POLS571 - Longitudinal Data Analysis
October 18, 2001
ARCH/GARCH Models

1 Introduction

So far, we've concentrated on modeling the mean of time series. As with other models (e.g. heteroskedastic probit, etc.), however, there are times when the variance is also interesting to us: this is particularly true when we have some interest in the reasons a series is more or less volatile.

Consider a simple, stationary autoregressive model of the mean of Y:

$$Y_t = \rho Y_{t-1} + \beta X_t + u_t \tag{1}$$

We typically treat the variance of $u_t = \sigma^2$ as a constant; this then determines the variability of Y. Another possibility, however, is to allow the variance to change over time (that is, to consider σ_t^2). One simple way to do this is to decompose the u term into a systematic part and a random part:

$$u_t = \nu_t \sqrt{h_t} \tag{2}$$

where ν_t is a mean-zero, variance-one, white-noise process and h_t is a scaling factor. In this setup, how we define h becomes very important, and yields a number of different possibilities.

2 The Basic ARCH(1) Model

One very simple alternative for h in (2) is:

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 \tag{3}$$

This is often known as the ARCH(1) model (for "autoregressive conditional heteroscedasticity"), and is due to Engle (1982). Its the model that started the ARCH craze in economics. The process for Y_t is now:

$$Y_t = \rho Y_{t-1} + \beta X_t + \nu_t \sqrt{\alpha_0 + \alpha_1 u_{t-1}^2} \tag{4}$$

Its easy to see that the expected value of u_t is zero (because $E(\nu_t) = 0$). Additionally, the unconditional long-run variance of the error term ut is:

$$Var(u_t) = E(\nu_t^2)E(h_t) = \frac{\alpha_0}{1 - \alpha_1} \tag{5}$$

In turn, this means that we need to impose the constraints $\alpha_0 > 0$ and $0 < \alpha_1 < 1$ in order to keep the variance of the us positive and stationary.

2.1 Intuition behind the ARCH(1) model

- The short-run (conditional) variance ("volatility") of the series is a function of the immediate past values of the (square of the) error term.
- This means that the effect of each new shock ν_t depends, in part, on the size of the shock in the previous period: A large shock in period t increases the effect (on Y) of shocks in periods $T + 1$, $t + 2$, etc.
- So: Large shocks tend to cluster together – the series goes through periods of large volatility, and some of less volatility.

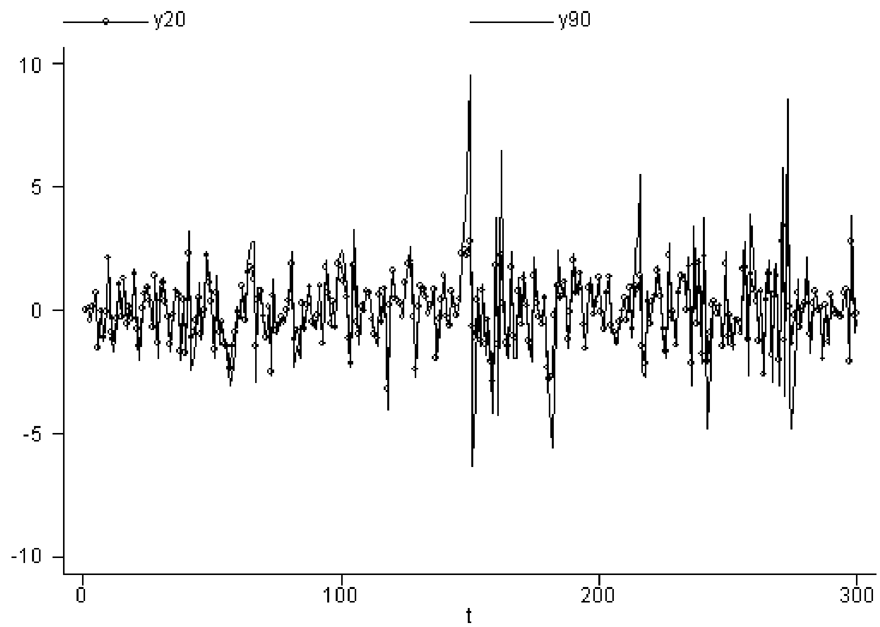
2.2 An Example

Consider an ARCH series where we have:

$$\begin{aligned}
Y_t &= 0.5X_t + u_t \\
u_t &= \nu_t\sqrt{1 + \alpha_1 u_{t-1}^2} \\
X_t &\sim i.i.d.N(0, 1) \\
\nu_t &\sim i.i.d.N(0, 1) \\
u_0 = Y_0 &= 0
\end{aligned}$$

where we have $\alpha_1 = 0.2$ and 0.9 and the same νs in each. These series are plotted for $T = 300$ in Figure 1.
Note that the more volatile parts of the series tend to cluster together, and that the model with $\alpha_1 = 0.9$ has greater influences of shocks than does the one with a smaller ARCH parameter.

Figure 1: ARCH series, $\alpha_1 = 0.2$ and 0.9



2.3 Higher-Order ARCH Models

An easy generalization of the ARCH(1) is to add additional, higher-order ARCH parameters in the variance of the us:

$$h_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j u_{t-j}^2 \tag{6}$$

Higher-order ARCH models are useful when the variability of the series is expected to change more slowly than in the ARCH(1) model (which is often the case). As with the ARCH(1) model, we need to impose some constraints on the αs to ensure that the series is variance stationary.

ARCH(p) models are often difficult to estimate, since high-order models of this sort (i.e., large p) often yield negative estimates of the αs. To address this issue, some smart people (notably Bollerslev 1986) came up with...

2.4 GARCH Models

The general GARCH ("generalized autoregressive conditional heteroscedasticity") model has:

$$h_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j u_{t-j}^2 + \sum_{k=1}^{q} \gamma_k h_{t-k} \tag{7}$$

that is, the value of the variance scaling parameter h_t now depends both on past values of the shocks, and on past values of itself. The simplest of such models is the GARCH(1,1) model:

$$h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \gamma_1 h_{t-1} \tag{8}$$

Successive substitution into the right-hand side of (8) gives:

$$\begin{aligned}
h_t &= \alpha_0 + \alpha_1 u_{t-1}^2 + \gamma_1 h_{t-1} \\
&= \alpha_0 + \alpha_1 u_{t-1}^2 + \gamma_1(\alpha_0 + \alpha_1 u_{t-2}^2 + \gamma_1 h_{t-2}) \\
&= \alpha_0 + \alpha_1 u_{t-1}^2 + \gamma_1 \alpha_0 + \gamma_1 \alpha_1 u_{t-2}^2 + \gamma_1^2 h_{t-2} \\
&= ... \\
&= \frac{\alpha_0}{1-\gamma_1} + \alpha_1(u_{t-1}^2 + \gamma_1 u_{t-2}^2 + \gamma_1^2 u_{t-3}^2 + ...) \qquad (9)
\end{aligned}$$

The current variance thus can be seen to depend on all previous squared disturbances u; but the effect of those disturbances declines exponentially over time.

As in the ARCH case, we need to impose some parameter restrictions on this model to insure that the series is variance-stationary: in the GARCH(1,1) case, we require that $\alpha_0 > 0$, $\alpha_1, \gamma_1 \geq 0$ and $\alpha_1 + \gamma_1 < 1$.

2.5 Other ARCH variants

There have been a bunch of variations on ARCH/GARCH models introduced in the last 20 years. Most are esoteric; some are useful, others less so. Among the more prominent ones are:

- **ARCH-in-mean (ARCH-M)**

- Model is of the form $Y_t = \beta X_t + \delta \sigma_t + u_t$.
- ARCH effects appear in the mean of Y as well as its variance.
- May be appropriate where, e.g., returns to investment depend on risk (as reflected in volatility).

- **Exponential ARCH/GARCH (E-(G)ARCH)**
  - The E-ARCH(1) model is $ln(h_t) = \alpha_0 + \alpha_1[\theta \frac{u_{t-1}}{h_{t-1}} + (|\frac{u_{t-1}}{h_{t-1}}| - \sqrt{2/\pi})]$
  - Here, h_t is an asymmetric function of past values of u_t.
  - Important in e.g. models of stock price volatility, which respond differently depending on whether shocks are positive or negative.

- Many others: Threshold ARCH (TARCH), simple asymmetric ARCH (SAARCH), power ARCH (PARCH), etc. etc.

3 Practical ARCH Modeling

3.1 Detection

there's a relatively straightforward way to assess whether ARCH is a problem in a particular dataset (the approach is due to Engle 1982):

1. Regress Y on X and obtain some residuals $\hat{u}_t$.
2. Regress $\hat{u}_t^2$ on p lags of $\hat{u}_t^2$; that is, $\hat{u}_t^2 = \alpha_0 + \alpha_1 \hat{u}_{t-1}^2 + \alpha_2 \hat{u}_{t-2}^2 + ... + \alpha_p \hat{u}_{t-p}^2$
3. Assess the joint significance of α_1 - α_p. If the coefficients are other than zero, the null of conditional homoscedasticity can be rejected.

Of course, as always, choosing the correct value of p is the tricky part...

3.2 ARCH Model Estimation

Thankfully, our old friend Stata^{TM} will estimate more ARCH and GARCH models than you can shake a stick at. The commands are, generally:

```
.  arch depvar indvar ...  , arch(#)
```

and

```
.  garch depvar indvar ...  , garch(#, #)
```

Maximization is almost always done via MLE; one *could* adopt a feasible GLS approach, using the estimated errors $\hat{u}_t$ as weights, but MLE is better. Stata defaults to the BHHH algorithm, which seems to work pretty well. There's also a good discussion of ARCH/GARCH models in the Stata manuals.

ARCH and GARCH models are notorious for being finicky – they often don't converge, or converge to local mimina, or give "wrong" parameter values (e.g. negative values on the variance terms), etc. Stata's procedures seem to have helped this quite a bit – they're fairly reliable.

I was going to do an example, but I couldn't come up with one...

Tuesday: Panel Data Models!