Cyree, Ken B.; DeGennaro, Ramon P.

Working Paper

A generalized method for detecting abnormal returns and changes in systematic risk

Working Paper, No. 2001-8

Provided in Cooperation with:
Federal Reserve Bank of Atlanta

Suggested Citation: Cyree, Ken B.; DeGennaro, Ramon P. (2001) : A generalized method for detecting abnormal returns and changes in systematic risk, Working Paper, No. 2001-8, Federal Reserve Bank of Atlanta, Atlanta, GA

This Version is available at:
http://hdl.handle.net/10419/100708

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
A Generalized Method for Detecting Abnormal Returns and Changes in Systematic Risk

Ken B. Cyree and Ramon P. DeGennaro

Working Paper 2001-8
April 2001
A Generalized Method for Detecting Abnormal Returns and Changes in Systematic Risk

Ken B. Cyree and Ramon P. DeGennaro

Federal Reserve Bank of Atlanta
Working Paper 2001-8
April 2001

Abstract: The authors generalize traditional event-study techniques to allow for event-induced parameter shifts, shifting variances, and firm-specific event periods. Their method, which nests traditional methods, also permits systematic risk to change gradually during the event period and exit the period at higher or lower levels. The authors use their approach to study 132 banks that acquired other institutions between 1989 and 1995. The authors find a significant change in the systematic risk of the acquiring firms, significant ARCH effects, and an event period that ends before the date of the announcement. None of these results is detectable using conventional methods. These results imply that (1) event studies that cannot account for information leakage may be biased, and (2) changes in systematic risk can occur in the absence of abnormal returns, and (3) regulators, investors and bank managers must evaluate each acquisition on its own merits; reliance on averages can mask important distinctions across acquisitions.

JEL classification: G14, G21, G34

Key words: event studies, mergers, acquisitions, banking
A Generalized Method for Detecting Abnormal Returns and Changes in Systematic Risk

Introduction

Traditional event-study methods use a three-step procedure. First, the researcher selects a model of returns. Second, he computes abnormal returns during some event interval as the difference between realized returns and the expected returns conditioned on this model. Finally, he evaluates the statistical significance of these abnormal returns in any of several ways.

The value of this method is indisputable. Although many researchers have challenged traditional approaches on various grounds, empirical evidence in several studies (e.g. Brown and Warner (1980, 1985), Malatesta (1986) and Henderson (1990)) has concluded that the event-study approach is quite robust for detecting abnormal mean returns.

We argue, though, that this perspective of event studies is limited. That is, it focuses only on mean returns during the event, while ignoring a wide range of other interesting financial events. When researchers have studied changes in volatility during the event period, they do so primarily to improve the third step of the event-study process, that of determining the statistical significance of abnormal returns (e.g. Boehmer, Musumeci and Poulsen (1991)). We believe that changes in risk (both systematic and unsystematic) are interesting in their own right. In addition, the choice of event interval is not obvious. Researchers implicitly acknowledge this by providing results over several different intervals.

We propose a new methodological framework that relaxes certain restrictions imposed by typical event study approaches, thereby letting us explore these issues. Our approach, which nests traditional methods, offers several advantages. These are: 1) allowing for heteroskedastic variances, 2) using firm-specific event periods (rather than assuming that all events have the same beginning and
ending points for all firms) that need not be symmetrical around the announcement date; indeed, they need not even contain the announcement date, 3) allowing for multiple economic events, 4) computing abnormal returns using measures of systematic risk that may be unique to each firm, and which may change during the event period, and 5) allowing for permanent changes in systematic risk. These last two permit detection of a class of event that is beyond the scope of traditional approaches; in fact, they are undetectable without multistep, ad hoc procedures. We call these risk-shifting events.

Others have advocated some of these improvements. Ibbotson (1975) uses a procedure that permits changes in beta, but imposes the restriction that betas are identical across firms. In addition, it uses only a small fraction of the available data. Ball and Torous (1988) suggest a method to estimate the probability that an event occurs on any given day within a pre-specified period, and which allows changes in the conditional mean and variance. This is a useful advance, but is limited in important ways. First, the method permits at most a single event during the event period. Second, the abnormal return is constrained to be proportional to the individual security’s return variance. This restriction is very problematic. In addition, the method imposes a further restriction: this is the only mechanism by which abnormal returns may vary across firms. Finally, Ball and Torous do not consider changes in systematic risk due to the event. Brown, Lockwood and Lummer (1985) are probably closest in spirit to our approach, but to our knowledge, none of these other methods estimates all parameters simultaneously, none generalizes the approach to include ARCH processes, and none have been applied to the banking industry.

It remains to be demonstrated that ignoring these issues is economically important. Brown and Warner (1980, 1985), Malatesta (1986) and Henderson (1990) all have established the simpler
methods’ reputation for robustness. None of these papers, however, has explored the implications of relaxing the implicit restrictions that our method permits. Malatesta, for example, finds “...no evidence that joint generalized least squares is superior to simpler procedures” (p. 27). However, he does not explore the impact of imposing an arbitrary, constant event interval for all firms, and does not explore the importance of changes in systematic risk. Still, it is only fair to ask whether economists who use traditional methods are susceptible to incorrect inferences.

To explore this, we use a sample of 132 banks that acquired other institutions. Madura and Wiant (1994) study acquirer abnormal returns for the 36 months after a bank merger. They find significantly negative CAARs for their sample of bank acquirers. Houston and Ryngaert (1994) use daily returns and find negative and significant average abnormal returns for acquiring banks. Both Houston and Ryngaert (1994) and Madura and Wiant (1994), however, use the same event interval for all firms, and all of the above ignore ARCH effects. If the findings of these studies are correct, then acquiring banks decrease shareholder wealth. Possibly, however, the findings could be due to biased event study methods.

We find that although the traditional methods’ reputation for robustness is not destroyed, it is tarnished; our method does indeed lead to different inferences. Traditional results indicate negative cumulative average abnormal returns (CAARs) on acquiring banks for the day prior to the takeover announcement. However, our new method finds insignificant CAARs for all event days and periods. More important, there is no reason for the residuals from our approach and those from traditional models to be identical, which means that the corresponding abnormal returns differ substantially. Even if a researcher were satisfied with the limitations of standard event-study approaches, it seems
unreasonably optimistic to expect acceptable results in a *cross-sectional* regression. We view this as especially important, given the increasing frequency with which financial economists conduct such analysis.¹

Although our model overturns traditional models’ finding of negative CAARs for day (-1), the most important results of our new approach are undetectable by traditional methods. First, the evidence suggests that the actual event period begins *and ends* well before the usual announcement date. On average, the event period begins almost three weeks before the earliest announcement in the *Wall Street Journal (WSJ)*, and ends nearly a week before that announcement. Finally, we uncover substantial changes in systematic risk during the event period, and find that for many acquirers, these shifts are permanent.

The paper is organized as follows. Section I develops our model. Section II describes the data and empirical results using a sample of 132 bank acquisitions. Section III reports cross section results. Section IV contains implications of our findings. Section V summarizes.

**I. Model Formulation**

Our goal is a generalized procedure for detecting abnormal returns that explicitly controls for heteroskedastic variances, while permitting measures of systematic risk to change during the event period, perhaps permanently. Further, this procedure must explicitly identify (possibly) unique event periods for each firm, which may be asymmetrical.

A typical event study uses the market model to obtain abnormal returns for firm *i*:

---

¹ For example, from March 1995 through July 1997, about two-thirds of the event studies published in the *Journal of Finance* involved some sort of cross sectional analysis, although not all used regressions. For the *Journal of Financial and Quantitative Analysis*, the corresponding figure is almost 90%.
\[ R_{i,t} = \gamma_{i,0} + \gamma_{i,1}R_{m,t} + \nu_{i,t}. \]  

with \( \text{var}\nu_{i,t} = \sigma_i^2 \). Thus, event studies using this approach assume that market model parameters are stationary. However, research shows that this assumption is tenuous, at best. For example, de Jong et al. (1992) estimate a model that allows for time dependent betas and reject the hypothesis of a constant beta for individual stocks. Scholes and Williams (1977) show that the issue of stationarity is somewhat mitigated by estimating parameters using returns before and after the event. However, this approach does not allow for changes in beta during the event period when calculating abnormal returns and interpreting results.

Typical event study methodology also assumes homoskedasticity. At least as early as Mandlebrot (1963), though, researchers have known that, “Large changes tend to be followed by large changes---of either sign---and small changes tend to be followed by small changes.”² It seems likely that significant corporate events, such as acquisitions, could induce further clustering of price changes.

Lockwood and Kadiyala (1988) develop an approach that allows different systematic risk parameters before, during, and after events. Additionally, their model permits firm-specific event periods. Each firm in a typical event study has a unique event window instead of all firms having, for example, a window extending from day(-10) to day(+10). Using monthly returns, Lockwood and Kadiyala find that the endpoints of event periods differ across firms. They reject the hypothesis of an event window of month (-30) to month (+30) for all firms. Lockwood and Kadiyala also confirm the existence of stochastic betas for firms involved in stock splits, and find time dependence in the error

\[^2\]Engle (1982) provides a useful approach for handling these changes, the now-familiar ARCH model. For a summary of ARCH models and their application in finance, see Bollerslev, Chou, and Kroner (1992).
variance (i.e., time induced heteroskedasticity). However, for reasons discussed below, they use a two-step procedure. This introduces errors in variables in the second step. In addition, parameters are not estimated jointly.

To generalize the return generating model in Equation (1), we begin by relaxing the restriction that systematic risk, captured by $\gamma_1$, is constant during the entire sample period. To do this, we first define $T_1$ and $T_2$ as the beginning and end of the event period, respectively. We also define two indicator variables, $D_{1,i,t} = 1$ if $T_1 < t < T_2$ and zero otherwise, and $D_{2,i,t} = 1$ if $t > T_2$ and zero otherwise. Thus, $D_{1,i,t} = 1$ during the event period, and $D_{2,i,t} = 1$ after the event period. Writing $\Delta \beta_{i,1}^E$ and $\Delta \beta_{i,1}^p$ to signify the changes in systematic risk during the event period and during the post-event period, respectively, we can substitute into Equation (1) to obtain:

$$R_{i,t} = \beta_{i,0} + [\beta_{i,1} + \Delta \beta_{i,1}^E D_{1,i,t} + \Delta \beta_{i,1}^p D_{2,i,t}] R_{m,t} + \epsilon_{i,t},$$

where we change notation of the coefficients and residuals to recognize that they may differ from Equation (1) if, in fact, Equation (1) incorrectly restricts the true return-generating process. Unlike the model in Equation (1), the model in Equation (2) permits $\beta_i$ to vary during and after the event period.

In principle, $\Delta \beta_{i,1}^E$ and $\Delta \beta_{i,1}^p$ can take any form. This subject is worth studying in its own right, but is beyond the scope of this paper. However, it is important to provide flexibility and to permit forms that other researchers have found. For example, Bar-Yosef and Brown (1977) find that systematic risk first rises and then falls around stock splits. They confirm this for both rising and falling markets. Therefore, our functional form must encompass this quadratic form. A plausible alternative stems from the insight that information about post-event covariances might trickle into prices through
insider trading, or through revisions in merger probabilities. This suggests a linear change, with an extremum at the end of the event period. Our specification must permit this, as well.

Although other functional forms are possible, we set:

\[ \Delta E_{i,1} = \beta_{i,2} (T_{1i} - t)(t - T_{2i}) + \beta_{i,3} (t - T_{1i}) \]  

This specification permits \( \beta_{i,1} \) to follow a continuous concave (or convex) function, and exit the sample period at a different level, thus permitting permanent changes in systematic risk.

It is worth noting that nothing prevents \( \beta_{i,2} = \beta_{i,3} \). That is, constant beta models are nested within our approach. Similarly, linear changes in beta are captured if \( \beta_{i,2} = 0 \) but \( \beta_{i,3} \) is nonzero. Thus, our functional form can handle linear changes in beta, quadratic changes in beta, and no changes in beta. This last item is important, for it is easy to construct scenarios producing abnormal returns with no change in systematic risk. Traditional methods are more likely to work well in such cases, but even then, the choice of event interval is problematic for those approaches. Event interval choice is particularly important in our information rich economy where rumor plays a key role in security pricing.

For example, Haw, Pastena, and Lilien (1990) find significant abnormal returns from four to eight weeks before the event was announced in the \textit{WSJ} or 750 other sources, such as the \textit{New York Times}.

Substituting Equation (3) and Equation (4) into Equation (2) yields:

\[ R_{i,t} = \beta_{i,0} + [\beta_{i,1} + [\beta_{i,2} (T_{1i} - t)(t - T_{2i}) + \beta_{i,3} (t - T_{1i})]D_{1i,t} + [\beta_{i,3} (T_{2i} - T_{1i})]D_{2i,t} ] R_{m,t} + \epsilon_{i,t} \]  

Rearranging, we obtain:
Equation (6) lets us study wealth effects while permitting event-induced parameter shifts. We generalize still further by permitting the conditional variance to follow an ARCH process:

$$\text{Var}(\varepsilon_{i,t}) = h_{i,t} = \alpha_{i,0} + \alpha_{i,1}\varepsilon_{i,t-1}^2.$$  \hfill (7)

Equation (7) is an ARCH(1) model of the conditional variance for acquiring bank $i$ at time $t$.

The mean and conditional variance [Equations (6) and (7)] are estimated jointly using maximum likelihood. Incorporating additional explanatory variables, such as size and the book-to-market ratio (see Fama and French, 1992) or an interest-rate index (see DeGennaro and Thomson, 1995) is straightforward. Extensions to higher-order ARCH or GARCH processes are also possible. Because this approach to event studies uses an ARCH specification, we call our model EVARCH. Note that if $\beta_2 = \beta_3 = \alpha_1 = 0$ for any firm $i$, then Equations (6) and (7) reduce to the standard market model.

To summarize, $\beta_{i,0}$ and $\beta_{i,1}$ are the intercept and systematic risk before the event period for firm $i$; $\beta_{i,2}$ and $\beta_{i,3}$ permit changes in systematic risk during and after the event period for firm $i$. $D_1$ is a binary variable equal to one during the announcement period, and $D_2$ is a binary variable indicating the post-announcement period; $T_1$ is the beginning and $T_2$ is the end of the event period for each bank.

In the absence of ARCH, $\beta_1$ in Equation (1) estimates beta during the estimation period, and corresponds to $\beta$ in the CAPM. Systematic risk during the event period (when $D_1=1$ and $D_2=0$) for firm $i$ is:

$$\beta_1 + \beta_2(T_1 - t)(t - T_2) + \beta_3(t - T_1) + \varepsilon_{i,t},$$  \hfill (8)
where the error terms in Equations (8) and (9) below show that systematic risk is an estimate and is not necessarily constant during the event period (and, therefore, that $\epsilon_{i,t}$ is a compound disturbance).

Because the current time ($time=t$) enters as a quadratic term, $\beta_2$ indicates the shape of the change in systematic risk during the event period. If $\beta_2$ is negative, systematic risk is convex during the event period; if $\beta_2$ is positive, systematic risk is concave. This would be consistent with Bar-Yosef and Brown (1977). If $\beta_2 = 0$, then systematic risk is linear. This handles the cases in which information about post-event covariances trickles in gradually, perhaps due to insider trading or revisions in merger probabilities, with systematic risk reaching an extreme at either the beginning or the end of the event period.

After the event period ends (when $D_1=0$ and $D_2=1$), systematic risk is:

$$\beta_1 + \beta_2(T_2 - T_1) + \nu_i$$

Thus, a non-zero $\beta_3$ coefficient implies a permanent change in systematic risk after the beginning of the event period.

Panel A of Figure 1 illustrates an event that increases systematic risk during the event period and permanently increases risk. Panel A also illustrates the case where the event period actually ends before the announcement date, perhaps due to information leakages or insider trading, as in Meulbroek and Hart (1997). Panel B of Figure 1 illustrates a decrease in systematic risk during the event period, with post-event systematic risk lower than the pre-event level. In the second panel, the announcement date is within the event period, although the event window is asymmetric. In each of these cases, traditional event study methods would be unable to capture the changes in systematic risk or the asymmetric
window around the true event period. Because these shifts reflect changes in systematic risk, they have important financial implications. Within the context of our sample of banks, for example, risk is important for regulatory issues such as pricing deposit insurance or determining the optimal frequency of regulatory examinations. In a more general sense, investors may need to rebalance their portfolios, and managers may wish to consider corporate restructurings, including adjustments in financial leverage.

This generalized approach offers four important advantages. First, if the model in Equation (1) incorrectly restricts the data generating process, then parameters are biased and the residuals are likely to be poor estimates of abnormal returns. Second, unless the researcher selects the correct event interval, perhaps simply by good luck, results are unlikely to be useful. Third, changes in systematic risk are important in themselves, yet commonly used traditional methods are incapable of detecting them. Finally, the approach allows for multiple economic events between $T_1$ and $T_2$, including sequential rumors or information leakages.

It is also likely that unsystematic risk changes during acquisitions. For example, the final price, presence of competitors, regulatory approval and other such factors are typically unknown when acquisitions are announced. One may reasonably assume that this takeover activity affects the unsystematic risk of firms involved, especially because information regarding a takeover tends to cluster in time. This is consistent with Lockwood and Kadiyala (1988). The clustering of price changes noted by de Jong et al. (1992) suggests that an ARCH model is appropriate for studying event-induced parameter shifts. This implies that $\alpha_i$ will be positive and significant in the conditional variance, Equation (7).
Equation (6) also allows for the work of Brown, Lockwood, and Lummer (1985). They show that using inaccurately imposed endpoints for the event period introduces a bias into parameter estimates that is proportional to the specification error. In the model from Equation (6), this amounts to inducing errors in the dummy variables used to define the beginning and ending points of the event window. That is, setting the dummy variables for the beginning and ending points in most event studies is ad hoc. Within the context of Equation (6), the practical problem is that \( D_1, D_2, T_1 \) and \( T_2 \) are jointly determined. That is, researchers must know \( D_1 \) and \( D_2 \) in order to estimate \( T_1 \) and \( T_2 \), but must know \( T_1 \) and \( T_2 \) in order to determine \( D_1 \) and \( D_2 \). Brown, Lockwood, and Lummer (1985) and Lockwood and Kadiyala (1988) address this with a two-step procedure. In the first step, they treat \( D_1 \) and \( D_2 \) as known and estimate the parameters of Equation (6) using generalized least squares. In the second step, they use these estimated parameters and maximize the likelihood with respect to the endpoints of the interval, \( T_1 \) and \( T_2 \).

While a useful contribution, this two-step procedure is prone to problems in the second step due to the use of generated regressors. Our method mitigates the problem by maximizing the likelihood over all possible event periods. The log-likelihood function for firm \( i \) is:

\[
L = -(1/2) \log 2\pi - (1/2) \log \sum_{r=1}^{T} h_r - (1/2) \sum_{r=1}^{T} \left( \epsilon_r^2 / h_r \right) .
\]  

(10)

Our procedure first maximizes the log-likelihood function with \( D_1 \) set equal to unity for event day(-30) and \( D_2 \) equal to unity on day(-29). In economic terms, this assumes that the event period begins on day(-30) and ends on day(-29). We repeat this with \( D_2 \) set to unity on both day(-29) and day(-28), while \( D_1 \) remains at unity for day(-30). This assumes that the event period begins on day(-}
30) and ends on day(-28). After $D_2$ has been incremented through day(+30), we repeat the process with $D_1$ set to one on day(-29) and $D_2$ equal to unity on day(-28). The log-likelihood function is estimated for all possible values of $D_1$ (day -30 to -29 in this case) and $D_2$ (day -29 to +30 in this case). In this way, we need not bias the results by imposing a constant, arbitrary event period, and we need not use estimated parameter values when determining endpoints. Note that our method permits leaks, rumors, or insider trading to release some or even all information before the announcement date. The largest of all maximized likelihoods indicates which combination of binary variables produces the best estimates. Thus, the procedure involves choosing the largest value of the log-likelihood function for each firm $i$:

$$\text{MAX} \left( \text{MAX}_{T_1<T_2} \ L[\beta_0, \beta_1, \beta_2, \beta_3, T_1, T_2, \alpha_0, \alpha_1] \right)$$

(10a)

We follow standard methods and interpret residuals from the model in Equations (6) and (7) as abnormal returns due to the event under study. But because our model allows firm-specific event periods and controls for shifts in systematic risk and ARCH effects, the residuals -- and therefore, the abnormal returns -- in general differ from those obtained using traditional methods. Thus, inferences drawn from the abnormal returns may also differ from the standard market model given in Equation (6).

Application of the model in Equation (1) assumes that parameters are stationary, homoskedastic, and that a single, symmetric event period is appropriate for all firms. If results differ from those of a more general model that relaxes these restrictions, then standard market model results are suspect. The ambiguity of prior findings could be due to methodological deficiencies.

---

3 We restrict the procedure to day(-30) through day(+30) due to computational considerations; even so, the number of likelihood maximizations for this scheme is 1830 per firm. To prevent mistaking a local maximum for a global maximum, we use different sets of starting values and different maximization algorithms.
II. Data and Empirical Results

The data consist of member banks that acquired other banks from 1989 to 1995, and are from the Board of Governors of the Federal Reserve System. Banks must have only one takeover during the 241 trading-day interval to remain in the sample. The final sample consists of 132 banks or bank holding companies that have returns on the Center for Research in Securities Prices (CRSP) tape. For these models, as well as the EVARCH model, the proxy for the market return is the CRSP equally weighted index. The geometric average return for a missing return interval replaces missing returns, and we delete any firm with more than 10 missing returns during the estimation period.

We first compute results from the standard market model and the Scholes-Williams model under the usual assumption of fixed, symmetrical event periods for each firm in the sample [Equation (1)]. As such, these results are the base case for comparison with the model that has time-specific parameters and ARCH effects [Equations (6) and (7)]. We calculate CAARs for these models using standard event-study methodology from Mikkelson and Partch (1986).

Results are in Table 1. Day(-1) CAARs are negative and although not quite significant at the 5% level, they are easily significant at the 10% level for both traditional models. This suggests that wealth effects for bidding banks are negative, on average, and that information is released one day prior to the announcement date. Researchers usually explain this as the news having been announced after the close of trading on the stock exchange, yet soon enough to make the newspapers and electronic media the next morning. CAARs are insignificant for day(0) and for event windows (-5,5) and (-10,10) for both estimation techniques. CAARs are mostly negative, except for day(0); however, Wilcoxon ranked sign tests cannot reject the null hypothesis of 50% positive residuals for either model in any event
period in Table 1. The findings for two-day CAARs contrast to the significantly negative CAARs of Houston and Ryngaert (1994), who use a sample from 1985-1991. However, they do find insignificant CAARs for 1990 and 1991. Though far from conclusive, this suggests the possibility that bank acquisitions have been better received by financial markets in recent years. Another interpretation is that the market began anticipating such acquisitions, and stock prices adjusted in advance of the actual announcements.

<table>
<thead>
<tr>
<th>Event Interval:</th>
<th>Day(-1)</th>
<th>Day(0)</th>
<th>Day(-1 to 0)</th>
<th>Day(-5 to +5)</th>
<th>Day(-10 to 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAAR’s from the Standard Market Model</td>
<td>-0.00194*</td>
<td>0.00168</td>
<td>-0.00026</td>
<td>0.00275</td>
<td>-0.00310</td>
</tr>
<tr>
<td></td>
<td>(-1.888)</td>
<td>(0.556)</td>
<td>(-0.941)</td>
<td>(0.168)</td>
<td>(-1.203)</td>
</tr>
<tr>
<td>CAAR’s from the Scholes-Williams model</td>
<td>-0.00187*</td>
<td>0.00169</td>
<td>-0.00018</td>
<td>0.00315</td>
<td>-0.00272</td>
</tr>
<tr>
<td></td>
<td>(-1.821)</td>
<td>(0.571)</td>
<td>(-0.884)</td>
<td>(0.061)</td>
<td>(-0.929)</td>
</tr>
</tbody>
</table>

* = significant at the ten-percent level.

We also calculate CAARs for the EVARCH model using residuals from Equations (6) and (7) during the event period. The average abnormal return is the mean of the residuals across the entire sample of 132 firms for the individual event period as defined by the maximum likelihood endpoints (\(T_{1i}\) and \(T_{2i}\)) for each firm \(i\). Our estimate of the standard deviation of the average abnormal return is:

$$\sigma_{EVARCH} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \sum_{t=T_{1i}}^{T_{2i}} ri}$$

(11)

where \(N = 1,\ldots,132\). This multi-period conditional standard deviation assumes that residuals are serially uncorrelated, a standard assumption in event studies.\(^4\) The t-statistic for the null hypothesis of zero abnormal returns is the CAAR computed from Equations (6) and (7) divided by its standard error:

\(^4\) Autocorrelation makes estimators inefficient and biases standard errors. Most event studies ignore this possibility. Note, too, that multi-period standard errors are unavailable for some other methods, such as Boehmer, Musumeci, and Poulsen (1991).
where \( K_t \) is the total number of residuals for all 132 firms from time \( T_1 \) to \( T_2 \). As is true for traditional methods, these test statistics assume that residuals are uncorrelated across firms. Because events occur at different times, the assumption of independence is appropriate.

Results are in Table 2. The estimated coefficient \( \beta_1 \), which measures systematic risk prior to the event, averages 0.83, with a standard error of 0.041. These correspond quite closely to the values for the market model (0.85 and 0.045) and the Scholes-Williams model (0.86 and 0.44). These latter models are unable to handle changes in systematic risk. EVARCH, though, indicates increased systematic risk during the event period, and that systematic risk increases at a decreasing rate (\( \beta_2 \) is positive). In addition, our model allows for abnormal returns without changes in beta and is therefore a more general case. Although the average value of \( \beta_2 \) for all 132 banks does not differ statistically from zero, the estimate is statistically significant at the 5% level in 51 cases. This is far above what might be expected due to chance. Thus, the insignificance of the average value might simply be due to our sample size. Differences in the behavior of systematic risk across firms during the event again highlight the importance of not relying on averages to prove parameter stationarity when doing cross sectional analysis. Turning to \( \beta_3 \), the parameter estimate is significant in 22 cases, and these changes are large enough to make the average for the entire sample statistically different from zero according to both the t-test and the Wilcoxon sign test. In addition, the estimates of \( \beta_2 \) and \( \beta_3 \) imply plausible changes in systematic risk. Figure 2 uses the estimates in Table 2 to construct the implied change in systematic risk during the sample. Systematic risk averages 0.83 for banks prior to the acquisition. During the event
period, systematic risk increases at a decreasing rate to a maximum of 1.34 before declining. The representative bank then exits the event period at a new, *permanent* level of systematic risk equal to 1.004.

Table 2: Average maximum likelihood estimates of EVARCH model parameters and beginning and ending points for the entire 1989 to 1995 sample of bidding bank stocks.

<table>
<thead>
<tr>
<th>Parameter (number out of 132 that are significant at 5%)</th>
<th>Entire Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Estimate</td>
<td>Standard Error</td>
</tr>
<tr>
<td>$\beta_0$ (6)</td>
<td>-0.00014**</td>
</tr>
<tr>
<td>$\beta_1$ (95)</td>
<td>0.83404***</td>
</tr>
<tr>
<td>$\beta_2$ (51)</td>
<td>0.02075</td>
</tr>
<tr>
<td>$\beta_3$ (22)</td>
<td>0.01935*</td>
</tr>
<tr>
<td>$\alpha_0$ (123)</td>
<td>0.00023**</td>
</tr>
<tr>
<td>$\alpha_1$ (64)</td>
<td>0.14908***</td>
</tr>
<tr>
<td>$T_1$</td>
<td>-14.01520***</td>
</tr>
<tr>
<td>$T_2$</td>
<td>-4.79545***</td>
</tr>
</tbody>
</table>

*=significant at the ten-percent level.     **=significant at the five-percent level.     ***=significant at the one-percent level.

a = Wilcoxon sign test is significant at the five-percent level.
b = Wilcoxon sign test is significant at the one-percent level.  The sample size is 132.

We argue that changes in systematic risk are a reasonable economic result of bank acquisitions. One rationale for interstate banking is risk reduction through diversification. However, this refers to *unsystematic* risk. The more the bank diversifies geographically and increases in size, the more its stock returns tend to reflect the entire economy. Also, banks are increasingly acquiring or merging with nonbanks, which tends to move systematic risk closer to the market level of risk. Therefore, *systematic* risk could well increase if the pre-merger bank has a systematic risk lower than the market, as is the case in our sample.

Turning to the conditional variance, we see that EVARCH identifies 64 significant $\alpha_1$ coefficients (the average $\alpha_1$ is 0.307 for these 64 firms). That is, about half of the time, the assumption
of constant variance is rejected in favor of ARCH. To the extent that this holds for other studies which assume constant variances, it may lead to biased t-tests and incorrect conclusions. If, as is likely, the variance increases during the event period, t-ratios will be systematically overestimated through use of a systematically understated variance from the pre-event period. Both Brown and Warner (1985) and Boehmer, Musumeci and Poulsen (1991) have shown that even small increases in variance can lead to substantial over-rejection of the null hypothesis of no abnormal returns.

Because EVARCH uses case-specific event periods, there is no general symmetrical event window (e.g., day(-5) to day(+5)) that is typical of event studies. How reasonable is the usual assumption of a fixed, symmetrical event window? The answer lies in the mean estimates of $T_1$ and $T_2$, the beginning and end of the event period. These results are also in Table 2. The mean estimate for the first day of the event period ($T_1$) for the entire sample is day(-14.01). The mean estimate of the last day of the event period ($T_2$) is day (-4.80).

We temporarily lay aside the location of the event window to discuss its other features. The estimated endpoints indicate that the average event window is much larger than the traditional (-1,0) interval, yet much smaller than other commonly used windows of (-20 to +20) or (-10 to +10). Even this conceals potentially important variation across firms. For example, the estimated start of the event period, $T_1$, ranges from day (-29) to day (7), and the estimated end of the event period, $T_2$, ranges from day(-23) to day(10). In fact, we can reject both the hypothesis that $T_1 = -1$ (t-statistic of -15.9) and that $T_2 = 0$ (t-statistic of -5.51) for the widely used day(-1,0) event period. In fact, any symmetrical event window centered on the announcement date would have the beginning, end, or both rejected in this sample of bank stocks. Hence, it is important that users of event-study methods recognize that
endpoints may conform to a presupposed window on average, but that case-specific event periods are
needed for proper analysis. Studies using smaller windows risk missing important economic effects,
while those using longer windows (but not case-by-case event periods) can bias results by combining
abnormal returns from the event period with those from outside of the event period. This is especially
ture if the abnormal returns are used in cross-sectional analysis to explore events or policy implications.

One could argue that $T_i$ could never fall after day(0), but this is not true in general. For this to
hold, investors first must be sufficiently rational, and the model must be well-specified. We do have
some confidence in these conditions. However, it is also necessary that the event be sufficiently large to
affect the return generating process, Equations (6) and (7). If not, then noise in the data could place $T_i$
virtually anywhere. Perhaps most important, though, we define day(0) as the day of the first
announcement. Conceivably, information about the distribution of returns to the combined firm or the
likelihood of regulatory approval might trickle in over a period of a few days. Indeed, unlike Ball and
Torous (1988) or traditional methods, our approach is well-suited for detecting multiple or cumulative
events of this type. In any event, only eight firms in our sample of 132 have $T_i > 0$, and only two of
these have $T_i > 2$.

To us, though, the most important feature of the estimated event window is its location. The
event window ends about a week before the first announcement in the WSJ. Several plausible
explanations exist for the event period ending before the announcement. The first reason applies to all
methods, whether they be traditional or new: the news media may fail to report the event, or the
researcher may fail to locate that report. We have no way of determining the frequency of media error,
but we can gauge the likelihood of researcher error. As is true for many studies, we use the WSJ as the
source of our event dates; this may not be the first media announcement. To explore this possibility, we spot-check several events using LEXIS. Two events do show evidence of earlier information release. Local newspapers reported that preliminary talks were taking place about six months before the *Journal’s* announcement in one case, and more than seven weeks in the second case. But even if the market reacted strongly to these local announcements, the impact on the location of the event window, which is restricted to begin at most 30 business days before the *Journal’s* announcement, is likely to be small. In such extreme cases, our method is no different from any other: the researcher misses the event.

A second, and much more interesting explanation for the location of the event window, is information leakage. Possibly, insiders learn of the coming announcement and act on that information before the media reveal it to the rest of the market. A third reason is that EVARCH admits the possibility of events not detectable using traditional approaches. We have labeled this a risk-shifting event. That is, our method allows the parameters measuring systematic risk to change, and the most likely values for the set of parameters utilize this advantage, even though the mean abnormal returns are zero. Lastly, large corporate events such as takeovers are often rumored, and sometimes market participants trade on these rumors before the actual announcement of the event, thereby creating a partially (or fully) anticipated event, as shown by Haw, Pastena, and Lilien (1990).

The importance of allowing for risk-shifting events depends both on the goals of the research and the nature of the data. Researchers interested solely in wealth effects might conceivably ignore risk-shifting events, although we argue that attempts to measure abnormal returns without accurate risk measures are problematic. Other researchers and regulators, though, must consider changes in risk.
For example, regulators interested in setting rates for risk-based capital standards or deposit insurance can hardly ignore changes in systematic risk. Similarly, researchers studying long-term stock returns cannot do so without considering systematic risk.

Results of the EVARCH event study are in Table 3. Day(-1) CAARs are negative but insignificant when accounting for ARCH and non-constant event risk. That is, negative wealth effects shown by traditional models are due to specification error, and not significant negative market reaction. Other results are similar to the standard approaches reported in Table 1, with EVARCH CAARs insignificantly negative on the event day itself. In all cases, the CAAR is insignificant for the entire event period using EVARCH. There is no convincing evidence that acquiring banks suffer wealth declines in our sample of acquisitions.

<table>
<thead>
<tr>
<th>Maximum Likelihood Event Period</th>
<th>Day(-1)</th>
<th>Day(0)</th>
<th>Day(-1) to Day(0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EVARCH Cumulative Average Abnormal Returns</td>
<td>0.02617</td>
<td>-0.00282</td>
<td>0.00244</td>
</tr>
<tr>
<td></td>
<td>(0.594)</td>
<td>(-0.140)</td>
<td>(0.119)</td>
</tr>
</tbody>
</table>

The t-statistics for EVARCH CAARs are dramatically smaller than those from traditional models. For example, the EVARCH t-statistics in Table 3 are less than a tenth of the size of the significant day(-1) CAARs in Table 1 using traditional market models. This result is intuitive because the standard error estimated in non-event periods is generally lower than the standard error during the event period. The lower standard errors create larger t-statistics and increase Type I error; they bias results towards rejecting the null hypothesis of no abnormal returns when the null is true. Hence, our

5 Note that the lower t-statistics are a result of higher variance and ARCH effects not captured by traditional models. A
sample illustrates how apparently significant results using traditional models could trace to changes in systematic risk or omitted ARCH effects.

These results collectively imply that for this sample, standard event study methods provide satisfactory results only on average, but more importantly, only for certain purposes. That is, mean abnormal returns from standard models and EVARCH models are similar, but standard methods miss important risk-shifting events. In addition, even if a researcher were satisfied with the limitations of standard event-study approaches, it seems unreasonably optimistic to expect acceptable results in cross-sectional regressions. We now explore the ramifications of this in our sample.

**III. Cross-Sectional Analysis**

Even though traditional event-study methods detect abnormal returns at only the 10% significance level, we conduct cross-sectional analysis to test for differences in acquiring banks’ CAARs around the takeover announcement in the interests of completeness. For the traditional models, we use two-day CAARs from the event day and the day preceding the event. Though ad hoc, the use of two-day CAARs is fairly standard in event study cross-sectional analysis (see Cornett and De (1991) or Cornett and Tehranian (1992)). The EVARCH model, however, is not restricted to a constant event period for every firm, so we use the event period determined by maximum likelihood for each firm. A plausible model for cross sectional analysis is:

\[
CAAR_i = a + b_1 FDICIA_i + b_2 LNASSETS_i + b_3 ROA_i + b_4 DEPS_i + b_5 LOANRAT_i + \eta_i ,
\]

(13)

where \(FDICIA_i\) is an indicator variable which equals zero if the takeover is before the passage of the FDIC Improvement Act and one if after, \(LNASSETS_i\) is the log of acquirer total asset size in thousands

comparison of individual AR's to Table 1 shows similar size to the traditional model findings.
of dollars, $ROA_i$ is acquirer return on assets, $DEPS_i$ is acquirer deposits to assets, and $LOANRAT_i$ is acquirer loan to assets. Data are from the Call and Income Reports for the year prior to the acquisition. Other specifications could be proposed. For example, Houston and Ryngaert (1997) demonstrate the importance of the financing medium. However, this model is illustrative of typical cross sectional analysis of banks, and these data are conveniently available -- we do not wish to reduce the sample size for our cross-sectional regressions more than necessary.\(^6\) Similarly, one might argue that researchers should use weighted least squares (WLS) or some other approach to estimate Equation (13). Karafiath (1994), though, shows that ordinary least squares (OLS) is both simpler and as good as other methods, concluding that OLS and WLS are as good as other methods. Although WLS may have some advantages in small samples, “...there is no advantage to weighted least squares at 75 securities” (p.296).\(^7\)

Table 4 presents results of the cross sectional analysis for the three competing methods. The sample size is reduced to 99 firms due to missing cross sectional data. Results are similar across all methods. This is not surprising: none of the methods finds significant abnormal returns, so in some sense we are asking Equation (13) to explain variation in noise. This is not quite accurate; for both the residuals from the normal market model and from the Scholes-Williams model, the pre-merger bidder ROA is significant and negatively related to CAARs. EVARCH reveals a similar relationship, but the EVARCH coefficient is a third larger. The FDICIA indicator, size, deposits-to-assets and loans-to-asset variables are insignificant for all models. The findings suggest, therefore, that banks acquiring other

\(^6\) We did experiment by adding the capital ratio. Results are nearly identical; if anything, they support the conclusions in the text more strongly.

\(^7\) Consistent with this, the results are very similar if we use WLS. They are therefore not reported. In fact, though, t-statistics are actually higher for the day(-5,+5) interval.
banks in the 1989 to 1995 period had significantly higher announcement CAARs if the bank’s pre-merger ROA was relatively low, although the average abnormal return in the sample is insignificantly different from zero.

Table 4: Cross sectional regressions with a dependent variable of the cumulative abnormal average residuals (CAARS) using the market model, the Scholes-Williams model, and the EVARCH model. For the normal market model and Scholes Williams model, the CAARs are the day -1 and day zero residuals. For the EVARCH model, the CAARs are the maximum likelihood event period as determined by the model. For independent variables, FDICIA is an indicator variable which equals zero before passage of the FDIC Improvement Act and one after; LNASSETS is the log of acquirer asset size in thousands of dollars; ROA is the acquirer’s return on assets; DEPS is the deposit to asset ratio for the acquirer; and LOANRAT is the acquirer’s loan to asset ratio. T-statistics are beneath the coefficient in parentheses.

<table>
<thead>
<tr>
<th>Interval for CAAR:</th>
<th>Normal Market Model</th>
<th>Scholes-Williams Model</th>
<th>EVARCH Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Days (-1,0)</td>
<td>Days (-1,0)</td>
<td>Days $(T_1,T_2)$ (from maximum likelihood)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.03980 (0.678)</td>
<td>0.03604 (0.621)</td>
<td>0.04920 (0.867)</td>
</tr>
<tr>
<td>FDICIA</td>
<td>-0.00009 (-0.015)</td>
<td>-0.00102 (-0.177)</td>
<td>-0.00102 (-0.182)</td>
</tr>
<tr>
<td>LNASSETS</td>
<td>-0.00161 (-0.576)</td>
<td>-0.00142 (-0.513)</td>
<td>-0.00215 (-0.796)</td>
</tr>
<tr>
<td>ROA</td>
<td>-3.65547*** (-3.455)</td>
<td>-3.48032*** (-3.331)</td>
<td>-4.82452*** (-4.723)</td>
</tr>
<tr>
<td>DEPS</td>
<td>0.04058 (0.459)</td>
<td>0.04460 (0.511)</td>
<td>0.04537 (0.532)</td>
</tr>
<tr>
<td>LOANRAT</td>
<td>0.02546 (0.798)</td>
<td>0.02348 (0.745)</td>
<td>0.04471 (1.451)</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.1131</td>
<td>0.1085</td>
<td>0.2139</td>
</tr>
<tr>
<td>(F-statistic)</td>
<td>(3.526)***</td>
<td>(3.410)***</td>
<td>(6.389)***</td>
</tr>
</tbody>
</table>

*The sample size is 99 for all models.

*** = significant at the one-percent level.

IV. Implications

What conclusions can we draw from our study? First, both of the standard approaches find low levels of significance for day(-1) abnormal returns in our sample of 132 acquiring banks, while EVARCH, which generalizes the approach to allow for event-induced parameter shifts, shifting (and in our case, larger) variances, and firm-specific event periods, finds none. The importance of this discrepancy is debatable, and fits well with the simpler methods’ reputation for robustness in detecting
abnormal returns (e.g., Brown and Warner (1980, 1985) and Henderson (1990)). However, the
evidence from our sample highlights the fact that for marginally significant results, the traditional models
incorrectly reject the null hypothesis of no significant abnormal returns. In our sample, there is relatively
little difference in the cross sectional results, as well, though an $R^2$ of double the traditional methods is
noteworthy. When cross-sectional results do differ, the EVARCH approach is able to detect more
subtle relationships.

EVARCH does indicate a much larger and more significantly negative relation between
abnormal returns and acquirer ROA than do traditional methods, suggesting that bidding banks with high
pre-acquisition ROA experience lower event-period returns. This negative coefficient for ROA might
be due to bidder overpayment or recognition that the acquisition could negatively affect future
performance. Finally, the EVARCH cross-sectional model indicates a much higher degree of
confidence in the results as shown by a higher F-statistic, significant at the 1% level.

The most important results of our research are, of course, undetectable by traditional methods:
we find not only changes in systematic risk, but also event periods that end before the announcement
period. Ignoring these two possibilities can cause incorrect inference, and also has policy implications
for corporate events. As an example, suppose a researcher is interested in day(0) ARs for an average
bank in our sample, and uses the pre-event traditional measures to calculate the abnormal returns. If the
true systematic risk for this firm is 1.34 (the maximum reached during the event interval by a
representative firm in our sample) and the beta from traditional models were used instead, then the AR
would be greatly overstated if $R_m$ is positive. For example, if the return for the bank is .01, the intercept
is zero, the market return is .01, and a beta of .87 is used, then the residual from the misspecified model

24
is: \[.01\times (.01)(.87) = .0013.\] By comparison, the true residual would be: \[.01\times (.01)(1.34) = .0034.\] Thus, the difference in this plausible example is 47 basis points for one day.

This example illustrates that erroneous results can occur, and highlights the problems of cross-sectional analysis of CAARs, even when traditional models work well on average. Of course, bank regulators and stockholders using traditional methods to study wealth effects could incorrectly measure wealth changes because they miss changes in systematic risk or use observations outside the event period.

V. Summary

This paper develops a new approach to event studies that relaxes some of the implicit restrictions imposed by traditional methods. Our approach allows for different model parameters before, during, and after the event period, explicitly accounts for changes in unsystematic risk and case-specific event periods, and estimates all parameters jointly, thus avoiding problems using generated regressors. In addition, our approach permits systematic risk to change gradually during the event period and to exit the period at higher or lower levels. We apply our new approach to explore acquiring banks’ stock returns during takeover activity, using a sample of 132 banks that acquired other institutions from 1989 through 1995. We find that a majority of institutions exhibit either significant ARCH effects or experience changes in systematic risk during the event period. Moreover, many changes in systematic risk are permanent. The systematic risk of a representative firm is 0.83 prior to the event, increases at a decreasing rate to a maximum of 1.34, then declines and exits the event window at a new, permanent level of 1.004.
How important is this? We cannot answer without knowing the context of the specific research or policy problem. The importance of allowing for risk-shifting events depends both on the goals of the research and the nature of the data. Traditional approaches have a good reputation for their ability to detect abnormal returns, so researchers interested solely in wealth effects might conceivably ignore risk-shifting events. Even in studies of this type, though, we advise caution, for measurements of abnormal returns without consideration of risk are always suspect. In other types of studies, researchers and regulators have much to gain by explicitly modeling changes in systematic risk. Studies of risk-based capital standards, deposit insurance, and long-term stock returns are obvious examples.

Although it proves relatively harmless in our study, we also argue that ignoring possible risk-shifting events is particularly important when conducting cross sectional analysis on abnormal returns. Even if a researcher were satisfied with traditional methods’ reputation for robustness, it seems unreasonably optimistic to expect accurate results when estimating a cross sectional regression, for there is no reason to expect the residuals from traditional approaches to be identical to those obtained using our generalized approach.

Indeed, our results show that although the traditional methods’ reputation is not destroyed, it is tarnished. Results using the market model and Scholes-Williams model are negative and significant at the ten-percent level for the day before an acquisition announcement. In contrast, EVARCH CAARs are not significant for any event period. Cross-sectional results also differ. EVARCH shows a stronger negative relation between pre-acquisition bidder ROA and ARs than do traditional methods, and adjusted $R^2$ is twice as large using EVARCH.
But most important, none of the event periods typically used with traditional methods even approximates the average interval determined by our generalized approach. Thus, differences in detecting abnormal returns between the traditional and generalized methods are masked by the absence of a significant economic event. To paraphrase an old tale, if there is no black cat in a dark room, it matters not which part of the room we search; regardless, we find no cat. We conclude that although event studies built around traditional market models detect abnormal returns well, on average, inference concerning CAARs for individual days is suspect. Failure to correct for the increased systematic risk during the event period and assuming that event periods are constant across firms can give erroneous results.

There is little doubt that continued consolidation within the banking industry will be the norm for the foreseeable future. In addition, we continue to see acquisitions of larger and larger institutions. The stakes are now much higher than in even the recent past. Our new approach suggests that bidder ROA is negatively related to abnormal returns and could indicate the expectation of lower future financial performance for the acquirer. In addition, we provide evidence that systematic risk increases by an average of 20%, from 0.83 to 1.0. The increase in systematic risk is likely due to banks acquiring riskier target banks in our sample, although data are unavailable to test this hypothesis explicitly. Regulators may prefer acquisition to liquidating banks, but if the risk increase makes larger institutions more susceptible to failure, then risk to the deposit insurance fund could be increased. This is especially true if large banks are deemed too-big-to-fail. Regulators, taxpayers, managers and investors would do well to examine each case carefully; reliance on industry averages is unwarranted.
Bibliography


Figure 1

Panel A

Panel B

Systematic Risk
Figure 2

The graph shows the systematic risk over event time, with the announcement day marked. The x-axis represents event time, ranging from -20 to 8, and the y-axis represents systematic risk, ranging from 0 to 1.5. The graph indicates an increase in systematic risk around the announcement day.