary measures for the empirical distribution of each test statistic, one for each sample, with sample sizes from 50 to 10 securities. The procedure to detect abnormal performance is the OLS market model and no abnormal performance is introduced. Each number in the table represents the simple average of 500 estimates.

Size	Test statistic	Mean	Standard deviation	Skewness	Kurtosis	Studentized range	Jarque-Bera test
50	Standardized	0,0997	1,01	0,04	3,00	6,24	0,14
	Cross-Sectional	0,0203	1,00	-0,31	2,70	5,47	9,73
	Portfolio	0,0606	0,98	0,00	3,02	6,11	0,01
40	Standardized	-0,0057	1,03	0,14	3,33	6,45	4,00
	Cross-Sectional	-0,0360	1,05	-0,16	3,05	6,03	2,06
	Portfolio	-0,0117	1,00	0,07	3,63	7,05	8,69
30	Standardized	0,0156	1,02	0,08	3,11	7,12	0,82
	Cross-Sectional	-0,0404	1,02	-0,20	2,78	6,10	4,44
	Portfolio	0,0127	0,97	0,21	3,34	6,30	6,20
20	Standardized	0,0151	1,08	0,23	3,33	6,83	6,56
	Cross-Sectional	-0,0499	1,09	-0,17	3,01	6,86	2,27
	Portfolio	0,0074	1,00	0,26	3,21	5,86	6,71
10	Standardized	-0,0251	0,99	0,26	3,54	6,83	11,64
	Cross-Sectional	-0,0848	1,07	-0,08	3,40	6,80	3,84
	Portfolio	-0,0245	0,90	0,32	3,55	6,38	14,64

4.5 Comparing Alternative RGM

Table 4 also compares the power of detecting abnormal performances among three different RGM. In general, the rejection frequencies indicate that both the market-adjusted returns model and the OLS market model show somewhat better performance than the mean-adjusted return method. For example, using a standardized t-test of $\alpha = 5\%$ and 0.5% of abnormal performance, the mean-adjusted returns model rejects the null hypothesis 49% of the times while the OLS market model and the market-adjusted method register rejection rates of 59% and 64%, respectively. These findings also seem to be robust with respect to changes in the significance level.[4]

Thus, our results suggest that in terms of procedure to measure abnormal performance there is some evidence indicating a better performance of those methods that consider the systematic risk of each security. However, the improvement power of the tests using these two methods over the simpler mean-adjusted model is limited. When an abnormal performance of 2.5% is introduced, the three RGM allow us to identify this abnormal return always, regardless of the test size.

4.6 Sensitivity Analysis

In general, given that data is less available in Latin American economies, event studies conducted in these markets have to deal with two additional issues: (1) a small sample portfolio and (2) a more uncertain event date. Table 5 documents a summary with event studies conducted in the Latin American stock markets from 1990 to 2005. Some of these studies involve Latin American samples with less than 50 securities. Moreover, they also use a procedure which, besides the fact that it may affect the specification and power of the method, it is difficult to reconcile with the market efficiency hypothesis: event windows of 21 or more days. To illustrate these two critical points, we analyze how different sample sizes and also different lengths of the event window can affect the performance of the event study method.

[4] However, as Brown and Warner (1980, 1985) point out, it is possible that these results depend significantly on the fact that in this simulation work the precise time at which the abnormal return occurs is known with certainty.

TABLE 4

A comparison of three alternative RGM and three test statistics for detecting abnormal excess return. Values in the table indicate the percentage of 500 samples where the null hypothesis of no abnormal performance on day 0 is rejected. Sample size is equal to 50 securities. Chilean stock securities and event dates are randomly selected (with replacement) from 01/02/85 through 07/31/03.

Panel A
Two tailed test, α=0.05

		Artificial level of abnormal performance (%) introduced at day 0					
Performance measure	Test statistic	0.0	0.5	1.0	1.5	2.0	2.5
OLS market model	Standardized	5.0	58.6	98.6	100	100	100
	Cross-sectional	4.8	44.0	90.4	99.2	99.8	100
	Portfolio	4.0	38.4	89.6	99.8	100	100
Market-adjusted	Standardized	5.6	64.4	98.8	100	100	100
	Cross-sectional	6.4	49.4	92.6	99.6	99.8	100
	Portfolio	6.0	42.6	92.2	100	100	100
Mean-adjusted	Standardized	5.0	49.4	94.2	100	100	100
	Cross-sectional	4.2	37.0	85.8	99.0	99.8	100
	Portfolio	3.2	32.0	83.8	99.6	100	100

Panel B
Two tailed test, α=0.01

		Artificial level of abnormal performance (%) introduced at day 0					
Performance Measure	Test Statistic	0.0	0.5	1.0	1.5	2.0	2.5
OLS Market Model	Standardized	2.4	44.2	96.0	100	100	100
	Cross-Sectional	1.4	30.0	82.6	98.2	99.6	100
	Portfolio	2.4	23.0	79.8	99.0	100	100
Market-Adjusted	Standardized	3.6	50.0	98.0	100	100	100
	Cross-Sectional	2.0	35.0	86.2	98.8	99.8	100
	Portfolio	2.8	29.8	85.6	99.4	100	100
Mean-Adjusted	Standardized	2.6	35.8	90.8	100	100	100
	Cross-Sectional	1.4	23.4	75.4	98.4	99.6	100
	Portfolio	1.4	18.8	73.6	98.6	100	100

A test of significance is well-specified under the null hypothesis at the 5% (1%) level if the percent of rejections falls between 2.7% and 7.3% (0.0 and 2.2%).

A. *Different Sample Sizes*

Results in Table 6 show that the specification of the tests is not particularly sensitive to the number of securities in the sample. When a test of α =5% is used, no special misspecification of any test statistic is found in samples from size 50 to 10 securities. Some symptoms of misspecification arise only at a 1% level of significance. For example, for a portfolio size of 20 securities the standardized and cross-sectional tests seems to be misspecified with error rates of 3.6% and 2.6%, respectively. As we should expect, results in Table 6 show that the power of the tests also falls strongly when the sample size decreases. Using the standardized test of size 5% and 1% level of abnormal performance, decreasing the sample size from 50 to 10 securities reduces the rejection frequency from 99% to 42%. Table 6 also indicates that the relative power of different test statistics also seems to be independent of the sample size. Thus, in terms of power, dominance of the standardized test over the cross-sectional and the portfolio test does not change.

B. *Multiday-Event Periods*

The simulations we have performed make the strong assumption that the date at which abnormal performance takes place is known with certainty. However, given that most of the times the calendar date of the event cannot be identified exactly, most event study settings involve multiday event periods where the date itself becomes a random variable. To analyze this, we also examine how uncertainty about the precise date of the abnormal performance affects the power of the event study technique.

TABLE 5
A summary of recent short-run event studies conducted in the Latin American stock markets

Study	Nature of event	Length (days) of event window
Kräusl (2005)	The impact of 302 sovereign rating changes on the financial stability of 8 Latin American economies (in a world sample of 28 stock indexes)	Various length from -10 to +10
Bin *et al.* (2004)	The valuation effect of 6 currency crisis on 21 Mexican and 10 Brazilian ADRs (in a world sample of 73 ADRs)	Various length from -30 to +30
Castillo (2004)	172 announcements of Chilean bonds (56) and equity offerings (116)	Various length from -10 to +10
Morán (2003)	The effect of adopting an optional provision (contained in a broader regulatory change) on 17 Chilean stocks.	Various length from -20 to +30
Tapia and Tokman (2003)	24 foreign exchange interventions of the Chilean Central Bank	-3 to +3
Mathur et al. (2002)	The contagion effect from the 1994 Mexican Peso Crisis on 15 Chilean ADRs	-1 to 0
Bhattacharya *et al.* (2000)	75 Mexican corporate news announcements	-80 to +10
Hensler (2000)	68 Mexican initial public offerings (IPO)	Various length from 0 to +300
Maquieira and Osorio (2000)	129 Chilean dividend announcements	Various length from -10 to +10
Parisi and Pérez (2000)	The impact of 88 bond rating changes on Chilean stock prices	Various length from -10 to +15
Wilson *et al.* (2000)	The effect of the Mexican Peso Crisis on six real sector indexes and ten financial groups	-30 to +2
Saens (1999)	19 Chilean ADR-IPO announcements	-10 to +10
Celis and Maturana (1998)	36 Chilean initial public offerings	0 to +10
Maquieira and Sierralta (1990)	The impact of the 1982 Chilean Devaluation on 23 stocks	-23 to +2

TABLE 6

The effect of different sample sizes for detecting abnormal excess return using the OLS market model. Values in the table indicate the percentage of 500 samples where the null hypothesis of no abnormal performance on day 0 is rejected. Sample sizes are equal to 50, 40, 30, 20 and 10 securities. Chilean securities and event dates are randomly selected (with replacement) from 01/02/85 through 07/31/03.

		Two tailed test, α=0.05						Two tailed test, α=0.01					
			Artificial level of abnormal performance (%) introduced at day 0						Artificial level of abnormal performance (%) introduced at day 0				
Portfolio Size	Test Statistic	0.0	0.5	1.0	1.5	2.0	2.5	0.0	0.5	1.0	1.5	2.0	2.5
50	Standardized	5.0	58.6	98.6	100	100	100	2.4	44.2	96.0	100	100	100
	Cross-Sectional	4.8	44.0	90.4	99.2	99.8	100	1.4	30.0	82.6	98.2	99.6	100
	Portfolio	4.0	38.4	89.6	99.8	100	100	2.4	23.0	79.8	99.0	100	100
40	Standardized	5.2	44.4	94.0	100	100	100	3.0	31.6	89.2	99.8	100	100
	Cross-Sectional	7.2	35.6	83.0	98.0	99.4	99.6	2.6	24.8	74.2	95.6	99.4	99.4
	Portfolio	4.4	28.4	80.0	98.4	99.8	100	2.6	16.2	67.2	97.4	99.8	100
30	Standardized	5.2	36.0	86.4	99.6	100	100	1.8	22.2	77.2	99.6	100	100
	Cross-Sectional	4.6	27.0	75.6	94.6	99.8	100	1.4	17.0	62.6	91.6	98.8	99.8
	Portfolio	4.0	23.8	67.8	96.0	100	100	2.2	11.6	53.8	90.8	99.0	100
20	Standardized	6.2	26.6	70.0	96.0	100	100						
	Cross-Sectional	7.0	23.6	61.2	86.2	95.6	99.0	3.6	17.4	56.6	90.8	99.0	100
	Portfolio	4.8	17.8	48.8	86.6	96.8	99.8	2.6	15.8	47.2	78.2	93.0	97.8
								1.8	9.0	36.0	73.4	93.6	99.0
10	Standardized	5.2	14.2	42.0	76.2	92.8	98.6	2.0	7.2	27.4	64.4	87.2	97.0
	Cross-Sectional	6.6	14.4	40.6	68.4	84.6	93.4	3.2	8.2	28.6	57.6	78.6	88.8
	Portfolio	3.2	8.4	24.0	58.2	79.2	92.8	1.6	4.2	15.2	40.0	68.8	87.2

A test of significance is well-specified under the null hypothesis if the percentage of rejections falls between 2.7 % and 7.3% for a test size of 5% level and between 0.0% and 2.0% for a test size of 1% level.

TABLE 7

A comparison of alternative test statistics when the precise date of abnormal performance is unknown. Using the OLS market model as RGM, abnormal performance for each security is introduced at random for one day during the event window. The numbers in the table show the percentage of 500 samples of 50 securities each where the null hypothesis of cumulative mean abnormal performance in the event period is rejected.

Panel A
Two tailed test, α=0.05

		Artificial level of abnormal performance (%) introduced at day 0					
Window	Test statistic	0.0	0.5	1.0	1.5	2.0	2.5
3 days	Standardized	8.8	22.0	58.0	90.0	99.4	100
	Cross-sectional	5.4	14.2	43.8	76.2	93.4	99.0
	Portfolio	6.4	15.2	42.0	73.0	93.2	99.2
5 days	Standardized	8.4	16.8	41.0	72.8	92.6	98.2
	Cross-sectional	5.4	8.6	27.2	51.4	76.8	90.2
	Portfolio	7.2	11.2	28.0	50.6	78.2	91.2
11 days	Standardized	7.8	10.8	28.0	45.0	64.8	82.2
	Cross-sectional	4.2	5.8	13.0	26.6	42.6	61.2
	Portfolio	5.0	8.4	17.0	27.4	44.0	63.2

Panel B
Two tailed test, α=0.01

		Artificial level of abnormal performance (%) introduced at day 0					
Window	Test statistic	0.0	0.5	1.0	1.5	2.0	2.5
3 days	Standardized	4.4	13.4	46.8	81.6	97.4	99.8
	Cross-sectional	2.8	7.8	29.4	63.0	87.8	97.2
	Portfolio	3.0	7.4	28.6	61.8	87.0	97.0
5 days	Standardized	4.4	8.6	29.2	60.4	87.6	96.4
	Cross-sectional	2.6	4.8	16.8	39.4	66.0	84.2
	Portfolio	3.8	5.8	18.2	38.4	64.8	84.6
11 days	Standardized	3.8	5.0	17.0	33.0	51.6	70.8
	Cross-sectional	1.4	1.8	5.8	14.8	30.0	46.0
	Portfolio	2.2	3.4	9.4	18.6	31.6	48.0

A test of significance is well-specified under the null hypothesis at the 5% (1%) level if the percent of rejections falls between 2.7% and 7.3% (0.0 and 2.2%).

Using the OLS market model as a RGM for each security in the 500 samples, we select one day of the event period at random and add a particular level of abnormal performance in one specific day in windows of 3, 5, and 11 days.[5] For example, for a window of 11 days we add a particular level of abnormal performance in one specific day (randomly selected) in the interval from day -5 through +5. Thus, this experiment simulates a situation where the abnormal performance occurs at some –unknown– date in the event period including the event day.

Table 7 reports results in the multiday setting for abnormal performance levels ranging from 0 to 2.5%. Similar to our findings involving a one-day setting, numbers in table 7 indicate that symptoms of misspecification arise for the three tests when a test of size 1% is used. However, results in Table 6 also indicate that misspecification is more severe for the standardized t-test when event periods are longer than one day. When a standardized t-test of size 5% is used, Type I error rates range from 7.8 to 8.8%. As should be expected, the power of the three tests are much lower than those where the precise date of abnormal performance was known with entire certainty. For example, for a window of 11 days, the portfolio t-test is able to detect the presence of 1.5% abnormal performance only 27% of the times, compared to the 100% from Table 4.

5. SUMMARY AND CONCLUSIONS

This paper examines the specification and power of the event studies technique using daily Chilean security return data. We test several procedures with which this methodology measures security price performance. Our findings show that although symptoms of nonnormality in security returns and security abnormal returns persist even at the portfolio level, methods based on the use of parametric tests for samples of 10 or more securities are well specified, at least for a significance level of 5%. However, for a significance level of 1% the three parametric tests we study seem to reject the null hypothesis too often when this hypothesis is true.

Comparison across test statistics indicates important power differences among the three tests. For Chilean daily security return data, the standardized t-test shows to be more powerful than both the cross-sectional and the portfolio t-tests. Moreover, these results are robust to changes in the portfolio size, the RGM or the length of the event period

One additional and interesting result is that in terms of procedure to measure abnormal performance there is some evidence indicating a better performance of those methods that consider the systematic risk of each security. However, the improvement in specification and power of the tests using these methods over the simpler mean-adjusted model is only limited.

[5] For each security, the event day is a drawing from a uniform distribution.

As should be expected, our results also indicate that the power of the three parametric tests falls strongly when the sample size decreases. For example, using the portfolio test with α =5% and 1% level of abnormal performance, through reducing the sample size from 50 to 10 securities reduces the rejection frequency between three and four times.

Finally, the specification and power of the tests also depends on the length of the event period. We find a more severe misspecification for the standardized t-test when the event period is longer than one day. As expected, the power of the three tests also decreases strongly the longer is the event window. Results for 11-day event periods reveal that in samples of 50 securities, when we introduce a 1% of abnormal performance with α =5%, the null hypothesis of no abnormal performance is rejected only 17 to 28% of the time. Thus, conclusions from event studies conducted in the Chilean equity market involving multiday event periods require some caution. As Brown and Warner (1980) point out, even if the researcher conducting an event study can take advantage today of a more advanced pool of computational and statistical techniques, it is still a good use of his time to read old issues of the newspapers to determine event dates more accurately.

REFERENCES

Barber, B. and J. Lyon (1997), “Detecting Long-Run Abnormal Stock Returns: The Empirical Power and Specification of Test Statistics,” *Journal of Financial Economics*, 43: 341-72.

Batarce, J. (1990), “El Efecto del Plan Brady en la Banca Comercial,” *Paradigmas en Administración,* 17: 67-94.

Bhattacharya, U., H. Daouk, B. Jorgenson and C. Kehr (2000), “When an Event is not an Event: The Curious Case an Emerging Market,” *Journal of Financial Economics,* 55: 69-101.

Bin, F-S., L.P. Blenman and D.-H. Chen (2004), “Valuation Impact of Currency Crisis: Evidence from the ADR Market. *International Review of Financial Analysis,* 13: 411-32.

Brav, A. (2000), “Inference in Long Horizon Event Studies: A Bayesian Approach with Application to Initial Public Offerings,” *The Journal of Finance,* 55: 1979-2016.

Brown, S. and J. Warner (1980), “Measuring Security Price Performance,” *Journal of Financial Economics,* 8: 205-58.

Brown, S. and J. Warner (1985), “Using Daily Stock Returns: The Case of Events Studies,” *Journal of Financial Economics,* 14: 3-31.

Callaghan, J., R. Kleiman and A. Sahu (1999), “The Market-Adjusted Investment Performance of ADR IPOs and SEOs,” *Global Finance Journal,* 10(2): 123-145.

Campbell, J., A. Lo and C. MacKinlay (1997), *The Econometrics of Financial Markets*, Princeton University Press.

Campbell, C. and C. Wasley (1993), “Measuring Security Price Performance Using Daily NASDAQ Returns,” *Journal of Financial Economics,* 33: 73-92.

Castillo, A. (2004), "The Announcement Effect of Bond and Equity Issues: Evidence from Chile," *Estudios de Economía,* 31(2): 177-205.

Celis, C., and G. Maturana (1998), "Initial Public Offerings in Chile," *Abante,* 1: 7-31.

Cowan. A. R. (1992), "Nonparametric event study tests." *Review of Quantitative Finance and Accounting,* 2: 343-358.

Cowan, A.R. and A. Sergeant (1996), "Trading Frequency and Event Study Test Specification," *Journal of Banking and Finance,* 20: 1731-57.

Dodd, P. and J. Warner (1983), "On Corporate Governance: A Study of Proxy Contests," *Journal of Financial Economics,* 11: 401-438.

Dyckman T., D. Philbrick and J. Stephan (1984), "A Comparison of Event Study Methodologies Using Daily Stock Returns: A Simulation Approach," *Journal of Accounting Research,* 22: 1-33.

Fama, E. (1976), *Foundations of Finances*, New York: Basic Books.

Hensler, D., M. Herrera and L. Lockwood (2000), "The Performance of Initial Public Offerings in the Mexican Stock Market, 1987-1993," *Journal of International Money and Finance,* 19: 93-116.

Jegadeesh, N. and J. Karceski (2004), "Long-Run Performance Evaluation: Correlation and Heteroskedasticity-Consistent Tests," Working Paper, Goizueta Business School, Emory University.

Kothari, S. (2001), "Capital Markets Research in Accounting," *Journal of Accounting and Economics,* 31: 105-231.

Kothari, S. and J. Warner (1997), "Measuring Long-Horizon Security Price Performance," *Journal of Financial Economic,* 43: 301-39.

Kothari, S. and J. Warner (2005), "Econometrics of Event Studies," Working Paper, Center for Corporate Governance, Tuck School of Business at Dartmouth.

Kräusl (2005), "Do Credit Rating Agencies add to the Dynamics of Emerging Market-Crisis?" *Journal of Financial Stability,* 1: 355-385.

MacKinlay, C. (1997), "Event Studies in Economics and Finance," *Journal of Economics Literature,* 35: 13-39.

Maquieira C. y D. Osorio (2000), "Anuncios de Cambios en Pagos de Dividendos y su Impacto en la Riqueza de los Accionistas: Evidencia Empírica en Chile," *Estudios de Administración,* 7 (1): 1-26.

Maquieira C. and M. I. Sierralta (1990), "El Efecto de la Devaluación en los Retornos Accionarios Chilenos," *Paradigmas en Administración,* 17: 1-18.

Mathur, I., K.C. Gleason, S. Dibooglu and M. Singh (2002), "Contagion Effects from the 1994 Mexican Peso Crisis: Evidence from Chilean Stocks," *The Financial Review,* 37: 17-34.

McWilliams, A. and D. Siegel (1997), "Event Studies in Management Research: Theoretical and Empirical Issues," *Academy of Management Journal,* 40: 626-57.

Marshall P. y E. Walker (2000), "Day-Of-The-Week and Size Effects in Emerging Markets: Evidence from Chile," *Revista de Análisis Económico,* 15 (2): 89-108.

Maynes, E. and J. Rumsey (1993), "Conducting Event Studies with Thinly Traded Stocks," *Journal of Banking and Finance,* 17: 145-157.

Morán, P. (2003), "Looking Back at the Controversy: Unexpected Wealth Effects of a Transitory Cause," *Abante,* 6 (2): 117-47.

Parisi, F., and D. Pérez (2000), "Cambios en el Rating de Bonos y su Efecto en los Precios Accionarios: El Caso Chileno," *Abante,* 3(2): 249-73.

Patell, J. (1976), "Corporate Forecasts Earnings per Share and Stock Price Behavior: Empirical Tests," *Journal of Accounting Research,* 14: 246-76.

Peterson, P. (1989), "Event Studies: A Review of Issues and Methodology," *Quarterly Journal of Business and Economics,* 28 (3): 36-66.

Roll, R. (1977), "A Critique of the Asset Pricing Theory's Tests," *Journal of Financial Economics,* 2: 129-176.

Rouwenhorst, K.G. (1999), "Local Return Factors and Turnover in Emerging Stock Markets," *The Journal of Finance,* 54: 1439-64.

Saens, R. (1999), "Premia in Emerging Market ADR Prices: Evidence from Chile," *Abante,* 2 (1): 51-70.

Tapia, M. y A. Tokman (2003), "Efectos de las Intervenciones Cambiarias en el Mercado Cambiario: El Caso de Chile," *Estudios de Economía,* 30 (1): 21-53.

Urrutia, J. (1995), "Test of Random Walk and Market Efficiency for Latin American Emerging Equity Markets," *The Journal of Financial Research,* 18: 299-309.

Wilson, B., A. Saunders and G. Caprio (2000), "Mexico's Financial Sector Crisis: Propagative Linkages to Devaluation," *The Economic Journal,* 110: 292-308.