Lecture 10: Modeling Volatility

Dr. Joao B. Duarte[1]

[1]Nova School of Business and Economics
University of Cambridge

Masters, Economics: Macroeconometrics
Lisbon
Spring 2017

Lecture Objectives:

- Intro to ARCH and GARCH models.
- ARCH and GARCH porperties.
- How to estimate the GARCH model.
- Learn ARCH and GARCH applications.

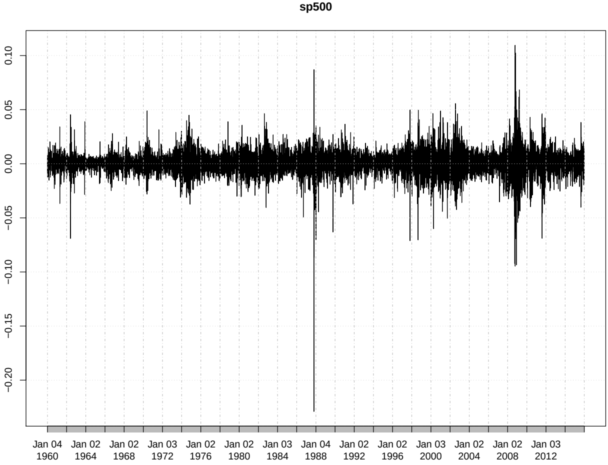
Secondary Readings:

- Chapter 3, Applied Econometric Time Series, Enders, Walter, Fourth Edition.

Intro

- As we saw before, many macroeconomic series are not stationary.
- Part of the reason why that is the case is the fact that time series often do not exhibit constant means. In other words they have trends. We saw in lecture 5 how to deal with trends.
- Besides the non-stationary nature of the macroeconomic data, there is also evidence of moments of high volatility and relatively tranquility. This is particularly true in financial markets. Hence, the assumption of conditional constant variance is violated.
- In this lecture will show to deal with conditional heteroskedasticity.

Example of heteroskedasticity: SP500 returns.



Intro to ARCH Models

- As it is clear from the two previous examples, some economic time series exhibit periods of unusually high volatility and some other moments of low volatility.
- In such cases, the assumption of constant variance (homoskedasticity) is inappropriate.
- Engle (1982) proposed the Autoregressive conditional heteroskedastic (ARCH) model to deal with such cases.

Intro to ARCH Models

- We saw in lecture 3 ARMA models that allowed for a flexible characterization of conditional mean of a series.
- In the ARMA, model we assumed stationarity which means that the unconditional mean and variance were both constant.
- The idea of Engle (1982) was to model volatility in a similar fashion using a model for the variance that yielded a constant unconditional variance, but a time varying conditional variance.

Intro to ARCH Models

- Suppose we have the following AR(1) model:

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$$

The unconditional mean of y_t assuming stationarity is:

$$E[y_t] = \frac{a_0}{1 - a_1}$$

and the conditional mean is given by:

$$E_t[y_t] = a_0 + a_1 y_{t-1}$$

Intro to ARCH Models

- We have also showed before that the unconditional variance of y_t assuming stationarity is:

$$var(y_t) = \frac{\sigma^2}{1 - a_1^2}$$

The conditional variance of y_t is given by:

$$var(y_{t+1}|y_t) = E_t[\varepsilon_{t+1}^2]$$

So far we have assumed that it is constant as well equal to σ^2. What if it is not constant?

Intro to ARCH Models

- One way to proceed is, just like for the conditional mean, to model the conditional variance as a AR(q) process:

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + ... + \alpha_q \varepsilon_{t-q}^2 + v_t \tag{1}$$

where v_t is white-noise. Hence the conditional variance forecast is:

$$E_t[\varepsilon_{t+1}^2] = \alpha_0 + \alpha_1 \varepsilon_t^2 + ... + \alpha_q \varepsilon_{t+1-q}^2$$

This is why the model is called ARCH (autoregressive conditional heteroskedastic).

ARCH Models

- The model for y_t and the conditional variance are estimated together via maximum likelihood (details later). Moreover, (1) is not the most convenient representation of the conditional variance since it complicates to show the model properties.
- Hence, Engle (1982) proposed the following model for the conditional variance with the lag $q = 1$:

$$\varepsilon_t = v_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2} \tag{2}$$

  where v_t is white-noise with variance $\sigma_v^2 = 1$, and v_t and ε_{t-1} are independent of each other. Moreover, we assume $\alpha_0 > 0$ and $0 \leq \alpha_1 \leq 1$ to ensure the conditional variance is always positive and stationary.

ARCH Properties

- Lets start by looking at the moments properties of the ARCH model. Since v_t and ε_{t-1} are independent we have that the unconditional mean of ε_t is given by:

$$\begin{aligned} E[\varepsilon_t] &= E[v_t(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)^{1/2}] \\ &= E[v_t]E[(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)^{1/2}] = 0 \end{aligned}$$

  Since $E[v_t] = 0$. Moreover, since $E[v_t v_{t-i}] = 0$ we have that $E[\varepsilon_t \varepsilon_{t-i}] = 0$ for $i \neq 0$ as well.

- Hence, we have that the unconditional mean and autocovariances are zero for ε_t.

ARCH Properties

- Now turning into the variance, we will first derive the unconditional variance of ε_t:

$$
\begin{aligned}
E[\varepsilon_t^2] &= E[v_t^2]E[\alpha_0 + \alpha_1 \varepsilon_{t-1}^2] \\
&= \frac{\alpha_0}{1 - \alpha_1}
\end{aligned}
$$

Since $E[v_t^2] = 1$ and the process is stationary (i.e. $E[\varepsilon_t^2] = E[\varepsilon_{t-1}^2]$).

ARCH Properties

- It seems that the new structure we gave to the error term in the ARMA model does not have any different properties form what we saw before.
- The difference falls entirely on the conditional variance:

$$E[\varepsilon_t^2|\varepsilon_{t-1}] = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 \tag{3}$$

- The conditional variance will depend on the last periods variance. This is the ARCH(1) model.
- Hence, the error term continues to be a white-noise process but with a conditional variance that is now explicitly modelled as opposed to be assumed constant.

ARCH Properties

- They key point is that the errors are not independent as they are connected through their second moments even though their first moments are uncorrelated.
- The conditional heteroskedasticity of the error term translates into y_t being heteroskedastic as well.
- Periods of volatile errors translate into volatility in y_t though the ARMA equation.
- The following example from Enders help illustrate this point.

ARCH Simulated Examples

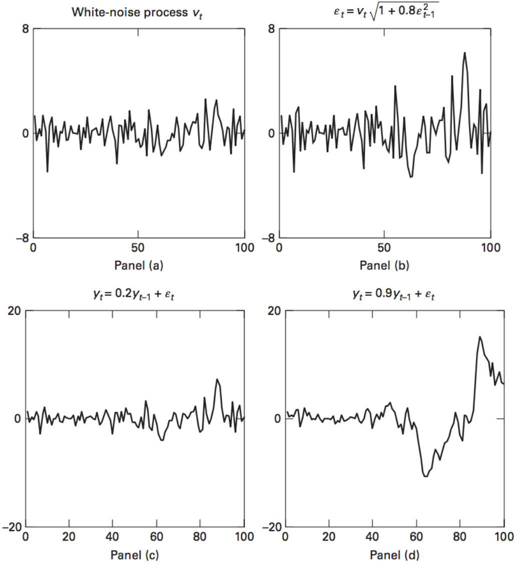


FIGURE 3.7 Simulated ARCH Processes

y Properties in ARCH Models

- Lets take a formal look at the properties of y_t. Remember that we assumed that it followed a simple AR(1) process. In fact, in general it can be any ARMA(p,q) model.

- Now we will do the opposite and start instead with the conditional mean and variance of y_t:

$$E_t[y_t] = a_0 + a_1 y_{t-1}$$

  and

$$var(y_t|y_{t-1}) = E_{t-1}[\varepsilon_t^2] = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

- Note that the variance of y_t is then heteroskedastic. Remember that α_1 cannot be negative. Hence, the smallest conditional variance possible is α_0.

y Properties in ARCH Models

- Now, lets compute the unconditional moments of y_t. To to that, we will solve for y_t and iterate back:

$$y_t = \frac{a_0}{1 - a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i} \tag{4}$$

- The unconditional mean is the same as in the regular AR process since $E[\varepsilon_t] = 0$. Hence, $E[y_t] = \frac{a_0}{1-a_1}$.

y Properties in ARCH Models

- Since $E[\varepsilon_t \varepsilon_{t-i}] = 0$ for $i \neq 0$, then the unconditional variance of y_t is :

$$var(y_t) = \sum_{i=0}^{\infty} a_1^{2i} var(\varepsilon_{t-i})$$

So,

$$var(y_t) = \left(\frac{\alpha_0}{1-\alpha_1}\right)\left(\frac{1}{1-a_1^2}\right)$$

- Hence, the unconditional variance of y_t is also constant and depends on both the persistence of the conditional variance process (α_1) and the y_t process (a_1).

ARCH Models with q lags

- The simple ARCH(1) model can be easily extended for any lag q:

$$\varepsilon_t = v_t \sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2}$$

- And the properties are similar to the ARCH(1) in the sense that the unconditional moments are constant, and only the conditional variance of ε_t and consequently y_t vary over time.

GARCH Models

- Bollerslev (1986) extended the ARCH model by allowing the conditional variance to be modelled as a general ARMA model. This is called the generalized ARCH(p,q) model - GARCH(p,q).
- The error structure of an initial ARMA for y_t is given by:

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

where again $\sigma_v^2 = 1$ and

$$h_t = \alpha_0 \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i}$$

- The conditional variance $E_{t-1}[\varepsilon_t^2] = h_t$. And so we have an ARMA model for the conditional variance. Note that a GARCH(0,1) is an ARCH(1).

GARCH Models Estimation

- We now turn to the estimation of GARCH models. First lets rewrite the full GARCH model. Let x_t be any independent variable that could include autoregressive terms and/or moving average terms. For simplicity lets assume a GARCH(0,1):

$$
\begin{aligned}
y_t &= \beta x_t + \varepsilon_t & (5) \\
\varepsilon_t &= v_t \sqrt{h_t} & (6) \\
h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 & (7)
\end{aligned}
$$

- One way to estimate the GARCH model is to estimate (5) using OLS, take its residuals and square them. Then we can use its squared residuals asd use OLS on: $\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + v_t$

GARCH Models Estimation

- However, there can be a large loss of efficiency by doing in two steps. Hence, typically GARCH models estimate (5), (6) and (7) jointly using maximum likelihood.

- Suppose that ε_t follows a normal distribution. Then the joint likelihood of $\varepsilon_1, ..., \varepsilon_T$ is given by:

$$L = \prod_{t=1}^{T} \left(\frac{1}{\sqrt{2\pi h_t}} \right) \exp \left(\frac{-\varepsilon_t^2}{2h_t} \right)$$

so that the log-likelihood is:

$$\ln L = -\frac{T-1}{2} \ln 2\pi - \frac{1}{2} \sum_{t=1}^{T} \ln h_t - \frac{1}{2} \sum_{t=1}^{T} \frac{\varepsilon_t^2}{h_t}$$

GARCH Models Estimation

- Finally, just replace h_t with (7) and ε_t with (5):

$$\ln L = -\frac{T-1}{2}\ln 2\pi - \frac{1}{2}\sum_{t=1}^{T}\ln \alpha_0 + \alpha_1\varepsilon_{t-1}^2 - \frac{1}{2}\sum_{t=1}^{T}\frac{\varepsilon_t^2}{\alpha_0 + \alpha_1\varepsilon_{t-1}^2}$$

$$\begin{aligned}\ln L = -\frac{T-1}{2}\ln 2\pi - \frac{1}{2}\sum_{t=1}^{T}\ln\left(\alpha_0 + \alpha_1(y_{t-1} - \beta x_{t-1})^2\right) \\ -\frac{1}{2}\sum_{t=1}^{T}\frac{(y_t - \beta x_t)^2}{\alpha_0 + \alpha_1(y_{t-1} - \beta x_{t-1})^2}\end{aligned}$$

- We can then maximize $\ln L$ with respect to α_0, α_1 and β.

GARCH Models Estimation

- The solution involves a complicated non-linear system, thus there is no solution analytically. However, software packages are generally able to deliver reliable numerical solutions.
- There are many packages in R that estimate GARCH models. In this course we will use the fGarch package.

GARCH Effects Identification

- How do we test for GARCH effects in our series?
- We will cover 4 different approaches that can help identifying GARCH effects. All of them use the residuals of an estimated ARMA model to construct their respective tests:
  1. ACF of the squared residuals $\hat{\varepsilon}_t^2$. Recall, that in large samples the standard deviations of the autocorrelation coefficient can be approximated by $1/\sqrt{T}$. If we have significant lags for the ACF, then this is indicative of GARCH effects.
  2. Ljung-Box Q-statistics can again be used for groups of ACF coefficients. The statistic:

     $$Q = T(T+2)\sum_{i=1}^{n} \rho_i^2/(T-i)$$

     Which follows a χ^2 distribution with n degrees of freedom if the squared residuals are serially uncorrelated. If we reject the null, then there is evidence of GARCH effects.

GARCH Effects Identification

3. Regress these squared residuals on a constant and on the q lagged values $\hat{\varepsilon}_{t-1}^2, \hat{\varepsilon}_{t-2}^2, ..., \hat{\varepsilon}_{t-q}^2$:

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + ... + \alpha_q \hat{\varepsilon}_{t-q}^2$$

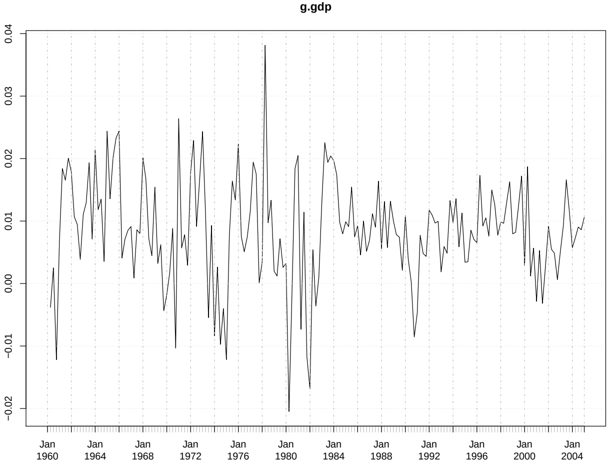
   If there are no GARCH effects, then this regression should explain very little of the squared residuals. Hence, the coefficient of determination R^2 should be very low. Under the null of no ARCH effects, the statistic TR^2 follows a χ^2 distribution with q degrees of freedom. If TR^2 is large we reject the null of no ARCH effects. This test is also called a Lagrange multiplier test (LM).

4. F-test on the same regression under the null that $\alpha_1 = \alpha_2 = ... = \alpha_q = 0$. In small samples the F-test outperforms the TR^2 test. The F-test has q degrees of freedom in numerator and $T - q$ in the denominator. If we reject, then there is evidence of GARCH effects.

GARCH Methodology

1. Estimate a regression model, typically an ARMA model in time series, but could be any regression model.
2. Test the residuals for autocorrelation and then for GARCH effects
3. If GARCH effects are detected, re-estimate the GARCH model with maximum likelihood.
4. Check if the model is adequate. In other words look at the significance of paramenters, check if the model is stable and finally look at the standardize residuals to test whether they seria correlated and if there are any remaining GARCH effects.

Example: Great Moderation



Example: Great Moderation

- Enders example show that a ARCH(1) model offers a good description of the data.
- Moreover the dummy for post-1984 is significant.

$$y_t = \underset{(7.50)}{0.004} + \underset{(6.76)}{0.398} y_{t-1} + \varepsilon_t$$

$$h_t = \underset{(7.87)}{1.10 \times 10^{-4}} + \underset{(2.89)}{0.182} \hat{\varepsilon}^2_{t-1} - \underset{(-6.14)}{8.76 \times 10^{-5}} D_t$$