8

EMPIRICAL EVIDENCE: CAPM AND APT

Aims

- Demonstrate how the CAPM (and zero-beta CAPM) can be tested using time-series and cross-section approaches.
- Show that a *cross section* of average stock market returns is best explained by a multifactor model. The cross section of *stock* returns depends on the betas for book-to-market and size variables but not on the CAPM-beta.
- Demonstrate that the CAPM-beta only explains the different cross-section average returns on bonds (or T-bills) *versus* stocks.
- Outline how the APT can be viewed as an *equilibrium* multifactor model.

8.1 CAPM: Time-Series Tests

The 'standard' Sharpe–Lintner CAPM is a direct implication of mean-variance efficiency, under the assumption of homogeneous expectations and the existence of a (known non-stochastic) risk-free rate. The most straightforward test is that $\alpha_i = 0$, in the excess returns regression

$$ER_{it} - r_t = \alpha_i + \beta_i (ER_{\mathrm{m}} - r)_t \tag{1}$$

We assume individual returns are temporally *iid*, although we allow contemporaneous correlation across assets $E(\varepsilon_{it}\varepsilon_{jt}) \neq 0$. Then, under the assumption of joint normality of returns, (1) can be estimated by maximum likelihood using panel data (N assets,

$t = 1, 2, \ldots, T$ time periods). But because we have the same (single) independent variable in all equations, maximum likelihood estimates for (α, β) are equivalent to OLS equation-by-equation: the OLS residuals can be used to form the (contemporaneous) covariance matrix for the N assets $\sum = \hat{e}\hat{e}'/T$. But the asymptotic results from Wald and likelihood ratio tests may have substantial size distortions in finite samples (see Campbell, Lo and MacKinlay 1997 Chapter 5). However, there are various tests for $\alpha(N \times 1) = 0$, valid in small samples. The power (i.e. the probability of rejecting the null, given that an alternative hypothesis is true) of the exact tests is found to be increasing in T (as we might expect) but is very sensitive to N, which should probably be kept small (i.e. less than 10). When there is conditional heteroscedasticity or serial correlation, then GMM can be used, although less is known about the small sample properties of exact tests of $\alpha = 0$. In this case, MCS can be used (e.g. with alternative distributions) or we can bootstrap the residuals, in order to obtain an empirical distribution for $\hat{\alpha}$, in finite samples.

Tests of Black's *zero-beta* CAPM use *real* returns in the regression

$$\boldsymbol{R}_t = \boldsymbol{\alpha} + \boldsymbol{\beta} R_{mt} + \boldsymbol{\varepsilon}_t \tag{2}$$

where $\boldsymbol{R} \sim (N \times 1)$ vector of N-asset returns etc. and $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ are both $N \times 1$. The null hypothesis is

$$H_0 : \boldsymbol{\alpha} = (\boldsymbol{e} - \boldsymbol{\beta})\gamma$$

where $\boldsymbol{e}$ is an $(N \times 1)$ unit vector and $\gamma = E(R_{0B})$ the expected return on the zero-beta portfolio, which is treated as an unobservable parameter. Panel data regression using an iterative systems estimator gives estimates of the constrained $\boldsymbol{\beta}, \gamma$ (and the covariance matrix Σ) and then standard (asymptotic) Wald and likelihood ratio tests are possible.

Early studies in the 1970s find $\alpha_i = 0$ and hence tend to favour the Sharpe–Lintner CAPM but subsequent work (e.g. Campbell, Lo and MacKinlay 1997) finds against. Cochrane (1996) directly estimates a 'conditional-CAPM' where the impact of the excess market return is 'scaled' by the dividend–price ratio or the term premium. Hence, the impact of the market return on the return on assets (or portfolio of assets) depends on variables that reflect the 'state of the business cycle'. He finds that for size-sorted portfolio returns, the pricing error (i.e. Jensen's alpha) is halved compared to the standard (unconditional) CAPM.

8.2 CAPM: Cross-Section Tests

The CAPM states that differences in average returns in a cross section of stocks depends linearly (and solely) on asset-betas. (This is the Security Market Line, SML.) The first problem in testing this hypothesis is that *individual* stock returns are so volatile that one cannot reject the hypothesis that *average* returns across different stocks are the same. For individual stocks, $\sigma \approx 30 - 80\%$ p.a. and hence $\sigma/\sqrt{T}$ can still be very large even if T is large. One answer is to sort stocks into portfolios where the sorting attempts to maximise differences in average returns – without differences in average returns, we cannot test the CAPM. In principle, any grouping is permissible, but a grouping based on, say, ticker symbols A–E, F–J, and so on, may not produce a good spread of average returns. However, grouping according to 'size' and 'book-to-market' are popular methods that produce a good spread of average returns. (Although, in early work, the grouping was based on asset-betas.)

The second major problem is that the betas are measured with error. Several methods have been used to minimise this problem. One is to assign individual stocks into a small number of 'portfolio betas'. These portfolio betas are estimated using a time-series regression of just a small number of portfolio returns (e.g. around 10) – this grouping is thought to minimise the error in estimating betas. To allow a firm to have a different beta over time, the above approach has been extended and used in rolling regressions. These issues are discussed below.

Cross-section tests take the form of a two-stage procedure (on which, more below). Under the assumption that β_i is constant over the whole sample, a *first-pass time-series regression* for each asset i, taken in turn is

$$R_{it} - r_t = \alpha_i + \beta_i(ER_{\text{m}} - r)_t + \varepsilon_{it} \tag{3}$$

The estimates of β_i *for each security* may then be used in a *second-pass cross-section* regression. Here, the *sample average* monthly returns $\overline{R}_i$ (usually over the whole sample) on all k-securities are regressed on the $\hat{\beta}_i$'s from the first-pass regression

$$\overline{R}_i = \lambda_0 + \lambda_1\hat{\beta}_i + v_i \tag{4}$$

Comparing (4) with the standard CAPM relation

$$R_i = r + \beta_i(R_{\text{m}} - r) + \varepsilon_i \tag{5}$$

we expect $\lambda_0 = \overline{r}$, $\lambda_1 = \overline{R}_{\text{m}} - \overline{r} > 0$ where a bar indicates the sample mean values. An even stronger test of the CAPM in the second-pass regression is to note that *only the betas* $\beta_i(i = 1$ to $k)$ should influence $\overline{R}_i$, so no other cross-section variables should be significant in (5) (e.g. the own variance of returns $\sigma_{\varepsilon i}^2$).

Note that in (4), the λ_0 and λ_1 are constant across all assets (in the cross section). If there is no risk-free rate, then λ_0 must be estimated and as it is the expected return on a portfolio when all betas are zero, λ_0 is the *zero-beta rate*.

An acute econometric problem is that in the first-pass time-series regression, the estimate $\hat{\beta}_i$ may be unbiased but it is measured with error. Hence, in the second-pass regression (4), we have a classic 'errors-in-variables' problem, which means that the OLS coefficient of λ_1 is downward biased. Also, if the true β_i is positively correlated with the security's error variance $\sigma_{\varepsilon i}^2$, then the latter serves as a proxy for the true β_i and hence if $\hat{\beta}_i$ is measured with error, then $\sigma_{\varepsilon i}^2$ may be significant in the second-pass regression. Finally, note that if the error distribution of ε_{it} is non-normal (e.g. positively skewed or fat-tailed), then incorrect inferences will also ensue. In particular, positive skewness in the residuals of the cross-section regressions will show up as an association between residual risk and return, even though in the true model there is no association.

As an illustration of these early studies, consider that of Black, Jensen and Scholes (1972), who use monthly rates of return 1926–1966 in the first-pass time-series regressions. They minimised the heteroscedasticity problem and the error in estimating the betas by grouping all stocks into a set of 10 portfolios based on the size of the betas for *individual* securities (i.e. the time-series estimates of β_i for individual securities over a rolling five-year estimation period are used to assemble the portfolios). For each of the 10 beta-sorted portfolios, the monthly return R_{it}^{p} is regressed on R_{mt} over a period of 35 years

$$R_{it}^{\mathrm{p}} = \hat{\alpha}_i + \hat{\beta}_i R_{mt} \tag{6}$$

This gives the *portfolio* betas, and each individual stock is then *allocated* to one of these 10 beta values. In the second-pass cross-section regressions, the average monthly excess return $\overline{R}_i - \overline{r}$ for all *individual* stocks is regressed on their 'allocated' portfolio betas. A statistically significant positive coefficient λ_1 in (4) supports the CAPM.

Cochrane (2001) updates the above studies and provides a clear simple test of the CAPM. First, he sorts all stocks on the NYSE into 10 portfolios on the basis of size (i.e. market capitalisation) and also includes a portfolio of corporate bonds and government bonds (i.e. 12 portfolios in all). Next, he uses separate time-series regressions to estimate the 12 portfolio betas. Finally, he takes the sample average returns $(t = 1, 2, \ldots, T)$ for each portfolio $\overline{R}_i$ $(i = 1, 2, \ldots, 12)$ and regresses these against the 12 estimated portfolio betas $\hat{\beta}_i$. The bond portfolios have low betas and low average returns, and the size-sorted stock returns are also positively related to beta (Figure 1). Things are looking good for the CAPM. However, there are two problems. First, if we draw the SML, $\overline{R}_i = \overline{r} + (\overline{R}_m - \overline{r})\beta_i$, the points for the *stock* returns tend to lie above it rather than on it – this is perhaps not too damning. Second, the OLS regression has an average return for the smallest stock decile way above the estimated SML (Figure 1). This is the 'small firm effect' (Banz 1981) – very small firms (i.e. low market cap) earn average returns above their SML risk-adjusted value. The CAPM/SML seems to do a reasonable job of explaining bond returns *relative to* stock returns but not such a good job in explaining the cross section of average stock

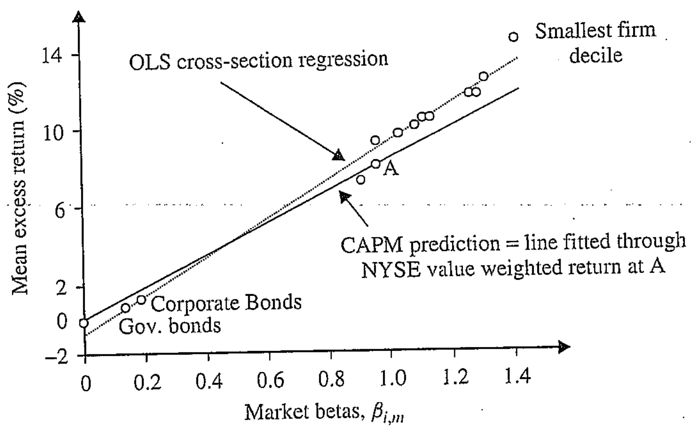


Figure 1 Size-sorted value-weighted decile portfolio (NYSE) : US data, post-1947

returns sorted by 'size'. As we see below, this apparent moderate 'success' of the CAPM is fragile since when stocks are sorted into deciles according to book-to-market value, these decile returns are *not* explained by market betas.

Fama–MacBeth Rolling Regression

Fama and MacBeth (1973) provide a much-used regression methodology that involves *'rolling' cross-section regressions*. Essentially, one undertakes a separate cross-section regression for each time period, hence obtaining a time-series of the coefficient on the chosen cross-section variable (e.g. the betas), on which we can then perform various tests. To illustrate, suppose we have N industry returns for any single month t, $\boldsymbol{R}_t$ and their associated cross-section betas *are known with certainty*. The simplest test of the CAPM is to run a cross-section OLS regression *for any single month* t.

$$\boldsymbol{R}_t = \alpha_t' \boldsymbol{e} + \gamma_t \boldsymbol{\beta} + \theta_t Z + \boldsymbol{\varepsilon} \tag{7}$$

where $R_t = (R_1, R_2, \ldots, R_N)_t \sim N \times 1$ vector of *cross-section* excess monthly returns at time t

$\beta = (\beta_1, \beta_2, \ldots, \beta_N) \sim N \times 1$ vector of CAPM-betas

$\gamma_t =$ *scalar* cross-section coefficient for time t (the 'price' of beta-risk)

$\alpha_t =$ *scalar* estimate of intercept for time t

$\boldsymbol{e} = N \times 1$ vector of ones

$Z =$ additional cross-section variables (e.g. book-to-market value)

$\boldsymbol{\varepsilon} = N \times 1$ vector of cross-section error terms

For any t, the CAPM implies $\alpha_t = \theta_t = 0$ and $\gamma_t > 0$. The Fama–MacBeth procedure 'averages' these parameter estimates of over all time periods as follows. We repeat regression (7) for each month $t = 1, 2, \ldots, T$ and obtain T estimates of α_t, γ_t and θ_t. In the second step, the time-series of these parameters are tested to see if $\alpha \equiv E(\alpha_t) = 0$, $\theta \equiv E(\theta_t) = 0$ and $\gamma \equiv E(\gamma_t) > 0$ (i.e. positive risk premium on the betas). If the returns are *iid* over time and normally distributed, then the following statistic is distributed as Student's-t (with $T - 1$ degrees of freedom)

$$t_\gamma = \tilde{\gamma}/\tilde{\sigma} \tag{8}$$

where

$$\tilde{\gamma} = \sum_{t=1}^{T} \hat{\gamma}_t / T \quad \text{and} \quad \tilde{\sigma}^2 = \frac{1}{T(T-1)} \sum_{t=1}^{T} (\hat{\gamma}_t - \tilde{\gamma})^2$$

A similar procedure gives t_α and t_θ and hence the CAPM restrictions are easily tested. One problem with the above is that the 'true' betas are unknown, and this introduces an errors-in-variables problem. One could use instrumental variables for the betas, but the popular method of minimising this problem is to group individual returns into, say, 100 portfolios (e.g. sorted by size and book-to-market) and calculate 100 *portfolio* betas. We then assign each *individual* stock to have one of these 100 betas (see Appendix).

Fama and MacBeth (1974) used the above procedure and also included β_i^2 and $\sigma_{\varepsilon i}^2$(from the 1st-pass regression) when estimating the cross-section portfolio returns equation separately for *each month* over 1935 to 1968.

$$\overline{R}_i^{\text{p}} = \lambda_0 + \lambda_1\beta_i + \lambda_2\beta_i^2 + \lambda_3\sigma_{\varepsilon i}^2 + \eta_i \tag{9}$$

They find that the (time series) *average* of λ_2 and λ_3 are *not* significantly different from 0 and $\lambda_1 > 0$, thus supporting the standard CAPM.

A practical use of the CAPM is in estimating the cost of equity for a firm, which is then used in the weighted average cost of capital (WACC) to provide a discount rate in evaluating investment projects (see Cuthbertson and Nitzsche 2001a). If the firm has an estimate of beta that is biased upwards, then it will set its CAPM hurdle rate too high and may therefore forego profitable investment projects (i.e. where the internal rate of return does not exceed the upward biased hurdle rate). There are various methods in use to calculate 'beta' (Blume 1975), but Bartholdy and Peare (2003) point out that if the beta *estimated* using a proxy for the market index (e.g. S&P500) is then multiplied by a *different proxy* for the average excess market return (NYSE index), then a biased estimate of the cost of equity ensues. They then show how the Fama–MacBeth procedure can be used to obtain an unbiased estimate of the 'true' cost of equity, even if we use a proxy for the 'market index'. First, use equation (3) with a *proxy* for the market index (e.g. S&P500), on (say) 60 months of data (up to time t) to obtain an estimate of β_i^{pr} for each firm. Now use monthly data for the next year ($t + 1$) to run the cross-section regression (7) to obtain the first estimate of γ_t. Roll the regressions forward every year and calculate the average value $\bar{\gamma}$. They show that the cost of equity capital $E_t R_{i,t+1} - r_t = \beta_i E_t(R_{\text{m},t+1} - r_t) = \beta_i^{\text{pr}}[E(R_{\text{m}}^{\text{pr}} - r)/\rho_{\text{pr},m}^2] = \beta_i^{\text{pr}}\bar{\gamma}$ and therefore an unbiased estimate is given by $\beta_i^{\text{pr}}\bar{\gamma}$. Hence, *any* proxy for the market portfolio can be used (e.g. S&P500, NYSE index), and we obtain an unbiased estimate of the cost of equity capital, if the above procedure is followed. Of course, the method still relies on the CAPM being the correct model of expected returns, and if the latter does not 'fit the data' well, an unbiased estimate of the cost of capital but with a large standard error may be of little comfort.

Roll's Critique

Roll (1977) demonstrated that for *any* portfolio that is efficient *ex-post* (call it q), then in a *sample* of data, there is an exact linear relationship between the mean return and beta. It follows that there is really only one testable implication of the CAPM, namely, that the market portfolio is mean-variance efficient. If the market portfolio is mean-variance efficient, then the CAPM/SML *must* hold in the sample. Hence, violations of the SML in empirical work may be indicative that the portfolio chosen by the researcher is not the true 'market portfolio'. Unless the researcher is confident he has the true market portfolio (which may include land, commodities, human capital, as well as stocks and bonds), tests based on the SML are largely superfluous and provide no *additional* confirmation of the CAPM. Despite this, critique researchers have continued to explore the empirical validity of the CAPM even though their proxy for the market portfolio could be incorrect. This is because it is still of interest to see how far a particular empirical model, even if an imperfect one, can explain equilibrium returns, and we can always see if the results in the second-pass regression are robust to alternative choices for the market portfolio.

8.3 CAPM, Multifactor Models and APT

One direct way to test the APT is to use some form of factor analysis. Roll and Ross (1984) applied factor analysis to 42 groups of 30 stocks using daily data between 1962 and 1972. In their first-pass regressions, they find for most groups, about five 'factors' provide a sufficiently good statistical explanation of R_{it}. In the second-pass regression, they find that three factors are sufficient. However, Dhrymes, Friend and Gultekin (1984) show that one problem in interpreting results from factor analysis is that the number of statistically significant factors appears to increase as we include more securities in the analysis. Also, the 'factors' being linear combinations of economic variables are also impossible to interpret. Hence, research over the past 20 years has focused on multifactor regression-based approaches to explain the cross section of average returns.

As we have seen, more recent studies try to sort portfolios in such a way as to minimise the errors in measuring betas and to get as large a spread in average cross-section returns across the chosen portfolio. Indeed, since the classic Fama and French (1993) paper, attention has shifted towards models with multiple factors. To maximise the spread in the cross-section returns, individual stocks are sorted into portfolios. A frequently chosen grouping is by *quintiles* of book-to-market value (i.e. 'value') and equity market value (i.e. 'size'), giving a cross section of 25 average returns in all. The use of portfolio betas reduces measurement error and also mitigates the problem that *individual* betas may change over time because of changes in leverage, firm size, business risks, and so on.

United States: Cross-Section Data

Fama and French (1993) find that the 25 'size and value' sorted, monthly *time-series* returns on US stocks are explained by a three-factor model where the factors are the market return, the return on a 'size' portfolio (i.e. small minus big portfolio R_{SMB}) and the return on a high minus low (HML) book-to-market portfolio, R_{HML}. The time-series and cross-section regressions are

$$R_{it} = \beta_{mi} R_{mt} + \beta_{\text{SMB},i} R_{\text{SMB},t} + \beta_{\text{HML},i} R_{\text{HML},t} \tag{10}$$

$$\overline{R}_i = \lambda_m \beta_{mi} + \lambda_{\text{SMB}} \beta_{\text{SMB},i} + \lambda_{\text{HML}} \beta_{\text{HML},i} \tag{11}$$

Using a time-series regression on each of the 25 portfolios in turn, we obtain estimates of the three betas in equation (10) for each of the 25 portfolios. These cross-section betas can then be used in (11) with the 25 average monthly returns $\overline{R}_i$ (averaged

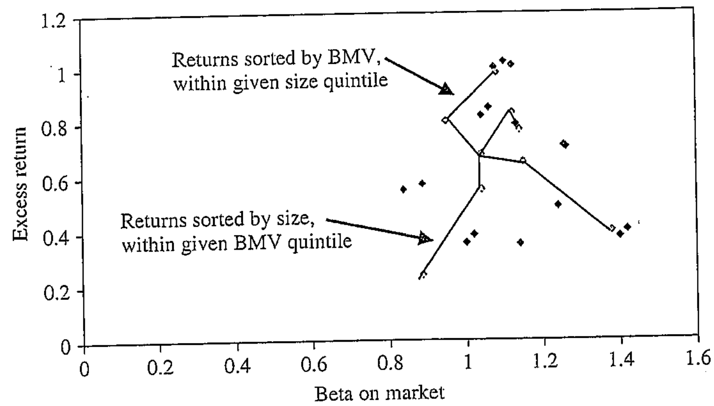
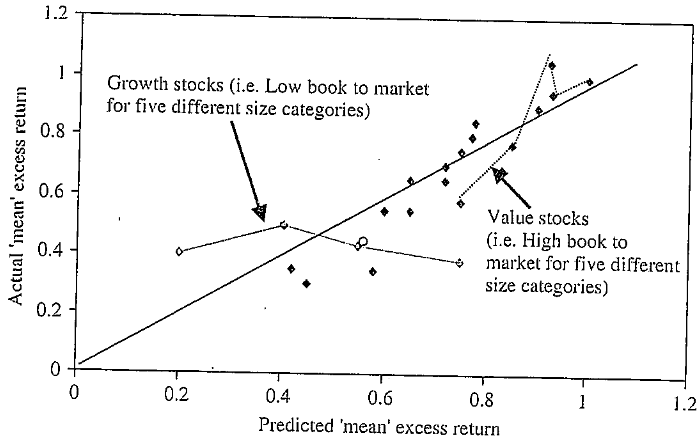


Figure 2 Average excess returns and market beta (25 size- and BMV-sorted portfolios).

over, say, 240 months, given the noise in the data – see Cochrane (2001), Chapter 15) to estimate the λs. Fama and French (1993) find that the market betas β_{mi} for the 25 size-value sorted portfolios are all clustered in the range 0.8 to 1.5, whereas the 25 average monthly returns have a large spread, from 0.25 to 1 (Figure 2). If the CAPM were correct, then the average returns and the market beta would be perfectly positively correlated and *all* of the points in Figure 2 would lie close to the 45° line. Hence, the CAPM-beta explains hardly any of the cross-section variability in average returns $\overline{R}_i$ across stocks (although, as we have noted above, it does help explain the different average cross-section returns on stocks versus T-bills and corporate bonds). As we see below, most of the variation in the cross-section of average stock returns is explained by the SMB and HML betas – the 'factor mimicking' portfolios and not by the CAPM-betas.

Note that if we join up points for different 'size' sorted returns (but within a given book-to-market value BMV quintile), then the positive relationship between size-sorted returns and market beta β_{mi} reappears. (For clarity, we have done this for only one of the book-to-market quintiles in Figure 2, but this positive relationship applies to the other four quintiles with constant book-to-market values.) Hence, higher average returns of smaller firms can partly be explained by their higher market betas β_{mi}. However, it is the BMV 'sorting' that is rejecting the CAPM, since if we look at returns with size held constant but for *different* book-to-market values, then these average returns are *negatively* related to their market betas β_{mi} (see Figure 2-this general negative relationship holds for the remaining quintiles sorted by different book-to-market value within a given size quintile – we have not 'joined up' these points in Figure 2). Hence, it matters how one sorts returns in deciding on the validity of the CAPM. But of course, the CAPM should hold for returns based on any sorting criteria, since (average) returns, whether *individual* or in *any* portfolios, should all show a linear relation between average return and market beta. So, the real anomaly (for the CAPM) is the fact that returns sorted using book-to-market 'value' cannot be explained by the CAPM-betas.

The success of the Fama–French three-factor model is demonstrated in Figure 3 where the *predicted* returns (based on equation (11)) and actual average returns for the



"The two lines connect portfolios of different size categories, within a given book-to-market category. We only connect the points within the highest and lowest BMV categories. If we had joined up points for the other BMV quintiles, the lines would show a positive relationship, like that for the value stocks – showing that the predicted returns from the Fama–French three-factor model broadly predict average returns on portfolios sorted by size and BMV".

Figure 3 Actual and predicted average returns Fama–French three-factor model

25 'size and value' sorted portfolios are graphed and are much closer to the 45° line than in Figure 2. The dotted line joining the five points in the upper right corner of Figure 3 shows stocks that have different sizes within the highest book-to-market quintile (i.e. 'low price' or 'value stocks'), and these provide a good fit to the Fama–French model. The solid line joining the five points in the lower left corner of Figure 3 are those stocks that have different sizes within the lowest book-to-market quintile (i.e. 'high price' or 'growth stocks'), and these provide the worst fit for the Fama–French model. Except for the latter 'growth stocks', we can see from Figure 3 that the predicted returns for the Fama–French three-factor model fit the actual average returns on our 25 portfolios rather well – since the points lie reasonably close to the 45° line and much closer than in Figure 2.

Also, note that the 'size effect' probably disappeared in the mid-1980s. Using data from 1979 to around 1998, Cochrane (2001) shows that for the size-sorted decile portfolios, average monthly returns are all clustered around point A in Figure 1. So, the small firm premium has disappeared as the points lie randomly around the SML, and all of the market betas of the decile portfolios are around 1.0, so there is no positive relationship between size-sorted returns and the market betas, *post*-1979 that is, after the size anomaly appeared in the literature (Banz 1981).

We can view the Fama–French model as an APT model. If the R-squared of the 25 time-series regressions is 100%, the three factors can perfectly mimic the 25 returns, without allowing arbitrage opportunities. Fama and French (1993) find the R-squareds of the time-series regressions are in excess of 90%, so the APT is a possible candidate model.

Fama and French (1996) extend their earlier analysis and find that the HML and SMB factor mimicking portfolios explain the cross-section of average returns based on sorting by 'new' categories such as price multiples (e.g. price–earnings ratios) and five-year sales growth. But the average returns on these new portfolios are not explained by their CAPM market betas. However, there is one well-known 'failure' of the Fama–French model – the HML and SMB betas do *not* explain the average returns on stocks sorted according to their recent return performance (i.e. 'momentum stocks') – see Chapter 18.

We are still left wondering about the *economic* causes of why the HML book-to-market factor is priced. There has been some progress in this area. Fama and French (1995) do find that the typical 'value firm', where price (and hence market-to-book) is low, is a generally distressed or near-bankrupt firm. Hence, HML could signal aggregate risk over the business cycle and there is weak evidence that the return on HML portfolios do help explain movements in GDP (Liew and Vassalon 1999, Heaton and Lucas 1996). It is also true that small firms tend to exhibit characteristics that indicate financial distress (e.g. high gearing or leverage), but note that this risk must be pervasive to qualify as a risk factor (since specific risk of individual firms can be diversified away at near zero cost). Also, the small-firm effect is a 'low price' phenomenon, since other measures of size, such as book value alone or the number of employees, do not predict the cross section of returns (Berk 1997). Finally, note that macro-economic factors that help predict stock returns, such as the dividend (earnings) price ratio, the term spread and the default spread, also help predict recessions, which is suggestive that a recession-type explanation may lie behind the three-factor Fama–French results using factor mimicking portfolios.

UK: Cross-Section Data

Miles and Timmermann (1996) use the Fama–MacBeth procedure on UK monthly returns from 1979 to 1991 (12 years). For each month, the cross-section regression on 457 firms is

$$R_i = \lambda_0 + \lambda_1\beta_i + \lambda_2 BMV_i + \lambda_3 \ln MV_i + \varepsilon_i \tag{12}$$

where R_i = monthly return and the independent variables book-to-market value BMV_i and market value ('size') $\ln(MV)_i$ are lagged by one year. This regression is repeated for each month, over a 12-year period (1979–1991), so the cross-section regressions give a time-series of 144 values of each cross-section parameter λ_i. They find that BMV is the key cross-section variable in explaining average returns (average $\lambda_2 = 0.35$, average t-statistic = 2.95) and the relationship is positive – that is, high BMV companies have high average returns. They find no effect on the cross section of returns from the CAPM-beta, or betas on dividend yields, P/E ratio or leverage. They find weak evidence of 'size effect' but only for the lowest size decile.

Miles and Timmermann (1996) also report a non-parametric bootstrap test of the influence of the 'factors' on average returns. This avoids having to make the linearity and normality assumptions imposed by regression tests. In April, for each of the 12 years, they *randomly* split the sample of 457 shares into 10 equally weighted portfolios. The average mean monthly return for each of the decile portfolios is then calculated for the whole 144 months, and these are ranked, giving the maximum and minimum mean returns as well as the mean return spread (= 'max' – 'min'). The procedure is repeated 5000 times, giving three separate distributions for the 'spread' 'max' and 'min' variables and their 1%, 5% and 10% critical values under the null of random returns. When actual portfolio returns are then sorted by BMV and the 'max', 'min' and 'spread' calculated, they are significant at the 1% significance level, but this generally does not apply to other accounting variables tried.

For example, the 1% critical value for the minimum (maximum) return in the bootstrap where portfolio rankings are random is 0.83% p.m. (1.50% p.m.). When sorted by BMV, the lowest (highest) decile has a mean return of 0.75% p.m. (1.54% p.m.), which implies that sorting by BMV gives a statistically significant positive relationship (not necessarily linear) between BMV and average return at a 1% significance level. Of the other 'sorting keys', only the lowest decile by 'size' has an average return lower than the 1% critical value from the bootstrap, indicating a 'size effect' only for very small firms. The results from the non-parametric approach are therefore consistent with the regression approach, namely that BMV is a key determinant of the cross section of average returns.

Alternative 'risk' factors

Alternative macro-economic variables (e.g. inflation, investment growth, labour income, consumption–wealth ratio) have also been used as factors in the APT interpretation (e.g. Chen, Roll and Ross 1986, Cochrane 1996, Lettau and Ludvigson 2001b) and often these macro-factors are priced (i.e. in the cross-section regression (11), the λ's are statistically significant), thus supporting a multifactor APT model. However, in general, these macro-variables do not explain the cross section of returns sorted on value and size, as well as do the Fama–French, SMB and HML factors. The exception here is the Lettau and Ludvigson (2001a) model where the macro-variable is $z_t = (c/w)_t \Delta c_{t+1}$. For any level of consumption growth, the factor z_t depends on a recession variable, the consumption–wealth ratio c/w, which provides a type of time-varying risk aversion. Here, the marginal utility of consumption tomorrow depends not just on consumption growth (as in the standard C-CAPM) but also on whether you are in a recession or not, that is the c/w variable. The beta on the z_t factor explains the cross-section returns of the 25 size-value sorted portfolios as well as do the SMB and HML, betas of the Fama–French three-factor model.

Clearly, the above empirical work suggests that more than one factor is important in determining the cross section of stock returns, and the work of Fama and French (1993, 1996) is currently the 'market leader' in explaining returns sorted by 'value' and 'size'.

Cross-Equation Restrictions and the APT

In the original Fama and MacBeth (1973) article, their second-stage regression consists of over 2000 individual stocks that are assigned to a limited number of 100 'portfolio' betas, but there are 2000 observations on other cross-section variables such as 'size' and BMV. Hence, the larger cross-section variation in 'size' and 'value' may bias the results against the *limited* number of portfolio β_i variables. Also, the Fama and French (1993) study uses OLS, which implicitly assumes (at any time t) that the cross-section idiosyncratic risks (i.e. error terms ε_{it} and ε_{jt}) have zero correlation. The APT implies cross-equation restrictions between the time-series and cross-section parameters that are not investigated in the Fama–MacBeth and Fama–French 'single-equation' studies. The APT can be represented:

$$R_{it} = E(R_{it}) + \sum_{j=1}^{k} b_{ij} F_{ij} + \varepsilon_{it} \tag{13}$$

where R_i are the asset returns and F_j are the factors. Expected returns are given by

$$E(R_{it}) = \lambda_0 + \lambda_1 b_{i1} + \cdots + \lambda_k b_{ik} \tag{14}$$

Hence,

$$R_{it} = \lambda_0 + \sum_{j=1}^{k} b_{ij}\lambda_j + \sum_{j=1}^{k} b_{ij} F_{jt} + \varepsilon_{it} \tag{15}$$

A regression with an unrestricted constant term α_i (and setting $\lambda_0 = r$) is

$$(R_{it} - r_t) = \alpha_i + \sum b_{ij} F_{jt} + \varepsilon_{it} \tag{16}$$

Comparing (15) and (16), we see that there are non-linear cross-equation restrictions, $\alpha_i = \sum_{j=1}^{k} b_{ij}\lambda_j$ in (16), if the APT is the correct model. In the combined time-series cross-section regressions, we can use NLSUR, which takes into account the contemporaneous covariances between the error terms (and gives more efficient estimators). Hence, the λ's and b_{ij}'s are jointly estimated in (15).

With US portfolio returns, McElroy, Burmeister and Wall (1985) find the APT restrictions do not hold. However, Clare et al. estimate (15) on UK data, imposing the APT parameter restrictions and allowing for the contemporaneous cross-section correlation of the error terms. The price of CAPM beta-risk (i.e. the λ_β coefficient on the market return beta) is found to be positive and statistically significant. In contrast, when the variance–covariance matrix of errors is restricted to be diagonal (i.e. closer to the Fama–MacBeth procedure), λ_β is not statistically significant. Also, the price of beta-risk λ_β using the NLSUR estimator is reasonably stable over time, and they find no additional explanatory power from other cross-section accounting variables such as betas on 'size', book-to-market and price–earnings ratios. It is not entirely clear why these results, using NLSUR on UK data, are so very different from the Fama–MacBeth two-step approach using US data, since all NLSUR does is improve efficiency and does not correct for any bias (errors-in-variables bias is potentially present in both approaches). Maybe it is the imposition of APT restrictions that makes a difference, or returns in the United Kingdom behave differently from those in the United States. So, it seems as if the CAPM is 'alive' in the United Kingdom but 'dead' in the United States. But this conflicting evidence is extremely puzzling.

In general, the key issues in testing and finding an acceptable empirical APT model are whether the set of factors F_{jt} and the resulting values of λ_j are constant over different sample periods and across different portfolios (e.g. sorted by 'size' and by 'value'). If the λ_j are different in different sample periods, then the price of risk for factor j is time-varying (contrary to the theory). Although there has been considerable progress in estimating and testing the APT, the empirical evidence on the above issues is far from definitive.

MCS and Bootstrap

In the previous section, there is a rather bewildering array of econometric techniques used in testing the CAPM and APT models and unfortunately space constraints limit our ability to go into these important matters in great detail – that is why we have econometrics texts, some of which do tell the truth about the limitations of these techniques in practical situations. In most econometrics texts, maximum likelihood assuming *niid* errors and a correct model specification (e.g. linear) is usually at 'the top of the tree', since it delivers asymptotically *efficient* estimates (OLS and GLS are of course special cases of ML). However, first-stage GMM with a correction to the covariance matrix of errors, for serial correlation and heteroscedasticity, (e.g. White 1980, Newey-West 1987), which does not necessarily assume a particular parameterisation of the error term or that the distribution is normal, is probably more popular in the asset pricing literature as it is thought to give estimates that are more robust to misspecification of the model. Also because many distributional results (e.g. 'the test statistic is chi-squared under the null') only apply asymptotically or exact statistics require the assumption of normality, there is increasing use of Monte Carlo simulation and bootstrapping. (The best and most accessible account of alternative estimation procedures used in asset pricing models is Cochrane (2001), Chapters 15 and 16.)

For example, to assess alternative estimation techniques when testing the CAPM, one might proceed as follows. The CAPM implies $R^e_{i,t} = \alpha_i + \beta_i R^e_{m,t} + \varepsilon_{it}$ where $R^e_{i,t} \equiv R_{i,t} - r_t$, $R^e_{m,t} \equiv R_{m,t} - r_t$ and if the CAPM is true, we expect $\alpha_i = 0$. Suppose you have results from OLS time-series regressions on 10 size-sorted portfolios using T monthly observations and hence have estimates of α_i, β_i and the residuals e_{it} for each of the 10 portfolios and hence the sample (contemporaneous) covariance matrix Σ. We can now generate artificial data under the assumption that the error terms are *iid* normal using the sample covariance matrix Σ. Then we generate the 10 size-portfolio returns (of length T) under the null that the CAPM is true: $R^e_{i,t} = 0 + \beta_i R^e_{m,t} + \varepsilon_{it}$, where the β_i are the OLS estimates. $R^e_{m,t}$ is *assumed to be normally distributed* and T values are drawn from $R^e_{m,t} \sim niid(\overline{R}^e_m, \sigma(R^e_m))$, the sample estimates, and are independent of the error term. These are the standard ML assumptions. With this artificial data, we can then estimate the CAPM, $R^e_{i,t} = \alpha_i + \beta_i R^e_{m,t} + \varepsilon_{it}$ using a variety of techniques (e.g. ML, GLS, OLS, one- or two-step GMM) and test $\alpha_i = 0$. We can repeat the above, say, 10,000 times and obtain 10,000 values of our parameters and test statistic for $\alpha_i = 0$. If the econometric technique is 'good', then we would expect to reject, the null $\alpha_i = 0$ at a 5% significance level (say) around 5% of the time (this is the 'size' of the test). Having generated artificial data on the 10 size portfolios over time (of length T), we can take sample averages and also run the cross-section regression $\overline{R}^e_i = \lambda_1 \hat{\beta}_i + v_i$ using OLS or cross-section GLS (= ML with covariance matrix $E(vv')$) or the Fama–MacBeth rolling regression. We can repeat this for our 10,000 simulations and test whether $\lambda_1 > 0$ and is equal to the excess market return.

We can also repeat all of the above but instead of drawing the error terms from a normal distribution, we can draw (randomly with replacement) from the original OLS *residuals* e_{it} and from the original sample data on $R^{e}_{m,t}$ to generate the artificial data series $R^{e}_{i,t} = 0 + \beta_i R^{e}_{m,t} + e_{it}$. In addition, if we believe the residuals are serially correlated, we can use a block bootstrap (e.g. drawing the original data $\{R^{e}_{m,t}, e_{it}\}$ in, say, blocks of length 3, to capture MA(2) errors). Our bootstrap distributions and tests of $\alpha_i = 0$ for the alternative estimation techniques will then reflect possible serial correlation, heteroscedasticity, non-normality and non-independence of the market return and residuals found in the real data set.

Cochrane (2001) does just this for 10 size-sorted NYSE portfolios with two alternative post-WWII monthly data sets of length $T = 876$ and a shorter period $T = 240$ months. He finds that for the cross-section regressions, the results of the alternative techniques are nearly identical. Now consider the time-series tests of $\alpha_i = 0$. Under the MCS assuming *niid* residuals, although GMM (with three lags) corrects for MA errors that are not there, the rejection frequency of the null are about the same as for ML, although both reject at around 6–7% level rather than the nominal size of 5%. (It is well known that GMM assuming a long-lag *unweighted* spectral density matrix rejects the null far too often and here it rejects about 25–40% of the time, for lag length 24.) However, the bootstrap results using the real data demonstrates the usefulness of this technique when the residuals may be non-normal. The ML χ^2 test has about *half* the correct size, rejecting 2.8% of the time at a 5% nominal size, while the first-stage GMM estimator (with correction for heteroscedasticity) corrects this size distortion-GMM is therefore more robust. In a later chapter, we consider further examples of the use of MCS and bootstrap techniques in examining the finite sample properties of alternative test statistics. All models are incorrect, so we should always compare the relative performance of alternative models (e.g. plausibility of assumptions, internal consistency, parsimony) as well as how they perform against the data, using a variety of techniques and tests (See *inter alia*, Cochrane 2001 and Hendry 1995 for interesting discussions of econometric testing and its relationship to economic theory, which is a much-debated topic, given that we cannot 'repeat our experiments' under (near) identical conditions, as natural scientists are able to do.)

8.4 Summary

- CAPM-betas explain the difference in average (cross-section) returns between stocks and bonds but not the spread of average returns *within* different portfolios of stocks.
- The Fama–French high minus low 'book-to-market returns' and 'size factor' largely explain the cross-section average returns across stock portfolios (for the United States and United Kingdom), sorted by size and book-to-market value. These 'factors' can be loosely interpreted as indicating that investors require high average returns to compensate for the risk caused by recessions – they represent 'distress premia'.
- The three-factor Fama–French model does not explain the cross section of returns where portfolios are sorted according to recent performance (i.e. momentum portfolios).

Appendix: Fama–MacBeth Two-Step Procedure

The Fama–MacBeth cross-section regression requires an *estimate* of the stock's beta. The CAPM may be represented by

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it} \quad \text{(A1)}$$

where R_i and R_m are the *excess* returns on the stocks and the market respectively. Fama and MacBeth (1973) have around 2000 stocks but they calculate a set of 100 *portfolio* betas to which they assign the individual stocks. Allocation of the 2000 stocks to a particular portfolio beta proceeds as follows.

(i) For each stock in turn, estimate β_i in equation (A1) using $t = 1 - 60$, monthly observations.

(ii) At $t = 61$, form 10 portfolios based on market value ('size') and then subdivide each of these portfolios into 10 further sub-samples according to the stock's estimated beta. Calculate the average monthly return on these 100 portfolios over $t = 61$ to 72 (i.e. one year). We now have 100 average returns, $\overline{R}_p$ ($p = 1, 2, \ldots, 100$) sorted by size and beta.

(iii) The above procedure is now repeated for *each year*. This gives an adequate spread in returns with which to estimate the betas. We can now either take the average betas (over time) for each of the 100 sorted portfolios or run a time-series regression for each of the 100 elements of R_p taken separately, on the market return to obtain 100 portfolio betas.

(iv) In each year, *individual* stocks are then assigned a portfolio beta based on the sorted 'size-beta' *portfolio* to which they belong. This does not imply that individual company betas are constant over time. If an individual firm switches from one of the 100 'size-beta' groups to another, then the (portfolio) beta assigned to that firm will also change.

The second stage then involves using these 100 *portfolio betas* in the *cross-section* regression (A2) for the 2000 firms. This cross-section regression is repeated for all months ($t = 1, 2, \ldots,$ T) of the sample giving a time-series for $\lambda_0, \lambda_1, \gamma$, which can be analysed as indicated in the text.

$$\overline{R}_i = \lambda_0 + \lambda_1 \hat{\beta}_{pi} + \gamma Z_i + v_i \quad \text{(A2)}$$

where Z_i is any cross-section company variable.