Eco 4306 Economic and Business Forecasting

Lecture 26
Chapter 14: Forecasting Volatility II

Motivation

- we saw that the conditional variance of several economic variables is time-varying
- important for forecasting - when we construct interval forecasts, e.g. $f_{t,h} \pm 1.96\sigma_{t+h|t}$, the time-varying standard deviation of the process will make the interval forecast either wider or narrower
- moving average (MA) and exponentially weighted moving average (EWMA) specification of time varying volatility are easy to calculate, but have limitations - they are not designed to model time dependence in volatility

- autocorrelation functions of the squared variable of interest are a good starting point is to to analyze the time dependence in volatility
- for example, as we saw last time, autocorrelograms of weekly squared returns to the SP500 index, daily squared returns to the yen/U.S. dollar exchange rate, and to the 10-year Treasury note all show significant positive autocorrelation coefficients
- autocorrelation functions show a slow decay toward zero, indicating that the squared returns may be modeled as autoregressive processes

14.1 The ARCH Family

- our main objective is to estimate and forecast the volatility of the stochastic process

$$r_t = \mu_{t|t-1} + \varepsilon_t$$

  where $\mu_{t|t-1}$ is the conditional mean (that can follow for example an AR or an MA or an ARMA specification)
- innovation ε_t is a white noise process, which by definition is uncorrelated
- we define the conditional variance at time t, as the expectation of the squared process in deviation from its mean given the information set up to $t-1$

$$\sigma^2_{t|t-1} = E[(r_t - \mu_{t|t-1})^2 | I_{t-1}]$$

14.1 The ARCH Family

- we let $\varepsilon_t = \sigma_{t|t-1} z_t$ where z_t is an independent innovation with zero mean and unit variance
- error term ε_t is thus conditionally heteroscedastic because its conditional variance is $\sigma^2_{t|t-1}$ which is time varying:

$$var(\varepsilon_t | I_{t-1}) = E(\varepsilon_t^2 | I_{t-1}) = E(\sigma^2_{t|t-1} z_t^2 | I_{t-1}) = \sigma^2_{t|t-1} E(z_t^2 | I_{t-1}) = \sigma^2_{t|t-1}$$

- although the conditional variance of ε_t is time varying, the *unconditional* variance is constant
- this is analogous to the conditional mean vs unconditional mean for AR/MA/ARMA models case: conditional mean is a function of the information set but the unconditional mean is a constant - for example, for AR(1) model $\mu_{t|t-1} = \phi_0 + \phi_1 y_{t-1}$ but $\mu = \frac{\phi_0}{1-\phi_1}$

14.1 The ARCH Family

- in the **autoregressive conditional heteroscedasticity (ARCH)** model, the conditional variance is assumed to follow an autoregressive process
- $\sigma^2_{t|t-1}$ follows an ARCH process of order p, ARCH(p) if

$$\sigma^2_{t|t-1} = \omega + \alpha_1 \varepsilon^2_{t-1} + \ldots + \alpha_p \varepsilon^2_{t-p}$$

- in ARCH model, conditional variance is thus a function of of previous shocks ε_{t-i} for $i = 1, 2, \ldots, p$
- conditional variance $\sigma^2_{t|t-1}$ is predetermined, known as of time $t-1$
- because ε_{t-i} are squared, the sign of the shocks is irrelevant, only the magnitude matters
- to guarantee that the conditional variance is positive, we need to impose conditions on the parameters: $\omega > 0$ and $\alpha_i \geq 0$ for all i

14.1 The ARCH Family

- consider the simplest possible case, the ARCH(1) process

$$r_t = \mu_{t|t-1} + \varepsilon_t \qquad \text{where } \varepsilon_t = \sigma_{t|t-1} z_t$$
$$\sigma^2_{t|t-1} = \omega + \alpha_1 \varepsilon^2_{t-1}$$

- innovation z_t is independent and identically distributed with $z_t \sim N(0,1)$
- if ε_{t-1} is large (in absolute value), then $\sigma_{t|t-1}$ is large and so ε_t is also expected to be large (in absolute value)
- we next simulate and examine several ARCH processes to better understand their properties
- conditional mean $\mu_{t|t-1}$ can in general follow an AR/MA/ARMA model, for example in AR(1) case we have

$$\mu_{t|t-1} = \phi_0 + \phi_1 y_{t-1}$$

  and so the conditional mean equation in the model above becomes

$$r_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$$

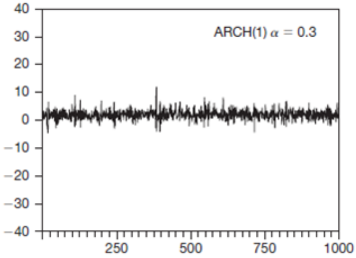
- but to simplify the exposition and focus just on the dynamics of the conditional variance, consider the case where $\mu_{t|t-1}$ is just a constant value $\mu_{t|t-1} = 2$

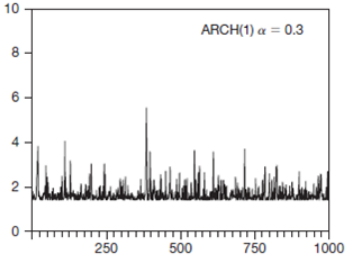
14.1.1 ARCH(1)

- three ARCH(1) processes with $\omega = 2$ and with $\alpha_1 =$ 0.3, 0.6, and 0.9
- conditional mean $\mu_{t|t-1}$ can in general follow an AR/MA/ARMA model, but to simplify the exposition and focus just on the dynamics of the conditional variance, consider the case where $\mu_{t|t-1}$ is just a constant value $\mu_{t|t-1} = 2$

14.1.1 ARCH(1)

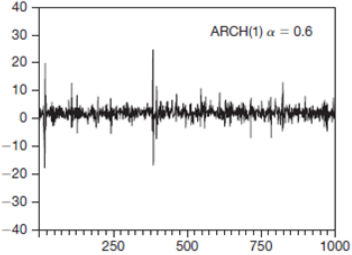
- left figure shows the simulated time series r_t, and the right figure shows the corresponding conditional standard deviation $\sigma_{t|t-1}$

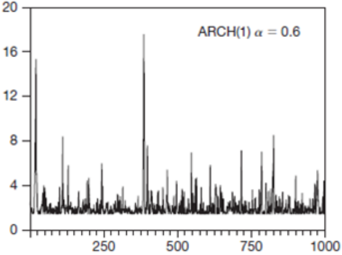




14.1.1 ARCH(1)

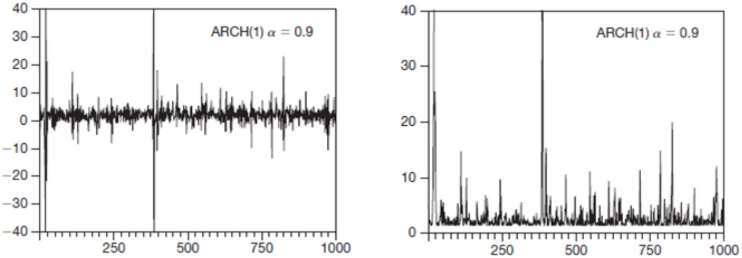
- left figure shows the simulated time series r_t, and the right figure shows the corresponding conditional standard deviation $\sigma_{t|t-1}$





14.1.1 ARCH(1)

- left figure shows the simulated time series r_t, and the right figure shows the corresponding conditional standard deviation $\sigma_{t|t-1}$



14.1.1 ARCH(1)

- unconditional variance of $\{r_t\}$ is $\sigma_\varepsilon^2 = \frac{\omega}{1-\alpha_1}$
- thus when α_1 becomes larger the time series of returns becomes more volatile
- unconditional distribution of r_t is not normal, kurtosis is much higher than 3, and the Jarque-Bera test rejects normality very strongly (p-values of the test are zero)
- kurtosis increases for high values of the parameter α_1, other things being equal

Panel A

Descriptive Statistics of an ARCH(1) process (returns)
Sample: 1 1000

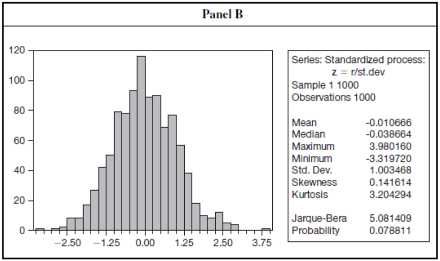
	$\alpha = 0.3$	$\alpha = 0.6$	$\alpha = 0.9$
Mean	1.975935	1.947474	1.881775
Median	1.930632	1.924136	1.906267
Maximum	11.78598	24.60907	55.02234
Minimum	-4.482281	-17.69567	-49.36718
Std. Dev.	1.766631	2.611512	4.810237
Skewness	0.208300	-0.065491	-0.863349
Kurtosis	4.847531	19.48612	56.65335
Jarque-Bera	149.4553	11325.38	120069.3
Probability	0.000000	0.000000	0.000000

14.1.1 ARCH(1)

- once the conditional mean and the conditional standard deviation are estimated, we can construct the estimated standardized residuals $\hat{z}_t$ are obtained as

$$\hat{z}_t = \frac{r_t - \hat{\mu}_{t|t-1}}{\hat{\sigma}_{t|t-1}}$$

- if the dynamics of the conditional mean and variance are correctly specified, then $\hat{z}_t$ should a standard normal random variable



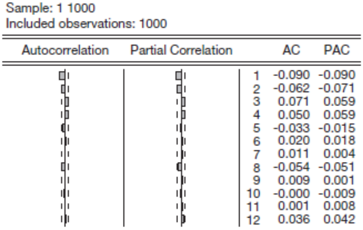
14.1.1.2 What Do the Corresponding Autocorrelation Functions Look Like?

- consider now ARCH(1)

$$r_t = \mu_{t|t-1} + \varepsilon_t \qquad \text{where } \varepsilon_t = \sigma_{t|t-1} z_t$$
$$\sigma^2_{t|t-1} = \omega + \alpha_1 \varepsilon^2_{t-1}$$

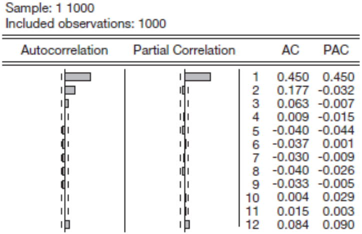
with $\mu_{t|t-1}$, $\omega = 2$, $\alpha_1 = 0.3$

- autocorrelograms of returns r_t (left panel) and squared returnsr_t^2 (right panel): time dependence in r_t^2 reveals time dependence in conditional variance



Sample: 1 1000
Included observations: 1000

Autocorrelation	Partial Correlation		AC	PAC
		1	-0.090	-0.090
		2	-0.062	-0.071
		3	0.071	0.059
		4	0.050	0.059
		5	-0.033	-0.015
		6	0.020	0.018
		7	0.011	0.004
		8	-0.054	-0.051
		9	0.009	0.001
		10	-0.000	-0.009
		11	0.001	0.008
		12	0.036	0.042



Sample: 1 1000
Included observations: 1000

Autocorrelation	Partial Correlation		AC	PAC
		1	0.450	0.450
		2	0.177	-0.032
		3	0.063	-0.007
		4	0.009	-0.015
		5	-0.040	-0.044
		6	-0.037	0.001
		7	-0.030	-0.009
		8	-0.040	-0.026
		9	-0.033	-0.005
		10	0.004	0.029
		11	0.015	0.003
		12	0.084	0.090

14.1.1.2 What Do the Corresponding Autocorrelation Functions Look Like?

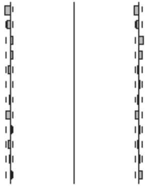
- consider now ARCH(1)

$$r_t = \mu_{t|t-1} + \varepsilon_t \qquad \text{where } \varepsilon_t = \sigma_{t|t-1} z_t$$
$$\sigma^2_{t|t-1} = \omega + \alpha_1 \varepsilon^2_{t-1}$$

with $\mu_{t|t-1}$, $\omega = 2$, $\alpha_1 = 0.3$

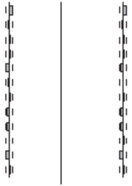
- autocorrelograms of standardized residuals $\hat{z}_t = (r_t - \hat{\mu}_{t|t-1})/\hat{\sigma}_{t|t-1}$ (left panel) and squared standardized residuals $\hat{z}_t^2 = ((r_t - \hat{\mu}_{t|t-1})/\hat{\sigma}_{t|t-1})^2$ (right panel): if the dynamics of the conditional mean and variance are correctly specified, both $\hat{z}_t$ and $\hat{z}_t^2$ should show no time dependence

Sample: 1 1000
Included observations: 1000



Autocorrelation	Partial Correlation		AC	PAC
		1	-0.072	-0.072
		2	-0.034	-0.039
		3	0.079	0.074
		4	0.050	0.061
		5	-0.027	-0.014
		6	0.024	0.018
		7	-0.005	-0.012
		8	-0.059	-0.060
		9	0.013	0.003
		10	-0.011	-0.015
		11	-0.002	0.008
		12	0.046	0.051

Sample: 1 1000
Included observations: 1000



Autocorrelation	Partial Correlation		AC	PAC
		1	0.067	0.067
		2	-0.012	-0.017
		3	0.018	0.020
		4	0.007	0.004
		5	-0.048	-0.049
		6	-0.008	-0.001
		7	0.033	0.032
		8	-0.033	-0.036
		9	-0.037	-0.031
		10	0.012	0.012
		11	-0.052	-0.055
		12	0.018	0.031

14.1.1.3 What Is the Optimal Forecast Corresponding to an ARCH(1) Process?

forecasting under a quadratic loss function

- the one-step-ahead variance forecast is $\sigma^2_{t+1|t} = \omega + \alpha_1 \varepsilon^2_t$
- the two-step-ahead forecast is $\sigma^2_{t+2|t} = E(\omega + \alpha_1 \varepsilon^2_{t+1}) = \omega + \alpha_1 \sigma^2_{t+1|t}$
- the three-step-ahead forecast is $\sigma^2_{t+3|t} = E(\omega + \alpha_1 \varepsilon^2_{t+2}) = \omega + \alpha_1 \sigma^2_{t+2|t}$
- this implies that in general for h step ahead forecast we have $\sigma^2_{t+h|t} = \omega + \alpha_1 \sigma^2_{t+h-1|t}$ or after substituting in

$$\sigma^2_{t+h|t} = \omega(1 + \alpha_1 + \alpha_1^2 + \ldots + \alpha_1^{h-2}) + \alpha_1^{h-1} \sigma^2_{t+1|t}$$

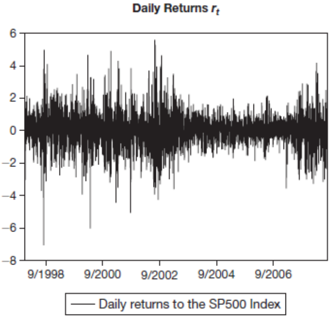
- thus as $h \to \infty$ then $\sigma^2_{t+h|t} \to \frac{\omega}{1-\alpha_1}$ which is the unconditional variance of r_t
- convergence toward the unconditional variance is slow when α_1 is large, and fast when α_1 is small
- this is analogous to the result that we saw when forecasting the conditional mean using an AR model: as the forecast horizon increases, the memory of the model is lost, and the forecast converges to the unconditional mean

14.1.2 ARCH(p)

- properties of the ARCH(1) extend to ARCH(p) processes with small modifications
- unconditional variance is $\sigma_{\varepsilon}^2 = \frac{\omega}{1-\alpha_1-\ldots-\alpha_p}$
- first p lags in PACF for demeaned squared returns are significantly different from zero

14.1.2 ARCH(p)

- consider daily returns for S&P 500
- autocorrelograms of squared daily returns shows more persistence in the squared daily returns than in weekly or monthly data
- we could entertain an ARCH(8) or ARCH(9) for these data



Autocorrelograms of the Squared Returns r_t^2

Sample: 5815 8471
Included observations: 2657

Autocorrelation	Partial Correlation		AC	PAC
		1	0.152	0.152
		2	0.196	0.177
		3	0.196	0.153
		4	0.136	0.067
		5	0.194	0.124
		6	0.144	0.062
		7	0.168	0.083
		8	0.160	0.066
		9	0.117	0.019
		10	0.147	0.049
		11	0.136	0.043
		12	0.130	0.033
		13	0.103	-0.000
		14	0.098	0.004
		15	0.080	-0.014

14.1.2 ARCH(p)

steps involved in building an autoregressive conditional heteroscedasticity (ARCH) model:

1. specify the conditional mean equation - use ACF, PACF, Q-statistics to identify serial dependence in the data
2. estimate the model for the conditional mean
3. use residuals of the mean equation from step 2 to specify a volatility model - use ACF, PACF, Q-statistics for squared residuals
4. perform a joint estimation of the mean and volatility equations
5. check the fitted model for adequacy - both standardized residuals and squared standardized residuals should be white noise

after the joint estimation in step 4 it is possible that the conditional mean equation from steps 1 and 2 has to be modified - some terms can for example become insignificant

14.1.2 ARCH(p)

- to estimate an ARCH(9) model for daily returns of S&P 500 index in EViews enter the following information in the specification window
  - estimation settings: choose “ARCH - Autoregressive Conditional Heteroscedasticity”" instead of “LS - Least Squares”
  - in the mean equation: r c AR(1)
  - in the variance equation select ARCH 9 and GARCH 0

14.1.2 ARCH(p)

SP500 daily returns—ARCH(9)

Dependent Variable: R
Method: ML - ARCH (BHHH) - Normal distribution
Sample: 5815 8471
Included observations: 2657
Convergence achieved after 16 iterations
Bollerslev-Wooldrige robust standard errors & covariance
Variance backcast: ON
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*RESID(-2)^2 + C(5)*RESID(-3)^2 + C(6)*RESID(-4)^2 + C(7)*RESID(-5)^2 + C(8)*RESID(-6)^2 + C(9)*RESID(-7)^2 + C(10)*RESID(-8)^2 + C(11)*RESID(-9)^2

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.037003	0.018214	2.031594	0.0422
Variance Equation				
C	0.271763	0.040891	6.645982	0.0000
RESID(-1)^2	0.029949	0.028081	1.066510	0.2862
RESID(-2)^2	0.149370	0.044623	3.347391	0.0008
RESID(-3)^2	0.095260	0.026377	3.611510	0.0003
RESID(-4)^2	0.101684	0.027620	3.681607	0.0002
RESID(-5)^2	0.082439	0.023397	3.523482	0.0004
RESID(-6)^2	0.060298	0.021251	2.837387	0.0045
RESID(-7)^2	0.090927	0.030511	2.980119	0.0029
RESID(-8)^2	0.142659	0.029601	4.819476	0.0000
RESID(-9)^2	0.082659	0.023815	3.470870	0.0005

R-squared	-0.000565	Mean dependent var	0.009761
Adjusted R-squared	-0.004346	S.D. dependent var	1.146761
S.E. of regression	1.149251	Akaike info criterion	2.910013
Sum squared resid	3494.776	Schwarz criterion	2.934377
Log likelihood	-3854.952	Durbin-Watson stat	2.079077

14.1.2 ARCH(p)

- the estimated ARCH(1) model is

$$r_t = 0.037 + \varepsilon_t$$
$$\sigma^2_{t|t-1} = 0.271 + 0.030\varepsilon^2_{t-1} + 0.149\varepsilon^2_{t-2} + 0.095\varepsilon^2_{t-3} + \ldots + 0.083\varepsilon^2_{t-9}$$

- note that the sum of the coefficients α_i is very high, $\alpha_1 + \alpha_2 + \ldots + \alpha_9 \approx 0.835$
- this sum is also known as the **persistence in variance**
- when the persistence is high, the conditional variance will tend to be high (or low) for many consecutive days

14.1.2 ARCH(p)

advantages of ARCH models

- simple, but able to generate volatility clustering

weaknesses of ARCH models

- large number of lags often required to adequately describe the volatility process
- positive and negative shocks have same effects on volatility (no asymmetry)
- conditional standard deviation process tends to have low persistence and high-frequency oscillations with high volatility coming in short bursts

14.1.3 GARCH(1,1)

- in **generalized autoregressive conditional heteroscedasticity (GARCH)** innovations follows a process similar to an ARMA model
- GARCH(1,1) model with normally distributed innovations has the following specification
  - the dynamics of the conditional mean is given by

$$r_t = \mu_{t|t-1} + \varepsilon_t \qquad \text{where } \varepsilon_t = \sigma_{t|t-1} z_t \text{ and } z_t \sim N(0,1)$$

  - the dynamics of the conditional variance is given by

$$\sigma^2_{t|t-1} = \omega + \alpha_1 \varepsilon^2_{t-1} + \beta_1 \sigma^2_{t-1|t-2}$$

  parameters satisfy $\omega > 0$, $\alpha_1 \geq 0$, and $\beta_1 \geq 0$

- the difference compared to ARCH(1) specification is that in GARCH(1,1) model the conditional variance at time t depends not only on the past innovation ε_{t-1} but also on the most recent level of volatility $\sigma^2_{t-1|t-2}$
- for instance, if $\beta_1 = 0.80$, we say that 80% of yesterday's variance carries over to today's variance
- unconditional first and second moments are $E(\varepsilon_t) = 0$, $var(\varepsilon_t) = \frac{\omega}{1-\alpha_1-\beta_1}$
- conditional first and second moments are $E_{t-1}(\varepsilon_t) = 0$, $var_{t-1}(\varepsilon_t) = \omega + \alpha_1 \varepsilon^2_{t-1} + \beta_1 \sigma^2_{t-1}$

14.1.3 GARCH(m, s)

- a general GARCH(m, s) model assumes that $\sigma^2_{t|t-1}$ is given by

$$\sigma^2_{t|t-1} = \omega + \sum_{i=1}^{m} \alpha_i \varepsilon^2_{t-i} + \sum_{i=1}^{s} \beta_i \sigma^2_{t-i|t-i-1}$$

- lower order GARCH models, GARCH(1,1), GARCH(2,1), GARCH(1,2), are used in most applications
- main advantage of GARCH models is that they can generate similar volatility dynamics as high order ARCH models with fewer parameters
- for example, with a GARCH (1,1) model we have three parameters $\omega, \alpha_1, \beta_1$ to estimate, while with an ARCH(9) we have 10 parameters $\omega, \alpha_1, \ldots, \alpha_9$ to estimate

14.1.3 GARCH(1,1)

- a GARCH(1,1) process is equivalent to an ARCH process of infinite order, ARCH(∞), with exponentially decreasing weights $\alpha_1, \alpha_1\beta_1, \alpha_1\beta_1^2, \alpha_1\beta_1^3, \ldots$ on squared past innovations $\varepsilon_{t-1}^2, \varepsilon_{t-2}^2, \varepsilon_{t-3}^2, \varepsilon_{t-4}^2, \ldots$
- to understand this equivalence, note that we could use backward substitution in the GARCH(1,1) conditional variance equation to obtain

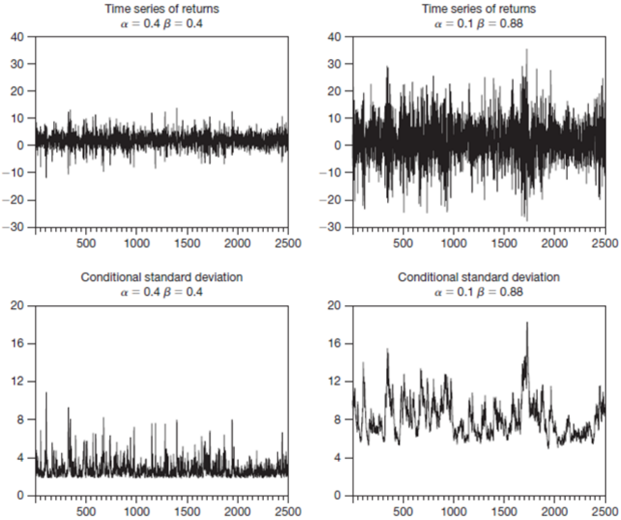
$$
\begin{aligned}
\sigma_{t|t-1}^2 &= \omega + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1|t-2}^2 \\
&= \omega + \alpha_1\varepsilon_{t-1}^2 + \beta_1\left(\omega + \alpha_1\varepsilon_{t-2}^2 + \beta_1\sigma_{t-2|t-3}^2\right) \\
&= \omega(1+\beta_1) + \alpha_1\varepsilon_{t-1}^2 + \beta_1\alpha_1\varepsilon_{t-2}^2 + \beta_1^2\sigma_{t-2|t-3}^2 \\
&= \ldots \\
&= \omega(1+\beta_1+\beta_1^2+\ldots) + \alpha_1\sum_{i=1}^{\infty}\beta_1^{i-1}\varepsilon_{t-i}^2 \\
&= \frac{\omega}{1-\beta_1} + \alpha_1\sum_{i=1}^{\infty}\beta_1^{i-1}\varepsilon_{t-i}^2
\end{aligned}
$$

- persistence of the GARCH(1,1) process - how permanent are the shocks - is given by $\alpha_1\sum_{i=1}^{\infty}\beta_1^{i-1} = \frac{\alpha_1}{1-\beta_1}$

14.1.3.1 What Does a Time Series of GARCH(1,1) Process Look Like?

- to assess the contribution of the α_1 and β_1 parameters we'll compare two GARCH(1,1) processes
- as in the case of ARCH(1) process before, we set the conditional mean to a constant value, $\mu_{t|t-1} = 2$, to focus just on the dynamics of the conditional variance
- in addition, as before we set $\omega = 2$, and we again assume that innovation z_t is normally distributed $z_t \sim N(0,1)$ image
- consider two simulated time series with 2,500 observations for the return process r_t and for the conditional variance $\sigma^2_{t|t-1}$
  1. low persistence process: $\alpha_1 = 0.4$, $\beta_1 = 0.4$, with persistence $\frac{\alpha_1}{1-\beta_1} = 0.667$
  2. high persistence process: $\alpha_1 = 0.1$, $\beta_1 = 0.88$, with persistence $\frac{\alpha_1}{1-\beta_1} = 0.83$

14.1.3.1 What Does a Time Series of GARCH(1,1) Process Look Like?



14.1.3.1 What Does a Time Series of GARCH(1,1) Process Look Like?

- the unconditional variance is larger in the high persistence process than in the low persistence process
- this is because the unconditional variance of a GARCH(1,1) process is $\sigma_\varepsilon^2 = \frac{\omega}{1-\alpha_1-\beta_1}$, and so in the low persistence process $\sigma_\varepsilon^2 = 10$, and in the high persistence process $\sigma_\varepsilon^2 = 100$
- in addition, in the high persistence process, 88% of the past volatility is transferred to the current volatility, while in the low persistence process, it is only 40%
- thus, in a high persistence process, high (low) volatility is followed by high (low) volatility over longer periods of time than in a low persistence process

14.1.3.2 What Do the Corresponding Autocorrelation Functions Look Like?

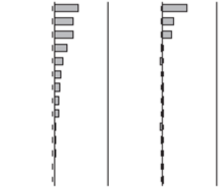
- correlograms of the squared time series r_t^2 for the low and high persistence cases have profiles that correspond to those from autoregressive processes
- degree of persistence in variance is evident in the speed of the decay toward zero - decay is faster in the low persistence process than in the high persistence process
- the correlogram for daily returns to the S&P500 index shows pattern similar to those of the simulated high persistent process

$$r_t = 2 + \varepsilon_t$$
$$\sigma^2_{t|t-1} = 2 + \alpha\varepsilon^2_{t-1} + \beta\sigma^2_{t-1|t-2}$$

Time series r_t^2

(1) $\alpha = 0.4$, $\beta = 0.4$ (low persistence)

Sample: 1 20000
Included observations: 20000

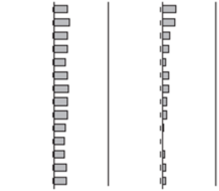


	AC	PAC
1	0.470	0.470
2	0.383	0.208
3	0.376	0.182
4	0.249	-0.022
5	0.167	-0.046
6	0.139	-0.006
7	0.113	0.019
8	0.089	0.018
9	0.086	0.024
10	0.048	-0.029
11	0.034	-0.014
12	0.026	-0.006
13	0.015	0.002
14	0.019	0.017

Time series r_t^2

(2) $\alpha = 0.1$, $\beta = 0.88$ (high persistence)

Sample: 1 20000
Included observations: 20000



	AC	PAC
1	0.263	0.263
2	0.301	0.249
3	0.249	0.143
4	0.254	0.129
5	0.209	0.065
6	0.255	0.119
7	0.277	0.139
8	0.248	0.079
9	0.260	0.085
10	0.215	0.025
11	0.199	0.009
12	0.212	0.041
13	0.233	0.064
14	0.245	0.073

14.1.3 GARCH(1,1)

- to estimate a GARCH(1,1) enter the following information in the specification window:
  - estimation settings: choose “ARCH - Autoregressive Conditional Heteroscedasticity”" instead of “LS - Least Squares”
  - mean equation: r c
  - variance and distribution specification: ARCH 1, GARCH 1

14.1.3 GARCH(1,1)

Dependent Variable: R
Method: ML - ARCH (BHHH) - Normal distribution
Sample: 5815 8471
Included observations: 2657
Convergence achieved after 10 iterations
Bollerslev-Wooldrige robust standard errors & covariance
Variance backcast: ON
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.036267	0.017439	2.079665	0.0376
		Variance Equation		
C	0.010421	0.005245	1.987099	0.0469
RESID(-1)^2	0.065649	0.011338	5.790038	0.0000
GARCH(-1)	0.927400	0.011045	83.96233	0.0000

R-squared	-0.000534	Mean dependent var	0.009761
Adjusted R-squared	-0.001666	S.D. dependent var	1.146761
S.E. of regression	1.147716	Akaike info criterion	2.888638
Sum squared resid	3494.671	Schwarz criterion	2.897498
Log likelihood	-3833.556	Durbin-Watson stat	2.079139

14.1.3 GARCH(1,1)

- the estimated model is

$$r_t = 0.036 + \varepsilon_t$$
$$\sigma^2_{t|t-1} = 0.010 + 0.066\varepsilon^2_{t-1} + 0.927\sigma^2_{t-1|t-2}$$

- the estimated persistence is thus $\frac{\alpha_1}{1-\beta_1} = \frac{0.066}{1-0.927} = 0.9$
- note that both information criteria for GARCH(1,1) model are lower than those for the ARCH(9) model
- to check whether the equation for the conditional mean is adequately specified we need to examine the autocorrelations of the standardized residuals $\hat{z}_t = \frac{\hat{\varepsilon}_t}{\hat{\sigma}_{t|t-1}}$
- to check whether the equation for the conditional variance is adequately specified we need to examine the autocorrelations of the standardized squared residuals $\hat{z}^2_t = \frac{\hat{\varepsilon}^2_t}{\hat{\sigma}^2_{t|t-1}}$
- in both cases the autocorrelations should be close to zero and insignificant

14.1.3 GARCH(1,1)

- autocorrelation function of the standardized squared residuals $\hat{z}_t^2 = \frac{\hat{\varepsilon}_t^2}{\hat{\sigma}_{t|t-1}^2}$ from GARCH(1,1) model for daily S&P 500 returns
- GARCH(1,1) specification of the model for volatility is adequate - there is no remaining time dependence left in squared residuals $\hat{z}_t^2$

Sample: 5815 8471
Included observations: 2657

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.031	-0.031	2.5244	0.112
		2	0.033	0.032	5.4669	0.065
		3	0.007	0.009	5.6164	0.132
		4	0.005	0.005	5.6914	0.223
		5	0.006	0.005	5.7756	0.329
		6	-0.016	-0.016	6.4731	0.372
		7	0.000	-0.001	6.4731	0.486
		8	0.023	0.024	7.8948	0.444
		9	0.001	0.002	7.8965	0.545
		10	0.024	0.023	9.4796	0.487
		11	0.009	0.010	9.7071	0.557
		12	-0.010	-0.012	9.9971	0.616
		13	0.002	0.000	10.011	0.693
		14	-0.004	-0.003	10.063	0.758

14.1.3.3 What Is the Optimal Forecast Corresponding to a GARCH(1,1) Process?

- under symmetric quadratic loss function, with information set I_t the optimal variance forecasts $\sigma^2_{t+h|t}$ are
- the 1-step-ahead forecast

$$\sigma^2_{t+1|t} = \omega + \alpha_1 \varepsilon^2_t + \beta_1 \sigma^2_{t|t-1}$$

- the 2-step-ahead forecast is

$$\sigma^2_{t+2|t} = \omega + \alpha_1 E(\varepsilon^2_{t+1}|I_t) + \beta_1 \sigma^2_{t+1|t} = \omega + (\alpha_1 + \beta_1)\sigma^2_{t+1|t}$$

- the 3-step-ahead forecast is:

$$\sigma^2_{t+3|t} = \omega + \alpha_1 E(\varepsilon^2_{t+2}|I_t) + \beta_1 \sigma^2_{t+2|t+1} = \omega + (\alpha_1 + \beta_1)\sigma^2_{t+2|t} = \omega(1 + \alpha_1 + \beta_1) + (\alpha_1 + \beta_1)$$

- in general, the h-step-ahead forecast of the conditional variance is

$$\sigma^2_{t+h|t} = \omega(1 + (\alpha_1 + \beta_1) + (\alpha_1 + \beta_1)^2 + \ldots + (\alpha_1 + \beta_1)^{h-2}) + (\alpha_1 + \beta_1)^{h-1} \sigma^2_{t+1|t}$$

14.1.3.3 What Is the Optimal Forecast Corresponding to a GARCH(1,1) Process?

- note that if $\alpha_1 + \beta_1 < 1$ the forecast converges to the unconditional variance of the process, as $h \to \infty$ we have

$$\sigma^2_{t+h|t} = \omega(1 + (\alpha_1 + \beta_1) + (\alpha_1 + \beta_1)^2 + \ldots) \to \frac{\omega}{1 - (\alpha_1 + \beta_1)}$$

- if $\alpha_1 + \beta_1 = 1$, that is in the IGARCH(1,1) case, the forecast is a linear function of the forecast horizon

$$\begin{aligned}\sigma^2_{t+h|t} &= \omega(1 + (\alpha_1 + \beta_1) + (\alpha_1 + \beta_1)^2 + \ldots + (\alpha_1 + \beta_1)^{h-2}) + (\alpha_1 + \beta_1)^{h-1}\sigma^2_{t+1|t} \\ &= \omega(h-1) + \sigma^2_{t+1|t}\end{aligned}$$

TARCH model

- reaction of volatility to positive vs negative shocks is likely not symmetric
- intuitively, negative shocks increase volatility more than positive shocks of the same size
- this is called the **leverage effect**
- **Threshold ARCH model, TARCH(1,1)**

$$\sigma^2_{t|t-1} = \omega + \alpha_1 \varepsilon^2_{t-1} + \gamma_1 \varepsilon^2_{t-1} I_{\{\varepsilon_{t-1}<0\}} + \beta_1 \sigma^2_{t-1|t-2}$$

  where γ_1 represents the 'leverage' term is one way how to capture this asymmetric reaction of volatility to positive vs negative shocks
- to see this note that the above model can be written as

$$\begin{aligned} \sigma^2_{t|t-1} &= \omega + \alpha_1 \varepsilon^2_{t-1} + \beta_1 \sigma^2_{t-1|t-2} && \text{if } \varepsilon_{t-1} \geq 0 \\ \sigma^2_{t|t-1} &= \omega + (\alpha_1 + \gamma_1) \varepsilon^2_{t-1} + \beta_1 \sigma^2_{t-1|t-2} && \text{if } \varepsilon_{t-1} < 0 \end{aligned}$$

  and so if $\gamma_1 > 0$ negative shocks lead to a larger increase in volatility than positive shocks

TARCH model

- in EViews, to estimate a TARCH(1,1) model instead of a GARCH(1,1), simply change "Threshold order" in "Variance and distribution specification" from 0 to 1
- result below are for TARCH(1,1) model for daily returns of Google stock
- $\alpha_1 = 0.0296$ and $\alpha_1 + \gamma_1 = 0.0296 + 0.0641 = 0.0937$ so the effect of a negative shock on volatility is about three times as large as the effect of a positive shock

Dependent Variable: R
Method: ML ARCH - Normal distribution (OPG - BHHH / Marquardt steps)
Date: 05/02/17 Time: 19:06
Sample (adjusted): 8/29/2004 4/30/2017
Included observations: 662 after adjustments
Convergence achieved after 75 iterations
Coefficient covariance computed using outer product of gradients
Presample variance: backcast (parameter = 0.7)
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*RESID(-1)^2*(RESID(-1)<0) +
C(5)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.430126	0.163007	2.638702	0.0083
Variance Equation				
C	0.789046	0.291020	2.711309	0.0067
RESID(-1)^2	0.029587	0.012420	2.382237	0.0172
RESID(-1)^2*(RESID(-1)<0)	0.064134	0.024918	2.573803	0.0101
GARCH(-1)	0.891791	0.029447	30.28423	0.0000

R-squared	-0.000005	Mean dependent var	0.420322
Adjusted R-squared	-0.000005	S.D. dependent var	4.315342
S.E. of regression	4.315353	Akaike info criterion	5.660740
Sum squared resid	12309.32	Schwarz criterion	5.694692
Log likelihood	-1868.705	Hannan-Quinn criter.	5.673899
Durbin-Watson stat	2.093189		

PARCH model

- reaction of volatility to positive vs negative shocks is likely not symmetric
- intuitively, negative shocks increase volatility more than positive shocks of the same size
- this is called the **leverage effect**
- **Power ARCH model, PARCH(1,1)**

$$\sigma^{\delta}_{t|t-1} = \omega + \alpha_1 \left(|\varepsilon_{t-1}| - \gamma_1 \varepsilon_{t-1} \right)^{\delta} + \beta_1 \sigma^{\delta}_{t-1|t-2}$$

  where γ represents the 'leverage' term and $\delta > 0$ is another way how to capture this asymmetric reaction of volatility to positive vs negative shocks
- to see this note that if $\delta = 2$ the above model can be written as

$$\begin{aligned} \sigma^2_{t|t-1} &= \omega + \alpha_1(1-\gamma_1)\varepsilon^2_{t-1} + \beta_1 \sigma^2_{t-1|t-2} \quad & \text{if } \varepsilon_{t-1} \geq 0 \\ \sigma^2_{t|t-1} &= \omega + \alpha_1(1+\gamma_1)\varepsilon^2_{t-1} + \beta_1 \sigma^2_{t-1|t-2} \quad & \text{if } \varepsilon_{t-1} < 0 \end{aligned}$$

  and so if $\gamma_1 > 0$ negative shocks lead to a larger increase in volatility than positive shocks

PARCH model

- in EViews, to estimate a PARCH(1,1) model instead of a GARCH(1,1), change "Threshold order" in "Variance and distribution specification" from 0 to 1, change "Model" from GARCH/TARCH to PARCH, and set "Fix power parameter" to 2
- result below are for TPARCH(1,1) model for daily returns of Google stock
- $\alpha_1 = 0.0570$, $\gamma_1 = 0.290$, thus $\alpha_1(1-\gamma_1) = 0.0404$ and $\alpha_1(1+\gamma_1) = 0.0735$, the effect of a negative shock is almost twice as large as the effect of a positive shock

Dependent Variable: R
Method: ML ARCH - Normal distribution (OPG - BHHH / Marquardt steps)
Date: 05/02/17 Time: 19:06
Sample (adjusted): 8/29/2004 4/30/2017
Included observations: 662 after adjustments
Convergence achieved after 75 iterations
Coefficient covariance computed using outer product of gradients
Presample variance: backcast (parameter = 0.7)
GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*RESID(-1)^2*(RESID(-1)<0) +
C(5)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.430126	0.163007	2.638702	0.0083
Variance Equation				
C	0.789046	0.291020	2.711309	0.0067
RESID(-1)^2	0.029587	0.012420	2.382237	0.0172
RESID(-1)^2*(RESID(-1)<0)	0.064134	0.024918	2.573803	0.0101
GARCH(-1)	0.891791	0.029447	30.28423	0.0000

R-squared	-0.000005	Mean dependent var	0.420322
Adjusted R-squared	-0.000005	S.D. dependent var	4.315342
S.E. of regression	4.315353	Akaike info criterion	5.660740
Sum squared resid	12309.32	Schwarz criterion	5.694692
Log likelihood	-1868.705	Hannan-Quinn criter.	5.673899
Durbin-Watson stat	2.093189		

GARCH-M model

- return of an asset can depend on volatility - risk premium argument: investors need to be compensated by higher return if they face larger risk
- **GARCH(1,1)-in mean model, GARCH(1,1)-M**

$$r_t = \mu + \xi \sigma^2_{t|t-1} + \varepsilon_t$$
$$\varepsilon_t = \sigma_{t|t-1} z_t$$
$$\sigma^2_{t|t-1} = \omega + \alpha_1 \varepsilon^2_{t-1} + \beta_1 \sigma^2_{t-1|t-2}$$

- parameter ξ captures the risk premium
- variants of the model can have the term $\sigma_{t|t-1}$ or $\log \sigma^2_{t|t-1}$ instead of $\sigma^2_{t|t-1}$ in the conditional mean equation