Macroeconometrics, Spring 2022
Bent E. Sørensen
March 3, 2022

I will talk about models of volatility ("ARCH" and "Stochastic Volatility") in finance. This is for three reasons. a) you may want see financial models, b) I will talk about a very detailed Monte Carlo study of GMM and EMM estimation of these models, with conclusions that generalize to other settings, and c) model with shocks to uncertainty has become quite the cottage industry in macro since the publications of "The Impact of Uncertainty Shocks" by Nicholas Bloom in Econometrica 2009 (I see over 5000 Google references by March 1, 2022). Here, I do not talk about macro models, but you might use the models below as components of such.

1 ARCH and generalizations.

For further or alternative readings,there are numerous treatments of this material in textbooks and surveys.

Many financial time series exhibit *volatility clustering*, which means that the series have periods where volatility is low and other periods where volatility is high. As econometricians we will understand "volatility" to mean (conditional) variance. The conditioning set will always be the behavior of relevant variables up to "time t" — the time when we observe the series. There have recently been developed many different models for data that shows volatility clustering, and it is still a *very* active research area. The main differences between competing models will often be the choice of which variables to condition on, and (as usual) the choice of functional forms. The problems of modeling series with varying volatility is of course well known in econometrics under the heading of heteroskedasticity; but most of the interest in time-varying volatility models comes from finance.

One of the main ideas of asset-pricing is that the variability of an asset should be reflected in its price. One would expect an asset with high variance ("risk") to give a higher return (for investors to want to hold it). With some polishing and generalizations this is the main point of the CAPM-, and APT-models that are common in finance. But standard economic reasoning also says that risk should not be measured as the unconditional variance but rather as the conditional variance, and therefore finance characters are interested in modeling conditional heteroskedasticity in its own right (e.g. for the purpose of pricing options).

1.1 Engle's ARCH model

Robert Engle got the Nobel price in 2003 for "for methods of analyzing economic time series with time-varying volatility (ARCH)," shared with Clive Granger, also of UCSD, "for methods of analyzing economic time series with common trends (cointegration).")

Consider the following simple scalar model

$$
\begin{aligned}
y_t &= \mu + e_t \ , \ t = 1, ..., T \\
e_t &= z_t \sigma_t \qquad (1) \\
\sigma_t^2 &= \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 = \omega + a(L)\epsilon_t^2
\end{aligned}
$$

where $\alpha_1, .., \alpha_q$, μ and ω are scalar parameters to be estimated. (The constant mean is for simplicity of exposition, if you have a vector of regressors x_t it is simple to replace μ with $x_t\beta$.) z_t is supposed to have mean zero and variance one, and will usually (but not always) be assumed to be normally distributed. One has to assume that ω and

α_i are all positive in order to obtain positive values for the estimate of the condition variance. In practice you have to assume this by either penalize the likelihood by setting it to a large negative number when negative values are met or by parameterizing it for example as the square of the parameter. I tend to prefer the later method since the former method potentially will create problems because the likelihood then will not be differentiable. Engle used the model on inflation data but for financial traders in options, conditional estimates of volatility are super important and the field took off with asset pricing applications.

1.2 The GARCH model

Bollerslev (1986) suggested the following natural generalization of the ARCH model, which [with extensions] has been very successful in fitting financial series. Let the $e_t = \sigma_t$ as before, but now let

$$\sigma_t^2 \;=\; \omega \;+\; \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \;+\; \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2$$

which is a natural generalization corresponding to an ARMA model for the variance. This model is called a GARCH(p,q) model. Also in the GARCH model one will restrict the parameters to be positive, which will ensure a positive estimate of the conditional variance (even though I am not sure whether this is also a necessary condition in higher order models). More compactly we write

$$\sigma_t^2 \;=\; \omega \;+\; b(L)\sigma_t^2 \;+\; a(L)\epsilon_t^2 \;.$$

Bollerslev and Engle (1986) looks at the case where the variance process follows the equivalent of an ARIMA model, allowing for unit roots in the lag polynomials. In the case where

$$\alpha_1 \;+\; ... \;+\; \alpha_q \;+\; \beta_1 \;+\; ... \;+\; \beta_p \;=\; 1$$

they refer to the model as an IGARCH(p,q) model. Whether there is a unit root in the autoregressive polynomium for σ_t^2 may be hard to determine, but typically there is very high persistence in the conditional variance (which, again, is why it is powerful for forecasting beyond the very short run)

1.3 The E-GARCH model

One limitation of the GARCH models is that it a priori restricts the shocks to the model to have the same effect on the conditional variance whether the shocks are negative or positive. This may or may not be a reasonable assumption but one would like to be able to test this. The positivity constraints on the parameters can also be viewed as restrictive, since it rules out cyclical behavior in the conditional variance. For those reasons (among others) Nelson (1991) suggests the EGARCH(p,q) model:

$$log(\sigma_t^2) \; = \; \omega \; + \; \sum_{i=1}^{p} \beta_i log(\sigma_{t-i}^2) \sum_{i=1}^{q} \alpha_i (\phi z_{t-i} + \gamma [|z_{t-i}| - E|z_{t-i}|]) \; .$$

In the EGARCH the parameters are not restricted to be positive. Note that the term $|z_{t-i}| - E|z_{t-i}|$ is positive if the error term is larger than its expected value and negative otherwise. One can use other models for $log(\sigma_t^2)$, but for the particular model suggested by Nelson, he shows that the model seems to behave well asymptotically. We will look a little bit at the issue of stationarity of *ARCH models.

1.4 Stationarity of ARCH models

The GARCH model is covariance stationary if $A(1) + B(1) < 1$. It turns out that if $A(1) + B(1) = 1$ then the process is still stationary; but not covariance stationary since the variance is infinite. Notice that this is very different from the ARMA models where strict stationarity and covariance stationarity coincides (if the intial conditions are chosen properly). One can also show for the standard GARCH model that if ω is equal to zero then the conditional variance of the process will converge to zero almost surely. I will not go into details of the stochastic properties of *ARCH processes, but you should be aware that they can be quite tricky. It is obvious that the form of the ARCH models is chosen to give convenient estimations, and not to give convenient theoretical properties.

1.5 ARCH-M models

As mentioned in the introduction, one of the major motivations for looking at conditional variances is that financial theory says that the expected return on an asset should be correlated with its conditional variance. (The expected return will actually depend on the covariance with other assets as well as on the variance, but we will not go into finance theory here). Therefore one may want to combine the models for σ_t^2, whether one prefers ARCH, GARCH, or whatever, with a model for the mean (f.eks. a mean return). So now one would model

$$y_t \ = \ g(\sigma_t, b) \ + \ e_t \ , \ t \ = 1, ..., T \ ,$$

where b is a parameter. (Of course one will usually also want to include regressors but that is supressed here). The ARCH-M model was first suggested by Engle, Lillien and Roberts (1987).

It is usually more complicated to estimate ARCH-M models, because of the fact that the model for the conditional mean now depends on the conditional variance, making the model a lot more non-linear.

1.6 Estimation of ARCH models

The most commonly used estimation strategy is Maximum Likelihood, with an assumption of normality of the error terms. (One may also use the normal likelihood function without wanting to claim that the error terms are normally distributed, in which case one speaks of Quasi Maximum Likelihood estimation). For financial data this is often not a reasonable assumption and there has been articles in the literature that performs Maximum Likelihood using distributions like the t-distribution, that has heavier tails than the normal distribution.

There are also articles in the literature that estimates ARCH models using GMM.

1.7 Other ARCH models

A lot of research is still being devoted to ARCH models. Some other of the newer research concerns factor-ARCH models, non-parametric ARCH models (you can have a nonparametric representation of the conditional variance of the probability density), multivariate ARCH, and all possible combinations. There is also STARCH (structural ARCH), and threshold ARCH and probably a lot of others. We will not go there.

2 Stochastic Volatility models

ARCH-type models are very hard to analyze. The assumption that volatility depends on the observed shock makes them easy to estimate but it is hard to characterize behavior. A conceptually much simpler model to analyze is the stochastic volatility model. Together with Torben Andersen, I have analyzed GMM and Efficient Method of Moments in details for the simplest stochastic volatility model of form: We investigate the following simple version of the lognormal stochastic volatility model:

$$y_t = \sigma_t Z_t$$

$$\ln \sigma_t^2 = \omega + \beta \ln \sigma_{t-1}^2 + \sigma_u u_t \ ,$$

where (Z_t, u_t) is i.i.d. $N(0, I_2)$, i.e. the error terms are mutually independent standard normals. The parameter vector is $\theta = (\omega, \beta, \sigma_u)$. For $0 < \beta < 1$ and $\sigma_u \geq 0$, the return innovation series, y_t, is strictly stationary and ergodic, and unconditional moments of any order exists. Throughout, we work with parameter values that satisfy these inequalities.

In the model, returns display zero serial correlation but dependency in the higher order moments is induced through the stochastic volatility term, σ_t, which follows an AR(1) in logarithms. The volatility persistence parameter, β, is estimated to be less than, but quite close to unity in most empirical studies in finance.

The stochastic volatility model is simple to analyze, but because the uncertainty shock is not directly observed, it is harder to estimate. It is not linear, so a simple application of the Kalman Filter is not feasible. I will show you have to estimate this model by GMM.

In macro, following Bloom (2009), model with random shocks to uncertainty, as in the stochastic volatility model, has become common. (To my taste, a bit too much because some authors pay little attention to what causes such increases in uncertainty. I think there is room for structural empirical work in that area.) As a macro economist, you may think of the stochastic volatility setup as a component that you can include in your model. Of course, you may have asset prices in your model directly, although this is somewhat less common, I believe mainly because people try to limit the number of state variables.