Economics 144
Economic Forecasting

Lecture 14

Autoregressive Conditional Heteroscedasticity Models (ARCH/GARCH)

Dr. Randall R. Rojas

Today's Class

- The ARCH Family
- ARCH Models
  - The ARCH(1) Process
  - The ARCH(p) Process
- GARCH Models
  - The GARCH(1,1) Process
  - The GARCH(p,q) Process

The ARCH Family

- Def: ARCH(1) = Autoregressive Conditional Heteroscedasticity Process of order 1:

ARCH - ar with longer lag

$Y_t = \mu_{t|t-1} + \varepsilon_t$, where $\mu_{t|t-1}$ is the conditional mean and ε_t is a white noise process s.t:

normal - shock

$\varepsilon_t = \sigma_{t|t-1} z_t$ and $z_t \sim WN(0,1)$, where

constant main factor: alpha

$\sigma^2_{t|t-1} = E[\varepsilon^2_t | I_{t-1}] = \omega + \alpha_1 \varepsilon^2_{t-1}$

variance model dependent on everything that has happened before

= Conditional variance of ε_t

subject to $\omega > 0$, $\alpha_1 \geq 0$

ARCH Models

For every ARCH process, we need to address the following 3 questions:

1. What does a time series of an ARCH process look like?
2. What do the corresponding ACFs and PACFs look like?
3. What is the optimal forecast?

ARCH Models

Example: ARCH(1) Process

- (1) What does a time series of an ARCH(1) process look like?

- Consider the following ARCH(1) process:

  $Y_t = \mu_{t|t-1} + \varepsilon_t = \mu_{t|t-1} + \sigma_{t|t-1} z_t,\ \sigma^2_{t|t-1} = \omega + \alpha\varepsilon^2_{t-1}$

  Divide epsilon across - get constant

  model for variance - model volatility & forecast volatility

- We can plot this process for $\mu_{t|t-1} = 2$, $z_t \sim N(0,1)$, $\omega = 2$, and different values of α, e.g., for α=0.3, α=0.6, α=0.9.

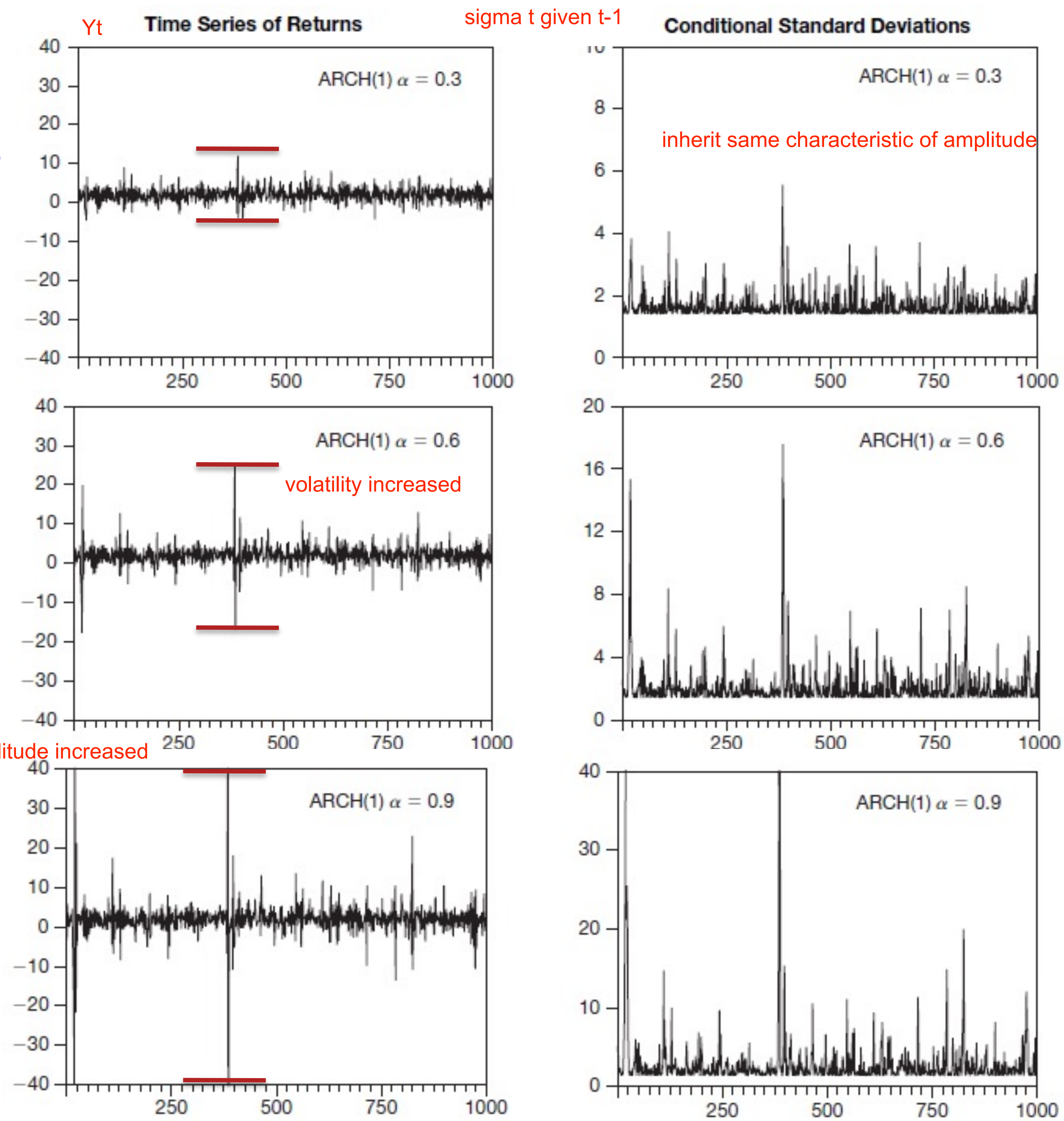
  should get white noise for data

  vary the value of alpha, see what happens to volatility

- We can show that as α increases, the series becomes more volatile.

Simulated ARCH(1) Process

$$Y_t = 2 + \varepsilon_t$$
$$\sigma^2_{t|t-1} = 2 + \mathbf{0.3}\varepsilon^2_{t-1}$$



$$Y_t = 2 + \varepsilon_t$$
$$\sigma^2_{t|t-1} = 2 + \mathbf{0.6}\varepsilon^2_{t-1}$$

still considered white noise, just amplitude increased

$$Y_t = 2 + \varepsilon_t$$
$$\sigma^2_{t|t-1} = 2 + \mathbf{0.9}\varepsilon^2_{t-1}$$

Descriptive Statistics of an ARCH(1) Process and Standardized Process

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

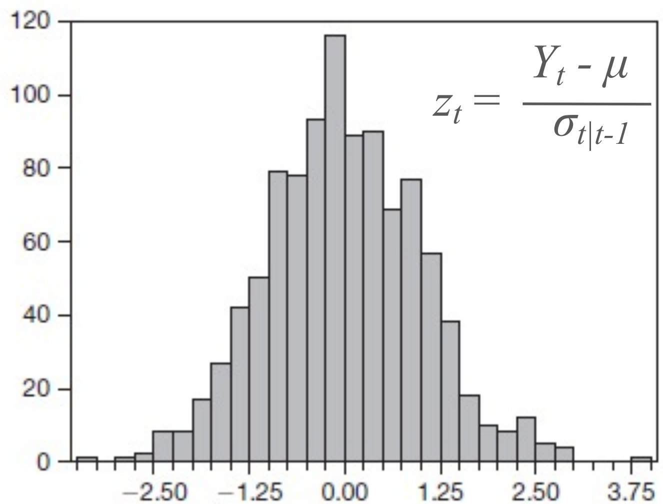
yt = ut/t-1 + epsilon t

μ (unconditional) remains constant ~ 2

σ (unconditional) increases with increasing α

render transform observation as standard normal

Panel B



$$z_t = \frac{Y_t - \mu}{\sigma_{t|t-1}}$$

Series: Standardized process: z = r/st.dev
Sample 1 1000
Observations 1000

Mean	-0.010666
Median	-0.038664
Maximum	3.980160
Minimum	-3.319720
Std. Dev.	1.003468
Skewness	0.141614
Kurtosis	3.204294
Jarque-Bera	5.081409
Probability	0.078811

If this is a true ARCH(1) process, then we expect the distribution of the standardized process to $\sim N(0,1)$

yt = ut|t-1 + epsilon t

epsilont = sigmat|t-1 zt

Fail to reject the normality hypothesis

Autocorrelation Functions of Simulated ARCH(1) Processes

$$Y_t = 2 + \varepsilon_t$$
$$\sigma^2_{t|t-1} = 2 + \mathbf{0.3}\varepsilon^2_{t-1}$$

Times series r_t $\quad Y_t$

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

data looks like white noise

- The series Y_t is uncorrelated
- No dynamics in the conditional mean →ACF and PACF →0

Times series r_t^2 $\quad Y^2_t$

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

know whether u can improve it or not

take residuals of project 2(already white noise) - squared it and feed it into acf/pacf may exhibit something like this - AR1

- The series Y^2_t is correlated with ACF and PACF →AR process
- Observe dynamics implied by the ARCH(1) Process

instead of ar, it would be arch

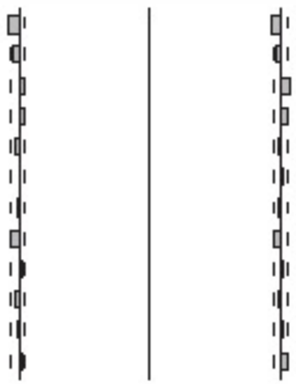
Autocorrelation Functions of Simulated ARCH(1) Processes

$$Y_t = 2 + \varepsilon_t$$

$$\sigma^2_{t|t-1} = 2 + \mathbf{0.3}\varepsilon^2_{t-1}$$

Times series $r_t / \sigma_{t|t-1}$ — $Y_t / \sigma_{t|t-1}$

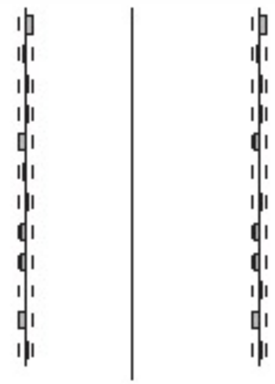
Sample: 1 1000
Included observations: 1000



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

Times series $r_t^2 / \sigma_{t|t-1}^2$ — $Y^2_t / \sigma^2_{t|t-1}$

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

No more correlations left. Both the ACF and PACF are 'clean'. Therefore, the ARCH(1) is accurate in capturing the dynamics of the Y_t process.

ARCH Models
Forecasting in an ARCH(1) Process

- Consider first the 1-step-ahead variance forecast, $h = 1$, ARCH(1):   inherit everything about the AR process

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

$\rightarrow h=2$: $\sigma^2_{t+2|t} = \omega + \alpha\sigma^2_{t+1|t}$

$\vdots$

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

as $h \rightarrow \infty$, $\sigma^2_{t+h|t} = \omega/(1-\alpha)$ following directly from the ar atchitecture

ARCH Models

- In general, for an ARCH(p) process:

$Y_t = \mu_{t|t-1} + \varepsilon_t = \mu_{t|t-1} + \sigma_{t|t-1} z_t$, where

$$\sigma^2_{t|t-1} = \omega + \alpha_1 \varepsilon^2_{t-1} + \alpha_2 \varepsilon^2_{t-2} + \dots + \alpha_p \varepsilon^2_{t-p}$$

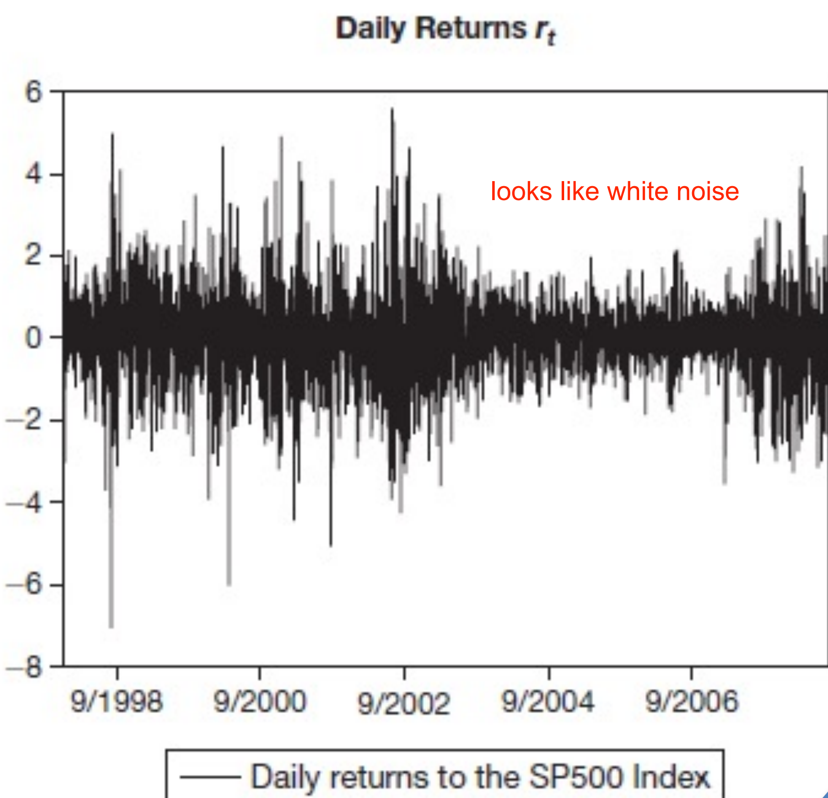
for $\omega > 0,\ \alpha_i \geq 0, i = 1,\ 2,\ \dots,\ p.$

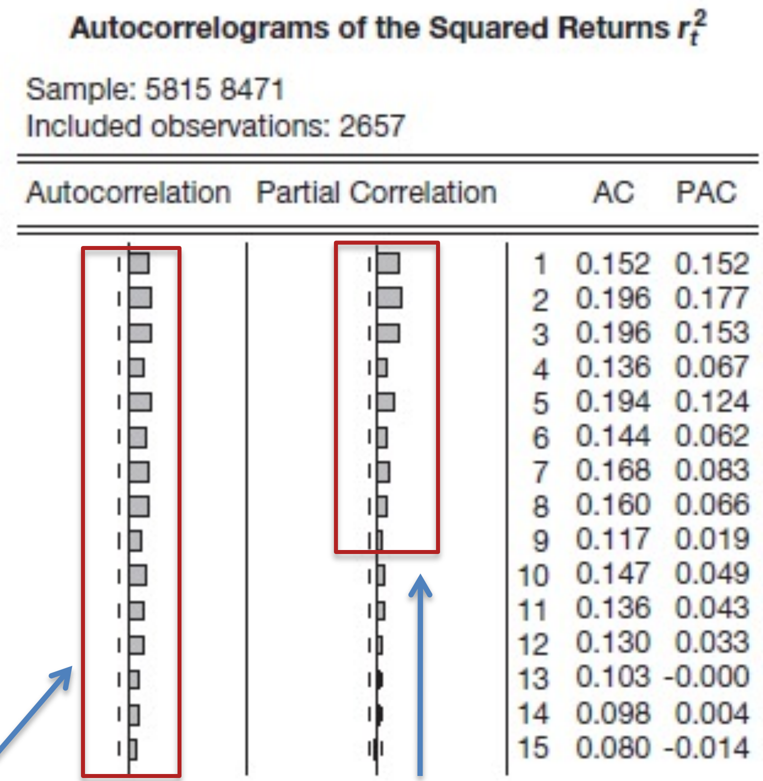
Note: $\alpha_1 + \alpha_2 + \dots + \alpha_p$ = Persistence in Variance

Example: ARCH(p) Process

Daily SP500 Returns and Autocorrelations of Squared Returns

Q: What type of a process is this?





$$\widehat{\alpha}_1 + \widehat{\alpha}_2 + \cdots + \widehat{\alpha}_9 \approx 0.83$$

High persistence

Try an ARCH(8) or ARCH(9)

would fit arma normally - too high order - switch to garch

GARCH Models

Example: GARCH(1,1) Process

- (1) What does a time series of a GARCH(1,1) process look like?

- Consider the following GARCH(1,1) process:

$$Y_t = \mu_{t|t-1} + \varepsilon_t = \mu_{t|t-1} + \sigma_{t|t-1} z_t,$$

carry over yesterday's volatility to today

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

Also depends on the most recent level of volatility

if no need for this, stick to arch

where $\omega > 0$, $\alpha \geq 0$, and $\beta \geq 0$.

For example, if β=0.7, we interpret this as saying that 70% of yesterday's variance carries over to today's variance.

GARCH Models

Example: GARCH(1,1) Process

- The key advantage of introducing the term $\beta\sigma^2_{t-1|t-2}$ is parsimony! → Fewer parameters to estimate than ARCH.
  - For example, the previous ARCH(9) model for S&P500 returns suggested an ARCH(9) –need 10 estimates- yet we can do the same with a GARCH(1,1) -only need 3 estimates.

- A GARCH(1,1) process is equivalent to an ARCH(∞) process with exponentially decreasing weights $\{\alpha, \alpha\beta, \alpha\beta^2, \ldots\}$.

- The Persistence of the GARCH(1,1) Process is equal to $\alpha + \alpha\beta + \alpha\beta^2 + \ldots = \alpha/(1-\beta)$.

beta approaching one - alpha not zero - infinity

GARCH Models

Example: GARCH(1,1) Process

$$Y_t = 2 + \varepsilon_t \quad \text{and} \quad \sigma^2_{t|t-1} = 2 + \alpha\varepsilon^2_{t-1} + \beta\sigma^2_{t-1|t-2}$$

- Case 1: Low Persistence process: $\alpha=0.4\ \beta=0.4$,

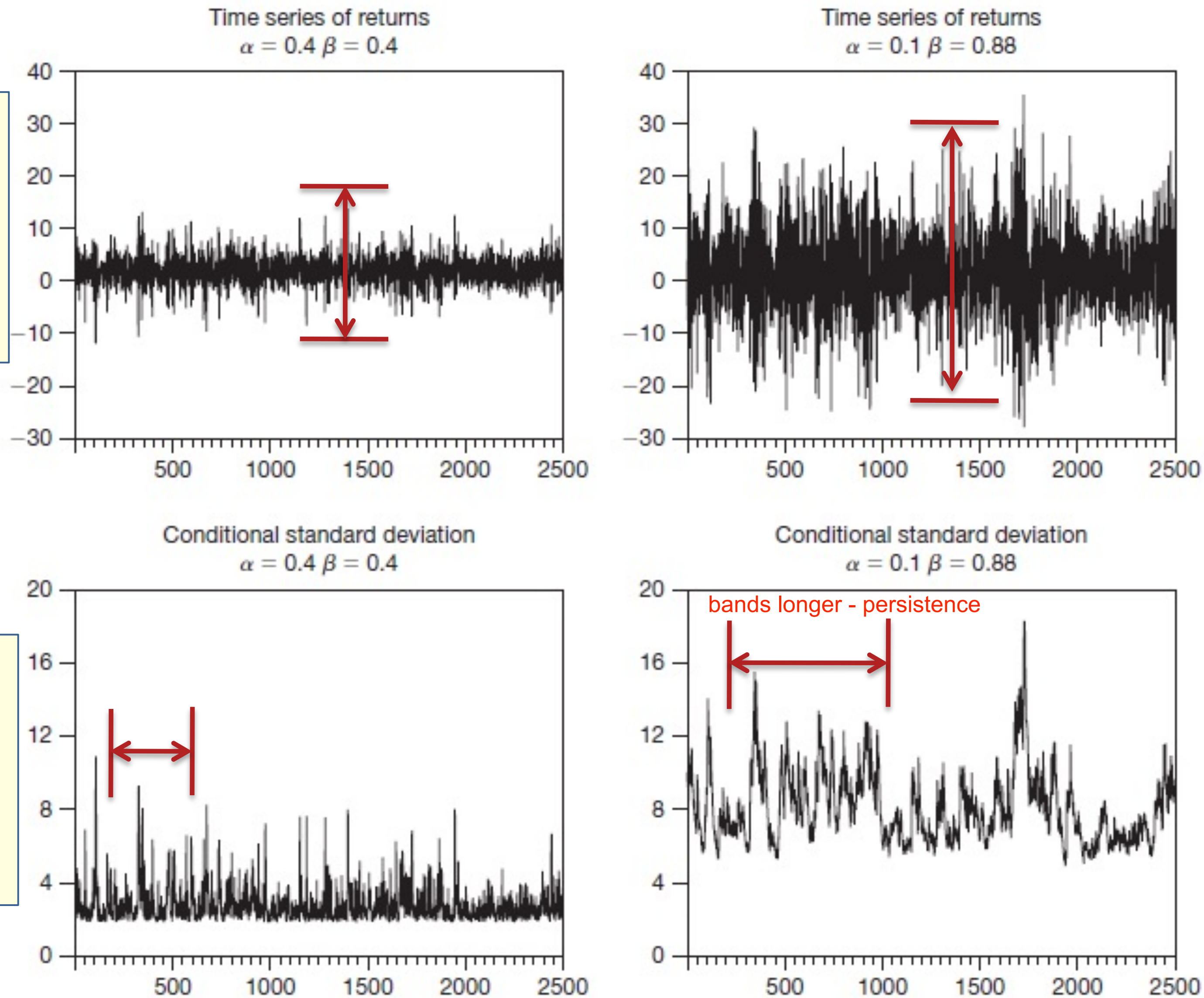
  $\rightarrow$ $persistence = \alpha/(1-\beta) = 0.67$

- Case 2: High Persistence process: $\alpha=0.1,\ \beta=0.88$,

  $persistence = \alpha/(1-\beta) = 0.83$

Simulated GARCH(1,1) Process

Low Persistence | **High Persistence**



For a high (low) persistence process, once the volatility is high, it tends to remain high (low).

For the low persistence process, only 40% of the past volatility is transferred to the current volatility.

Simulated GARCH(1,1) Process

Descriptive Statistics of a GARCH(1,1) process (returns) Sample: 1 20000		
	$\alpha = 0.4, \beta = 0.4$	$\alpha = 0.1, \beta = 0.88$
Mean	1.992061	2.019090
Median	1.998115	1.993319
Maximum	67.24605	84.91548
Minimum	-46.49060	-80.57394
Std. Dev.	3.281838	9.920762
Skewness	0.160152	0.116894
Kurtosis	26.86840	5.859899
Jarque-Bera	474835.8	6861.400
Probability	0.000000	0.000000

Simulated GARCH(1,1) Process

$$Y_t = 2 + \varepsilon_t \quad \text{and} \quad \sigma^2_{t|t-1} = 2 + \alpha\varepsilon^2_{t-1} + \beta\sigma^2_{t-1|t-2}$$

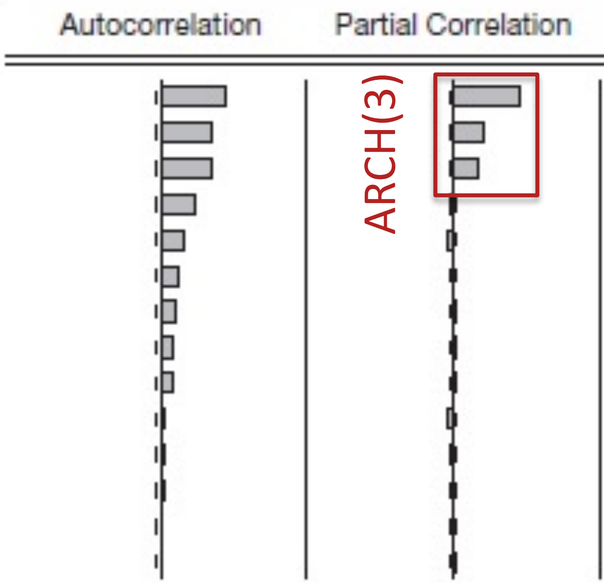
Time series r_t^2

square - significant spike in PACF - garch process - decay in ACF

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

Sample: 1 20000
Included observations: 20000

lower persistence - less dynamic to worry about



	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

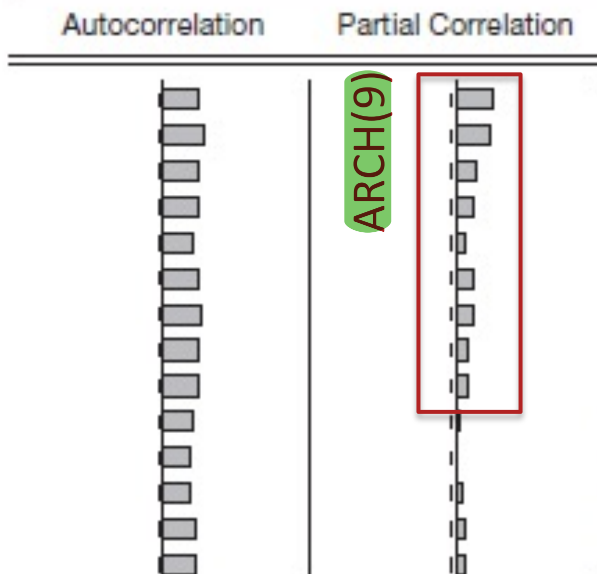
Low persistence → faster decay

go away with arch - economic interpretability

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

High persistence → slower decay

Example: GARCH(1,1) Process

Daily SP500 Returns and Autocorrelations of Squared Returns

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

Q: Was GARCH(1,1) a good model fit?

Example: GARCH(1,1) Process

Autocorrelation Function of the Standardized Squared Residuals from GARCH(1,1) for S&P500 Daily Returns

why use this to test residuals?

Sample: 5815 8471
Included observations: 2657

epsilon not dependent on sigma?

$$\widehat{\varepsilon}_t^2 / \widehat{\sigma}_{t|t-1}^2$$

looked pretty clean

A: Yes!

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

GARCH Models

Forecasting in a GARCH(1,1) Process

- Consider first the 1-step-ahead variance forecast, *h* = 1,

GARCH(1,1): $\sigma^2_{t+1|t} = \omega + \alpha\varepsilon^2_t + \beta\sigma^2_{t|t-1}$

$\rightarrow h=2$: $\sigma^2_{t+2|t} = \omega + (\alpha+\beta)\sigma^2_{t+1|t}$

$$\vdots$$

$$\sigma^2_{t+h|t} = \omega(1 + (\alpha+\beta) + (\alpha+\beta)^2 + \cdots + (\alpha+\beta)^{h-2}) + (\alpha+\beta)^{h-1}\sigma^2_{t+1|t}$$

as $h \rightarrow \infty$, $\sigma^2_{t+h|t} = \omega/(1-(\alpha+\beta)) = \sigma^2$(unconditional variance).

GARCH Models
GARCH(p,q)

$$Y_t = \varepsilon_t$$

$$\varepsilon_t | \Omega_{t\text{-}1} \sim N(0, \sigma^2{}_t)$$

$$\sigma^2{}_t = \omega + \alpha(L)\varepsilon^2{}_t + \beta(L)\sigma^2{}_t$$

higher order - do that in R

$$\omega > 0, \sum \alpha_i + \sum \beta_i < 1$$

$$\alpha(L) = \sum_{i=1}^{p} \alpha_i L^i \qquad \beta(L) = \sum_{i=1}^{q} \beta_i L^i$$

$$\boxed{\sigma_t^2 = \frac{\omega}{1 - \sum_i \beta_i} + \frac{\alpha(L)}{1 - \beta(L)}\varepsilon_t^2 = \frac{\omega}{1 - \sum_i \beta_i} + \sum_{i=1}^{\infty} \delta_i \varepsilon_{t-i}^2}$$

GARCH(p,q) ≈ ARCH(∞)

For Next Class

- Readings about today's class:
  Chapter 14[a] & 14[b]

- Review Exercises / Problems:
  Chapter 13[a]: 1, 2, 3, 4

- Readings for next class:
  Chapter 12 (Cointegration)