Available online at www.sciencedirect.com
ScienceDirect
Journal of Corporate Finance 13 (2007) 129 – 145
Journal of CORPORATE FINANCE
www.elsevier.com/locate/jcorpfin

Event studies with a contaminated estimation period

Nihat Aktas [a,*], Eric de Bodt [a,b], Jean-Gabriel Cousin [b]

[a] *IAG Louvain School of Management, Université catholique de Louvain, Belgium*
[b] *ESA, Université de Lille 2, France*

Received 14 September 2005; received in revised form 13 September 2006; accepted 27 September 2006
Available online 13 November 2006

Abstract

Event studies are an important tool for empirical research in Finance. Since the seminal contribution of Fama et al. [Fama, E., Fisher, L., Jensen, M., Roll, R., 1969. The adjustment of stock prices to new information. International Economic Review 10, 1–21], there have been many enhancements to the classical test methodology. Somewhat surprisingly, the estimation period has attracted less interest. It is usually routinely determined as a fixed window prior to the event announcement day. In this study, we propose a test that reduces the impact of potentially unrelated events during the estimation period. Our proposition is based on a two-state version of the classical market model as a return-generating process. We present standard specification and power analyses. The results highlight the importance of explicitly controlling for unrelated events occurring during the estimation window, especially in the presence of event-induced increase in return volatility.

JEL classification: G14; G34
Keywords: Event study; Unrelated events; Markov switching regression model

1. Introduction

Since the seminal contribution of Fama, Fisher, Jensen and Roll (1969) (hereinafter referred to as FFJR), event studies have become a standard empirical methodology in research in Finance. Applications are so numerous that it is impractical to try to list them exhaustively. Many suggestions have been put forward to improve the basic empirical methodology. Brown and Warner (1980, 1985) analyzed the specification and power of several modifications of the FFJR

* Corresponding author. Place des Doyens 1, B-1348 Louvain-la-Neuve, Belgium. Tel.: +32 10 478439; fax: +32 10 478324.
E-mail address: aktas@fin.ucl.ac.be (N. Aktas).

doi:10.1016/j.jcorpfin.2006.09.001

approach. Ball and Torous (1988) explicitly took into account the uncertainty about event dates. Corrado (1989) introduced a non-parametric test of significance. Boehmer et al. (1991) proposed an adaptation of the standard methodology to tackle an event-induced increase in return volatility. Since this contribution, most methodological papers have explicitly controlled for this important phenomenon. Salinger (1992) suggested an adjustment of the abnormal returns standard errors robust to event clustering. Savickas (2003) recommended the use of a GARCH specification to control for the effect of time-varying conditional volatility. Aktas et al. (2004) advocated the use of a bootstrap method as an alternative to Salinger's (1992) proposition. Recently, Harrington and Shrider (in press) have argued that all events induce variance, and therefore tests robust to cross-sectional variation should always be used.

The estimation period has attracted less attention. It is most often defined as a period preceding the event, which is sufficiently long to enable the parameters of the chosen return-generating process to be properly estimated. In studies using daily data, a window going from day −250 to day −30 relative to the event date is usually (somewhat arbitrarily) chosen. This mechanical choice is, however, not free of complications. In particular, unrelated events may be present during the chosen estimation window, which bias the estimation of the return-generating process parameters. A natural solution seems to be to choose, on a case-by-case basis, an estimation window free of such contaminating events. This solution is, however, unreasonable for large-sample analyses. When compiling data for several hundred (or several thousand) observations (e.g., in the field of mergers and acquisitions (M&A), see Fuller et al., 2002; Mitchell and Stafford, 2000; Moeller et al., 2003), using such a "brute force" approach quickly becomes intractable.

It is worth emphasizing that, in many research areas, the presence of contaminating events during the estimation window is not just a presumption. Let us take the case of M&As again. Imagine that a specific bidder has, during the months preceding the transaction being studied, undertaken other operations, as frequently it appears to be the case (see, e.g., Asquith et al., 1983; Schipper and Thompson, 1983; Malatesta and Thompson, 1985; Fuller et al., 2002; Aktas et al., 2006). For example, out of the 4135 deals comprising the M&A sample used by Fuller et al. (2002),[1] 2721 (66%) would have been contaminated if the classical definition of the estimation window had been used.[2] The existence of such firm-specific events in the estimation window will most likely affect the estimation of the return-generating process and, in particular, the estimated variance of the parameters.

The approach we introduce in this paper to solve this contaminating event problem is essentially based on a combination of the well-established market model (Sharpe, 1963) and the more recent Markov switching regression models, largely introduced and developed by Hamilton (1989, 1994) and significantly extended by Krolzig (1997). Using a two-state market model, the estimated parameters of the model are less subject to the influence of contaminating events. This can be understood as a statistical filtering of the data. Another way to interpret our proposition is to see it as a better-specified return-generating model, which takes into account the probability of the occurrence of firm-specific events. From this perspective, our approach is in line with Roll's (1987) results. According to Roll, the true return-generating process seems to be better described by a mixture of two distributions: one corresponding to a state of information arrival, and the other to the normal return behavior.

[1] We thank the authors for providing us with access to their data set.

[2] This is the reason why the authors use the BETA-1 return-generating process, which is not affected by unrelated events during the estimation window, since there are no parameters to be estimated. However, according to our analyses, this approach is clearly less powerful than the other alternatives (see Section 5).

The analysis that we develop in this paper is now classical in the field of event study methodology (Brown and Warner, 1980, 1985). Using daily CRSP data, we carried out specification and power analyses while simulating a contaminated estimation period. We compare our approach to a classical set of alternatives (such as Corrado (1989), Boehmer et al. (1991), Savickas (2003)). The results show that (i) our approach is robust to the estimation window contamination and that (ii), in the context of an event-induced increase in return volatility, it dominates competing methods. Given the results of Harrington and Shrider (in press), following which event-induced increase in return volatility must be taken into account, we recommend the use of our approach.

The paper is organized as follows: Section 2 presents a simple model to show that ordinary least squares (OLS) methods overestimate the standard error of an individual firm's abnormal return when the true process is state dependent. Section 3 is devoted to a short review of the classical event study approaches and to the presentation of our test. Section 4 describes our experimental design. In Section 5, we present simulation results comparing the specification and the power of the test statistics being considered. Section 6 contains our summary and conclusions.

2. State dependent return-generating process and OLS inferences

In this section we show that OLS estimators overestimate the standard error of an individual firm's abnormal returns when the true return-generating process has two-states. We use the following notation:

- $\mathbf{X}_j$ is the matrix of explanatory variables for firm j;
- $\boldsymbol{R}_j$ and $\boldsymbol{R}_m$ are vectors of returns for firm j and for a market portfolio proxy;
- D_j is a dummy variable equal to 1 at the event date for firm j, and 0 otherwise;
- $\boldsymbol{b}_j$ is the vector of coefficient estimates for firm j.

When estimating firm j's abnormal returns using the market model (MM) as the return-generating process (Sharpe, 1963), $\mathbf{X}_j$ has three columns (the first column is set to 1 for the constant, the second to $\boldsymbol{R}_m$ and the third to D_j). The vector of coefficients $\boldsymbol{b}_j$ is composed of the intercept of MM, the β coefficient for firm j and a third coefficient (denoted by γ) capturing the firm j's abnormal returns at the announcement date:

$$\boldsymbol{R}_j = \mathbf{X}_j\boldsymbol{b}_j + \varepsilon_j = [1\boldsymbol{R}_m D_j]\begin{bmatrix}\alpha\\ \beta\\ \gamma\end{bmatrix} + \varepsilon_j. \tag{1}$$

Assuming homoskedascity, the covariance matrix of the OLS estimator is well known:

$$\mathrm{COV}_j^{\mathrm{OLS}}(\boldsymbol{b}_j|\mathbf{X}_j) = \sigma_j^2(\mathbf{X}_j{}'\,\mathbf{X}_j)^{-1}. \tag{2}$$

Assume now that the residuals are state dependent. S_t denotes the state variable. We consider the case of a two-state regime model. More precisely, we have a low variance regime (S_t=1) and a high variance regime (S_t=2):[3]

$$\begin{aligned}\boldsymbol{R}_j &= \mathbf{X}_j'\,\boldsymbol{b}_j + \varepsilon_{j,1} \quad \text{if } S_t = 1\\ \boldsymbol{R}_j &= \mathbf{X}_j'\,\boldsymbol{b}_j + \varepsilon_{j,2} \quad \text{if } S_t = 2.\end{aligned} \tag{3}$$

[3] Note that the market-model parameters $\boldsymbol{b}_j$ are identical in the two states. However, this setup could be easily extended to take regime dependent parameters into account.

The variance of the residuals for each state is set to:

$$E[\varepsilon_{j,1}\varepsilon_{j,1}'|X] = \sigma_{j,1}^2 I \quad \text{if } S_t = 1$$
$$E[\varepsilon_{j,2}\varepsilon_{j,2}'|X] = \sigma_{j,2}^2 I \quad \text{if } S_t = 2' \tag{4}$$

where $\sigma_{j,2}^2 > \sigma_{j,1}^2$. The high variance regime allows us to explicitly incorporate the presence of unrelated events into the statistical model.

The transition between the two regimes is governed by a Markov chain of order 1, for which the transition matrix is given by:

$$\mathbf{P} = \begin{bmatrix} p_{11} & 1-p_{22} \\ 1-p_{11} & p_{22} \end{bmatrix}, \tag{5}$$

where $p_{m,n}=p(S_t=m|S_{t-1}=n)$ corresponds to the probability of changing from state n to state m. The unconditional probability of the regime is given by (Hamilton, 1994, p. 683):

$$p(S_t = 1) = \frac{1-p_{22}}{2-p_{11}-p_{22}}; p(S_t = 2) = \frac{1-p_{11}}{2-p_{11}-p_{22}}. \tag{6}$$

Assuming strict exogeneity and that the transition probabilities are deterministic, the covariance matrix of the OLS estimator is a weighted average of the covariance matrix in the two-states, the unconditional probability being the weights:

$$\mathrm{COV}^{\mathrm{OLS}}(\boldsymbol{b}_j|\mathbf{X}_j) = p(S_t = 1)\sigma_{j,1}^2(\mathbf{X}_j'\,\mathbf{X}_j)^{-1} + p(S_t = 2)\sigma_{j,2}^2(\mathbf{X}_j'\,\mathbf{X}_j)^{-1}. \tag{7}$$

Considering the null hypothesis of no event, we are interested in the low variance regime covariance matrix, which is $\sigma_{j,1}^2(\mathbf{X}_j'\mathbf{X}_j)^{-1}$, and since $\sigma_{j,2}^2 > \sigma_{j,1}^2$, this leads to the inequality

$$\sigma_{j,1}^2(\mathbf{X}_j'\,\mathbf{X}_j)^{-1} \le p(S_t = 1)\sigma_{j,1}^2(\mathbf{X}_j'\,\mathbf{X}_j)^{-1} + p(S_t = 2)\sigma_{j,2}^2(\mathbf{X}_j'\,\mathbf{X}_j)^{-1}. \tag{8}$$

Eq. (8) shows that the standard error of the OLS estimates of the abnormal return (γ_j) is overestimated when the true generating process is a two-state process. This result offers a clear econometric foundation for the loss of power of classical event study methodology when the estimation window is contaminated by unrelated events.

3. Event study methodology

The seminal contribution of FFJR has been the starting point for an impressive diffusion of event study methodology in finance, accounting and economics. Its component steps are well known. In this section, we focus on the choice of the return-generating process and the construction of the statistical test of significance, two key points of concern in this area. The set of approaches on which we focus has been chosen either because they are used in classical empirical studies (e.g., the standardized cross-sectional test (Boehmer et al., 1991) and the RANK test (Corrado, 1989)) or because they contain features making them potentially well-suited to controlling for the presence of unrelated events during the estimation period (e.g., the BETA-1 approach and the GARCH approach developed by Savickas (2003)). Many other propositions have been made in the literature (e.g., Ball and Torous, 1988; Nimalendran, 1994) but they do not relate directly to our work.

3.1. Return-generating processes

Abnormal returns (AR) correspond to the forecast errors of a specific normal return-generating model (in Section 2, to simplify the exposition, we expressed them as the coefficient values of a dummy variable). Using MM as return-generating process[4] and employing the notation described in Section 2, we obtain

$$\boldsymbol{R}_j = \alpha_j + \beta_j \boldsymbol{R}_m + \varepsilon_j. \tag{9}$$

The residuals, ε_j, provide the estimates of abnormal returns. Classically, the residuals are supposed to be identically and independently (normally) distributed (IID). Numerous contributions have dealt with violations of this hypothesis. For example, Ruback (1982) suggests a way of coping with the existence of first-order auto-correlation in asset returns.

3.2. Statistical tests of significance

To introduce the different tests to be analyzed, we expand the notation used in Section 2 following Boehmer et al. (1991):

- N: number of firms in the sample;
- AR_{jE}: abnormal return of firm j on the event date;
- AR_{jt}: abnormal return of firm j on date t;
- T: number of days within the estimation period;
- TE: number of days within the event period;
- $\bar{\boldsymbol{R}}_m$: average return on the market portfolio during the estimation period;
- $\hat{S}_j$ standard deviation of firm j's AR during the estimation period;
- SR_{jE}: standardized AR of firm j on the event date, calculated as:

$$\mathrm{SR}_{jE} = \mathrm{AR}_{jE} \Bigg/ \left[\hat{S}_j \sqrt{1 + \frac{1}{T} + \frac{(R_{m,E} - \bar{R}_m)^2}{\sum_{t=1}^{T} (R_{m,t} - \bar{R}_m)^2}} \right]. \tag{10}$$

3.2.1. The BMP test

For each of the significance tests, we consider the null hypothesis of no cross-sectional average (cumulative) abnormal returns around the event date. The BMP (Boehmer, Musumeci and Poulsen) test is similar in spirit to Patell's (1976) test. However, Boehmer et al. (1991) use the estimated cross-sectional variance of the standardized abnormal returns instead of the theoretical

[4] Brown and Warner (1980, 1985) show that the results of short-term event studies are not sensitive to the choice of a specific return-generating process. Aktas et al. (2004) reported similar results in a European context.

value. This adaptation captures the event-induced increase in return volatility. The BMP test takes the following form:

$$Z_{\mathrm{BMP}} = \frac{\frac{1}{N}\sum_{j=1}^{N} \mathrm{SR}_{jE}}{\sqrt{\frac{1}{N(N-1)}\sum_{j=1}^{N}\left(\mathrm{SR}_{jE}-\sum_{i=1}^{N}\frac{\mathrm{SR}_{iE}}{N}\right)^{2}}}. \tag{11}$$

3.2.2. The BETA-1 test

The BETA-1 test may be viewed as a drastic simplification of the BMP test. The assumed return-generating process is the market-adjusted model, which amounts to imposing $\beta=1$ and $\alpha=0$ in Eq. (9). The test is based on cross-sectional estimates of the standard deviation of the event-day abnormal returns (ARE). We therefore obtain:

$$Z_{\mathrm{BETA-1}} = \frac{\frac{1}{N}\sum_{j=1}^{N} \mathrm{AR}_{jE}}{\sqrt{\frac{1}{N(N-1)}\sum_{j=1}^{N}\left(\mathrm{AR}_{jE}-\sum_{j=1}^{N}\frac{\mathrm{AR}_{jE}}{N}\right)^{2}}}. \tag{12}$$

The BETA-1 test constitutes, in our framework, an interesting alternative. $Z_{\mathrm{BETA\text{-}1}}$ does not use data from the estimation window and is therefore free of the potential biases generated by contaminating events.

3.2.3. The Corrado (1989) RANK test

Corrado (1989) introduced a test based on the ranks of abnormal returns. The RANK test merges the estimation and event windows in a single time series. Abnormal returns are sorted and a rank is assigned to each day. If K_{jt} is the rank assigned to firm j's abnormal return on day t, then the RANK test is given by

$$T_{\mathrm{CORRADO}} = \frac{\frac{1}{N}\sum_{j=1}^{N}(K_{jE}-\bar{K})}{S(K)}, \tag{13}$$

where $\bar{K}$ is the average rank and $S(K)$ is the standard error, calculated as

$$S(K) = \sqrt{\frac{1}{T+TE}\sum_{t=1}^{T+TE}\left(\frac{1}{N}\sum_{j=1}^{N}(K_{jt}-\bar{K})\right)^{2}}. \tag{14}$$

The use of ranks neutralizes the impact of the shape of the AR distribution (e.g., its skewness and kurtosis and the presence of outliers). It should therefore represent an attractive alternative way of neutralizing contaminating events within the estimation window. Corrado (1989), Corrado and Zivney (1992), and Campbell and Wasley (1993) provide an in-depth analyses of this approach.

3.2.4. The GARCH test

The conditional time-varying behavior of the variance of returns has been widely recognized in finance since it was first pointed out by Engle (1982). Building on the Bollerslev (1986) generalized autoregressive conditional heteroskedastic (GARCH) approach, Savickas (2003) suggested the use of the return-generating process

$$\begin{aligned}&\boldsymbol{R}_{j,t}=\alpha_j+\beta_j\boldsymbol{R}_{m,t}+\gamma_jD_{j,t}+\eta_{j,t}\\&\eta_{j,t}\sim N(0,h_{j,t}),\\&h_{j,t}=aj+b_jh_{j,t-1}+c_j\eta^2_{j,t-1}+d_jD_{j,t}\end{aligned} \tag{15}$$

where $h_{j,t}$ is the conditional time-varying variance and a_j, b_j, c_j and d_j are the coefficients of the GARCH(1,1) specification. $D_{j,t}$ is a dummy variable equal to 1 at the event date for firm j, and 0 otherwise. As in Section 2, the γ_j coefficient captures the abnormal return at the announcement date.

The conditional variance $h_{j,t}$ provides a natural estimator of the AR variance. Savickas (2003) used it to standardize the AR before proceeding with the BMP test. In this setting, Eq. (10) is replaced by

$$\mathrm{SR}_{jE}=\hat{\gamma}j\Big/\sqrt{\hat{\mathrm{h}}_{j,E}}. \tag{16}$$

Savickas (2003) shows that the GARCH test allows him to control for the time-varying variance of AR and the event-induced increase in return volatility. It is worth mentioning that the simulation results he provides rely on abnormal returns generated on five consecutive days. This very specific simulation procedure must be kept in mind when interpreting the specification and power results that we present in Section 5. As the main effect of contaminating events is an increase in variance during the estimation window, the GARCH test also constitutes an attractive alternative to the BMP test. The GARCH(1,1) specification might indeed provide a way of controlling for the impact of unrelated events on the estimate of the variance in AR (probably at the cost of an increase in persistence).

3.2.5. The two-state market model test (TSMM)

The idea underpinning the TSMM test is simple: the presence of unrelated events within the estimation window has an impact on the (*ex post*) estimation of the AR variance. Classical tests, such as the BMP, will overestimate the variance of the residuals during the estimation period in such conditions, leading to a downward bias in the significance test (i.e. less likelihood of rejecting the null hypothesis) during the event window.

To deal with this bias, the TSMM test relies on the Markov switching regression framework developed by Hamilton (1989, 1994). As discussed in Section 2, we assume that the return-generating process can be adequately modeled by a two-state process,[5] in which one regime has

[5] This hypothesis is supported by unpublished results. Using the approach developed by Krolzig (1997), we found that a two-state model was an adequate representation of the return-generating process in most cases. Three-regime decomposition appeared to be justified only in the presence of strong outliers.

normal variance and the other high variance. Note that the MM parameters are assumed to be the same in the two regimes.[6] Eq. (3) then becomes

$$\begin{aligned} \boldsymbol{R}_{j,t} &= \alpha_j + \beta_j \boldsymbol{R}_{m,t} + \gamma_j D_{j,t} + \varepsilon_{j,S,t} \\ \varepsilon_{j,S,t} &\sim N(0, \sigma^2_{j,s}), \end{aligned} \tag{17}$$

where S is a state variable taking value 1 in the low variance state and value 2 in the high variance state ($\sigma^2_{j,2} > \sigma^2_{j,1}$). The proposed model is a direct and parsimonious extension of the classical MM. As for the previous GARCH model, the γ_j coefficient corresponds to the estimated event-day abnormal return. We use the estimated standard error of γ_j to standardize the AR. Our standardized abnormal return is therefore

$$\mathrm{SR}_{jE} = \hat{\gamma}_j / \mathrm{SE}(\gamma_j), \tag{18}$$

where $\mathrm{SE}(\gamma_j)$ corresponds to the standard error of the γ_j coefficient. The test can then be constructed using the same approach as for ZBMP:

$$Z_{\mathrm{TSMM}} = \frac{\frac{1}{N}\sum_{j=1}^{N} \mathrm{SR}_{jE}}{\sqrt{\frac{1}{N(N-1)}\sum_{j=1}^{N}\left(\mathrm{SR}_{jE} - \sum_{i=1}^{N}\frac{\mathrm{SR}_{iE}}{N}\right)^2}} \tag{19}$$

The estimation of Eq. (17) is based on a maximum-likelihood approach.[7] The estimated probability of being in a specific state on a specific date is one of the interesting byproducts of this approach. In some specific cases, it allows the reasons for an increase in variance to be explored.[8]

4. Experimental design

Our investigation of the specification and power of the TSMM test follows the procedure introduced by Brown and Warner (1980, 1985) and used repeatedly since then (see, e.g., Corrado, 1989; Boehmer et al., 1991; Corrado and Zivney, 1992; Cowan, 1992; Cowan and Sergeant, 1996; Savickas, 2003).

4.1. Data and sample selection

Our universe of firms is composed of companies included in the Center for Research in Security Prices (CRSP) daily returns file from January 1, 1973 through December 31, 2004.[9] As

[6] We have also implemented a more general version of the TSMM approach, where the MM parameters are regime dependent. Since the results are similar, we have not reported them in this paper. The assumption that the MM parameters are the same for both regimes does not rely on a specific economic/financial foundation. It is simply a product of the principle of parsimony, which allows for much more efficient numerical estimation. We thank an anonymous referee for having stressed this point.

[7] All the estimations presented in this paper have been performed with the Ox econometric software, using the Krolzig MSVAR package. We thank Professor J. Hamilton for advising us on the use of this package.

[8] In a recent paper, Cousin and de Launois (2006) produce empirical evidence supporting that the high variance state is generated by intensive information arrival.

[9] We chose to start the sampling process in 1973 in order to draw the samples from an homogenous population, Nasdaq stocks being only included in the CRSP files from 1973 onwards.

our market portfolio we used the CRSP value weighted index. All firms and event dates were randomly chosen with replacement such that each firm/date combination had an equal chance of being chosen at each selection. For each replication, we constructed 1000 samples of 50 firms. The estimation window length was 225 days and the event date was situated at day 250. Like Savickas (2003), our sampling process excluded securities with missing returns during the 250-day interval. Moreover, to be included in the samples, securities need to have at least 100 non-zero returns over the estimation window, and no zero return due to a 'reported price' on the event-day.

4.2. Contaminating the estimation window

The aim of our simulation was to study the specification and power of the TSMM test, as compared to other tests, when the data in the estimation window were deliberately contaminated. To simulate significant events, we injected AR into the estimation window (on a contaminated day) which was twice the standard deviation of the actual stock. To generate stochastic shocks, we followed the method proposed by Brown and Warner (1985) by adding another two demeaned returns randomly drawn from the estimation window. Moreover, we generated both positive and negative AR to represent the unknown nature of the events likely to affect the estimation window. The sign (±) of the simulated AR ($2^*\sigma_{\boldsymbol{R}}$) was determined by random sampling from a Bernoulli distribution. The transformed return for security j on a contaminated day, denoted $\boldsymbol{R}'_{j,t}$, is therefore computed as

$$\boldsymbol{R}'_{j,t} = \boldsymbol{R}_{j,t} \pm 2^* \sigma_{\boldsymbol{R}_j} + (\boldsymbol{R}_{j,X} - \overline{\boldsymbol{R}}_j) + (\boldsymbol{R}_{j,Y} - \overline{\boldsymbol{R}}_j), \tag{20}$$

where $\boldsymbol{R}_{j,t}$ is the actual return, $\boldsymbol{R}_{j,X}$ and $\boldsymbol{R}_{j,Y}$ are returns randomly selected from the estimation period, and $\overline{\boldsymbol{R}}_j$ is the average return in the estimation period.

The number and nature of the events during the estimation window was determined in two steps. First, a random sample was drawn from a Poisson distribution with a mean of two. This value, denoted by λ^*, represents the number of events during the estimation window. Events were then randomly assigned to specific days in the estimation window (by random sampling from a uniform distribution). Finally, the length (in number of days) of each event was again randomly sampled from a Poisson distribution, this time with a mean of four. This approach allowed random events to be generated in the estimation window.

Fig. 1 presents a typical result. The solid line is the time series of the initial returns, and the dotted line is the time series of the returns obtained after the generation of random, contaminating events. Three events were generated, at dates $T=2$, 51 and 119. The standard error of the initial estimation period was 3.73%, and after the generation of events it became 4.01%. Fig. 1 shows that the simulated shocks are not particularly uncommon events. They merely represent disruptions in the normal return-generating process, of the size and amplitude we might expect from various corporate event announcements (e.g., M&A operations, share buy-backs, earning announcements).

4.3. Simulating abnormal performance

We generated abnormal returns at the event date in the same way as Brown and Warner (1980 1985) by adding a constant to each stock return observed on day 0 (event date). The abnormal

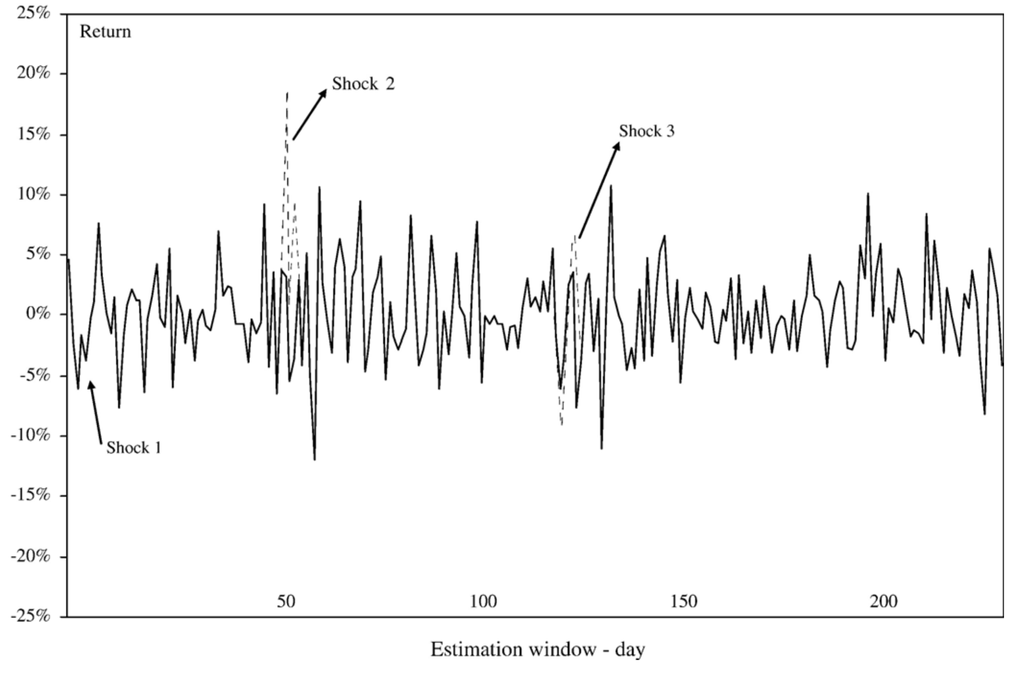

Fig. 1. This figure provides an example of estimation window contamination using the procedure described in Section 4.2. The solid line is the time series of initial returns and the dotted line is the time series obtained after event generation. Three events are generated, at dates $T=2$, 51 and 119. The standard error of the estimation window is 3.73% initially and 4.01% after event generation.

performance simulated is 0% for the specification analysis and +1% for the power analysis.[10] To produce stochastic abnormal returns (the event-induced variance phenomenon) we again followed Brown and Warner's approach. Each security's day 0 return, $R_{j,0}$, was transformed to triple its variance by adding two demeaned returns randomly drawn from the estimation window. The event-day transformed return was obtained using the procedure described in Section 4.2 above.

5. Empirical results

Our results are presented in Tables 1–4, which show the rejection rates for different cross-sectional test statistics under different conditions. We compare specifications (Tables 1 and 3) and powers (Tables 2 and 4) without (Tables 1 and 2) and with (Tables 3 and 4) event-induced increase in volatility.[11] In each case, we present the results without (Panel A) and with (Panel B) contaminating events during the estimation window. Results are presented for the BMP, RANK, GARCH, BETA-1 and TSMM tests.

[10] Unreported results show that, beyond a simulated abnormal performance of 2%, all the methods we compared are very powerful.

[11] According to Harrington and Shrider (in press) all events induce variance. We have retained in the paper all the analyses of the "no event-induced variance" cases (Tables 1, 2) to allow comparison with previous papers in the field.

Table 1
Rejection rates of test statistics: no event-induced returns–no event-induced variance

	Significance level		
	1%	5%	10%
Panel A. Without contaminating events			
BMP	1.60%	5.20%	9.80%
RANK	0.90%	4.50%	8.40%
GARCH	0.70%	5.90%	11.40%
BETA-1	0.80%	4.80%	9.70%
TSMM	1.50%	4.60%	9.50%
Panel B. With contaminating events			
MANUAL	0.80%	4.50%	10.50%
BMP	1.60%	6.70%	12.10%
RANK	0.40%	3.80%	8.10%
GARCH	2.50%	6.70%	10.80%
BETA-1	0.90%	5.20%	11.00%
TSMM	0.70%	4.30%	9.80%
Panel C. Confidence intervals for rejection rate			
95% Confidence interval	0.4%–1.6%	3.7%–6.4%	8.1%–11.9%
99% Confidence interval	0.2%–1.8%	3.2%–6.8%	7.6%–12.4%

The specification analysis with no event-induced increase in return volatility, showing the rejection rates for different cross-sectional test statistics when an event creates no abnormal returns and no increase in variance. BMP corresponds to Boehmer et al.'s (1991) test, RANK is as in Corrado (1989), GARCH is the test studied by Savickas (2003) and BETA-1 is the cross-sectional test using the constrained version of the market model. TSMM and MANUAL are, respectively, the two-state market model extension and the manually filtered version of Boehmer et al.'s (1991) standardized cross-sectional approach. Panel A provides the analysis when the estimation window is not contaminated, and Panel B when it has been contaminated using the procedure described in Section 4.2. Panel C provides the confidence limits for rejection frequency in 1000 binomial trials.

For the B panels (with contaminating events) we added, in the spirit of Thompson (1988), a manually filtered approach (MANUAL). In this approach, days contaminated by simulated events during the estimation window were manually removed from the sample of observations before the implementation of the Boehmer et al. (1991) cross-sectional test of significance. This cleaning approach provides a clear benchmark against which the robustness of the set of tests used in the presence of contaminating events can be evaluated.[12]

5.1. Tests with no change in the return variance

Table 1, Panel A shows that, in the absence of contaminating events and event-induced increase in return volatility, all the tests we survey are (reasonably) well specified. We provide in Table 1 Panel C confidence intervals[13] to check whether actual rejection rates differ from the nominal ones. None of the tests present a rejection rate outside the 95% and 99% confidence

[12] It is worth noting that, using real returns, the manually filtered approach only controls for simulated events. Any real contaminating events are not neutralized by the MANUAL approach.
[13] To estimate the confidence intervals, we use the same methodology as in Cowan and Sergeant (1996).

Table 2
Rejection rates of test statistics: event-induced returns–no event-induced variance

	Significance level		
	1%	5%	10%
Panel A. Without contaminating events			
BMP	52.40%	74.40%	83.50%
RANK	71.40%	88.10%	92.20%
GARCH	33.10%	52.10%	62.80%
BETA-1	29.30%	51.00%	61.40%
TSMM	56.70%	75.50%	81.90%
Panel B. With contaminating events			
MANUAL	52.30%	74.30%	81.80%
BMP	44.70%	67.40%	76.40%
RANK	65.00%	82.50%	89.00%
GARCH	26.70%	47.40%	59.90%
BETA-1	28.80%	49.60%	60.40%
TSMM	57.90%	76.30%	82.70%

The power analysis with no event-induced increase in return volatility, showing the rejection rates for different cross-sectional test statistics when an event creates an increase in returns of 1% and no increase in variance. BMP corresponds to Boehmer et al.'s (1991) test, RANK is as in Corrado (1989), GARCH is the test studied by Savickas (2003) and BETA-1 is the cross-sectional test using the constrained version of the market model. TSMM and MANUAL are, respectively, the two-state market model extension and the manually filtered version of Boehmer et al.'s (1991) standardized cross-sectional approach. Panel A provides the analysis when the estimation window is not contaminated, and Panel B when it has been contaminated using the procedure described in Section 4.2.

interval. The impact of contaminating events is analyzed in Table 1 Panel B. We get almost the same results as in Panel A. None of the rejection rates seem to be statistically different from the nominal ones.

Table 2 presents the results of the power analysis, still in the absence of event-induced increase in return volatility. Without contaminating event (Panel A), the RANK test clearly outperform the other alternatives, and the power of the BMP and the TSMM tests are quite comparable. With contaminating events (Panel B), the RANK test is still the most powerful test. However, in this setting, the TSMM test is more powerful than the BMP one, and it has a power very close to the benchmark approach, which is the MANUAL test. With and without contaminating events the GARCH and the BETA-1 are the less powerful tests. Nevertheless the results for the GARCH test must be interpreted with caution, keeping in mind that we have only simulated a one-day shock (as in, for example, Brown and Warner (1980, 1985) and Boehmer et al. (1991)) and not five consecutive shocks (as in Savickas (2003)).

Tables 1 and 2 show that the GARCH and BETA-1 approaches allow the presence of contaminating events to be controlled for. However this comes at the cost of a severe reduction in power. The RANK test appears to be the most attractive alternative: it is both less sensitive to the presence of contaminating events (in terms of specification error) and more powerful. The TSMM test is, in the absence of event-induced increase in return volatility, a second best approach.

Since in practice all events induce variance (Harrington and Shrider, in press), the most realistic case to analyze is the one where we simulate also a variance increase on the event date, and in this setting, we know from the literature that the RANK test is clearly miss-specified (see, e.g., Cowan and Sergeant, 1996; Serra, 2002; Savickas, 2003). The next sub-section is devoted to this analysis.

Table 3
Rejection rates of test statistics: no event-induced returns–but event-induced variance

	Significance level		
	1%	5%	10%
Panel A. Without contaminating event			
BMP	1.40%	6.50%	11.10%
RANK	4.10%	14.40%	21.40%
GARCH	0.70%	5.60%	10.30%
BETA-1	1.10%	4.80%	10.80%
TSMM	0.80%	5.80%	10.20%
Panel B. With contaminating events			
MANUAL	1.30%	6.20%	9.50%
BMP	1.50%	5.90%	10.20%
RANK	4.40%	12.40%	20.40%
GARCH	1.10%	5.20%	11.20%
BETA-1	1.10%	3.60%	9.00%
TSMM	0.90%	4.60%	8.90%
Panel C. Confidence intervals for rejection rate			
95% Confidence interval	0.4%–1.6%	3.7%–6.4%	8.1%–11.9%
99% Confidence interval	0.2%–1.8%	3.2%–6.8%	7.6%–12.4%

The specification analysis with an event-induced increase in return volatility, showing the rejection rates for different cross-sectional test statistics when an event creates an increase in variance and no increase in returns. BMP corresponds to Boehmer et al.'s (1991) test, RANK is as in Corrado (1989), GARCH is the test studied by Savickas (2003) and BETA-1 is the cross-sectional test using the constrained version of the market model. TSMM and MANUAL are, respectively, the two-state market model version and the manually filtered version of Boehmer et al.'s (1991) standardized cross-sectional approach. Panel A provides the analysis when the estimation window is not contaminated, and Panel B when it has been contaminated using the procedure described in Section 4.2. Panel C provides the confidence limits for rejection frequency in 1000 binomial trials.

5.2. Tests with a variance increase on the event date

In Tables 3 and 4, we simulate an event-induced increase in return volatility. The results reported above for the specification analysis are qualitatively confirmed, except for the RANK test. Without any surprise, Panel A and Panel B of Table 3 indicate clearly that the RANK test is poorly specified in the presence of an event-induced increase in return volatility. This result has also been reported in previous studies.

With respect to power, Table 4 shows that the presence of an event-induced increase in return volatility strongly affects the power of all the tests (compare the rejection rates in Table 2 to those displayed in Table 4). The RANK and TSMM tests seem to be the most powerful approaches to detect the 1% simulated event-day abnormal returns. However, since these two tests differ significantly with respect to the specification error (see Table 3), it is not possible to compare their power as such. In order to overcome this problem, we resorted to a graphical method, denoted 'size–power curves', as proposed by Davidson and MacKinnon (1998).[14] These size–power curves allow the power of alternative test statistics that do not have the same size (specification) to be compared. Using the simulation techniques described in Section 4, we computed the power and size of the tests for 100 different theoretical significance levels (between 0% and 100%).

[14] Special thanks to Luc Bauwens for having suggested this method.

Table 4
Rejection rates of test statistics: event-induced returns–event-induced variance

	Significance level		
	1%	5%	10%
Panel A. Without contaminating events			
BMP	18.00%	35.70%	47.50%
RANK	28.90%	48.50%	58.50%
GARCH	8.80%	21.90%	30.90%
BETA-1	8.20%	21.20%	31.70%
TSMM	18.80%	37.50%	49.80%
Panel B. With contaminating events			
MANUAL	14.20%	32.30%	44.20%
BMP	11.30%	28.40%	39.00%
RANK	21.40%	39.30%	50.10%
GARCH	5.70%	16.80%	26.60%
BETA-1	7.60%	19.70%	29.50%
TSMM	14.80%	31.90%	44.30%

The power analysis with an event-induced increase in return volatility, showing the rejection rates for different cross-sectional test statistics when an event creates an increase in returns of 1% and an increase in variance. BMP corresponds to Boehmer et al.'s (1991) test, RANK is as in Corrado (1989), GARCH is the test studied by Savickas (2003) and BETA-1 is the cross-sectional test using the constrained version of the market model. TSMM and MANUAL are, respectively, the two-state market model extension and the manually filtered version of Boehmer et al.'s (1991) standardized cross-sectional approach. Panel A provides the analysis when the estimation window is not contaminated, and Panel B when it has been contaminated using the procedure described in Section 4.2.

Fig. 2 presents the results. The size–power curve of the TSMM test clearly predominates over the curves of the alternative tests, both without and with contaminating events. However, it is important to note that the supremacy of the TSMM approach is more important when the estimation window is contaminated (Panel B). In other words, holding the size constant (or for a comparable level of size), the TSMM approach is the most powerful; BETA-1 and GARCH are the less powerful tests. Another interesting result that is worth pinpointing is that the size–power curves of the BETA-1 and GARCH tests are almost the same in Panel A (without contaminating events) and Panel B (with contaminating events), suggesting that the power of these two approaches is not significantly affected by contaminating events. This result justifies somehow the use of the BETA-1 approach in recent empirical work (e.g., in the field of M&As, see Fuller et al., 2002; Moeller et al., 2003; Aktas et al., 2006).

5.3. Practical recommendations

It follows from our analysis that the choice of method depends on the conditions of the study. As already reported in the literature (e.g., Cowan and Sergeant, 1996; Serra, 2002; Savickas, 2003), if the return variance is unlikely to increase on the event date, the RANK test offers the best specification and power for general use. However, in the presence of an event-induced increase in return volatility, the TSMM offers the best compromise between robustness to the presence of contaminating events and power.

It is important to emphasize that, according to Boehmer et al. (1991), it is not uncommon for an event to be accompanied by an increase in the cross-sectional dispersion of stock returns. More recently, Harrington and Shrider (in press) have gone a step further by stressing the fact that all

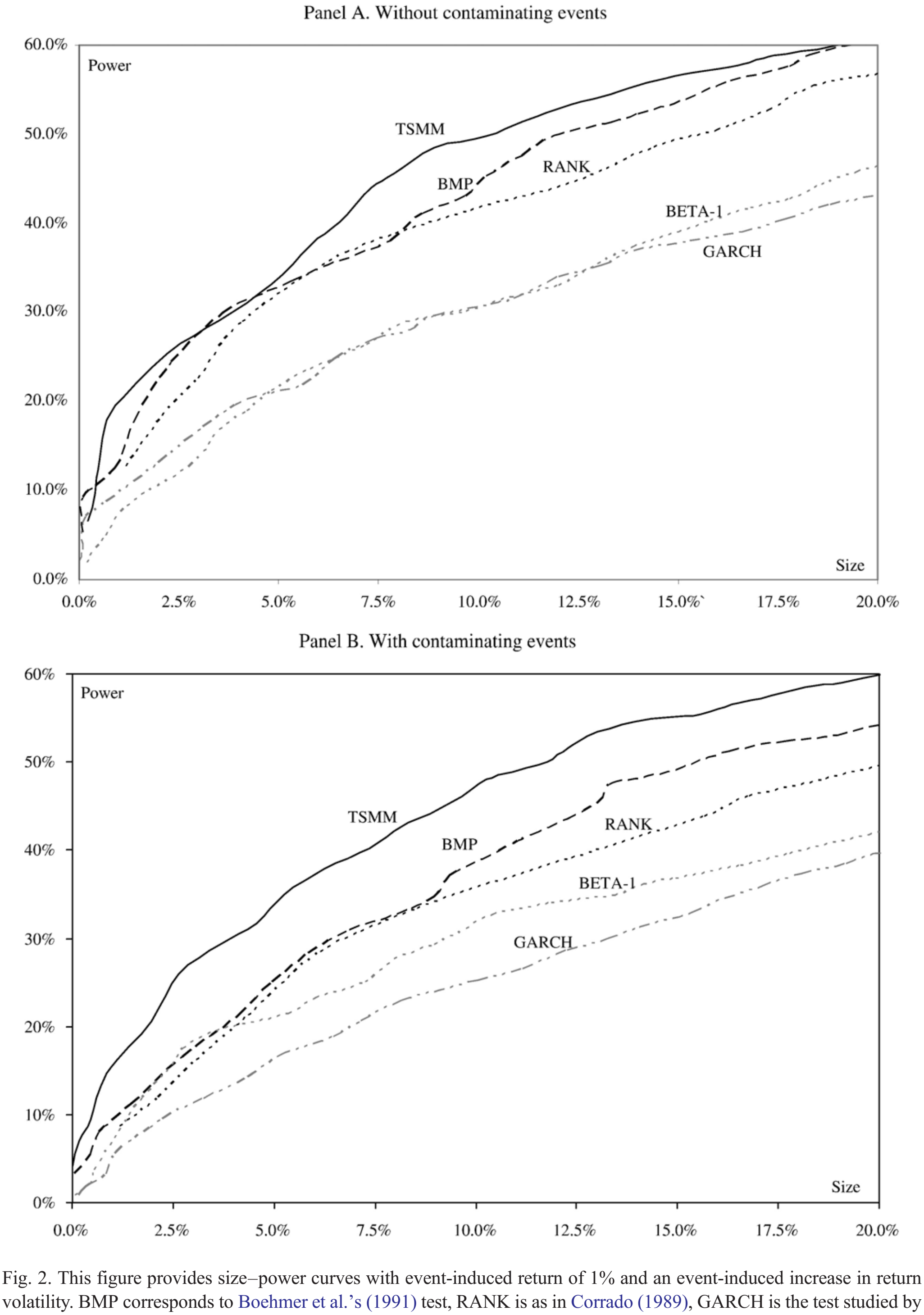

Fig. 2. This figure provides size–power curves with event-induced return of 1% and an event-induced increase in return volatility. BMP corresponds to Boehmer et al.'s (1991) test, RANK is as in Corrado (1989), GARCH is the test studied by Savickas (2003) and BETA-1 is the cross-sectional test using the constrained version of the market model. TSMM is the two-state market model extension of Boehmer et al.'s (1991) standardized cross-sectional approach.

events induce variance. The key contribution of their theoretical and empirical analyses of tests for non-zero mean abnormal returns is to show that tests which are robust to cross-sectional variation should always be used. According to them the BMP approach is a good candidate for a robust test. In this paper, we showed that the TSMM test is more robust than the BMP within the context of unrelated events in the estimation window.

6. Conclusion

Analysis of the estimation window has attracted less interest in the event study literature. In this paper we have shown that unrelated events during the estimation window do affect the specification and the power of standard event-study methods. To alleviate this problem we propose the use of the TSMM approach, which is built on a two-state market model extension of Boehmer et al.'s (1991) standardized cross-sectional approach.

We have compared the TSMM approach to four alternative cross-sectional tests. These are (1) the standardized cross-sectional test (Boehmer et al., 1991), (2) the RANK test (Corrado, 1989), (3) the BETA-1 approach, which does not require any estimation window data, and (4) the GARCH-based approach introduced by Savickas (2003).

The TSMM test studied here explicitly models the volatility of the residuals as a two-state model, one corresponding to a low variance regime, and the other to a high variance regime. Our simulation results show that, in the presence of an event-induced increase in return volatility, the TSMM approach provides the best compromise in terms of the specification error and power of the test. As event-induced variance is a well-established empirical fact (see Harrington and Shrider, in press), we suggest the use of the TSMM approach.

Acknowledgements

We are grateful for constructive comments from the participants at the SIFF 2002 (Rennes, September), AFFI 2003 (Lyon, June), EFMA 2003 (Helsinki, June), and MFS 2004 (Istanbul, July) meetings. Key suggestions by Luc Bauwens, Christophe Perignon, Evangelos Sekeris, and the participants at the Paris IX-Dauphine Research Seminar (Paris, March 2003), and more specifically from Edith Ginglinger, Myron Slovin and Marie Sushka are acknowledged. We especially wish to thank Kathleen Fuller, Jeffry Netter and Mike Stegemoller who give us access to their dataset and James Hamilton for his advice. All remaining errors are ours. Part of this research was carried out while the second author was a visiting scholar in the Finance Department of the Anderson School (University of California at Los Angeles).

References

Aktas, N., de Bodt, E., Roll, R., 2004. Market response to European regulation of business combinations. Journal of Financial and Quantitative Analysis 39, 731–758.

Aktas, N., de Bodt, E., Roll, R., 2006. Hubris, learning, and M&A decisions: empirical evidence. Finance Working Paper, vol. 02–06. University of California, Los Angeles.

Asquith, P., Bruner, R.F., Mullins, D.W., 1983. The gains to bidding firms from merger. Journal of Financial Economics 11, 121–139.

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

Boehmer, E., Musumeci, J., Poulsen, A., 1991. Event study methodology under conditions of event induced variance. Journal of Financial Economics 30, 253–272.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.
Brown, S., Warner, J., 1980. Measuring security price performance. Journal of Financial Economics 8, 205–258.
Brown, S., Warner, J., 1985. Using daily stock returns: the case of event studies. Journal of Financial Economics 14, 3–31.
Campbell, C., Wasley, C., 1993. Measuring security price performance using daily Nasdaq return. Journal of Financial Economics 33, 73–92.
Corrado, C., 1989. A nonparametric test for abnormal security price performance in event studies. Journal of Financial Economics 23, 385–395.
Corrado, C., Zivney, T., 1992. The specification and power of the sign test in event study hypothesis tests using daily stock returns. Journal of Financial and Quantitative Analysis 27, 465–478.
Cousin, J.G., de Launois, T., 2006. News intensity and conditional volatility on the French stock market. Finance 27, 7–60.
Cowan, A., 1992. Nonparametric event study tests. Review of Quantitative Finance and Accounting 2, 343–358.
Cowan, A., Sergeant, A., 1996. Trading frequency and event study test specification. Journal of Banking and Finance 20, 1731–1757.
Davidson, R., MacKinnon, J.G., 1998. Graphical methods for investigating the size and power of hypothesis tests. The Manchester School 66, 1–26.
Engle, R., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. Econometrica 50, 987–1008.
Fama, E., Fisher, L., Jensen, M., Roll, R., 1969. The adjustment of stock prices to new information. International Economic Review 10, 1–21.
Fuller, K., Netter, J., Stegemoller, M., 2002. What do returns to acquiring firms tell us? Evidence from firms that make many acquisitions. Journal of Finance 57, 1763–1793.
Hamilton, J., 1989. A new approach to the economic analysis of non-stationary time series and the business cycle. Econometrica 57, 357–384.
Hamilton, J., 1994. Time Series Analysis. Princeton University Press.
Harrington, S. Shrider, D. in press. All events induce variance: analyzing abnormal returns when effects vary across firms. Journal of Financial and Quantitative Analysis.
Krolzig, H.M., 1997. Markov-switching vector autoregressions. University of Oxford, UK Lecture Notes in Economics and Mathematical Systems, vol. 454. Springer. XIV.
Malatesta, P., Thompson, R., 1985. Partially anticipated events: a model of stock price reactions with an application to corporate acquisitions. Journal of Financial Economics 14, 237–250.
Mitchell, M., Stafford, E., 2000. Managerial decisions and long-term stock price performance. Journal of Business 73, 287–330.
Moeller, S., Schlingemann, F., Stulz, R., 2003. Firm size and the gains from acquisitions. Journal of Financial Economics 73, 201–228.
Nimalendran, M., 1994. Estimating the effects of information surprises and trading on stock returns using a mixed jump-diffusion model. Review of Financial Studies 7, 451–473.
Patell, J., 1976. Corporate forecasts of earnings per share and stock price behavior: empirical tests. Journal of Accounting Research 14, 246–276.
Roll, R., 1987. R-squared. Journal of Finance 43, 541–566.
Ruback, R., 1982. The effect of discretionary price control decisions on equity values. Journal of Financial Economics 10, 83–105.
Salinger, M., 1992. Standard errors in event studies. Journal of Financial and Quantitative Analysis 27, 39–53.
Savickas, R., 2003. Event-induced volatility and tests for abnormal performance. Journal of Financial Research 16, 165–178.
Schipper, K., Thompson, R., 1983. Evidence on the capitalized value of merger activity for acquiring firms. Journal of Financial Economics 11, 85–119.
Serra, A.P., 2002. Event study tests: a brief survey. Working Paper FEP, vol. 117. University of Porto.
Sharpe, W., 1963. A simplified model for portfolio analysis. Management Science 9, 277–293.
Thompson, J., 1988. More methods that make little difference in event studies. Journal of Business Finance and Accounting 15, 77–86.