

(GMM) – an important and increasingly widely used estimation process – which is particularly relevant for asset pricing models, is presented and discussed in detail.

14.1 Event Studies

Event studies are very useful in finance research and as a result they are extremely commonly employed in the literature. In essence, they represent an attempt to gauge the effect of an identifiable *event* on a financial variable, usually stock returns. So, for example, research has investigated the impact of various types of announcements (e.g., dividends, stock splits, entry into or deletion from a stock index) on the returns of the stocks concerned. Event studies are often considered to be tests for market efficiency: if the financial markets are informationally efficient, there should be an immediate reaction to the event on the announcement date and no further reaction on subsequent trading days.

MacKinlay (1997) argues that conducting event studies initially appears difficult but is in fact easy; my view is that exactly the reverse is true: in principle, event studies are simple to understand and to conduct, but to do so in a rigorous manner requires a great deal of thought. There is a bewildering array of approaches that can be deployed, and at first blush it is not at all clear which of them is appropriate or optimal. The main groundwork for conducting modern event studies was established by Ball and Brown (1968) and by Fama *et al.* (1969), but as MacKinlay notes, something like them was conducted more than three decades earlier.

While there are now many useful survey papers that describe the various aspects of event studies in considerable detail, unfortunately each has its own notation and approach which can be confusing. Corrado (2011) is a recent example, although Armitage (1995) and MacKinlay (1997) are particularly clearly explained and more closely resemble the treatment given here. A similar discussion is offered by Campbell *et al.* (1997) but using matrix notation.

14.1.1 Some Notation and a Description of the Basic Approach

We of course need to be able to define precisely the dates on which the events occur, and the sample data are usually aligned with respect to this date. If we have N events in the sample, we usually specify an ‘event window’, which is the period of time over which we investigate the impact of the event. The length of this window will be set depending on whether

we wish to investigate the short- or long-run effects of the event. It is common to examine a period comprising, say, ten trading days before the event up to ten trading days after as a short-run event window, while long-run windows can cover a month, a year, or even several years after the event.

A first question to ask once the event has been identified is what frequency of data should be employed in the analysis. MacKinlay (1997) shows that the power of event studies to detect abnormal performance is much greater when daily data are employed rather than weekly or monthly observations, so that the same power can be achieved with a much smaller N , or for given N , the power will be much larger. Although it would in some cases be possible to use intra-daily data, these are harder to collect and may bring additional problems including nuisance microstructure effects; this is perhaps why daily observations are the frequency of choice for most studies in the literature.¹

Define the return for each firm i on each day t during the event window as R_{it} . We can conduct the following approach separately for each day within the event window – for example, we might investigate it for all of ten days before the event up to ten days after (where $t = 0$ represents the date of the event and $t = -10, -9, -8, \dots, -1, 0, 1, 2, \dots, 8, 9, 10$). Note that we will need to exercise care in the definition of the reference day $t = 0$ if the announcement is made after the close of the market.

In most cases, we need to be able to separate the impact of the event from other, unrelated movements in prices. For example, if it is announced that a firm will become a member of a widely followed stock index and its share price that day rises by 4%, but on average the prices of all other stocks also rise by 4%, it would be unwise to conclude that all of the increase in the price of the stock under study is attributable to the announcement. This motivates the idea of constructing abnormal returns, denoted AR_{it} , which are calculated by subtracting an expected return from the actual return

$$AR_{it} = R_{it} - E(R_{it}) \tag{14.1}$$

There are numerous ways that the expected returns can be calculated, but usually this is achieved using a sample of data before the event window so that the nature of the event is not allowed to ‘contaminate’ estimation of the expected returns. Armitage (1995) suggests that estimation periods can comprise anything from 100 to 300 days for daily

observations and 24 to 60 months when the analysis is conducted on a monthly basis. Longer estimation windows will in general increase the precision of parameter estimation, although with it the likelihood of a structural break and so there is a trade-off.

If the event window is very short (e.g., a day or a few days), then we need be far less concerned about constructing an expected return since it is likely to be very close to zero over such a short horizon. In such circumstances, it will probably be acceptable to simply use the actual returns in place of abnormal returns.

It is further often the case that a gap is left between the estimation period and the event window, to be completely sure that anticipation (i.e., ‘leakage’) of the event does not affect estimation of the expected return equation. However, it might well be the case that in practice we do not have the luxury of doing this since the sample period available is insufficient. Clearly, what we would like to do is to calculate the return that would have been expected for that stock if the event did not happen at all so that we can isolate the impact of the event from any unrelated incidents that may be occurring at the same time.

The simplest method for constructing expected returns (apart from setting them to zero) is to assume a constant mean return, so that the expected return is simply the average return for each stock i calculated at the same frequency over the estimation window, which we might term \bar{R}_i . Brown and Warner (1980, 1985) conduct a simulation experiment to compare methods of estimating expected returns for event studies. They find that an approach simply using historical return averages outperforms many more complicated approaches because of the estimation error that comes with the latter.

A second, slightly more sophisticated approach, is to subtract the return on a proxy for the market portfolio that day t from the individual return. This will certainly overcome the impact of general market movements in a rudimentary way, and is equivalent to the assumption that the stock’s beta in the market model or the CAPM is unity.

Probably the most common approach to constructing expected returns, however, is to use the market model. This in essence works by constructing the expected return using a regression of the return to stock i on a constant and the return to the market portfolio

$$R_{it} = \alpha_i + \beta_i R_{mt} + u_{it} \quad (14.2)$$

The expected return for firm i on any day t during the event window would then be calculated as the beta estimate from this regression multiplied by the actual market return on day t .

An interesting question is whether the expected return should incorporate the α from the estimation period in addition to β multiplied by the market return. Most applications of event studies include this, and indeed the original study by Fama *et al.* (1969) includes an alpha. However, we need to exercise caution when doing so since if – either because of some unrelated incident affecting the price of the stock or in anticipation of the event – the alpha is particularly high (particularly low) during the estimation period, it will push up (down) the expected return. Thus it may be preferable to assume an expected value of zero for the alpha and to exclude it from the event period abnormal return calculation.

In most applications, a broad stock index such as the FTSE All-Share or the S&P500 would be employed to proxy for the market portfolio. This equation can be made as complicated as desired – for example, by allowing for firm size or other characteristics – these would be included as additional factors in the regression with the expected return during the event window being calculated in a similar fashion. An approach based on the arbitrage pricing models of Chen, Roll and Ross (1986) or of Fama and French (1993) could also be employed – more discussion is made of this issue in the following section.

A final further approach would be to set up a ‘control portfolio’ of firms that have characteristics as close as possible to those of the event firm – for example, matching on firm size, beta, industry, book-to-market ratio, etc. – and then using the returns on this portfolio as the expected returns. Armitage (1995) reports the results of several Monte Carlo simulations that compare the results of various model frameworks that can be used for event studies.

The hypothesis testing framework is usually set up so that the null to be examined is of the event having no effect on the stock price (i.e. an abnormal return of zero). Under the null of no abnormal performance for firm i on day t during the event window, we can construct test statistics based on the standardised abnormal performance. These test statistics will be asymptotically normally distributed (as the length of the estimation window, T , increases)

$$AR_{it} \sim N(0, \sigma^2(AR_{it}))$$

where $\sigma^2(AR_{it})$ is the variance of the abnormal returns, which can be

estimated in various ways. A simple method, used by Brown and Warner (1980) among others, is to use the time series of data from the estimation of the expected returns separately for each stock. So we could define $\hat{\sigma}^2(AR_{it})$ as being the equal to the variance of the residuals from the market model, which could be calculated for example using

$$\hat{\sigma}^2(AR_{it}) = \frac{1}{T-2} \sum_{t=2}^T \hat{u}_{it}^2 \quad (14.3)$$

where T is the number of observations in the estimation period. If instead the expected returns had been estimated using historical average returns, we would simply use the variance of those.

Sometimes, an adjustment is made to $\hat{\sigma}^2(AR_{it})$ that reflects the errors arising from estimation of α and β in the market model. Including the adjustment, the variance in the previous equation becomes

$$\hat{\sigma}^2(AR_{it}) = \frac{1}{T-2} \sum_{t=2}^T \left(\hat{u}_{it}^2 + \frac{1}{T} \left[1 + \frac{R_{mt} - \bar{R}_m}{\hat{\sigma}_m^2} \right]^2 \right) \quad (14.4)$$

where \bar{R}_m and $\hat{\sigma}_m^2$ are the average and variance of the returns on the market portfolio, respectively, during the estimation window. It should be clear that as the length of the estimation period, T , increases, this adjustment will gradually shrink to zero.

We can then construct a test statistic by taking the abnormal return and dividing it by its corresponding standard error, which will asymptotically follow a standard normal distribution²

$$S\hat{A}R_{it} = \frac{\hat{A}R_{it}}{[\hat{\sigma}^2(AR_{it})]^{1/2}} \sim N(0, 1) \quad (14.5)$$

where $S\hat{A}R_{it}$ stands for the standardised abnormal return, which is the test statistic for each firm i and for each event day t .

It is likely that there will be quite a bit of variation of the returns across the days within the event window, with the price rising on some days and then falling on others. As such, it would be hard to identify the overall patterns. We may therefore consider computing the time series cumulative abnormal return over a multi-period event window (for example, over ten trading days) by summing the average returns over several periods, say from time T_1 to T_2

$$\hat{C}AR_i(T_1, T_2) = \sum_{t=T_1}^{T_2} \hat{A}R_{it} \quad (14.6)$$

Note that the time from T_1 to T_2 may constitute the entire event window or it might just be a sub-set of it. The variance of this $\hat{C}AR$ will be given by the number of observations in the event window plus one multiplied by the daily abnormal return variance calculated in [equation \(14.4\)](#) above

$$\hat{\sigma}^2(CAR_i(T_1, T_2)) = (T_2 - T_1 + 1)\hat{\sigma}^2(\hat{A}R_{it}) \quad (14.7)$$

This expression is essentially the sum of the individual daily variances over the days in T_1 to T_2 inclusive.³

We can now construct a test statistic for the cumulative abnormal return in the same way as we did for the individual dates, which will again be standard normally distributed

$$S\hat{C}AR_i(T_1, T_2) = \frac{\hat{C}AR_i(T_1, T_2)}{[\hat{\sigma}^2(CAR_i(T_1, T_2))]^{1/2}} \sim N(0, 1) \quad (14.8)$$

It is common to examine a pre-event window (to consider whether there is any anticipation of the event) and a post-event window – in other words, we sum the daily returns for a given firm i for days $t - 10$ to $t - 1$, say, and separately from $t + 1$ to $t + 10$, with the actual day of the event, t , being considered on its own.

Typically, some of the firms will show a negative abnormal return around the event when a positive figure was expected, and this is probably not very useful. But if we have N firms or N events, it is usually of more interest whether the return averaged across all firms is statistically different from zero than whether this is the case for any specific individual firm. We could define this average across firms for each separate day t during the event window as

$$\hat{A}R_t = \frac{1}{N} \sum_{i=1}^N \hat{A}R_{it} \quad (14.9)$$

This firm-average abnormal return, $\hat{A}R_t$ will have variance given by $1/N$ multiplied by the average of the variances of the individual firm returns

$$\hat{\sigma}^2(\hat{A}R_t) = \frac{1}{N^2} \sum_{i=1}^N \hat{\sigma}^2(\hat{A}R_{it}) \quad (14.10)$$

Thus the test statistic (the standardised return) for testing the null hypothesis that the average (across the N firms) return on day t is zero will be given by

$$S\hat{A}R_t = \frac{\hat{A}R_t}{[\hat{\sigma}^2(\hat{A}R_t)]^{1/2}} = \frac{\frac{1}{N} \sum_{i=1}^N \hat{A}R_{it}}{[\frac{1}{N^2} \sum_{i=1}^N \hat{\sigma}^2(\hat{A}R_{it})]^{1/2}} \sim N(0, 1) \quad (14.11)$$

Finally, we can aggregate both across firms and over time to form a single test statistic for examining the null hypothesis that the average multi-horizon (i.e., cumulative) return across all firms is zero. We would get an equivalent statistic whether we first aggregated over time and then across firms or the other way around. The CAR calculated by averaging across firms first and then cumulating over time could be written

$$\hat{C}AR(T_1, T_2) = \sum_{t=T_1}^{T_2} \hat{A}R_t \quad (14.12)$$

Or equivalently, if we started with the $CAR_i(T_1, T_2)$ separately for each firm, we would take the average of these over the N firms

$$\hat{C}AR(T_1, T_2) = \frac{1}{N} \sum_{i=1}^N \hat{C}AR_i(T_1, T_2) \quad (14.13)$$

To obtain the variance of this $\hat{C}AR(T_1, T_2)$ we could take $1/N$ multiplied by the average of the variances of the individual $\hat{C}AR_i$.

$$\hat{\sigma}^2(\hat{C}AR(T_1, T_2)) = \frac{1}{N^2} \sum_{i=1}^N \hat{\sigma}^2(\hat{C}AR_i(T_1, T_2)) \quad (14.14)$$

And again we can construct a standard normally distributed test statistic as

$$S\hat{C}AR(T_1, T_2) = \frac{\hat{C}AR(T_1, T_2)}{[\hat{\sigma}^2(\hat{C}AR(T_1, T_2))]^{1/2}} \sim N(0, 1) \quad (14.15)$$

14.1.2 Cross-Sectional Regressions

The methodologies and formulae presented above provide various tools for examining whether abnormal returns are statistically significant or not. However, it will often be the case that we are interested in allowing for

differences in the characteristics of a sub-section of the events and also examining the link between the characteristics and the magnitude of the abnormal returns. For example, does the event have a bigger impact on small firms? Or on firms which are heavily traded, etc.? The simplest way to achieve this would be to calculate the abnormal returns as desired using something like [equations \(14.1\)](#) and [\(14.2\)](#) above and then to use these as the dependent variable in a cross-sectional regression of the form

$$AR_i = \gamma_0 + \gamma_1 x_{1i} + \gamma_2 x_{2i} + \dots + \gamma_M x_{Mi} + w_i \quad (14.16)$$

where AR_i is the abnormal return for firm i measured over some horizon, and x_{ji} , ($j = 1, \dots, M$) are a set of M characteristics that are thought to influence the abnormal returns, γ_j measures the impact of the corresponding variable j on the abnormal return, and w_i is an error term. We can examine the sign, size and statistical significance of γ_0 as a test for whether the average abnormal return is significantly different from zero after allowing for the impacts of the M characteristics. MacKinlay (1997) advocates the use of heteroscedasticity-robust standard errors in the regression.

The abnormal return used in this equation would typically be measured over several days (or perhaps even the whole event window), but it could also be based on a single day.

14.1.3 Complications When Conducting Event Studies and Their Resolution

The above discussion presents the standard methodology that is commonly employed when conducting event studies, and most of the time it will provide appropriate inferences. However, as always in econometrics, the use of test statistics requires a number of assumptions about the nature of the data and models employed. Some of these assumptions will now be highlighted and their implications explored.

Cross-Sectional Dependence

A key assumption when the returns are aggregated across firms is that the events are independent of one another. Often, this will not be the case, particularly when the events are clustered through time. For example, if we were investigating the impact of index recompositions on the prices of the

stocks concerned, these index constituents generally only change at specific times of the year. So, typically, a bunch of stocks will enter into an index on the same day, and then there may be no further such events for three or six months.

The impact of this clustering is that we cannot assume the returns to be independent across firms, and as a result the variances in the aggregates across firms (equations (14.10) and (14.14)) will not apply since these derivations have effectively assumed the returns to be independent across firms so that all of the covariances between returns across firms can be set to zero. An obvious solution to this would be not to aggregate the returns across firms, but simply to construct the test statistics on an event-by-event basis and then to undertake a summary analysis of them (e.g., reporting their means, variances, percentage of significant events, etc.).

A second solution would be to construct portfolios of firms having the event at the same time and then the analysis would be done on each of the portfolios. The standard deviation would be calculated using the cross-section of those portfolios' returns on day t (or on days T_1 to T_2 , as desired). This approach will allow for cross-correlations since they will automatically be taken into account in constructing the portfolio returns and the standard deviations of those returns. But a disadvantage of this technique is that it cannot allow for different variances for each firm as all are equally weighted within the portfolio; the standard method described above would do so, however.

Changing Variances of Returns

It has been argued in the literature that often the variance of returns will increase over the event window, but the variance figure used in the testing procedure will have been calculated based on the estimation window, which is usually some time before the event. Either the event itself or the factors that led to it are likely to increase uncertainty and with it the volatility of returns. As a result, the measured variance will be too low and the null hypothesis of no abnormal return during the event will be rejected too often. To deal with this, Boehmer, Musumeci, and Poulsen (1991), among others, suggest estimating the variance of abnormal returns by employing the cross-sectional variance of returns across firms during the *event* window. Clearly, if we adopt this procedure we cannot estimate separate test statistics for each firm (although arguably these are usually of little interest anyway). The variance estimator in equation (14.10) would be replaced by

$$\hat{\sigma}^2(AR_t) = \frac{1}{N^2} \sum_{i=1}^N (\hat{AR}_{it} - \hat{AR}_t)^2 \quad (14.17)$$

with the test statistic following as before. A similar adjustment could be made for the variance of the cumulative abnormal return

$$\hat{\sigma}^2(CAR(T_1, T_2)) = \frac{1}{N^2} \sum_{i=1}^N (C\hat{AR}_i(T_1, T_2) - C\hat{AR}(T_1, T_2))^2 \quad (14.18)$$

While this test statistic will allow for the variance to change over time, a drawback is that it does not allow for differences in return variances across firms and nor does it allow for cross-correlations in returns caused by event clustering.

Weighting the Stocks

Another issue is that the approach as stated above will not give equal weight to each stock's return in the calculation. The steps outlined above construct the cross-firm aggregate return (in [equation \(14.9\)](#)) and then standardise this using the aggregate standard deviation (as in [equation \(14.11\)](#)). An alternative method would be to first standardise each firm's abnormal return (dividing by its appropriate standard deviation) and then aggregating these standardised abnormal returns. If we take the standardised abnormal return for each firm, $S\hat{AR}_{i,t}$, from [equation \(14.5\)](#), we can calculate the average of these across the N firms

$$S\hat{AR}_t = \frac{1}{N} \sum_{i=1}^N S\hat{AR}_{it} \quad (14.19)$$

These $SARs$ have already been standardised so there is no need to divide them by the square root of the variance. If we take this SAR_t and multiply it by \sqrt{N} , we will get a test statistic that is asymptotically normally distributed and which, by construction, will give equal weight to each SAR (because we have taken an unweighted average of them)

$$\sqrt{N}SAR_t \sim N(0, 1)$$

We could similarly take an unweighted average of the standardised cumulative abnormal returns ($SCAR$)

$$SC\hat{A}R(T_1, T_2) = \frac{1}{N} \sum_{i=1}^N SC\hat{A}R_i(T_1, T_2) \quad (14.20)$$

and

$$\sqrt{N}SCAR(T_1, T_2) \sim N(0, 1)$$

If the true abnormal return is similar across securities, we would be better to equally weight the abnormal returns in calculating the test statistics (as in equations (14.19) and (14.20)), but if the abnormal return varies positively with its variance measure, then it would be better to give more weight to stocks with lower return variances (as in equation (14.15) for example).

Long Event Windows

Event studies are joint tests of whether the event-induced abnormal return is zero and whether the model employed to construct expected returns is correct. If we wish to examine the impact of an event over a long period (say, more than a few months), we need to be more careful about the design of the model for expected returns and also to ensure that this model appropriately allows for risk. Over short windows, discrepancies between models are usually small and any errors in model specification are almost negligible. But over the longer run, small errors in setting up the asset pricing model can lead to large errors in the calculation of abnormal returns and therefore the impact of the event.

A key question in conducting event studies to measure long-run impacts is whether to use cumulative abnormal returns (CARs), as described above, or buy-and-hold abnormal returns (BHARs). There are several important differences between the two. First, BHARs employ geometric returns rather than arithmetic returns (used in computing CARs) in calculating the overall return over the event period of interest. Thus the BHAR can allow for compounding whereas the CAR does not. The formula for calculating the BHAR is usually given by

$$BH\hat{A}R_i = [\prod_{t=T_1}^{T_2} (1 + R_{it}) - 1] - [\prod_{t=T_1}^{T_2} (1 + E(R_{it})) - 1] \quad (14.21)$$

where $E(R_{it})$ is the expected return. Usually, when constructing BHARs the expected return is based on a non-event firm or portfolio of firms that

is matched in some way to the event firm (e.g., based on size, industry, etc.). Alternatively, although less desirably, it could be obtained from a benchmark such as a stock market index.

If desired, we can then sum the $BHAR_i$ across the N firms to construct an aggregate measure. BHARs have been advocated, among others, by Barber and Lyon (1997) and Lyon, Barber and Tsai (1999) because they better match the 'investor experience' than CARs given the former's use of geometric rather than arithmetic averaging. CARs represent biased estimates of the actual returns received by investors. However, by contrast, Fama (1998) in particular argues in favour of the use of CARs rather than BHARs. The latter seem to be more adversely affected by skewness in the sample of abnormal returns than the former because of the impact of compounding in BHARs.⁴ In addition, Fama indicates that the average CAR increases at a rate of $(T_2 - T_1)$ with the number of months included in the sum, whereas its standard error increases only at a rate $\sqrt{(T_2 - T_1)}$. This is not true for BHARs where the standard errors grow at the faster rate $(T_2 - T_1)$ rather than its square root. Hence any inaccuracies in measuring expected returns will be more serious for BHARs as another consequence of compounding.

Event Time versus Calendar Time Analysis

All of the procedures discussed above have involved conducting analysis in *event time*. There is, however, an alternative approach that involves using *calendar time*, advocated by Fama (1998) and Mitchell and Stafford (2000) among others. In essence, using a calendar time methodology involves running a time series regression and examining the intercept from that regression. The dependent variable is a series of portfolio returns, which measure the average returns at each point in time of the set of firms that have undergone the event of interest within a pre-defined measurement period before that time. So, for example, we might choose to examine the returns of firms for a year after the event that they announce cessation of their dividend payments. Then, for each observation t , the dependent variable will be the average return on all firms that stopped paying dividends at any point during the past year. One year after the event, by construction the firm will drop out of the portfolio. Hence the number of firms within the portfolio will vary over time (as the number of firms ceasing dividend payment varies) and the portfolio will effectively be rebalanced each month. The explanatory variables may be risk

measures from the Carhart (1997) four-factor model for example – this will be discussed in detail below.

The calendar time approach will weight each time period equally and thus the weight on each individual firm in the sample will vary inversely with the number of other firms that have undergone the event during the observation period. This may be problematic and will result in a loss of power to detect an effect if managers time their events to take advantage of misvaluations.

Small Samples and Non-Normality

The test statistics presented in the previous section are all asymptotic, and problems may arise either if the estimation window (T) is too short, or if the number of firms (N) is too small when the firm-aggregated statistic is used. As we discussed earlier in the book, it is well known that stock returns are leptokurtic and tend to have longer lower tails than upper tails. And particularly with small samples, the presence of outliers – for example, very large returns during the estimation window affecting the market model parameter estimation or the residual variance estimates – may also be problematic. One possible remedy would be to use a bootstrap approach to computing the test statistics.

A second strategy for dealing with non-normality would be to use a non-parametric test. Such tests are robust in the presence of non-normal distributions, although they are usually less powerful than their parametric counterparts. In the present context, we could test the null hypothesis that the proportion of positive abnormal returns is not affected by the event. In other words, the proportion of positive abnormal returns across firms remains at the expected level. We could then use the test statistic, Z_p

$$Z_p = \frac{[p - p^*]}{[p^*(1 - p^*)/N]^{1/2}} \quad (14.22)$$

where p is the actual proportion of negative abnormal returns during the event window and p^* is the expected proportion of negative abnormal returns. Under the null hypothesis, the test statistic follows a binomial distribution, which can be approximated by the standard normal distribution. Sometimes p^* is set to 0.5, but this may not be appropriate if the return distribution is skewed, which is typically the case. Instead, it is better to calculate p^* based on the proportion of negative abnormal returns during the estimation window. The Wilcoxon signed-rank test can also be

used.

Event Studies: Some Further Issues

A further implicit assumption in the standard event test methodology is that the events themselves occur involuntarily. In practice, however, firms often have discretion about the extent, timing and presentational forms of the announcements that they make. Thus they are likely to use any discretion they have to make announcements when market reactions are going to be the most favourable. For example, where the local regulatory rules allow discretion, firms may release bad news when the markets are closed or when the media and investors are preoccupied with other significant news items. Prabhala (1997) discusses the implications of and solutions to the endogeneity of the firm's decision about when (and perhaps even whether) to make an announcement. When a firm chooses not to announce at a particular time, we have a sort of truncated sample since we can observe events only for firms who choose to make an announcement.

A way of simultaneously dealing with a number of the issues highlighted above (i.e., differing return variances across firms, changing return variances over time, and clustering of events across firms) is to use what has been termed generalised least squares (GLS) in constructing the test statistics. In essence this works by constructing a variance–covariance matrix from the abnormal returns and using this to weight the returns in computing the aggregate test statistic – see Armitage (1995) for further details.

We can see from the above that a range of procedures exists for conducting event studies. The core of the approach is the same in each case, but they differ according to how the aggregation is performed over time and across firms and this affects the method of calculation of the standard deviations. So how do we choose which approach to use? Hopefully, given the context and the nature of the events under consideration, we can gain a reasonable idea of which approach is likely to be the most justifiable. For example, is clustering an issue? Is it expectable that the return variances will have changed over time? Is it important to allow for the variances of returns to vary between firms? By answering these questions, we can usually select the appropriate procedure. But if in doubt, it is always advisable to examine a range of methods and to compare the results as a robustness check. With luck, the various calculation techniques will lead to the same conclusion.

14.1.4 Conducting an Event Study Using Excel

This section will now use the core of the approaches described above in order to conduct an event study. While this ought to be enough to get started and to obtain some indicative results, it is important to note that there is far more that can be done with event studies to make them more rigorous than the approach presented here and readers are encouraged to consult the papers cited above for further details.

The first step would be to decide which event to consider the impact of, and there is certainly no shortage of possibilities (dividend announcements; stock split announcements; index composition changes; merger announcements; CEO turnover; new contract announcements; macroeconomic announcements, etc.). Once this is done and the data are collected, the time-consuming part is to then organise them in a way to make them easy to work with. It would be possible to conduct the analysis in any software package for data analysis, including EViews. However, since the bulk of the task involves data arrangement and the econometric part is usually not sophisticated (in most cases, a regression will not even be conducted), it probably makes sense to revert back to Microsoft Excel or a similar spreadsheet package.⁵

The starting point for the analysis conducted here are the abnormal returns for $N = 20$ firms, which are given in the Excel file 'Event.xls', and have already been calculated using the market model using [equations \(14.1\) and \(14.2\)](#). The returns are given for days -259 to $+263$. The raw data are on the sheet 'abnormal returns'. The spreadsheet has been set up with the data aligned on the event day, so while the firms underwent the event on different days, the spreadsheet is constructed so that day '0' is the event day in the same row for all firms. The estimation period is from day -259 to day -10 inclusive (249 days), while the event periods examined are $(T - 10, T - 1)$, day T itself, $(T + 1, T + 10)$ and $(T + 1, T + 250)$. The first of these windows allows us to examine whether there was any leakage of information that affected stock returns prior to the event. Whether there is an immediate effect on the day that the event occurs will depend on whether the announcement is made in advance or it is a 'surprise' to the markets. If the event was known in advance to be happening on day T then the impact on the markets that day may be muted since it could have already been reflected in prices. Note that in this case the adjustment in [equation \(14.4\)](#) is not employed since the estimation period ($T = 249$) is quite long and would render the adjustment term negligible.

We first calculate the average return across all twenty firms for each day

during the estimation and event windows in column V of the ‘abnormal returns’ sheet using the Excel AVERAGE formula in the usual way. All of the calculations of the key statistics are done on a separate sheet which I have called ‘summary stats’. The sheet first calculates the AR for day T and the CARs for the date ranges using equations (14.1) and (14.6), respectively, for each individual firm and also for the average across all firms.

The next step is to calculate the variances of the abnormal returns or cumulative abnormal returns. For day T , this is done using equation (14.3), which is simply the time series variance of returns during the estimation window and placed in row 2 (and copied directly into row 11). For the multi-day event windows, the one-day variance from equation (14.3) is scaled up by the number of days in the event window (10 or 250) using equation (14.7). Then the test statistics are calculated by dividing the AR by its respective standard deviation (i.e. the square root of the variance) using equation (14.5) or its CAR equivalent in equation (14.8). Finally, the easiest way to obtain p -values for the tests is to use the TDIST function in Excel for a two-sided test and with a large number of degrees of freedom (say, 1,000), so that it approximates the normal distribution.

As discussed in the previous section, there are several potential issues with the fairly simple event study methodology just described. So, for robustness, it is a good idea to examine different ways of tackling the problem, and two possible checks are given in columns X and Y of the ‘summary stats’ sheet. Both of these procedures can only be undertaken based on the average across firms and not at the individual firm level. The first tweak is to calculate the standard deviation used in the test statistics cross-sectionally in order to allow for the possibility that the return variances may have changed (typically, risen) around the time of the event. Thus we simply take the variance across firms for the abnormal return or cumulative abnormal return of interest, divide this by N (i.e. 20) and then proceed in the usual way.

A further possibility examined in column Y is to equally weight firms by calculating the average of the standardised abnormal returns as in equations (14.19) or (14.20). Then the test statistic is simply this average multiplied by the square root of the N .

If we now consider the results on this sheet, it is clear that there is little evidence of a short-term reaction to the event. During the two trading weeks before the event, ($T - 10$ to $T - 1$), only one firm has a significant abnormal return at the 5% level (firm 20 has a CAR of 15.43% with a test

statistic of 2.02). None of the individual firms have significant returns on the event date (T), and neither do any of them show significance in the short post-event window ($T + 1$ to $T + 10$). It is over the longer term – the next trading year – where there is some action. Now five firms have statistically significant returns together with economically quite large cumulative abnormal returns of 20% to 55%.

Examining the aggregate-level results, it is reassuring that the three slightly different approaches in columns W to Y yield very similar conclusions. Here the null hypothesis is that the average abnormal return (or average cumulative abnormal return) is zero. There is again no discernible market reaction before, on, or in the short-run after, the event. However, the long-run abnormal return is positive and highly statistically significant whichever of the three approaches is considered. Interestingly, the variance estimates before the event (at times $t - 10$ to $T - 1$) are higher for the cross-sectional approach in [equation \(14.18\)](#), although they are lower for cross-sectional approach during and after the event.

Finally, in the third sheet of the Event.xls workbook, labelled ‘non-parametric test’, the non-parametric statistic Z_p of [equation \(14.22\)](#) is calculated and then the p -value is obtained using the TDIST function as above. This examines the null hypothesis that the proportion of abnormal returns around the event is the same as it was during the estimation window. So the first calculation row of the sheet (row 2) calculates p^* , the expected proportion of negative returns based on data from the estimation window. Then for each event period range, we calculate p , the actual proportion of negative returns.⁶

The expected proportion of negative returns varies from 0.43 for firm 18 to 0.55 for firm 8, but the actual proportions for the short pre- and post-event windows are often much lower than that. For example, for firm 1, p was 0.3 (i.e., negative returns on only three days from ten) before the event. Pre-event, six of the twenty firms have significant differences between p and p^* , while for the two weeks immediately after the event, only three of them show significant differences. Over the long-run, however, there are no significant differences between the expected and actual proportions of negative return days – either for any of the individual firms or for the average.

14.2 Tests of the CAPM and the Fama–French Methodology