*Notes Section 11*

**ARCH: Modelling volatility**

Consider the $AR(p)$ model:

$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + u_t$$

with $u$ i.i.d. Covariance stationarity requires that the $AR$ polynomial has roots outside the unit circle. The optimal linear forecast of the level of $y_t$, given covariance stationarity, is:

$$E\left(\hat{y}_t | y_{t-1}, y_{t-2}, ...\right) = c + \sum \hat{\phi}_i y_{t-i}$$

So that the conditional mean is changing as the process evolves. However, given covariance stationarity, the unconditional mean is constant:

$$Ey_t = c\left(1 - \phi_1 - \phi_2 - ... - \phi_p\right)^{-1}$$

What if we wanted to forecast the variance of the series, rather than its mean? We consider $u_t$ as a process with a fixed unconditional variance, $\sigma^2$. But the conditional variance of $y_t$ could change over time. If it followed a systematic pattern, we might have something like:

$$u_t^2 = \eta + \sum_{i=1}^{m} \alpha_i u_{t-i}^2 + \omega_t$$

where $\omega$ is an i.i.d. error process. This law of motion implies that:

$$E\left(u_t^2 | u_{t-1}^2, u_{t-2}^2, ...\right) = \eta + \sum_{i=1}^{m} \alpha_i u_{t-i}^2$$

which is then the $m^{th}-$order model of Autoregressive Conditional Heteroskedasticity ($ARCH$), as proposed by Engle (1982). This conditional expectation must be non–negative for all realizations of the $u_t$ process; a necessary condition is that $\alpha_i > 0$ for all $i$. For $u_t^2$ itself to be covariance stationary, we further require that all the roots of the $\alpha$ polynomial lie outside the unit circle. If the $\alpha_i$ are all non–negative, this condition may be written as $\sum \alpha_i < 1$.

An alternative way to write such a model is in the form:

$$u_t = \sqrt{h_t} v_t$$

where $v_t$ is distributed with mean zero and variance of unity. If we then write:

$$h_t = \eta + \sum_{i=1}^{m} \alpha_i u_{t-i}^2$$

then the conditional expectation gives us the same expression in terms of $\eta$ and the $\alpha_i$ terms.

This $ARCH$ model may then be used to augment a regression equation, as a way of modelling the conditional variance of that equation's errors. The presence of $ARCH$ effects (detected, for instance, by the Lagrange Multiplier test for $ARCH$ :`archlm` in Stata) does not invalidate the use of OLS to estimate the equation, but if there are systematic movements in the conditional variance, we might want to be able to model them jointly with the level of the series. The "mean equation" and the $ARCH$ equation for the conditional variance may be jointly estimated in a maximum likelihood context. Various solutions have been proposed to deal with the non–negativity constraints, which are quite difficult to impose in a ML estimation procedure.

Non–Gaussian distributions may also be used: to cope with the stylized facts of excess kurtosis in asset returns, it may be desirable to allow for this in the model. $ARCH$ models have often been fit using a $t$ distribution, where an addditional parameter: the degrees of freedom—is estimated in the process. An even more general solution relies upon the Generalized Error Distribution (GED), which encompasses both the Normal and $t$ distributions as special cases, allowing for both excess and less–than–normal kurtosis.

**The GARCH model of Bollerslev**

The primary extension of the $ARCH(m)$ methodology of Engle is the Generalized $ARCH$, or $GARCH(r,m)$ model of Bollerslev (1986). The extension from $ARCH$ to $GARCH$ considers:

$$h_t = \eta + \Pi(L)u_t^2$$

where $\Pi(L)$ is an infinite–order lag polynomial. Under appropriate conditions, we can rewrite this as a rational lag in two finite–order polynomials in the lag operator:

$$\Pi(L) = \alpha(L)\left[1 - \delta(L)\right]^{-1}$$

where the roots of $\delta(L)$ are outside the unit circle. This gives rise to the $GARCH(r,m)$ model:

$$h_t = \kappa + \sum_{i=1}^{r} \delta_i h_{t-i} + \sum_{j=1}^{m} \alpha_j u_{t-j}^2$$

where $\kappa = (1 - \sum \delta_i)\eta$. This is an $ARMA(p,r)$ process for the squared errors, where the $j^{th}$ AR coefficient is $(\delta_j + \alpha_j)$and the $j^{th}$ MA coefficient is $-\delta_j$, with $p = \max(r, m)$. The non–negativity requirement is that all parameters in this process are non–negative, with $\kappa > 0$. The process is CS if $(\sum \delta_i + \sum \alpha_i) < 1$. Just as a low–order $ARMA(p, q)$ process will often work as well as a high–order $AR(p)$, a low–order $GARCH(r, m)$ will often suffice to capture the dynamics of the conditional variance as well as an $ARCH(m)$ for large $m$. The ability to specify a more parsimonious model, especially given the non–negativity constraints on the maximum likelihood problem, is attractive.

An interesting special case is that of $IGARCH$, or integrated $GARCH$ : where $(\sum \delta_i + \sum \alpha_i) = 1$ (or cannot be distinguished from 1). This causes the unconditional variance of $u_t$ to be infiinite, so that neither $u_t$ nor $u_t^2$ is CS. The issue is essentially that of a unit root in the $ARMA$ process for $u_t^2$, and is often encountered in practice.

**The GARCH–in–mean model**

A very useful variation on the $GARCH$ model is $GARCH - in - mean$ : a specification where the conditional variance itself enters the mean equation. For assets, we might expect higher return and higher risk to be positively correlated, and thus a positive $ARCH - in - mean$ term would be expected. A similar rationale would apply if we confront the stylized fact that countries with higher levels of inflation often are observed to have higher variances of the inflation process. An example of this model is provided by Engle et al. (1987).

**Alternative GARCH specifications**

A huge literature on alternative $GARCH$ specifications exists; many of these models are preprogrammed in Stata's `arch` command, and references for their analytical derivation are given in the Stata manual. One of particular interest is Nelson's (1991) exponential $GARCH$, or $EGARCH$. He proposed:

$$\log h_t = \eta + \sum_{j=1}^{\infty} \pi_j \left( |\nu_{t-j}| - E\,|\nu_{t-j}| + \theta \nu_{t-j} \right)$$

which is then parameterized as a rational lag of two finite–order polynomials, just as in Bollerslev's $GARCH$. Advantages of the $EGARCH$ specification include the positive nature of $h_t$ irregardless of the estimated parameters, and the asymmetric nature of the impact of innovations: with $\theta \neq 0$, a positive shock will have a different effect on volatility than will a negative shock, mirroring findings in equity market research about the impact of "bad news" and "good news" on market volatility. Nelson's model is only one of several extensions of $GARCH$ that allow for asymmetry, or consider nonlinearities in the process generating the conditional variance: for instance, the threshold $ARCH$ model of Zakoian (1990) and the Glosten et al. model (1993).

The $ARCH$ and $GARCH$ models have also been extended in a multivariate context (although considering more than two variables is quite difficult, as the number of parameters to be estimated grows very rapidly).

Useful surveys of the literature (although now somewhat dated) are provided by Bollerslev et al. (1992, 1994).

References

[1] Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.

[2] Bollerslev, T., Chou, R. and K. Kroner, 1992. ARCH modeling in finance. Journal of Econometrics 52, 5–59.

[3] Bollerslev, T., Engle, R. and D. Nelson, 1994. ARCH models. In Handbook of Econometrics, Vol. IV, ed. R. Engle and D. McFadden. Elsevier.

[4] Engle, R., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. Econometrica, 50, 987–1008.

[5] Engle, R., Lilien, D., and R. Robins, 1987. Estimating time varying risk premia in the term structure: the ARCH–M model. Econometrica, 55, 391–407.

[6] Glosten, L. Jagannathan, R., and D. Runkle, 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance, 48, 1779–1801.

[7] Zakoian, J., 1990. Threshold heteroskedastic models. Unpublished manuscript, CREST, INSEE.