CHULALONGKORN BUSINESS SCHOOL
FLAGSHIP FOR LIFE

MSF Chula

Financial Econometrics

Lecture IX:

Serial Correlation & Time-varying Volatility Models

Narapong Srivisal, Ph.D.

Outline

Serial Correlation

- **Test for Serial Correlation**
  - AR Test
  - Durbin-Watson Test
- **Newey-West Standard Error**

Time-varying Volatility Models

- **ARCH**
- **GARCH**

$$Var[\hat{\beta}|X] = (X'X)^{-1}X' \; Var[U|X] \; X(X'X)^{-1}$$

$$\Downarrow$$

SKEDASTICITY

$$\begin{pmatrix} Var(U_1|X_1) & & & \\ Cov(U_2, U_1|X) & Var(U_2|X) & & \\ Cov(U_2, U_1|X) & Cov(U_3, U_2|X) & Var(U_3|X) & \dots \end{pmatrix}$$

Serial correlation

correlation of error term across observation

no problem if iid

but time series have this problem — cannot use ~~OLS~~ white-robust

* cannot newey west for cross section with serial correlation problem

ใช้ได้แค่กับ time-serie

SKEDASTICITY

Serial correlation	Homo	Hetero
without	reg y x reg y x, robust → consistent newey y x → may not efficient (hard to use)	reg y x, robust newey y x
with	• F/GLS → transform until get Homo then run OLS • new $se(\hat{\beta})$ that take problem into account ↓ known as " newey-west $se(\hat{\beta})$ newey y x	

Review (3 Nov 2023)

Time Series

* should declare first that this is time-series variable that index your time

1) STATA → tsset $\boxed{t}$

→ l. □ (r.v.) l2. □ (2 periods ago)

d. □

2) Check order of integration — check how many times you need to do difference before it becomes a stationary process

test whether we have stationary process or not

dfuller y ⇒ reject : I(0) — order of integration is 0

= stationary of itself

fail

dfuller d.y → reject : I(1)

fail

dfuller d2.y → reject

operator

run regression :

$\underline{y_t} = \beta_0 + \beta_1 \underline{\underline{y_{t-1}}} + u_t$ — x

$\underbrace{\Delta y_t}_{y} = \gamma_0 + \gamma_1 \underbrace{\Delta y_{t-1}}_{y} + u_t$

cannot use any function, but you can use this operation

reg y l.y

reg d.y l.d.y

or get x = l.y first

serial correlation problem → use Newey y x instead

you have $\{y_t\}$, $\{x_t\}$ after you check stationary process

1) Both I(0) ⇒ regular OLS with command reg

$y_t = \beta_0 + \beta_1 x_t + u_t$

allow effect of x many periods

$y_t = \beta_0 + \gamma_0 x_t + \gamma_1 x_{t-1} + \dots + \gamma_p x_{t-p} + u$

Finite distributed lag (FDL)

If you believe effect can last p period e.g. 4Q sales ⇒ use 4 lags

compare aic, bic

quarterly back up 1 yr

2) Both I(1) ⇒ try to find cointegration factor (to find relationship)

✓ ⇒ it has long run relationship

✗ ⇒ cannot do regression (no long run relationship)

3) difference order of integration ⇒ no relationship

Var [U | X]

covariance of error term & itself in the past

Cov [U_i, U_j)

cov(U_t, U_{t-k})

"autocorrelation"

skedasticity

serial correlation

serial correlation

Homoskedastic	without	serial correlation	→ default case → use reg (most efficient)
Heteroskedastic	without	"———"	→ use reg, robust
Homoskedastic	with	serial correlation	} ⇒ newey
Heteroskedastic	with	"———"	

to correct autocorrelation

newey standard error

1) assume fading serial correlation

2) either reg, robust or newey ⇒ doesn't change $\hat{\beta}^{OLS}$, change only standard error (inference)

3) has nothing to do with consistency / unbiased of $\hat{\beta}$

→ If you use OLS or F/OLS to solve problem will change $\hat{\beta}$

→ have to defend functional form (if don't know don't use!!)

Serial Correlation

- Serial correlation means that **the error terms are correlated across observations.**
- Serial correlation usually happens with time-series data, since each observation represents each period that is not independent of the other period.
- Some people refer to serial correlation problem in time-series as autocorrelation problem which means correlation with itself in other periods.
- Mathematically for a linear regression model $Y_t = X_t\beta + U_t$, there is serial correlation if for some period $s \neq t$, $Cov(U_t, U_s) \neq 0$
- In matrix notation, the variance of the vector of OLS estimator $\hat{\beta}$ is

$$Var\left[\hat{\beta}|\boldsymbol{X}\right] = (\boldsymbol{X}\boldsymbol{X}')^{-1}\boldsymbol{X}(Var[\boldsymbol{U}|\boldsymbol{X}])\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{X}')^{-1}$$

- Note: that $\boldsymbol{U}$ is the vector of error terms of all the observations, and $\boldsymbol{X}$ is the matrix of all the regressors of all the observations.

- As the case of heteroscedasticity, when there is serial correlation, the variance $Var[\boldsymbol{U}|\boldsymbol{X}]$ cannot be simplified:
  - **For the iid homoskedastic case,** $Var(\boldsymbol{U}|\boldsymbol{X}) = \sigma^2 I_n$
  - **For the iid heteroskedastic case,** $Var(\boldsymbol{U}|\boldsymbol{X})$ **is a diagonal matrix but each entry are not the same.**
  - **For serial correlation,** $Var(\boldsymbol{U}|\boldsymbol{X})$ **is not diagonal.**
- Therefore, we cannot use the simple OLS standard error or White heteroskedasticity robust standard error like before.
- Note: it is also possible to have homoskedasticity with serial correlation.
- As for heteroskedasticity, **serial correlation does not affect unbiasedness or consistency of the OLS estimator** as long as the error term is uncorrelated with regressors.
- However, **it has impact on inference**, as we need correct standard error of the estimator.

① AR Tests for Serial Correlation

check whether error term correlate with itself in the past

- Since serial correlation means that there is correlation between the error term and its lags, the simplest way is to hypothesize that the error term follows an AR(1) process: $U_t = \rho U_{t-1} + \epsilon_t$
- Since the actual process of the error terms $\{U_t\}$ is unobserved, we need to make a test on the estimated residual.
  - Run OLS: $Y_t = \beta X_t + U_t$ and keep the residuals $\{\widehat{U}_t\}$
  - If the regressor is exogenous, run OLS regression:

$$\widehat{U}_t = \rho \widehat{U}_{t-1} + \epsilon_t$$

up to 1 lag ⇒ AR(1)

  - If the regressor is not exogenous, run OLS regression:

$$\widehat{U}_t = \delta X_t + \rho \widehat{U}_{t-1} + \epsilon_t$$

  - Test the hypothesis: $H_0: \rho = 0 \; vs \; H_1: \rho \neq 0$ using the standard t-test.

- Note: the AR(1) test can also capture serial correlation of higher order, e.g. if the error term process follows AR(p); $p > 1$ or other kinds of serial correlation, as long as $cov(U_t, U_{t-1}) \neq 0$
- Note: we can also use the heteroskedasticity robust standard error in the t-test if suspecting that $Var[\epsilon_t | \widehat{U}_{t-1}]$ is not constant.
- We can as also extend the test to higher order AR(p) process:

$$\widehat{U}_t = \rho_1 \widehat{U}_{t-1} + \rho_2 \widehat{U}_{t-2} + \cdots + \rho_p \widehat{U}_{t-p} + \epsilon_t, \text{ or}$$

$$\widehat{U}_t = \delta X_t + \rho_1 \widehat{U}_{t-1} + \rho_2 \widehat{U}_{t-2} + \cdots + \rho_p \widehat{U}_{t-p} + \epsilon_t$$

AR(p)

- Then, doing the usual F-test for the hypothesis:

- no serial correlation ⇒ use F test or LM test

$$H_0: \rho_1 = \rho_2 = \cdots = \rho_p = 0 \; vs \; H_1: \rho_1 \neq 0 \; or \ldots or \; \rho_p \neq 0$$

\# obs in data set

- We can also use LM statistic: $LM = (n - p) R^2_{\widehat{U}} \sim \chi^2_p$

include p lags → lose p obs

- The LM test here is called **Breusch-Godfrey Test** for AR(p)

\# obs × R^2 = test statistics

check

Breusch-Godfrey Test for Serial Correlation

```
. tsset time
        time variable:  time, 1 to 30
                delta:  1 unit

. reg y x1 x2, nohead
------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |  -.6670978   .0655501   -10.18   0.000    -.8015955   -.5326002
          x2 |  -.7792461   .1072015    -7.27   0.000    -.9992055   -.5592867
       _cons |   66.21805    6.31436    10.49   0.000     53.26206    79.17405
------------------------------------------------------------------------------

. estat bgodfrey , lag ( )

Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
    lags(p)  |          chi2               df                 Prob > chi2
-------------+-------------------------------------------------------------
       1     |          5.932               1                   0.0149
---------------------------------------------------------------------------
                        H0: no serial correlation
```

LM stat

p-value

< 5%

reject H_0 at 5%, 10%

there is serial correlation at 5%, 10% significant level

Durbin-Watson Test

X not recommend !
too old → based on 1 lag only
cannot test serial correlation more than 1 lag

- This test proposed by Durbin and Watson (1950) and is also based on the AR(1) model, but against the one-sized alternative hypothesis

$$H_0: \rho = 0 \; vs \; H_1: \rho > 0$$

- The Durbin-Watson test statistics is

$$DW = \frac{\sum_{t=2}^{T}\left(\hat{U}_t - \hat{U}_{t-1}\right)^2}{\sum_{t=1}^{T} \hat{U}_{t-1}^2} \approx 2(1-\hat{\rho})$$

- Under the null that $\rho = 0, DW \approx 2$, and we will reject the null if the DW test statistic is significantly less than 2.
- A drawback of the Durbin-Watson test is that the distribution of DW under H_0 is not in a form of well-known distributions, but numerically calculated and provided by the authors.

- The DW distribution also depends on the number of regressors used in the residual regression.
- So, if a statistical software does not contain a command for the Durbin-Watson test that automatically calculate the p-value, we must use the table for DW statistics, which provides two sets of critical values: the upper bound dU and lower bound dL.
- We can reject the null if $DW < dL$
- We cannot reject the null if $DW > dU$

two kind of critical value

- **If $dL < DW < dU$, then we cannot conclude whether we can reject the null hypothesis.**

Durbin-Watson Test for Serial Correlation

```
. reg y x1 x2, nohead
------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |  -.6670978   .0655501   -10.18   0.000    -.8015955   -.5326002
          x2 |  -.7792461   .1072015    -7.27   0.000    -.9992055   -.5592867
       _cons |   66.21805    6.31436    10.49   0.000     53.26206    79.17405
------------------------------------------------------------------------------

. estat dwatson

Durbin-Watson d-statistic(  3,    30) =  1.116148
```

t-stat

- critical value

Durbin-Watson Statistic: 5 Per Cent Significance Points of dL and dU

	k'*=1		k'=2		k'=3		k'=4		k'=5		k'=6		k'=7		k'=8	
n	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU
28	1.328	1.476	1.255	1.560	1.181	1.650	1.104	1.747	1.028	1.850	0.951	1.959	0.874	2.071	0.798	2.188
29	1.341	1.483	1.270	1.563	1.198	1.650	1.124	1.743	1.050	1.841	0.975	1.944	0.900	2.052	0.826	2.164
30	1.352	1.489	1.284	1.567	1.214	1.650	1.143	1.739	1.071	1.833	0.998	1.931	0.926	2.034	0.854	2.141
31	1.363	1.496	1.297	1.570	1.229	1.650	1.160	1.735	1.090	1.825	1.020	1.920	0.950	2.018	0.879	2.120
32	1.373	1.502	1.309	1.574	1.244	1.650	1.177	1.732	1.109	1.819	1.041	1.909	0.972	2.004	0.904	2.102
33	1.383	1.508	1.321	1.577	1.258	1.651	1.193	1.730	1.127	1.813	1.061	1.900	0.994	1.991	0.927	2.085
34	1.393	1.514	1.333	1.580	1.271	1.652	1.208	1.728	1.144	1.808	1.079	1.891	1.015	1.978	0.950	2.069
35	1.402	1.519	1.343	1.584	1.283	1.653	1.222	1.726	1.160	1.803	1.097	1.884	1.034	1.967	0.971	2.054

Newey-West Standard Error

- Recall that for the OLS estimator, the covariance matrix of $\hat{\beta}$ is

$$Var[\hat{\beta}|\boldsymbol{X}] = (\boldsymbol{X}\boldsymbol{X}')^{-1}\boldsymbol{X}(Var[\boldsymbol{U}|\boldsymbol{X}])\,\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{X}')^{-1}$$

- In the case of heteroskedasticity without serial correlation, we can use White standard error which estimates $Var[\boldsymbol{U}|\boldsymbol{X}]$ by

$$S = \sum_{t=1}^{T} \hat{U}_t^2 X_t X_t'$$

- With the presence of serial correlation, this S is not enough to provide a consistent estimator for $Var[\boldsymbol{U}|\boldsymbol{X}]$, since it ignores non-zero correlations among $\hat{U}_t$

$cov(U_t, U_{t-k})$ not y!!!

self error correlated with itself in the past

- A better estimator is the **Heteroskedasticity and Autocorrelation Consistent (HAC) standard error** or well-known as **Newey-West Standard Error**, proposed by Newey and West (1987) up to 4 period

e.g. $H=4 \Rightarrow j=1 \Rightarrow 1-\frac{j}{H+1} = 1-\frac{1}{5} = \frac{4}{5}$

$j=2 \Rightarrow \quad = 1-\frac{2}{5} = \frac{3}{5}$

$j=3 \Rightarrow \quad = \frac{2}{5}$

$j=4 \Rightarrow \quad = \frac{1}{5}$

- The Newey-West Standard Error:

capture heteroskedasticity — gap btw period — capture serial correlation / autocorrelation — cov across obs

$$\boxed{\sum_{t=1}^{T} \widehat{U}_t^2 X_t X_t'} + \sum_{j=1}^{H} \left\{ \left(\mathbf{1} - \frac{\boldsymbol{j}}{\boldsymbol{H}+\mathbf{1}} \right) \sum_{s=j+1}^{T} \widehat{\boldsymbol{U}}_s \widehat{\boldsymbol{U}}_{s-j} \left[\boldsymbol{X}_s \boldsymbol{X}_{s-j}' + \boldsymbol{X}_{s-j} \boldsymbol{X}_s' \right] \right\}$$

$1 - \frac{1}{5} = \frac{4}{5}$

- Where the second term captures the covariances of $\widehat{U}X$ in each period and its j-period lags.

- Note that this standard error assume that there are serial correlations up to H lags. That is $E[X_t U_t] = E[U_t U_{t-j}] = 0$ for $j \neq 1, 2, \dots, H$.

(after that, no serial correlation anymore

- The expression $1 - \frac{j}{H+1}$ is a weight of impact of each covariance term. Technically, this is to ensure that standard error is positive definite. Intuitively, this means that the impact of autocorrelation of j periods apart diminishes when j increases.

MSF Chula

CHULALONGKORN BUSINESS SCHOOL
FLAGSHIP FOR LIFE

Estimating Static Model with Newey-West Standard Error

```
. reg y x1 x2
------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |  -.6670978   .0655501   -10.18   0.000    -.8015955   -.5326002
          x2 |  -.7792461   .1072015    -7.27   0.000    -.9992055   -.5592867
       _cons |   66.21805    6.31436    10.49   0.000     53.26206    79.17405
------------------------------------------------------------------------------
```

H
more lag means lose obs
e.g. monthly data → up to 12 lags or 1.5 yr (18 lags)

```
. newey y x1 x2, lag(2)

Regression with Newey-West standard errors          Number of obs     =        30
maximum lag: 2                                      F(  2,        27) =     48.81
                                                    Prob > F          =    0.0000

------------------------------------------------------------------------------
             |             Newey-West
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |  -.6670978   .0722839    -9.23   0.000    -.8154121   -.5187836
          x2 |  -.7792461   .1169575    -6.66   0.000    -1.019223   -.5392691
       _cons |   66.21805   7.017444     9.44   0.000     51.81945    80.61666
------------------------------------------------------------------------------
```

change only s.d. not estimator

MSF Chula

CHULALONGKORN BUSINESS SCHOOL
FLAGSHIP FOR LIFE

people concern about risk

Volatility

- Volatility is defined as **standard deviation of the unpredictable part of the series**.
- For example, if return r_t is modeled as

$$r_t = X_t\beta + U_t$$

Then, the return volatility is the standard deviation of U_t.

- In the mainstream Finance literature nowadays, volatility could be time-varying.
- This means that $Var[U_t|X_t]$ is not constant across observations t, i.e. the return data could potentially be heteroscedastic.

return : $\{r_t\}$

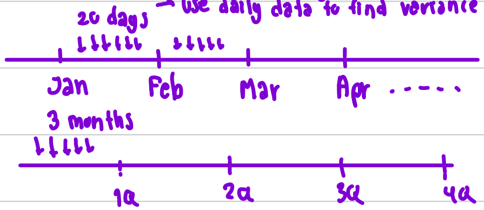
risk : standard deviation of r_t

$r_t = \bar{r}$

$\sigma_r^2 = \frac{1}{T}\sum_t [(r_t - \underbrace{\bar{r}}_{=0 \text{ if there's one obs}})^2]$ ⇒ need >1 obs

e.g. want to estimate volatility during a month

→ use daily data to find variance or s.d.



UNCONDITIONAL VOLATILITY
- let data speak for itself
- but, in finance, cannot find data with high frequency → e.g. want to find daily volatility
  - need hourly data

CONDITIONAL VOLATILITY

ARCH model of Volatility

Steps
1) construct model * hard
2) run test
3) inference

- $\boxed{\text{Engle (1982)}}$ proposed that the volatility is related to past periods' errors, as it's generally observed in financial markets that high volatility periods are followed by periods with low volatility and vice versa.

link past U

variance of error term change across period

use MLE maximum likelihood estimator

- Engle's **Autoregressive Conditional Heteroskedasticity of order q** or **ARCH(q)** model:

parameter for mean eq

Var(U_t) is not constant, order

shocks

* β & α run together bcs we cannot observe U

$$Y_t = X_t\beta + U_t$$

ARCH(q)

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{q} \alpha_j U_{t-j}^2$$

don't have to have many periods

variance today

depends on error term in the past → want to model variance of shocks

where $U_t = \sigma_t z_t$ and $z_t \sim N(0,1)$. So, σ_t^2 is the square of the volatility

* always have 2 equations in model

if you change Y & X → meaning of error term also change

- The first equation is called **Mean Equation**, and the second equation is called **Volatility Equation**.

functional form of volatility

- To ensure that variance is always positive, the model has restrictions that all the coefficients in the volatility equation are positive, i.e. $\alpha_0, \alpha_1, \dots, \alpha_q > 0$

model 1 :

$$R_t = \beta_0 + u_t$$

total shock that comprises both mkt movement & firm specific

u_t of ① > u_t of ②

need to have mean equation !!!

model 2 :

extra from mkt

idiosyncratic shock

$$R_t = \beta_0 + \underbrace{\beta_1 R_t^M + \beta_2 SMB_t + \beta_3 HML_t}_{\text{mkt movement}} + u_t$$

☆ Volatility model construction

① Think about mean equation → what is your Y & X ⇒ define U_t that you want to estimate

② After get "————" → form volatility equation

2.1) functional form — name of model

- ARCH (q)
- GARCH (p,q)
- EGARCH — exponencial
- TGARCH — threshold
- GJR -

2.2) # lags

↳ AIC, BIC

lowest info criteria help you to choose # lags

GARCH model of Volatility

- **Generalized Autoregressive Conditional Heteroskedasticity** or **GARCH** model is an extension of the ARCH model by allowing the volatility to also depend on its history.
- **GARCH(p, q)** model is

add lag

GARCH

ARCH

variance today depends on error-term in past

$$Y_t = X_t\beta + U_t$$

$$\sigma_t^2 = \alpha_0 + \boxed{\sum_{i=1}^{p} \gamma_i \sigma_{t-i}^2} + \sum_{j=1}^{q} \alpha_j U_{t-j}^2$$

- Similar to ARCH, all the coefficients in the volatility equation are restricted to be positive value.
- One of the popular specifications used in estimation of stock market volatility in the finance literature is GARCH(1,1)
- These ARCH or GARCH models could be estimated by MLE.

Property of

- ARCH (3) = $\sigma_t^2 = \alpha_0 + \alpha_1 U_{t-1}^2 + \alpha_2 U_{t-2}^2 + \alpha_3 U_{t-3}^2$

drawback of this model :

1) restrict $\alpha_0 > 0,\ \alpha_1, \alpha_2, \ldots \geq 0$ to make variance positive

2) only shock 3 periods ago can affect

⇒ lose #obs if we need to add many lag
cannot allow shock to have very long effect on volatility

3) Positive / negative shock or U ⇒ will be the same

people normally response to negative news more

Positive or negative U_{t-j} affect volality in the same way

GARCH

generalized of ARCH

e.g. GARCH (1, 1) — 1 lag for GARCH, 1 lag for ARCH

$\sigma_t^2 = \alpha_0 + \gamma \sigma_{t-1}^2 + \alpha_1 U_{t-1}^2$

allow effect of shocks for many period by just add this

$\alpha_0 + \gamma \sigma_{t-2}^2 + \alpha_1 U_{t-2}^2$

$\alpha_0 + \gamma \sigma_{t-3}^2 + \alpha_1 U_{t-3}^2$

estimate volality is not purpose, just want proxy of volatility

follow GARCH (1,1) to estimate volatility

not have to justify anything

EGARCH ($\sigma^2 = e$) — going up & down not same (also capture asymmetric shock

TGARCH — capture a asymmetric shock

GJR - GARCH

Lagrange Multiplier Test for ARCH

close price of Toyota Motor

keep error term

Mean Equation: $close_{tm_t} = \beta_0 + U_t$

regular OLS

```
. reg close_tm
```

close_tm	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_cons	104.9591	.5628959	186.46	0.000	103.8549	106.0632

```
. estat archlm, lag(3)
```

test whether error term have ARCH structure or not

heterodasticity

LM test for autoregressive conditional heteroskedasticity (ARCH)

lags(p)	chi2	df	Prob > chi2
3	1437.509	3	0.0000

→ reject H_0 → there's evidence for ARCH structure

ไม่ได้บอกว่า ต้องใช้ lag เท่าไร่

there is ARCH structure

H0: no ARCH effects vs. H1: ARCH(p) disturbance

Estimating ARCH(3) model with mean equation: $close_{tm_t} = \beta_0 + U_t$

mean equation   volatility equation   arch y x_1 x_2 ,

```
. arch close_tm, arch(1/3) nolog
ARCH family regression

Sample: 1 - 1468                                Number of obs     =      1,468
Distribution: Gaussian                          Wald chi2(.)      =          .
Log likelihood = -5745.079                      Prob > chi2       =          .
------------------------------------------------------------------------------
             |                 OPG
    close_tm |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
close_tm     |
       _cons |   117.4195   .1100404  1067.06   0.000     117.2038    117.6352
-------------+----------------------------------------------------------------
ARCH         |
        arch |
         L1. |   .8458274   .1695703     4.99   0.000     .5134758    1.178179
         L2. |   .2017088   .1010716     2.00   0.046     .0036121    .3998055
         L3. |  -.0494428   .0429668    -1.15   0.250    -.1336562    .0347705
             |
       _cons |   1.653013   .3041491     5.43   0.000     1.056892    2.249134
------------------------------------------------------------------------
```

MSF Chula

CHULALONGKORN BUSINESS SCHOOL
FLAGSHIP FOR LIFE

AACSB ACCREDITED · EFMD EQUIS ACCREDITED · EFMD EPAS ACCREDITED

e.g. close price persistence or not? ↗ check on β_1
do hypothesis testing

Estimating GARCH(1,1) model with mean equation:

$$close_{tm_t} = \beta_0 + \beta_1 close_{tm_{t-1}} + U_t$$

$\sigma_t^2 = \alpha_0 + \alpha_1 U^2_{t-1} + \gamma_1 \sigma^2_{t-1}$

y x_1 x_2 ...

```
arch close_tm l.close_tm, arch(1) garch(1) nolog nohead
```

mean equation   # of lags for ARCH   # of lags for GARCH

```
------------------------------------------------------------------------------
             |                 OPG
    close_tm |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
close_tm     |
    close_tm |
         L1. |   .9979838   .0016113   619.36   0.000     .9948257    1.001142
             |
       _cons |   .2420924   .1742509     1.39   0.165    -.0994331     .583618
-------------+----------------------------------------------------------------
ARCH         |
        arch |
         L1. |    .121934   .0173797     7.02   0.000     .0878704    .1559975
             |
       garch |
         L1. |   .7751161   .0353553    21.92   0.000      .705821    .8444111
             |
       _cons |   .2112655   .0494103     4.28   0.000     .1144232    .3081078
------------------------------------------------------------------------------
```

$\hat{\beta}_0$, $\hat{\beta}_1$ — mean equation

$\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\gamma}_1$ — volatility equation

ปกติคน ไม่ inference

σ^2 รู้อันนี้?

to generate new value and keep it

```
. predict vol_tmg11, var
. tsline vol_tmg11
```

→ tell STATA to estimate σ^2

σ^2

(plot r.v. against time

or

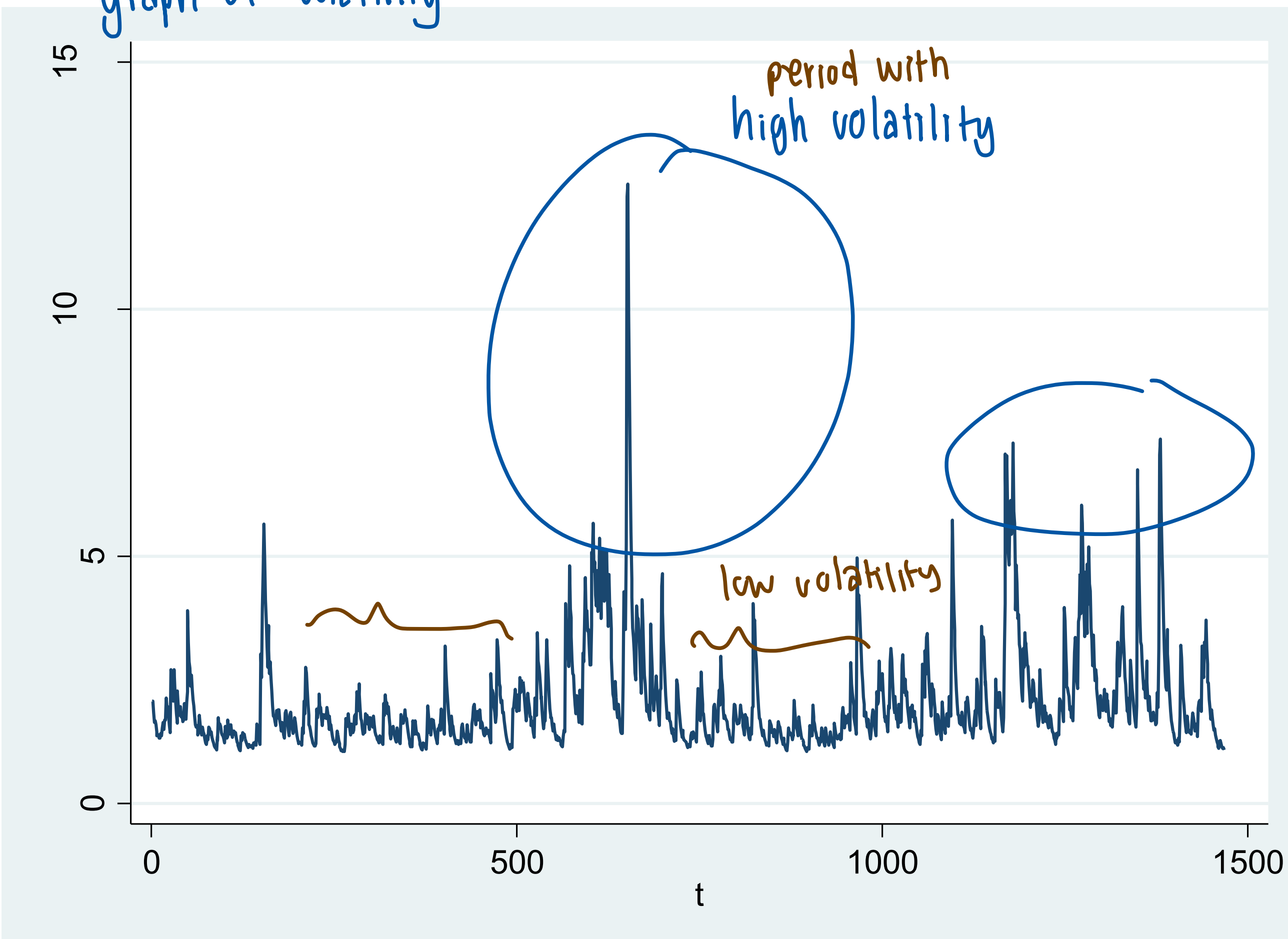
gen vol = (vol_tmg11)^(0.5)

= sqrt( )

กราฟ vol_tmg 11

graph of volatility

t vol
1 .
2 .
3 .
: :



time-series

create another r.v.
↙ run with this instead

↳ ใช้ excel ง่ายสุด

t	year	month	time
1	2000	1	2000 m1
2	2000	2	2000 m2
3			
4			
⋮			