Lecture 8

Stephen G. Hall

ARCH and GARCH

REFS

A thorough introduction
'ARCH Models' Bollerslev T, Engle R F and Nelson D B Handbook of Econometrics vol 4. or UCSD Discussion paper no 93.49. (available on my web site)

A quick survey
Cuthbertson Hall and Taylor

Until the early 80s econometrics had focused almost solely on modelling the means of series, ie their actual values. Recently however we have focused increasingly on the importance of volatility, its determinates and its effects on mean values.

A key distinction is between the conditional and unconditional variance.

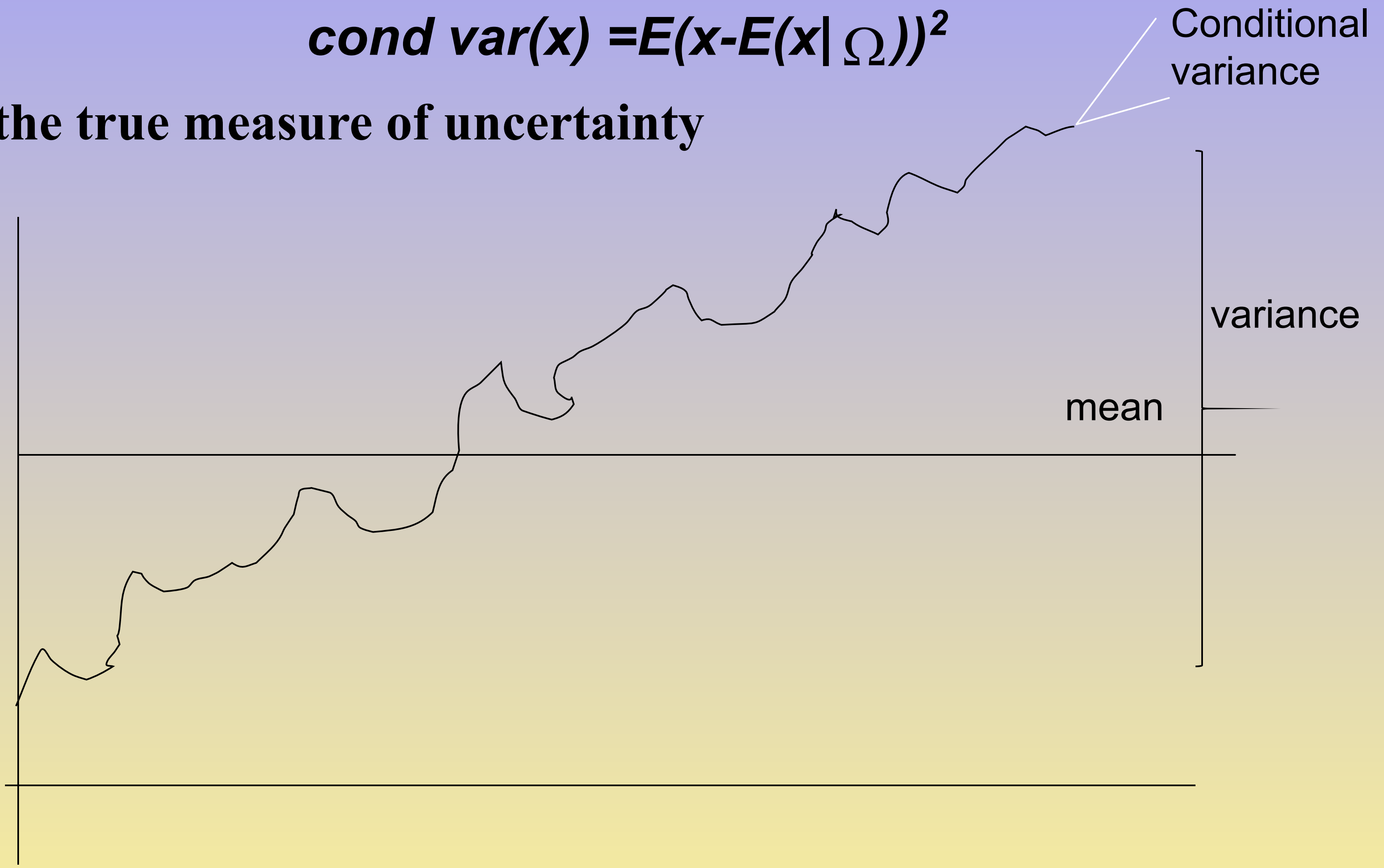
the unconditional variance is just the standard measure of the variance

$$var(x) = E(x - E(x))^2$$

the conditional variance is the measure of our uncertainty about a variable given a model and an information set.

$$cond\ var(x) = E(x - E(x|\Omega))^2$$

this is the true measure of uncertainty



Stylised Facts of asset returns

i) Thick tails, they tend to be leptokurtic

ii)Volatility clustering, Mandelbrot, ‘large changes tend to be followed by large changes of either sign’

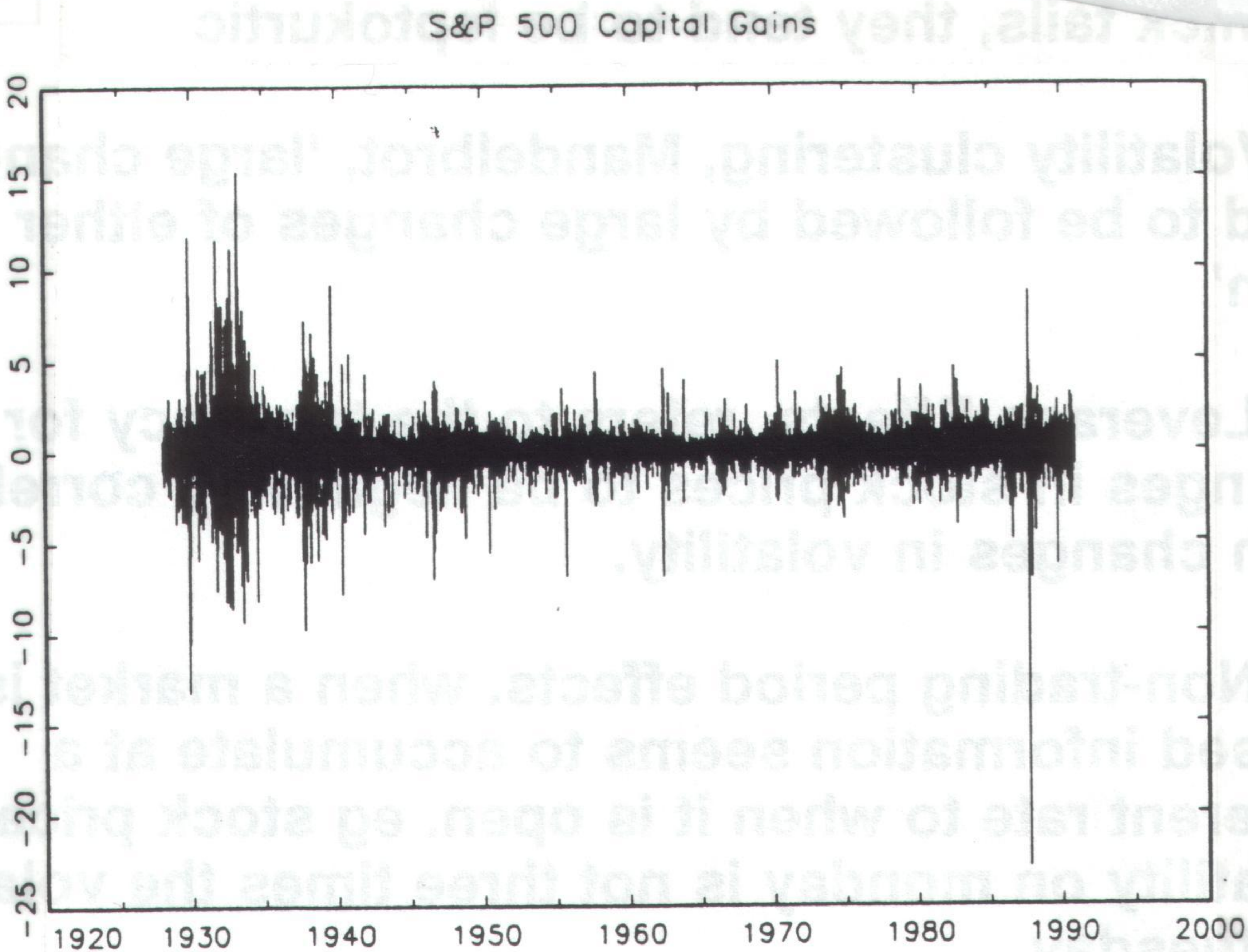
iii)Leverage Effects, refers to the tendency for changes in stock prices to be negatively correlated with changes in volatility.

iv)Non-trading period effects. when a market is closed information seems to accumulate at a different rate to when it is open. eg stock price volatility on Monday is not three times the volatility on Tuesday.

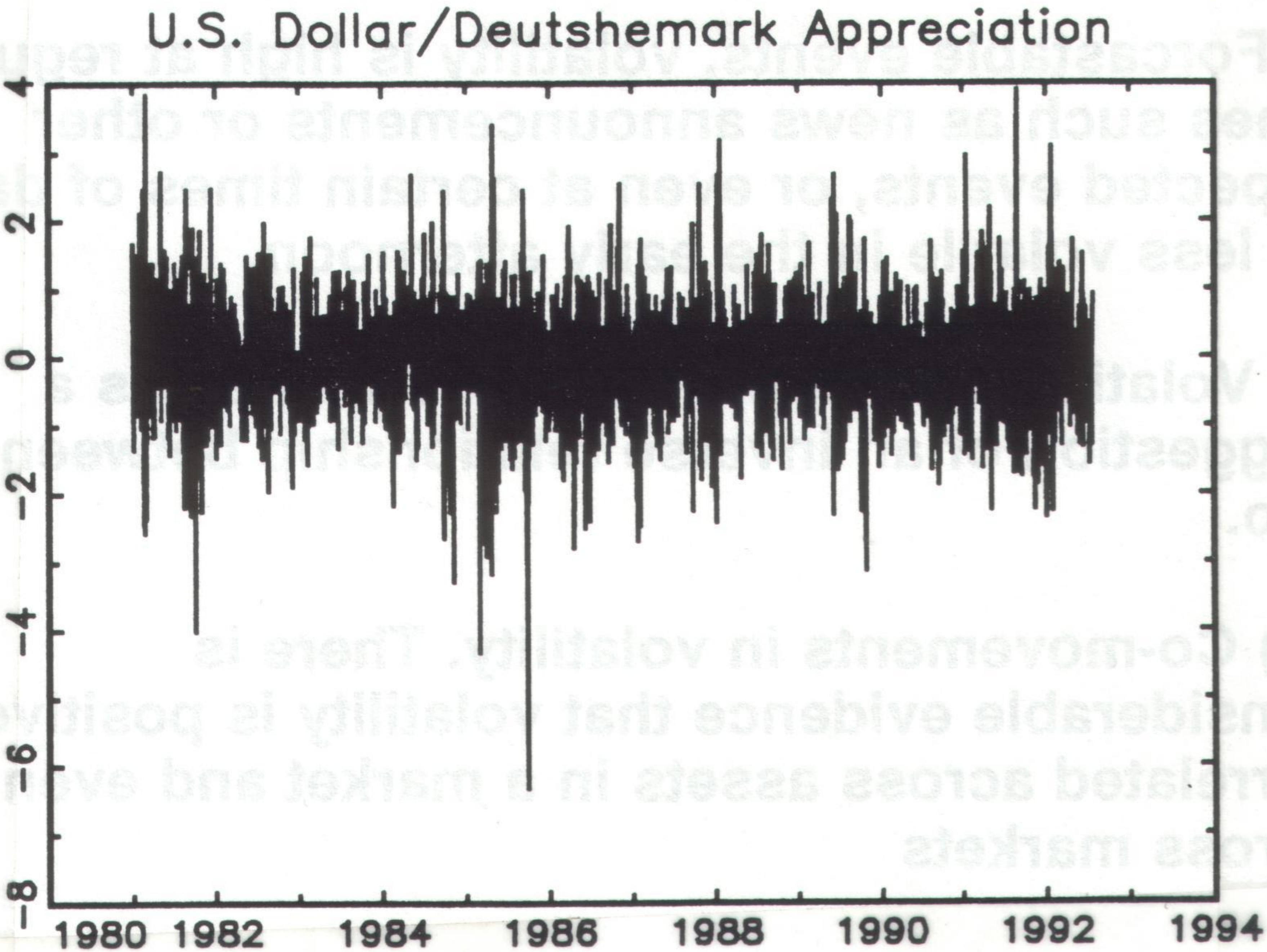
v) Forcastable events, volatility is high at regular times such as news announcements or other expected events, or even at certain times of day, eg less volatile in the early afternoon.

vi) Volatility and serial correlation. There is a suggestion of an inverse relationship between the two.

vii) Co-movements in volatility. There is considerable evidence that volatility is positively correlated across assets in a market and even across markets



S&P 500 Capital Gains
20
15
10
5
0
-5
-10
-15
-20
-25
1920
1930
1940
1950
1960
1970
1980
1990
2000


U.S. Dollar/Deutshemark Appreciation
4
2
0
-2
-4
-6
-8
1980
1982
1984
1986
1988
1990
1992
1994

Engle(1982) ARCH Model

Auto-Regressive Conditional Heteroscedasticity

$$Y_t = \beta X_t + \varepsilon_t \qquad \varepsilon_t \sim N(0, \sigma_t^{\,2})$$

$$\sigma_t^{\,2} = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon^2{}_{t-i} = \omega + \alpha(L)\varepsilon^2$$

$$\text{define} \qquad \nu_t \equiv \varepsilon_t^{\,2} - \sigma_t^{\,2}$$

$$\varepsilon_t^{\,2} = \omega + \alpha(L)\,\varepsilon^2 + \nu_t$$

an AR(q) model for squared innovations.

note as we are dealing with a variance

$$\omega > 0 \quad \alpha_i > 0 \; all \; i$$

even though the errors may be serially uncorrelated they are not independent, there will be volatility clustering and fat tails.

if the standardised residuals

$$z_t = \varepsilon_t / \sigma_t$$

are normal then the fourth moment for an ARCH(1) is

$$E(\varepsilon^4{}_t) / E(\varepsilon^2{}_t)^2 = 3(1-\alpha^2)/(1-3\alpha^2) \quad if\, 3\alpha^2 < 1$$

GARCH (Bollerslev(1986))

In empirical work with ARCH models high q is often required, a more parsimonious representation is the Generalised ARCH model

$$\sigma^2{}_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon^2{}_{t-i} + \sum_{j=1}^{p} \beta_j \sigma^2{}_{t-j}$$

$$= \omega + \alpha(L)\varepsilon^2 + \beta(L)\sigma^2$$

$$define \quad \upsilon_t \equiv \varepsilon^2{}_t - \sigma^2{}_t$$

$$\varepsilon^2{}_t = \omega + (\alpha(L) + \beta(L))\varepsilon^2 + \beta(L)\upsilon \;\; + \upsilon_t$$

which is an ARMA(max(p,q),p) model for the squared innovations.

This is covariance stationary if all the roots of

$$\alpha(L) + \beta(L) = 1$$

lie outside the unit circle, this often amounts to

$$\alpha(1) + \beta(1) < 1$$

If this becomes an equality then we have an Integrated GARCH model (IGARCH)

Nelsons’ EGARCH model

this captures both size and sign effects in a non-linear formulation

$$\log(\sigma^2{}_t) = \omega + \sum_{i=1}^{q} \alpha_i (z_{t-i} + \gamma(|\, z_{t-i}\,| - E\,|\, z_{t-i}\,|)) + \sum_{j=1}^{p} \beta_j \log(\sigma^2{}_{t-j})$$

Non-linear ARCH model NARCH

this then makes the variance depend on both the size and the sign of the variance which helps to capture leverage type effects.

$$\sigma^{\gamma}{}_{t} = \omega + \sum_{i=1}^{q} \alpha_i \, | \varepsilon_{t-i} - \kappa |^{\gamma} + \sum_{j=1}^{p} \beta_j \sigma^{\gamma}{}_{t-j}$$

Threshold ARCH (TARCH)

Large events to have an effect but no effect from small events

$$\sigma^2{}_t = \omega + \sum_{i=1}^{q} (\alpha^+{}_i I(\varepsilon_{t-i} > 0) + \alpha^-{}_i (\varepsilon_{t-i} < 0))\varepsilon^2{}_{t-i}) + \sum_{j=1}^{p} \beta_j \sigma^2{}_{t-j}$$

Many other versions are possible by adding minor asymmetries or non-linearities in a variety of ways.

All of these are simply estimated by maximum likelihood using the same basic likelihood function, assuming normality,

$$\log(L) = \sum_{i=1}^{T} (-\log(\sigma^2{}_t) - \varepsilon^2{}_t / \sigma^2{}_t)$$

ARCH in MEAN (G)ARCH-M

Many classic areas of finance suggest that the mean of a relationship will be affected by the volatility or uncertainty of a series. Engle Lilien and Robins(1987) allow for this explicitly using an ARCH framework.

$$y_t = \beta x_t + \delta \sigma^2{}_t + \varepsilon_t$$

$$\sigma^2{}_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon^2{}_{t-i} + \sum_{j=1}^{p} \beta_j \sigma^2{}_{t-j}$$

typically either the variance or the standard deviation are included in the mean relationship.

often finance stresses the importance of covariance terms. The above model can handle this if y is a vector and we interpret the variance term as a complete covariance matrix. The whole analysis carries over into a system framework

Non normality assumptions

While the basic GARCH model allows a certain amount of leptokurtic behaviour this is often insufficient to explain real world data. Some authors therefore assume a range of distributions other than normality which help to allow for the fat tails in the distribution.

t Distribution

The t distribution has a degrees of freedom parameter which allows greater kurtosis. The t likelihood function is

$$l_t = \ln(\Gamma(0.5(v+1))\Gamma(0.5v)^{-1}(v-2)^{-1/2}(1+z_t(v-2)^{-1})^{-(v+1)/2}) - 0.5\ln(\sigma^2{}_t)$$

***where* F *is the gamma function and v is the degrees of freedom* as $v \rightarrow \infty$ this tends to the normal distribution**

IGARCH.
The standard GARCH model

$$\sigma^2{}_t = \omega + \alpha(L)\varepsilon^2 + \beta(L)\sigma^2$$

is covariance stationary if

$$\alpha(1) + \beta(1) < 1$$

But Strict stationarity does not require such a stringent restriction (That is that the unconditional variance does not depend on t),in fact we often find in estimation that

$$\alpha(1) + \beta(1) = 1$$

this is then termed an Integrated GARCH model (IGARCH), Nelson has established that as this satisfies the requirement for strict stationarity it is a well defined model.

However we may suspect that IGARCH is more a product of omitted structural breaks than the result of true IGARCH behavior.

Multivariate Models

In general the Garch modelling framework may be easily extended to a multivariate framework where

$$E_t(\varepsilon_t \varepsilon'_t) = \Omega_t$$

however there are some practical problems in the choice of the parameterisation of the variance process.

A direct extension of the GARCH model would involve a very large number of parameters.

The conditional variance could easily become negative even when all the parameters are positive.

The chosen parameterisation should allow causality between variances.

Vector ARCH

let vech denote the matrix stacking operation

$$vech\begin{pmatrix} a & b \\ b & d \end{pmatrix} = (a\, b\, d)$$

a general extension of the GARCH model would then be

$$vech(\Omega_t) = W + A(L)vech(\varepsilon_{t-1}\varepsilon'_{t-1}) + B(L)vech(\Omega_{t-1})$$

this quickly produces huge numbers of parameters, for p=q=1 and n=5 there are 465 parameters to estimate here.

One simplification used is the Diagonal GARCH model where A and B are taken to be diagonal, but this assumes away causality in variances and co-persistence. We need still further complex restrictions to ensure positive definiteness in the covariance matrix.

A more tractable alternative is to state

$$\Omega_t = V'V + \sum_{i=1}^{q} A'_i \, \varepsilon_{t-i}\varepsilon'_{t-i} \, A_i + \sum_{j=1}^{p} B'_j \, \Omega_{t-j} B_j$$

we can further reduce the parameterisation by making A and B diagonal.

Factor ARCH

Suppose a vector of N series has a common factor structure. Such as;

$$y_t = B\xi_t + \varepsilon_t$$

where ξ **are the common factors and**

$$\varepsilon \sim N(0, \Psi) \qquad E(\xi'_t \xi) = \Lambda_t$$

then the conditional covariance matrix of y is given by

$$Cov_{t-1}(y_t) = \Omega_t = \Psi_t + B\Lambda_t B'$$

Or

$$\Omega_t = \Psi_t + \sum_{i=1}^{k} \beta_i \beta'_i \xi_k$$

So given a set of factors we may estimate a parsimonious model for the covariance matrix once we have parameterized

$$\xi$$

One assumption is that we observe a set of factors which cause the variance, then we can simply use these. E.G. GDP, interest rates, exchange rates, etc.

another assumption is that each factor has a univariate GARCH representation.

$$\Omega_t = \Psi + \sum_{k=1}^{K} \alpha_k (\beta_k \varepsilon_{t-1} \varepsilon'_{t-1} \beta'_k) + \sum_{k=1}^{K} \gamma_k (\beta_k \Omega_{t-1} \beta'_k)$$

6

A capital asset pricing model with time-varying betas: some results from the London Stock Exchange

S. G. HALL, D. K. MILES AND M. P. TAYLOR

$$E(R_i) - r = \beta_i\{E(R_m) - r\} \tag{6.1}$$

$$\beta_i = \frac{\operatorname{cov}(R_i, R_m)}{\operatorname{var}(R_m)} \tag{6.2}$$

where R_i and R_m are the one-period returns on the asset and the market portfolio respectively and r is the one-period safe rate of return. A natural extension of the model is to condition the moments of 6.1 and 6.2 on information available to agents at the end of period $t - 1$ when formulating required returns during period t:

$$E(R_{it}|\Omega_{t-1}) - r_{t-1} = \beta_{it}\{E(R_{mt}|\Omega_{t-1}) - r_{t-1}\} \tag{6.3}$$

$$\beta_{it} = \frac{\operatorname{cov}(R_{it}, R_{mt}|\Omega_{t-1})}{\operatorname{var}(R_{mt}|\Omega_{t-1})} \tag{6.4}$$

$$\lambda = \frac{E(R_{mt}|\Omega_{t-1}) - r_{t-1}}{\mathrm{var}(R_{mt}|\Omega_{t-1})} \tag{6.5}$$

is a constant, we have

$$E(R_{mt}|\Omega_{t-1}) - r_{t-1} = \lambda \,\mathrm{var}(R_{mt}|\Omega_{t-1}) \tag{6.6}$$

Using 6.3, 6.4 and 6.6, we can write

$$R_{it} = r_{t-1} + \lambda \operatorname{cov}(R_{it}, R_{mt} | \Omega_{t-1}) + \varepsilon_{it}$$

$$R_{mt} = r_{t-1} + \lambda \operatorname{var}(R_{mt} | \Omega_{t-1}) + v_t$$

$$\varepsilon_{it} = R_{it} - E(R_{it}|\Omega_{t-1})$$

$$v_t = R_{mt} - E(R_{mt}|\Omega_{t-1})$$

is clear from 6.9 and 6.10 that the relevant conditional seco
re themselves equal to the forecast error variances and covarian

$$\operatorname{var}(R_{mt}|\Omega_{t-1}) = E(v_t^2|\Omega_{t-1})$$

$$\operatorname{cov}(R_{it}, R_{mt}|\Omega_{t-1}) = E(\varepsilon_{it} v_t|\Omega_{t-1})$$

A straightforward multivariate extension of the ARCH-M model can be applied to the CAPM formulation of the previous section as follows. We establish the following notation:

$$R_t = (R_{it} - R_{mt})'$$

$$\omega_t = (\varepsilon_t \ v_t)'$$

$$\iota = (1 \ 1)'$$

$$e = (0 \ 1)'$$

$$H_t = \begin{bmatrix} \mathrm{var}(R_{it}|\Omega_{t-1}) & \mathrm{cov}(R_{it}, R_{mt}|\Omega_{t-1}) \\ \mathrm{cov}(R_{it}, R_{mt}|\Omega_{t-1}) & \mathrm{var}(R_{mt}|\Omega_{t-1}) \end{bmatrix}$$

Then the ARCH-M formulation of 6.7, 6.8, 6.11 and 6.12 is

$$R_t = r_{t-1}\iota + \lambda H_t e + \omega_t$$

$$\mathrm{vech}(H_t) = A_0 + \sum_{i=1}^{n} A_i \ \mathrm{vech}(\omega_{t-i}\omega'_{t-i})$$

A further extension of the ARCH formulation, which imposes smoother behaviour on the conditional second moments, has been suggested by Bollerslev (1986). In Bollerslev's GARCH formulation, the conditional second moments are functions of their own lagged values as well as the squares and cross-products of lagged forecast errors. Thus, for example, the GARCH-M(n, p) formulation of the above model would consist of 6.13 and

$$\operatorname{vech}(H_t) = A_0 + \sum_{i=1}^{n} A_i \operatorname{vech}(\omega_{t-i}\omega'_{t-i}) + \sum_{i=1}^{p} B_i \operatorname{vech}(H_{t-i}) \quad (6.15)$$

Stacking all the parameters of the system into a single vector

$$\mu = (\lambda, (A_0)', \text{vech}\,(A_1)', \ldots, \text{vech}\,(A_n)', \text{vech}\,(B_1)', \ldots, \text{vech}\,(B_p)')'$$

and applying Schweppe's (1965) prediction error decomposition form of the likelihood function, we obtain the log likelihood for a sample of T observations (conditional on initial values) as

$$L(\mu) = \sum_{t=1}^{T} \log|H_t(\mu)| - \sum_{t=1}^{T} \omega_t' H_t^{-1}(\mu)\omega_t \qquad (6.16)$$

Table 6.1 ARCH-M estimates of the four sectors

	Chemical		Electrical		Mechanical		Financial	
λ	5.0	(7.8)	4.9	(7.6)	3.1	(4.4)	5.1	(7.8)
$A_0(\varepsilon_t, \varepsilon_t)$	0.0033	(18.5)	0.0042	(23.8)	0.0002	(0.9)	0.0028	(16.3)
$A_0(\varepsilon_t, v_t)$	0.0022	(8.6)	0.0007	(10.5)	0.0001	(0.3)	0.00044	(7.5)
$A_0(v_t, v_t)$	0.0023	(18.6)	0.0024	(23.5)	0.0005	(6.4)	0.0022	(15.8)
A_1	0.047	(1.4)	0.0000001	(0.00)	0.985	(6.0)	0.06	(1.7)
SK_1	0.39		0.09		−0.33		−0.01	
$KURT_1$	0.35		0.18		0.24		0.66	
BJ_1	4.64		0.44		3.1		2.6	
$LB(1)_1$	0.19		1.8		1.9		0.04	
$LB(2)_1$	1.05		1.8		2.4		1.11	
$LB(4)_1$	1.54		3.6		5.5		1.98	
$LB(8)_1$	8.00		6.8		18.1		4.5	
$LB(16)_1$	12.73		18.6		21.6		8.23	
SK_2	−0.2		−0.19		−0.19		−0.2	
$KURT_2$	0.32		0.32		0.24		0.32	
BJ_2	1.63		1.68		1.34		1.63	
$LB(1)_2$	0.32		0.33		0.31		0.32	
$LB(2)_2$	0.70		0.69		0.81		0.71	
$LB(4)_2$	3.05		3.02		3.08		3.06	
$LB(8)_2$	9.62		9.67		8.34		9.59	
$LB(16)_2$	14.59		14.58		13.49		14.6	

When we allow unrestricted parameter matrices, the question of the positive semi-definiteness of H_t arises. This is discussed by Baba *et al.* (1987), who suggest the following restriction to impose positive semi-definiteness on the model:

$$H_t = A_0' A_0 + A_1' \omega_{t-1} \omega_{t-1}' A_1 + B_1' H_{t-1} B_1$$

where A_0, A_1 and B_1 are symmetric-parameter matrices. In our case A_1 and B_1 are scalars rather than vectors, so that we only need restrict A_0.

Estimation of this five-equation (four sectors and the market index) model by maximum likelihood then gives the following set of parameter estimates:

$$\hat{\lambda} = 3.24 \qquad (3.1)$$

$$\hat{A}_1 = 0.027 \qquad (4.1)$$

$$\hat{B}_1 = 0.956 \qquad (119.6)$$

$$\hat{A}_0 =$$

$$\begin{bmatrix} -0.006\ (7.5) & & & & \\ 0.001\ (0.9) & 0.005\ (4.0) & & & \\ -0.0004\ (0.1) & 0.0003\ (0.4) & 0.005\ (8.4) & & \\ -0.0009\ (0.5) & 0.000006\ (0.0) & -0.0005\ (1.1) & 0.003\ (2.6) & \\ -0.001\ (22.2) & 0.0002\ (0.9) & 0.0003\ (1.7) & -0.02\ (1.0) & -0.002\ (4.6) \end{bmatrix}$$

… for the restricted model are

$$\hat{\lambda} = 3.25\ (3.0)$$

$$\hat{A}_1 = 0.031\ (4.8)$$

$$\hat{B}_1 = 0.955\ (142.0)$$

$$\hat{A}_0 = \begin{bmatrix} 0.004\ (5.2) & & & & \\ & 0.002\ (4.4) & & & \\ & & 0.006\ (8.8) & & \\ & & & 0.005\ (5.5) & \\ & & & & 0.005\ (8.1) \end{bmatrix}$$

where the asymptotic t ratios are in parentheses. The diagnostics for the model residuals are given in Table 6.3.

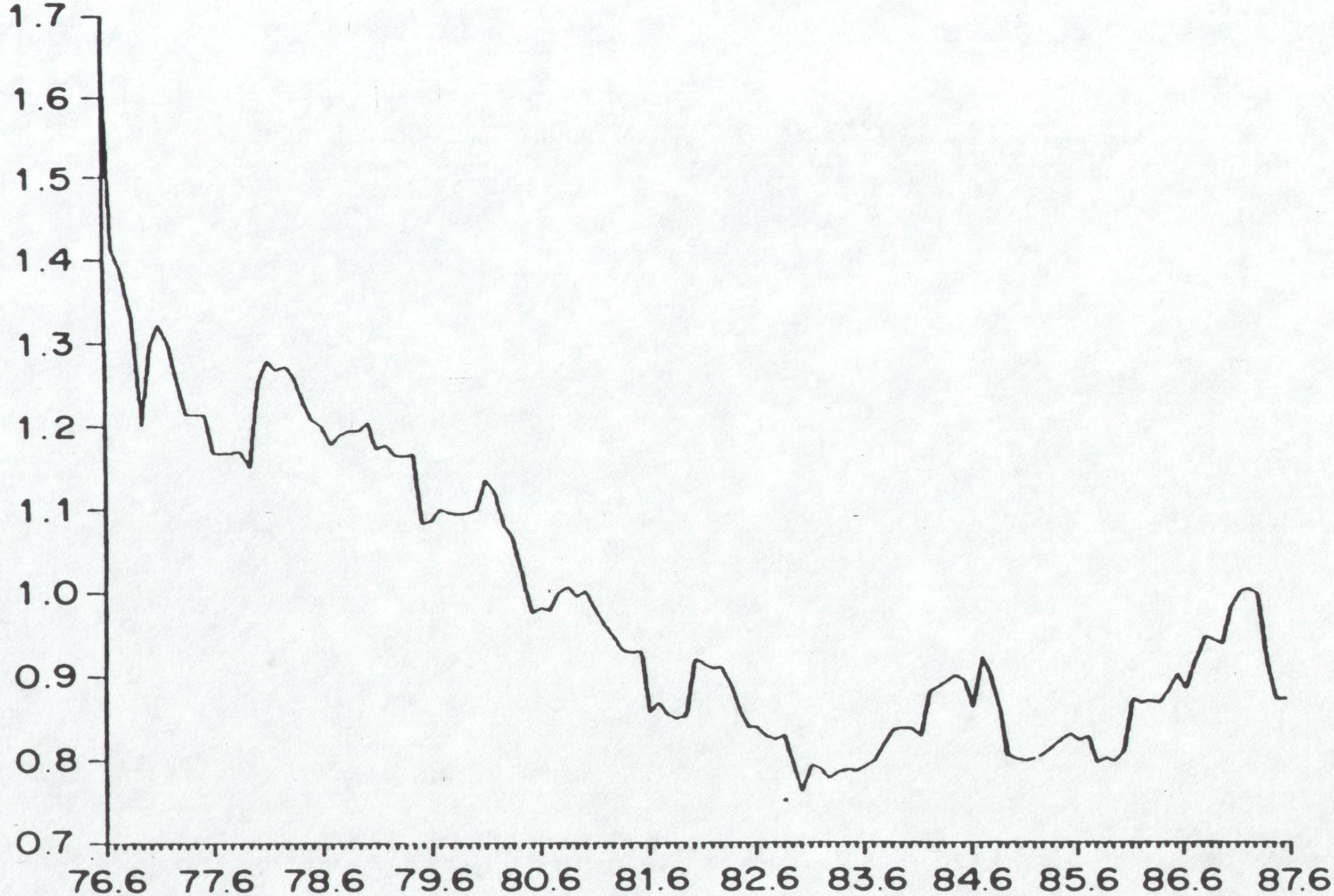


Fig. 6.4 Financial sector beta.