Chapter 14

Additional econometric techniques for financial research

Conducting an Event Study

What is an Event Study?

- Event studies are extremely common in finance and in research projects!
- 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
- The "modern" event study literature began with Brown (1968) and by Fama et al. (1969).

Event Studies: Background

- 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
- 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
- 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

Event Studies: Background (Cont'd)

- MacKinlay (1997) shows that the power of event studies to detect abnormal performance is much greater when daily data are employed rather than monthly, quarterly or annual data
- Intra-daily data are likely to be full of microstructure noise.

Event Studies: The Event Window

- 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 – e.g., we might investigate it for all of 10 days before the event up to 10 days after (where $t =$ 0 represents the date of the event and $t = -10, -9, -8, ..., -1, 0, 1, 2, ..., 8, 9, 10$)
- 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 stock index and its share price that day rises by 4%, but 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
- So we construct abnormal returns, denoted AR_{it}, which are calculated by subtracting an expected return from the actual

Event Studies: Abnormal Returns

- There are numerous ways that the expected returns can be calculated, but usually this is achieved using data before the event window so that 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

- If the event window is very short then we are far less concerned about constructing an expected return since it is likely to be very close to zero over such a short horizon

Event Studies: Abnormal Returns (Cont'd)

- In such circumstances, it will probably be acceptable to simply use the actual returns in place of abnormal return
- The simplest method for constructing expected returns is to assume a constant mean return, so the expected return is the average return for each stock i which we might term $\bar{R}_i$

Event Studies: The Market Model

- A 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 constructs the expected return using a regression of the return to stock i on a constant and the return to the market portfolio:

$$R_{it} = \alpha_i + \beta_i R_{mt} + u_{it}$$

Event Studies: The Market Model (Cont'd)

- 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.

Event Studies: The Market Model 2

- 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

- A final further approach would be to set up a '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

Event Studies: Hypothesis Testing

- 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 is to use the time-series of data from the estimation of the expected returns separately for each stock.

Event Studies: Hypothesis Testing 2

- We could define as being the variance of the residuals from the market model, which could be calculated for example using

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

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

Event Studies: Hypothesis Testing 2 (Cont'd)

- 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

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

Event Studies: Hypothesis Testing 3

- 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 distribution:

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

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

Event Studies: Cumulative Abnormal Returns

- It is likely that there will be quite a bit of variation of the returns across the days within the event window
- We may therefore consider computing the time-series cumulative average return (CAR) 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:

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

- The variance of this CAR will be given by the number of observations in the event window plus one multiplied by the daily abnormal return variance calculated previously:

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

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

Event Studies: A Test Statistic for the CAR

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

$$S\hat{CAR}_i(T_1, T_2) = \frac{C\hat{A}R_i(T_1, T_2)}{[\hat{\sigma}^2(CAR_i(T_1, T_2))]^{1/2}} \sim N(0, 1)$$

- 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.

Event Studies: Averaging Returns Across Firms

- Typically, some of the firms will show a negative abnormal return around the event when a positive figure was expected
- 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

$$\hat{AR}_t = \frac{1}{N}\sum_{i=1}^{N}\hat{AR}_{it}$$

- This firm-average abnormal return will have variance given by $1/N$ multiplied by the average of the variances of the individual firm returns:

$$\hat{\sigma}^2(AR_t) = \frac{1}{N^2}\sum_{i=1}^{N}\hat{\sigma}^2(AR_{it})$$

Event Studies: Averaging Returns Across Firms 2

- 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

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

Event Studies: Averaging Returns Across Firms and Time

- 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:

$$C\hat{A}R(T_1, T_2) = \sum_{t=T_1}^{T_2} \hat{AR}_t$$

Event Studies: Averaging Returns Across Firms and Time (Cont'd)

- 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:

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

Event Studies: Averaging Returns Across Firms and Time 2

- To obtain the variance of this $CAR_i(T_1, T_2)$ we could take $1/N$ multiplied by the average of the variances of the individual CAR_i:

$$\hat{\sigma}^2(CAR(T_1, T_2)) = \frac{1}{N^2}\sum_{i=1}^{N}\hat{\sigma}^2(CAR_i(T_1, T_2))$$

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

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

Event Studies: Cross-Sectional Regressions

- 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.?
- To do this, calculate the abnormal returns as desired and then to use these as the dependent variable in a cross-sectional regression of the form

$$AR_i = \gamma_0 + \gamma_1 x_{1i} + \gamma_2 x_{2i} + \ldots + \gamma_M x_{Mi} + w_i$$

  where AR_i is the abnormal return for firm i, x_{ji}, (j=1,..., M) are a set of characteristics 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

Event Studies: 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, 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 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 could be set to zero.

Event Studies: Cross-Sectional Dependence - Solutions

- An obvious solution 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)

Event Studies: Cross-Sectional Dependence - Solutions (Cont'd)

- 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.

Event Studies: Changing Variances of Returns

- Often the variance of returns will increase over the event window
- 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 et al. (1991), amongst others, suggest estimating the variance of abnormal returns by employing the cross-sectional variance of returns across firms during the event window

Event Studies: Changing Variances of Returns (Cont'd)

- Clearly, if we adopt this procedure we cannot estimate separate test statistics for each firm
- The variance estimator would be:

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

- The test statistic would be calculated as before.

Event Studies: 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 and then standardise this using the aggregate standard deviation
- An alternative method would be to first standardise each firm's abnormal return (dividing by its appropriate standard deviation) and then to aggregate these standardised abnormal returns

Event Studies: Weighting the Stocks (Cont'd)

- If we take the standardised abnormal return for each firm, we can calculate the average of these across the N firms:

$$\hat{SAR}_t = \frac{1}{N}\sum_{i=1}^{N}\hat{SAR}_{it}$$

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

Event Studies: 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 we need to be more careful about the design of the model for expected returns
- 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 is whether to use cumulative abnormal returns (CARs) or buy-and-hold abnormal returns (BHARs)
- There are important differences between the two:
  - BHARs employ geometric returns rather than arithmetic returns in calculating the overall return over the event period of interest
  - Thus the BHAR can allow for compounding whereas the CAR does not.

Event Studies: Buy-and-Hold Abnormal Returns

- A formula for calculating the BHAR is

$$\hat{BHAR}_i = [\Pi_{t=T_1}^{T_2}(1 + R_{it}) - 1] - [\Pi_{t=T_1}^{T_2}(1 + E(R_{it})) - 1]$$

- If desired, we can then sum the $BHAR_i$ across the N firms to construct an aggregate measure.
- BHARs have been advocated, amongst others, by Barber and Lyon (1997) and Lyon et al. (1999) because they better match the 'investor experience'
- 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.

Event Studies: Buy-and-Hold Abnormal Returns 2

- BHARs seem to be more adversely affected by skewness in the sample of abnormal returns than CARs because of the impact of compounding in BHARs
- 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 Studies: Event Versus Calendar Time

- All of the procedures discussed above have involved conducting analysis in event time
- An alternative approach involves using calendar time, which 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.

Event Studies: Event Versus Calendar Time 2

- 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 and the portfolio will effectively be rebalanced each month

Event Studies: Event Versus Calendar Time 2 (Cont'd)

- The explanatory variables may be risk measures from a factor model
- The calendar time approach will weight each time period equally and thus the weight on each individual firm in the sample will vary
- This may be problematic and will result in a loss of power if managers time events to take advantage of misvaluations.

Event Studies: 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
- Outliers may cause problems, especially in the context of small samples
- Bootstrapped standard errors could be used in constructing t-statistics
- Another strategy for dealing with non-normality would be to use a non-parametric test

Event Studies: Small Samples and Non-normality (Cont'd)

- Such tests are robust in the presence of non-normal distributions, although they are usually less powerful than their parametric counterparts
- 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.

Event Studies: A Non-parametric Test

- We could then use the test statistic, Z_p:

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

- where p is the actual proportion of negative abnormal returns during the event window and p^* is the expected proportion
- Under the null hypothesis, the test statistic follows a binomial distribution, which can be approximated by the standard normal distribution
- p^* is calculated based on the proportion of negative abnormal returns during the estimation window.
- Tests of the CAPM and the Fama-French Methodology

Tests of the CAPM and the Fama-French Methodology

Testing the CAPM: The Basics

- The most commonly quoted equation for the CAPM is

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f]$$

- So the CAPM states that the expected return on any stock i is equal to the risk-free rate of interest, R_f, plus a risk premium.
- This risk premium is equal to the risk premium per unit of risk, also known as the market risk premium, $[E(R_m) - R_f]$, multiplied by the measure of how risky the stock is, known as 'beta', β_i
- Beta is not observable from the market and must be calculated, and hence tests of the CAPM are usually done in two steps:
  - Estimating the stock betas
  - Actually testing the model
- If the CAPM is a good model, then it should hold 'on average'.

Testing the CAPM: Calculating Betas

- A stock's beta can be calculated in two ways – one approach is to calculate it directly as the covariance between the stock's excess return and the excess return on the market portfolio, divided by the variance of the excess returns on the market portfolio:

$$\beta_i = \frac{Cov(R_i^e, R_m^e)}{Var(R_m^e)}$$

  where the ^{e} superscript denotes excess return

- Alternatively, and equivalently, we can run a simple time-series regression of the excess stock returns on the excess returns to the market portfolio separately for each stock, and the slope estimate will be the beta:

$$R_{i,t}^e = \alpha_i + \beta_i R_{m,t}^e + u_{i,t}$$

Testing the CAPM: The Second Stage Regression

- Suppose that we had a sample of 100 stocks (N=100) and their returns using five years of monthly data (T=60)
- The first step would be to run 100 time-series regressions (one for each individual stock), the regressions being run with the 60 monthly data points
- Then the second stage would involve a single cross-sectional regression of the average (over time) of the stock returns on a constant and the betas:

$$\bar{R}_i = \lambda_0 + \lambda_1 \beta_i + v_i$$

where $\bar{R}_i$ is the return for stock i averaged over the 60 months

Testing the CAPM: The Second Stage Regression (Cont'd)

- Essentially, the CAPM says that stocks with higher betas are more risky and therefore should command higher average returns to compensate investors for that risk
- If the CAPM is a valid model, two key predictions arise which can be tested using this second stage regression: $\lambda_0 = R_f$ and $\lambda_1 = [R_m - R_f]$.

Testing the CAPM: Further Implications

- Two further implications of the CAPM being valid:
  - There is a linear relationship between a stock's return and its beta
  - No other variables should help to explain the cross-sectional variation in returns
- We could run the augmented regression:

$$\bar{R}_i = \lambda_0 + \lambda_1 \beta_i + \lambda_2 \beta_i^2 + \lambda_3 \sigma_i^2 + v_i$$

  where β_i^2 is the squared beta for stock i and σ_i^2 is the variance of the residuals from the first stage regression, a measure of idiosyncratic risk
- The squared beta can capture non-linearities in the relationship between systematic risk and return
- If the CAPM is a valid and complete model, then we should see that $\lambda_2 = 0$ and $\lambda_3 = 0$.

Testing the CAPM: A Different Second-Stage Regression

- It has been found that returns are systematically higher for small capitalisation stocks and are systematically higher for 'value' stocks than the CAPM would predict.

- We can test this directly using a different augmented second stage regression:

$$\bar{R}_i = \alpha + \lambda_1 \beta_i + \lambda_2 MV_i + \lambda_3 BTM_i + v_i$$

  where MV_i is the market capitalisation for stock i and BTM_i is the ratio of its book value to its market value of equity

- Again, if the CAPM is a valid and complete model, then we should see that $\lambda_2 = 0$ and $\lambda_3 = 0$.

Problems in Testing the CAPM

These are numerous, and include:

- Non-normality – e.g. caused by outliers can cause problems with inference
- Heteroscedasticity – some recent research has used GMM or another robust technique to deal with this
- Measurement errors since the betas used as explanatory variables in the second stage are estimated – in order to minimise such measurement errors, the beta estimates can be based on portfolios rather than individual securities
- Alternatively, the Shanken (1992) correction can be applied to adjust the standard errors for beta estimation error.

The Fama-MacBeth Approach

- Fama and MacBeth (1973) used the two stage approach to testing the CAPM outlined above, but using a time series of cross-sections
- Instead of running a single time-series regression for each stock and then a single cross-sectional one, the estimation is conducted with a rolling window
- They use five years of observations to estimate the CAPM betas and the other risk measures (the standard deviation and squared beta) and these are used as the explanatory variables in a set of cross-sectional regressions each month for the following four years

The Fama-MacBeth Approach (Cont'd)

- The estimation is then rolled forward four years and the process continues until the end of the sample is reached

- Since we will have one estimate of the lambdas for each time period, we can form a *t*-ratio as the average over *t* divided by its standard error (the standard deviation over time divided by the square root of the number of time-series estimates of the lambdas).

- The average value of each lambda over *t* can be calculated using:

$$\hat{\lambda}_j = \frac{1}{T_{FMB}} \sum_{t=1}^{T_{FMB}} \hat{\lambda}_{j,t}, \quad j = 1, 2, 3, 4$$

where T_{FMB} is the number of cross-sectional regressions used in the second stage of the test, the j are the four different

The Fama-MacBeth Approach (Cont'd)

parameters (the intercept, the coefficient on beta, etc.) and the standard deviation is

$$\hat{\sigma}_j = \sqrt{\frac{1}{T_{FMB} - 1} \sum_{t=1}^{T_{FMB}} (\hat{\lambda}_{j,t} - \hat{\lambda}_j)^2}$$

- The test statistic is then simply $\sqrt{T_{FMB}}\hat{\lambda}_j / \hat{\sigma}_j$, which is asymptotically standard normal, or follows a *t*-distribution with $T_{FMB} - 1$ degrees of freedom in finite samples.

Fama-MacBeth: Their Key Results

- We can compare the estimated values of the intercept and slope with the actual values of the risk-free rate (R_f) and the market risk premium $[R_m - R_f]$, which are, for the full-sample corresponding to the results presented in the table, 0.013 and 0.143 respectively.
- The intercept and slope parameter estimates (the lambdas) have the correct signs but they are too small
- Thus the implied risk-free rate is positive and so is the relationship between returns and beta
- Both parameters are significantly different from zero, although they become insignificant when the other risk measures are included as in the second row of the table
- It has been argued that there is qualitative support for the CAPM but not quantitative support.

Fama-MacBeth: A Results Table

- It is also worth noting from the second row of the table that squared beta and idiosyncratic risk have parameters that are even less significant than beta itself in explaining the cross-sectional variation in returns.

Fama and MacBeth's Results on Testing the CAPM

Model	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\lambda}_3$
Model 1: CAPM	0.0061*	0.0085*		
	(3.24)	(2.57)		
Model 2: Augmented CAPM	0.0020	0.0114	-0.0026	0.0516
	(0.55)	(1.85)	(-0.86)	(1.11)

Notes: t-ratios in parentheses; * denotes significance at the 5% level. *Source:* Fama and MacBeth (1973), numbers extracted from their Table 3.

The Fama-French Methodology

- The 'Fama-French methodology' is a family of related approaches based on the notion that market risk is insufficient to explain the cross-section of stock returns
- The Fama-French and Carhart models seek to measure abnormal returns after allowing for the impact of the characteristics of the firms or portfolios under consideration
- It is widely believed that small stocks, value stocks, and momentum stocks, outperform the market as a whole
- If we wanted to evaluate the performance of a fund manager, it would be important to take the characteristics of these portfolios into account to avoid incorrectly labelling a manager as having stock-picking skills.

The Fama-French (1992) Approach

- The Fama-French (1992) approach, like Fama and MacBeth (1973), is based on a time-series of cross-sections model
- A set of cross-sectional regressions are run of the form

$$R_{i,t} = \alpha_{0,t} + \alpha_{1,t}\beta_{i,t} + \alpha_{2,t}MV_{i,t} + \alpha_{3,t}BTM_{i,t} + u_{i,t}$$

  where $R_{i,t}$ are again the monthly returns, $\beta_{i,t}$ are the CAPM betas, $MV_{i,t}$ are the market capitalisations, and $BTM_{i,t}$ are the book-to-price ratios, each for firm i and month t
- So the explanatory variables in the regressions are the firm characteristics themselves
- Fama and French show that size and book-to-market are highly significantly related to returns
- They also show that market beta is not significant in the regression (and has the wrong sign), providing very strong evidence against the CAPM.

The Fama-French (1993) Approach

- Fama and French (1993) use a factor-based model in the context of a time-series regression which is run separately on each portfolio *i*

$$R_{i,t} = \alpha_i + \beta_{i,M} RMRF_t + \beta_{i,S} SMB_t + \beta_{i,V} HML_t + \epsilon_{i,t}$$

  where *R i,t* is the return on stock or portfolio *i* at time *t*, *RMRF*, *SMB*, and *HML* are the factor mimicking portfolio returns for market excess returns, firm size, and value respectively

- The excess market return is measured as the difference in returns between the S&P 500 index and the yield on Treasury bills (*RMRF*)

The Fama-French (1993) Approach (Cont'd)

- *SMB* is the difference in returns between a portfolio of small stocks and a portfolio of large stocks, termed 'Small Minus Big
- *HML* is the difference in returns between a portfolio of value stocks and a portfolio of growth stocks, termed 'High Minus Low'.

The Fama-French (1993) Approach 2

- These time-series regressions are run on portfolios of stocks that have been two-way sorted according to their book-to-market ratios and their market capitalisations
- It is then possible to compare the parameter estimates qualitatively across the portfolios i
- The parameter estimates from these time-series regressions are factor loadings that measure the sensitivity of each individual portfolio to the factors
- The second stage in this approach is to use the factor loadings from the first stage as explanatory variables in a cross-sectional regression:

$$\bar{R}_i = \alpha + \lambda_M \beta_{i,M} + \lambda_S \beta_{i,S} + \lambda_V \beta_{i,V} + e_i$$

The Fama-French (1993) Approach 2 (Cont'd)

- We can interpret the second stage regression parameters as factor risk premia that show the amount of extra return generated from taking on an additional unit of that source of risk.

The Carhart (1997) Approach

- It has become customary to add a fourth factor to the equations above based on momentum
- This is measured as the difference between the returns on the best performing stocks over the past year and the worst performing stocks – this factor is known as UMD – 'up-minus-down'
- The first and second stage regressions then become respectively:

$$\begin{aligned} R_{i,t} &= \alpha_i + \beta_{i,M} RMRF_t + \beta_{i,S} SMB_t + \beta_{i,V} HML_t \\ &\quad + \beta_{i,U} UMD_t + \epsilon_{i,t} \\ \bar{R}_i &= \alpha + \lambda_M \beta_{i,M} + \lambda_S \beta_{i,S} + \lambda_V \beta_{i,V} + \lambda_U \beta_{i,U} + e_i \end{aligned}$$

The Carhart (1997) Approach (Cont'd)

- Carhart forms decile portfolios of mutual funds based on their one-year lagged performance and runs the time-series regression on each of them
- He finds that the mutual funds which performed best last year (in the top decile) also had positive exposure to the momentum factor (UMD) while those which performed worst had negative exposure
- More recent research adds a variety of other factors using the same framework.

Extreme Value Theory

Extreme Value Theory: An Introduction

- Conventional statistical models are based upon estimates of the average or typical behaviour of a series
- Such models often perform poorly when the focus switches to extreme events
- In particular, statistics based on a normal distribution will systematically under-estimate the probability of extreme events in most financial time series as they cannot capture the fat tails
- Extreme value distributions are sufficiently flexible that they can capture these fat tails much more accurately

Extreme Value Theory: An Introduction (Cont'd)

- For example, a normal distribution would suggest the probability of a return of -10.84% or lower in a 30-year period on A-rated corporate bonds is 0.00008%, whereas based on an extreme value distribution it is 1.4% (source: Levine, D. (2009) Modelling tail behavior with extreme value theory Risk Management Society of Actuaries 17, 15-18).

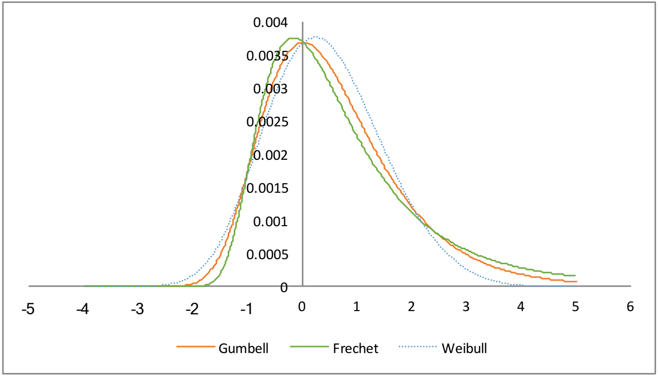
Extreme Value Theory: Two Frameworks

- There are two broad approaches to parameter estimation under extreme value theory: the block maximum framework and the peaks over threshold (POT) framework
- The block maximum framework involves separating the observations into blocks and taking the maximum from each block as constituting the extreme data points
- The POT framework specifies an arbitrary high threshold and any observed value of the series exceeding this is defined as being an extreme data point
- In each case there are three classes of distributions which normalised versions of these extremes could follow

Extreme Value Theory: Two Frameworks (Cont'd)

- For the block maximum they are called the Weibull, Gumbel, and Frechét distributions while the corresponding distributions under the POT approach are the ordinary Pareto, exponential and beta distributions respectively.

PDFs for the Weibull, Gumbel and Frechét Distributions



The Peaks Over Threshold Approach

- The block maxima approach is rarely used since splitting the data into blocks is inefficient unless the extremes happen to be evenly spaced across the blocks
- To implement the peaks over threshold approach, we specify a threshold U and calculate the values of the data points minus the threshold:
  $\tilde{y}_t = y_t - U | y_t > U$
- As the threshold tends to infinity, a normalised version of $\tilde{y}_t$ tends to the generalised Pareto distribution (GPD). The cdf of the GPD can be written:

$$G_{\xi,\sigma}(\tilde{y}_t) = \begin{cases} 1 - (1 + \xi \tilde{y}_t / \sigma)^{-1/\xi} & \text{if} \quad \xi \neq 0 \\ 1 - exp[-\tilde{y}_t / \sigma] & \text{if} \quad \xi = 0 \end{cases}$$

  where ξ is the shape parameter and σ is the scale parameter

The Peaks Over Threshold Approach (Cont'd)

- The key parameter is ξ, sometimes also known as the tail index, measuring how rapidly the tails decay.

The Peaks Over Threshold Approach 2

- When $\xi > 0$, this corresponds to the fat-tailed case
- The tail index is the inverse of the number of degrees of freedom in a t distribution so $\xi = 1/v$
- Typical estimates of v are of the order 4–6, suggesting plausible value of ξ would be 0.1–0.2
- The choice of the threshold U is tricky and involves a trade-off
- If the threshold is too big in absolute value (too far into the tail), the number of points classified as extremes will be too small leading to high standard errors for the parameter estimates (ξ and σ)

The Peaks Over Threshold Approach 2 (Cont'd)

- If the threshold is too small in absolute value (too near the centre of the distribution), many points will be classified as extremes when they are not, leading to biased parameter estimates

- One way to resolve this trade-off is to estimate the parameters for increasingly large values of U until they become stable.

Parameter Estimation for Extreme Value Distributions

- The parameters can be estimated by maximum likelihood
- The pdf for the value of yt over the threshold (for $\xi \neq 0$) is given by:

$$g_{\xi,\sigma}(\tilde{y}_t) = \frac{1}{\sigma}\left(1 + \frac{\xi \tilde{y}_t}{\sigma}\right)^{-\left(\frac{1}{\xi}+1\right)}$$

- Then we can form the joint density for all observations over the threshold, N_U:

$$LF(\xi, \sigma, \tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_{N_U}) = \prod_{i=1}^{N_U} g_{\xi,\sigma}(\tilde{y}_t) = \prod_{i=1}^{N_U} \frac{1}{\sigma}\left(1 + \frac{\xi \tilde{y}_t}{\sigma}\right)^{-\left(\frac{1}{\xi}+1\right)}$$

Parameter Estimation for Extreme Value Distributions (Cont'd)

- The log likelihood would then be given by taking the natural log of this and rearranging:

$$LLF(\xi, \sigma, \tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_{N_U}) = -N_U \ln(\sigma) - \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^{N_U} \ln\left(1 + \frac{\xi \tilde{y}}{\sigma}\right)$$

Parameter Estimation for Extreme Value Distributions 2

- Provided that $\xi > -0.5$, maximum likelihood estimators are consistent and asymptotically normal
- But no analytical solutions exist and thus a numerical search procedure is required
- Instead, it is common to use a non-parametric procedure to estimate the shape parameter directly from the data
- The simplest approach is the Hill (1975) estimator

Parameter Estimation for Extreme Value Distributions 2 (Cont'd)

- If we order the exceedences over the threshold from the largest to the smallest, $\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_T$, the Hill estimator of ξ is given by

$$\hat{\xi} = \frac{1}{k-1} \sum_{i=1}^{k-1} [\ln(\tilde{y}_{(i)}) - \ln(\tilde{y}_{(k)})]$$

where k is an integer to be selected equal to the number of observations in the tail.

Parameter Estimation for Extreme Value Distributions 3

- Two further non-parametric estimators of ξ are due to Pickands (1975)

$$\hat{\xi} = \frac{1}{\ln(2)} \ln \left(\frac{\tilde{y}_{(k)} - \tilde{y}_{(2k)}}{\tilde{y}_{(2k)} - \tilde{y}_{(4k)}} \right)$$

  and due to De Haan and Resnick (1980)

$$\hat{\xi} = \frac{\ln(\tilde{y}_{(1)}) - \ln(\tilde{y}_{(k)})}{\ln(k)}$$

- If the shape parameter is estimated using one of the above non-parametric techniques, we still need to estimate the scale parameter, σ

Parameter Estimation for Extreme Value Distributions 3 (Cont'd)

- This can be achieved by plugging the given value of ξ into the LLF and then estimating σ as the only free parameter using maximum likelihood.

Introduction to Value at Risk

- Value at Risk (VaR) is a method for measuring the financial risk inherent in a portfolio or securities position
- It can be defined as the loss in financial terms that is expected to occur over a given horizon with a given degree of confidence
- Extreme value distributions can be effective for calculating VaR as they estimate the probability of extreme events more accurately than approaches not allowing for fat tails
- The delta-normal model for calculating VaR simply takes the standard deviation of the portfolio returns data, σ, and multiples it by the relevant quantile from the normal distribution, Z_α at the α significance level
  $VaR_{notmal} = \sigma Z_\alpha$

Introduction to Value at Risk (Cont'd)

- A second straightforward approach to calculating VaR is to use *historical simulation*, which involves sorting the actual historical portfolio returns and selecting the appropriate quantile from the empirical distribution of ordered returns.

Calculating Value at Risk using EVT

- Assume that the peaks over threshold approach is used and that the parameters are ξ and σ. VaR is calculated as follows, with a significance level of α and where N and N_U are the total number of data points and the number exceeding the threshold U respectively

$$VaR = U + \frac{\hat{\sigma}}{\hat{\xi}}\left[\left(\frac{N}{N_U}\alpha\right)^{-\hat{\xi}} - 1\right]$$

- If the block maximum approach is used, VaR can be calculated using the following formula where m is the block length

$$VaR = \hat{\mu} + \frac{\hat{\sigma}}{\hat{\xi}}\left[1 - (-m\ln(\alpha))^{-\hat{\xi}}\right]$$

Calculating Value at Risk using EVT (Cont'd)

- EVT can be extended to the multivariate case to measure common dependence and spillovers between extreme events in time-series.

The Generalised Method of Moments

The Method of Moments

- The method of moments is an alternative to OLS and maximum likelihood used for estimating the parameters in a model
- It works by computing the moments of the sample data and then setting them equal to their population values based on an assumed probability distribution for the latter
- If we have k parameters to estimate, we need k sample moments
- If the observed data y follow a normal distribution with population mean μ and population variance σ^2, we need two moment restrictions to estimate the two parameters

The Method of Moments (Cont'd)

- The sample moments converge upon their population counterparts asymptotically. Thus:

$$\frac{1}{T}\sum_{t=1}^{T} y_t - \mu_0 \rightarrow 0 \quad \text{as} \quad T \rightarrow \infty$$

The Method of Moments 2

- The first sample moment is then found by taking the usual sample average, $\bar{y}$:

$$\frac{1}{T}\sum_{t=1}^{T} y_t - \mu_0 = 0$$

- We would then adopt the same approach to match the second moment

$$\sigma^2 = E[(y_t - \mu_0)^2]$$

- And thus:

$$\frac{1}{T}\sum_{t=1}^{T} y_t^2 - \sigma^2 = 0$$

The Method of Moments 2 (Cont'd)

- So we have:

$$\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T} y_t^2 - \bar{y}^2 = 0.$$

The Method of Moments 3

- If we had a more complex distribution than the normal, we would match the third, fourth, ..., moments until we had the same number as parameters to estimate

- In the context of estimating the parameters of a regression model, the method of moments relies on the assumption that the $T \times k$ matrix of observations on the explanatory variables is orthogonal to the disturbances:
$\mathrm{E}[u_t x_t] = 0$
- Here there would be the same number of moment restrictions as parameters to estimate

- If we let β^* denote the true value of β, the vector of parameters, then the moment conditions would be written as:

$$E[(y_t - x_t'\beta^*)x_t] = 0$$

The Generalised Method of Moments

- The conventional method of moments estimator requires us to have the same number of moment conditions as parameters to estimate
- This is unrealistic, and it is more likely that we will have an over-identified system that has more moment restrictions than parameters to estimate, in which case the method of moments cannot be used
- But the GMM was developed by Hansen (1982) for precisely this purpose
- When there are more moment restrictions than parameters to estimate, we will have multiple solutions, and GMM selects from among them the solution that minimises the variance of the moment conditions

The Generalised Method of Moments (Cont'd)

- Suppose that we have L moment conditions ($l = 1, \ldots, L$) to estimate k parameters in a vector β. The moment conditions are

$$E[m_l(y_t, x_t; \beta)] = 0$$

The Generalised Method of Moments 2

- We then estimate the parameters as those coming as close as possible to satisfying the moment conditions. The parameter vector estimator is

$$\hat{\beta}_{GMM} = \text{argmin}_{\beta}\ \hat{m}(\hat{\beta})' W \hat{m}(\hat{\beta})$$

  where $\hat{m}(\hat{\beta}) = (\hat{m}_1, \ldots, \hat{m}_L)$ are the L moment conditions (which will be a function of the estimated parameters, and W is the weighting matrix are the L moment conditions and W is the weighting matrix

- Is is possible to show that the optimal W is the inverse of the variance-covariance matrix of the moment conditions

- Often, a two-step approach is used:

The Generalised Method of Moments 2 (Cont'd)

- In the first stage, the weighting matrix is substituted by an arbitrary choice that does not depend on the parameters (such as the identity matrix)
- In the second stage, it is substituted by an estimate of the variance given the parameter estimates from the first stage

Over-identifying Restrictions and GMM

- For over-identified systems we can use the effective degrees of freedom to test the over-identifying restrictions through a Sargan J-test
- The null hypothesis is that all of the moment conditions are exactly satisfied
- The test statistic, which asymptotically follows a chi-squared distribution with $L - k$ degrees of freedom is given by

$$\hat{m}(\hat{\beta})'[\mathrm{EAV}(\hat{m}(\hat{\beta}))]^{-1}\hat{m}(\hat{\beta})$$

  where *EAV* is the estimated asymptotic variance
- If the null is rejected, it would indicate that the parameter estimates are not supported by the data.