Technical Appendix

1 Estimating the CAPM β using GARCH

The CAPM (treating the risk free rate as a constant) implies a relationship between return on a stock R_i and the market return R_M

$$E(R_i) = \alpha + \beta E(R_M)$$

where

$$\beta = \frac{Cov(R_i, R_M)}{Var(R_M)}$$

Given a time series of stock and market returns we can estimate β from a least squares regression of $R_{i,t}$ on $R_{M,t}$ ie

$$R_{i,t} = \alpha + \beta R_{M,t} + \varepsilon_{i,t}$$

This is usually done on a rolling window of two to five years of data (which can be daily, weekly, monthly or even quarterly). This allows for some time variation in β. From the formula for β we see that in theory any such time variation must arise from time varying covariances or variances, or both, in the underlying data.

Rather than using a rolling window an alternative is to model the time variation explicitly via a GARCH model (see for example "Good News, Bad News, Volatility, and Betas" P A Braun D B Nelson and A M Saunier *Journal of Finance*, Vol. 50, No. 5 (Dec., 1995), pp. 1575-1603).

We specify the following joint model for market and individual stock returns

$$\begin{pmatrix} R_{M,t} \\ R_{i,t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

but allow for time varying volatility and covariance as follows

$$Cov_t \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} = \begin{pmatrix} \sigma^2_{11,t} & \sigma_{12,t} \\ \sigma_{21,t} & \sigma^2_{22,t} \end{pmatrix}$$

A variety of different specifications have been proposed for the conditional variance process. Two practical difficulties with these models are first to ensure positive definiteness of the conditional covariance matrix and also the number of parameters can easily grow very large indeed causing computational issues.

The BEKK (Baba, Engle, Kraft and Kroner (1990) published as "Multivariate Simultaneous Generalized ARCH" by R F Engle and K F Kroner, *Econometric Theory*, Volume 11, Issue 1 February 1995 , pp. 122-150) provides a simple tractable model that ensures positive definiteness of the covariance matrix. The first order diagonal BEKK model for the two variable case, as above, is as follows.

$$\begin{pmatrix} \sigma^2_{11,t} & \sigma_{12,t} \\ \sigma_{21,t} & \sigma^2_{22,t} \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} +$$
$$+ \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix} \begin{pmatrix} u_{1t-1} \\ u_{2t-1} \end{pmatrix} \begin{pmatrix} u_{1t-1} & u_{2t-1} \end{pmatrix} \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix}$$
$$+ \begin{pmatrix} b_{11} & 0 \\ 0 & b_{22} \end{pmatrix} \begin{pmatrix} \sigma^2_{11,t-1} & \sigma_{21,t-1} \\ \sigma_{12,t-1} & \sigma^2_{22,t-1} \end{pmatrix} \begin{pmatrix} b_{11} & 0 \\ 0 & b_{22} \end{pmatrix}$$

where the returns have constant (conditional) means and time varying conditional variance and covariances.

In detail for the BEKK model we have the following equations for the conditional evolution of $Var(R_M)$ and $Cov(R_i, R_M)$

$$\begin{aligned} Var(R_{M,t}) &= \sigma^2_{11,t} = m_{11} + a^2_{11} u^2_{1t-1} + b^2_{11} \sigma^2_{11,t-1} \\ Cov(R_{i,t}, R_{M,t}) &= \sigma_{12,t} = m_{21} + a_{11} a_{22} u_{1t-1} u_{2t-1} + b_{11} b_{22} \sigma_{12,t-1} \end{aligned}$$

The implied long run (unconditional) variance and covariance are then (noting that the unconditional expectations are given by $E\left(u^2_{1t-1}\right) = Var(R_M)$ and $E(u_{1t-1} u_{2t-1}) = Cov(R_i, R_M)$)

$$\begin{aligned} Var(R_M) &= m_{11} / \left(1 - a^2_{11} - b^2_{11}\right) \\ Cov(R_i, R_M) &= m_{21} / (1 - a_{11} a_{22} - b_{11} b_{22}) \end{aligned}$$

which can then used to calculate a long-run, or unconditional β as the ratio of the unconditional covariance to the unconditional variance. In long-horizon forecasting this is the appropriate measure of β, as long as the GARCH process is both stable and converges reasonably rapidly.

There are a number of alternative measures of this long-run β. The first calculates the ratio of implied long-run covariance and variance, using estimated parameters from the model.

$$\hat{\beta}_{LR} = \frac{\hat{m}_{21} / \left(1 - \hat{a}_{11} \hat{a}_{22} - \hat{b}_{11} \hat{b}_{22}\right)}{\hat{m}_{11} / \left(1 - \hat{a}^2_{11} - \hat{b}^2_{11}\right)}$$

An alternative is to use the estimated values $\hat{\sigma}^2_{11,t}$ and $\hat{\sigma}^2_{12,t}$. The time series average of each of these should converge to their long run values ie $Var(R_M)$ and $Cov(R_i, R_M)$ respectively. So one can also estimate β as

$$\hat{\beta}_{avs} = \frac{\frac{1}{T} \sum \hat{\sigma}_{12,t}}{\frac{1}{T} \sum \hat{\sigma}^2_{11,t}}$$

Finally one can calculate short run conditional βs as

$$\hat{\beta}_{SR,t} = \frac{\hat{\sigma}_{12,t}}{\hat{\sigma}^2_{11,t}}$$

(The moving average of this series seems to track the rolling least squares reasonably well).

One could then estimate the long run β as a simple average of these short runs βs ie

$$\hat{\beta}_{SR} = \frac{1}{T}\sum_t \hat{\beta}_{SR,t} = \frac{1}{T}\sum_t \left(\hat{\sigma}_{12,t}/\hat{\sigma}^2_{11,t}\right)$$

This gives three methods to estimate β from the GARCH-BEKK specification.

Firstly note that for $\hat{\beta}_{LR}$ even if we have unbiased estimates of the coefficients $(m_{11}, m_{12}, m_{22}, a_{11}, a_{22}, b_{11}, b_{22})$ we do not necessarily get unbiased estimates of $Cov\left(R_i, R_M\right)$ and $Var\left(R_M\right)$ by plugging in these estimates.

To see this note

$$\begin{aligned} E\left[\hat{m}_{21}/\left(1-\hat{a}_{11}\hat{a}_{22}-\hat{b}_{11}\hat{b}_{22}\right)\right] \neq & E\left(\hat{m}_{21}\right)/\left(1-E\left(\hat{a}_{11}\right)E\left(\hat{a}_{22}\right)-E\left(\hat{b}_{11}\right)E\left(\hat{b}_{22}\right)\right) \\ = & m_{21}/\left(1-a_{11}a_{22}-b_{11}b_{22}\right) = Cov\left(R_i, R_M\right) \end{aligned}$$

and similarly for the denominator. We would however at least have have

$$plim\left(\hat{\beta}_{LR}\right) = \beta$$

as long as we have consistent estimates of $(m_{11}, m_{12}, m_{22}, a_{11}, a_{22}, b_{11}, b_{22})$ and $a^2_{11} + b^2_{11} \neq 1$ and $a_{11}a_{22} - b_{11}b_{22} \neq 1$.

If the GARCH specification implies stationary processes for the conditional moments $\sigma^2_{12,t}$ and $\sigma^2_{11,t}$ then time series averages of the estimated processes will converge in probability to the unconditional (long run) parameter (via Weak Law of Large Numbers). So again we will have

$$plim\left(\hat{\beta}_{avs}\right) = \beta$$

although if we are fairly close to (integrated) I-GARCH this convergence may be quite slow

Finally note also that for $\hat{\beta}_{SR,t}$ since

$$E\left(\frac{\hat{\sigma}_{12,t}}{\hat{\sigma}^2_{11,t}}\right) \neq \frac{E\left(\hat{\sigma}_{12,t}\right)}{E\left(\hat{\sigma}^2_{11,t}\right)} = \frac{\sigma_{12}}{\sigma^2_{11}} = \beta$$

an average of the short run $\hat{\beta}_{SR}$ will again not be unbiased for β. Here the direction of the bias is determined by the interaction of two terms, since

$$E\left(\frac{\hat{\sigma}_{12,t}}{\hat{\sigma}^2_{11,t}}\right) = E\left(\hat{\sigma}_{12,t}\right)E\left(\frac{1}{\hat{\sigma}^2_{11,t}}\right) + cov\left(\hat{\sigma}_{12,t}, \frac{1}{\hat{\sigma}^2_{11,t}}\right)$$

The first term will be larger than $\frac{E(\hat{\sigma}_{12,t})}{E\left(\hat{\sigma}^2_{11,t}\right)}$ due to the convexity of the function $f\left(X\right) = \frac{1}{X}$.The second term is however negative in the data used here since conditional covariances and covariances are correlated over time.

2 Possible Extensions

In principle it might interesting to investigate whether results are different for alternative multivariate GARCH estimation techniques.

Given the strong evidence of shared properties in the two stocks we consider, there may be some gains in efficiency from a joint estimation procedure in which both are modelled in the same system (at present we actually have two alternative models for the market return, although in practice results are virtually identical).

One alternative would be the constant conditional correlation model of Bollerslev (Bollerslev, T. (1990) "Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model" *Review of Economics and Statistics* , 72, 498–505). Here the conditional correlation is assumed to be constant while the conditional variances are varying. Given the quite strong evidence that correlations rise during times of high volatility (which we also see evidence of in our dataset) this seems unlikely to be suitable but may provide a useful cross-check.

Another alternative would be the Dynamic Conditional Correlation (DCC) model proposed by Engle in 2002 (Engle, R.F.: "Dynamic Conditional Correlation – A Simple Class of Multivariate GARCH Models" *Journal of Business and Economic Statistics* 20(3), 339–350 (2002)) which reduces the number of parameters relative to a BEKK model. The shortcoming of this model is that all conditional correlations follow the same dynamic structure which could be seen as too restrictive.

Finally, while we have followed standard practice in treating the risk-free rate as a constant in estimation, this is clearly not descriptively accurate: in principle CAPM regressions should be specified in terms of excess returns. While this simplification is likely to have vanishingly small effects at high frequencies, it is possible that at monthly and quarterly frequencies the effects may be of at least some consequence.