x is linear combination of other x (dummy trap) → drop some x

Problem: multicollinearity & β not unique — OLS estimator not consistent [$Cov(u,x) \neq 0$] — Endogeneity — Heteroskedasticity: affect $se(\hat{\beta})$, not estimation — "will mean independent" — causation may have this

Model construction → Estimation → Inference

Financial Econometrics Final :

3 interpretation:
① $E[y|x]$ linear conditional exp (unbiased ($E[u|x]=0$))
② BLP $[y|x]$ best linear prediction
③ Causation (pure effect x on y)
} consistent ($Cov(u,x)=0$)

I Skedasticity, GLS & FGLS Estimators

$y = \beta x + u \longrightarrow \hat{\beta}^{OLS} = (x'x)^{-1}x'y$

$Var[\hat{\beta}|x] = (x'x)^{-1}x'(Var[u|x])x(x'x)^{-1}$

vce matrix

$$\begin{pmatrix} var[\hat{\beta}_0|x] & cov(\ ) & \cdots & cov(\ ) \\ & var[\hat{\beta}_1|x] & & \\ & & \ddots & \\ cov(\ ) & & & var[\hat{\beta}_k|x] \end{pmatrix} \qquad \begin{pmatrix} Var[u_1|x] & cov & \cdots & cov \\ cov(u_2,u_1|x) & Var[u_2|x] & & \\ cov(u_3,u_1|x) & & \ddots & \\ & & & Var[u_n|x] \end{pmatrix}$$

σ^2_{11}, σ^2_{22}, σ^2_{kk}; cov = 0

* SERIAL CORRELATION — no this if cross sectional analysis (<assume iid) except time series

* $Var[u|x]$ doesn't depend on x

SKEDASTICITY: $Var[u_i|x_i]$ — doesn't affect $E[u|x]$ or $E[xu]$, no impact on unbiasedness & consistency of OLS

* $Var[u|x]$ depends on x

STATA assume this **Homoskedasticity = 1**

$Var[u|x] = Var[u] = \sigma^2 I_n$

simplified formula:

$\star\ Var[\hat{\beta}|x] = \sigma^2(x'x)^{-1}x'I_n x(x'x)^{-1}$ ($x'x$)

$= \sigma^2(x'x)^{-1}$

and asymptotic variance simplifies to $\Sigma = E[xx']^{-1}Var[u]$

- Near perfectly collinear, (xx') and $E[xx']^{-1}$ very large

→ variance and $se(\hat{\beta})$ very large as well → t-stat ↓ hard to reject or detect effect x on y

★ Homo $se(\hat{\beta}_j) = \sqrt{\hat{\sigma}_{jj}/n}$ — $se(\hat{\beta})$ < robust $se(\hat{\beta})$; t-stat $= \frac{\hat{\beta}-\beta}{se(\hat{\beta})}$ ↓ hard to reject H_0

more efficient than White's robust for homoske (ใช้ได้เฉพาะไม่ใช่ Homo)

Heteroskedasticity = more than 1

* cannot use simplified formula

★ $Var[u_i|x] = \sigma_i^2$ for each obs i

2 ways to proceed:
1) Construct & use another standard error
★ White's Robust SE ($\hat{\beta}$)
2) Use other estimator than OLS → FGLS estimator

★ Test for heteroskedasticity: (all interpretations)
① Breusch-Pagan test
as $Var[u|x] = E[(u-E[u])^2] = E[u^2]$

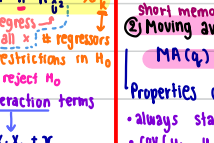
proxy $\hat{u}^2 = \delta_0 + \delta_1 x_1 + \dots + \delta_k x_k + v$

$\hat{u}^2 = \delta_0 + \delta_1 x_1 + \dots + \delta_k x_k + v$

$H_0: \delta_1 = \dots = \delta_k = 0$ (homo) vs. $H_1: \delta_1 \neq 0$ or … or $\delta_k \neq 0$ (reject H_0 = hetero)

STATA
H_0: constant variance (homo)
variable: $x_1\ x_2 \dots$
chi2 (df) = LM stat (# regressors)
Prob > chi2 = p-value
< α → reject H_0

★ test stat:
① F-test $\sim F_{k,n-k-1}$
② Lagrange multiplier stat: $LM = n \cdot R^2_{\hat{u}^2} \sim \chi^2_k$ (# of obs used to run regression; better x can explain $\hat{u}$ ... chi-square)



when regress $\hat{u}^2$ for all x' (# regressors or # restrictions in H_0)
further from H_0 → reject H_0

STATA
H_0: homoskedasticity
H_a: unrestricted hetero
chi(df) = LM stat
Prob > chi2 = p-value

③ White test: involves x x' interaction terms

$\hat{u}^2 = \delta_0 + \sum_{j=1}^{k}\delta_j x_j + \sum_{j=1}^{k}\sum_{r=1}^{j}\gamma_{jr}x_j x_r + v$

H_0: all the coefficients = 0 except δ_0

df = 5 = M

$LM = n \cdot R^2_{\hat{u}^2} \sim \chi^2_M$ (# of regressors (รวม inter) ที่เราต้ง ×100)

ถ้ามีใน $x_1 = x_2^2$ & dummy

e.g.
$y = \beta_0 + \beta_1 fe + \beta_2 edu + \beta_3 edu \cdot fe + u$
x: fe, edu, edu·fe
x^2: fe^2, edu^2, $(edu \cdot fe)^2$
xx: fe·edu, $fe^2 \cdot edu$, fe·edu^2
$\hat{u}^2 = \delta_0 + \delta_1 fe + \delta_2 edu + \delta_3 edu \cdot fe + \delta_4 edu^2 + \delta_5 fe \cdot edu^2$

★ GLS estimator (Generalized Least square / weighted LS): transform to homo

Given $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i$

$\frac{1}{x_{1i}} Var[u_i|x_i] = \frac{\sigma^2 x_{1i}}{x_{1i}} = \sigma^2 = h^2(x)$ (constant) — need to justify this

Thus, $Var\left[\frac{u_i}{\sqrt{x_{1i}}}\middle|x_i\right] = \sigma^2$ (1) — ★ assume functional form of hetero

$\frac{y}{h(x)} = \beta_0\left(\frac{1}{h(x)}\right) + \beta_1\left(\frac{x_1}{h(x)}\right) + \dots + \beta_k\left(\frac{x_k}{h(x)}\right) + \frac{u}{h(x)}$

★ FGLS estimator (Feasible)
estimate h rather than make assumption
e.g. $h = \exp(\alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k)$
$\ln h = \alpha_0 + \alpha_1 x_1 + \dots + \alpha_k x_k$
$\ln \hat{u}^2 = \alpha_0 + \alpha_1 x_1 + \dots + \alpha_k x_k + \varepsilon$
$\hat{h} = \exp(\hat{\alpha}_0 + \hat{\alpha}_1 x_1 + \dots + \hat{\alpha}_k x_k)$ → use this

II Endogeneity, Instrumental Variable, and 2SLS estimator:

$Cov(x_j, u) \neq 0$ ($E[xu] \neq 0$) — * x_j is exogenous if $Cov(x_j, u) = 0$

OLS estimator is inconsistent & biased

e.g. $y = \beta_0 + \beta_1 x + (\gamma w + u)$

$\hat{\beta}_1^{OLS} = \frac{Cov(y,x)}{Var(x)} \xrightarrow{WLLN} \frac{Cov(y,x)}{Var(x)} = \beta_1 + \frac{Cov(\gamma w, x) + Cov(u,x)}{Var(x)} = \beta_1 + \frac{\gamma\, Cov(w,x)}{Var(x)}$ — $\hat{\beta}_1^{OLS}$ is consistent iff $Cov(w,x) = 0$

★ Cause of endogeneity:

① Omitted variable bias: $y = \beta_0 + \beta_1 x + (\beta_2 w + u)$ — OLS is consistent if $E[xw] = 0$

$w = \alpha_0 + \alpha_1 x + \varepsilon$ — when x↑ 1 unit, w↑ α_1

omitted variable model → $y = \gamma_0 + \gamma_1 x + \eta$ — when x↑1, we will not get pure effect of x if not control for w (get only correlation)

$\hat{\gamma}_1 = \hat{\beta}_1 + \hat{\alpha}_1\hat{\beta}_2$ — contaminated effect from w

OLS is consistent if $Cov(\tilde{x}, \varepsilon) = 0$

② Measurement error:

reg y x (omitted model) — pure effect of x — get $\hat{\alpha}_1$ from reg w x — get $\hat{\beta}_1$ & $\hat{\beta}_2$ from reg y x w

Ex1 wage = $\beta_0 + \beta_1 edu + controls + (a + \varepsilon)$
a: innate capabilities of people — left in error term as it cannot be captured — and cause endogeneity
brokerage acct size

Ex2 $y = \beta_0 + \beta_1 x + u$
income + cannot be observed
reported income = income + noise (measurement error)
$\tilde{x} = x + \varepsilon$ (observed)

$y = \beta_0 + \beta_1(\tilde{x} - \varepsilon) + u$
$y = \beta_0 + \beta_1\tilde{x} + (u - \beta_1\varepsilon)$

NOTE: $Cov(x+\varepsilon, \varepsilon) = Cov(x,\varepsilon) + Var(\varepsilon)$
$Var(ax+b\varepsilon) = a^2 Var(x) + b^2 Var(\varepsilon) + 2ab\,Cov(x,\varepsilon)$

$Cov(\tilde{x}, \varepsilon) \neq 0$ → endogeneity problem — $\tilde{x} = x + \varepsilon$

reg y $\tilde{x}$ → $\hat{\beta}_1^{OLS} \xrightarrow{WLLN} \frac{Cov(y,\tilde{x})}{Var(\tilde{x})} = \beta_1 - \beta_1\left(\frac{Cov(\tilde{x},\varepsilon)}{Var(\tilde{x})}\right) = \beta_1\left(1 - \frac{Var(\varepsilon)}{Var(x)+Var(\varepsilon)}\right) = \beta_1\left(\frac{Var(x)}{Var(x)+Var(\varepsilon)}\right)$ — >1, <1, $\neq \beta_1$

$\hat{\beta} \downarrow$, β

* Biased toward zero: estimate is closer to zero than what it's suppose to be

If we do inference: $H_0: \beta = 0$ vs. $H_1: \beta \neq 0$

- Reject H_0 with $\hat{\beta}^{OLS}$ ✓ ∴ Safe to say that we reject H_0 in population as well as we know direction of biasedness ∴ $\beta \neq 0$

! fail to reject H_0 with $\hat{\beta}$ OLS — cannot make any conclusion → move to use new estimator

③ Reversed causality: $y = \beta_0 + \beta_1 x + u$ but reg x y → endogeneity bcs $Cov(y,u) \neq 0$ → solve by switch x & y (OLS not consistent)

④ Simultaneous equations: x & y satisfy various equations at the same time. card = $\gamma_0 + \gamma_1$ cash + controls + η

Ex $q^D = \beta_0 + \beta_1 x + u = \gamma_0 + \gamma_1 x + \eta = q^S$ (correlated) — cash holding in economy — $\beta_0 + \beta_1$ card wage + controls + u

$Cov(x,\eta) \neq 0$ and $Cov(x,u) \neq 0$

1st solution: move from transaction data → survey data (not straight forward) e.g. find willingness to pay for given price

* Resolution for endogeneity — $E[xu] \neq 0$ OLS x — consistent & biased!

▸ Instrument (z):
→ exogeneity: $Cov(z,u) = 0$ or $E[zu] = 0$
→ relevance: $Cov(z,x) \neq 0$

for bivariate model: use $E[u] = 0 \Rightarrow \beta_0 = E[y] - \beta_1 E[x]$; $E[zu] = 0 \Rightarrow \beta_1 = \frac{Cov(z,y)}{Cov(z,x)}$

Then use★ IV estimator: $\hat{\beta}_0^{IV} = \bar{y}_n - \left(\frac{\hat{\sigma}_{zy}}{\hat{\sigma}_{zx}}\right)\bar{x}_n \xrightarrow{p} \beta_0$; $\hat{\beta}_1^{IV} = \frac{\hat{\sigma}_{zy}}{\hat{\sigma}_{zx}} \xrightarrow{p} \frac{Cov(z,y)}{Cov(z,x)} = \beta_1$

1 endogenous variable but several z: $Cov(z,u) = 0$, $Cov(z,x) \neq 0$ (linear combination e.g. $\alpha_0 + \alpha_1 z_1 + \alpha_2 z_2$)

→ we need to check that $\alpha_1 Cov(z_1,x) + \alpha_2 Cov(z_2,x) + \dots + \alpha_m Cov(z_m,x) \neq 0$

→ Most efficient way is to use best linear projection of x: $x = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + \dots + \pi_m z_m + \eta$

BLP $[x, z_1, \dots, z_m] = \hat{x}^{OLS}$ as instrument

Two-step least squares
★ 2SLS or TSLS estimator: as $\hat{\sigma}_{\hat{x}x} = \hat{\sigma}^2_{\hat{x}} \Rightarrow \hat{\beta}_1^{IV} = \frac{\hat{\sigma}_{\hat{x}y}}{\hat{\sigma}_{\hat{x}x}} = \frac{\hat{\sigma}_{\hat{x}y}}{\hat{\sigma}^2_{\hat{x}}}$

STATA
• ivregress 2sls y exogenous regressors (endo regressor = instruments)
coef. $\hat{\beta}$, Std Err. $se(\hat{\beta})$, z $\frac{\hat{\beta}}{se(\hat{\beta})}$

Overidentifying test:
* we need $Cov(z,u) = 0$ but we cannot observe u

Suparat Patarasupanit 6684069326 MSF_FT27

3 cases: # endogenous variable (q), # instrumental variable (m)
① $q < m$: over identified → overidentify test
② $q = m$: just identified
③ $q > m$: under → cannot be cannot estimated

STATA
• estat overid
Sargan (score) chi2 (1) = test-stat (p-value)
Basmann chi2 (1) = " ——— "
Hansen
H_0: there're sufficient valid instruments
reject this if p-value < 10%, 5%, or 1%)
* df of both tests = # instruments − # endogenous variables in model

III Intro to Time Series Analysis: 1 individual, many consecutive periods without gaps

$\{y_t\} \Rightarrow y_1, y_2, y_3, \dots, y_{t-1}, y_t, y_{t+1}, \dots$ (correlated not iid)

▸ random time-series process is called a stochastic process.
▸ as it is not iid → cannot use WLLN and CLT from cross-sectional analysis
▸ STATIONARY:
• Strictly stationary: requires the joint dist of obs in adjacent periods does not change overtime.
• Weak Stationary (covariance): all t, k, $E[y_t]$, $E[y_t^2]$, and $Cov(y_t, y_{t+k})$ does not depend on t.
→ requires up to 2nd moment (mean & variance)

requires 3 things:
① $E[y_t] = \mu$
② $E[y_t^2]$ or $Var[y_t]$ → constant
③ $Cov(y_t, y_{t+k})$ → can depend on k but not t (gap btw period)
e.g. $Cov(y_1, y_3) = Cov(y_t, y_{t-2})$ (k=2, same gap should have same covariance)

• White-noise process: sequence of r.v. $\{e_t\}$ such that $E[e_t] = 0$, $E[e_t e_s] = 0$, $E[e_t^2]$ is constant
random shock happening in each period

▸ UNIVARIATE TIME SERIES MODELS: AR, MA, ARMA process

① Autoregressive (AR process): predicts future value Y_t from past outcomes $\{Y_{t-1}, Y_{t-2}, \dots\}$ (long memory → as k↑, ρ^k approach zero but not equal to zero)

AR(1): $y_t = \rho y_{t-1} + e_t$ (white noise)

3 things to check: Stationary
1) $E[y_t] = E[\rho y_{t-1} + e_t] = \rho E[y_{t-1}] + E[e_t] \Rightarrow \mu = \rho\mu \Rightarrow \mu = 0$
2) $Var[y_t] = Var[\rho y_{t-1} + e_t] \Rightarrow \sigma_y^2 = Var[\rho y_{t-1}] + \sigma_e^2 = \frac{\sigma_e^2}{1-\rho^2} \Rightarrow |\rho| < 1$
3) how much info in k periods still reflect in y_t
$Cov(y_t, y_{t-1}) = Cov(\rho y_{t-1} + e_t, y_{t-1}) = Cov(\rho y_{t-1}, y_{t-1}) + Cov(e_t, y_{t-1}) = \rho\sigma_y^2$ (bcs $Cov(e_t, e_{t-1}) = 0$)
$Cov(y_t, y_{t-2}) = Cov(\rho^2 y_{t-2} + \rho e_{t-1} + e_t, y_{t-2}) = \rho^2\sigma_y^2$
$\rho[\rho y_{t-2} + e_{t-1}] + e_t$
$Cov(y_t, y_{t-k}) = \rho^k\sigma_y^2$
$Corr(y_t, y_{t-k}) = \frac{\rho^k\sigma_y^2}{\sigma_{y_t}\sigma_{y_{t-k}}} = \rho^k$ if stationary

② Moving average (MA) Process: weighted average of white noises of q periods
$MA(q): y_t = e_t + \alpha_1 e_{t-1} + \dots + \alpha_q e_{t-q}$ — implies magnitude of shock but not tell about persistent — as $|\rho| < 1 \Rightarrow \rho^{k1}$ ↓ fading memory

Properties of MA(q): 1) $E[y_t] = 0$ | 2) $Var[y_t] = (1 + \alpha_1^2 + \alpha_2^2 + \dots + \alpha_q^2)\sigma_e^2$
• always stationary
• $Cov(y_t, y_{t-k}) = 0$ when $k > q$ → short-memory process
3) $Cov(y_t, y_{t-k})$ doesn't depend on t but on k (no 2 cov( ) terms & $Var(e_t)$ is constant (white noise))

▸ MA(1): $Cov(y_t, y_{t-1}) = Cov(e_t + \alpha_1 e_{t-1}, e_{t-1} + \alpha_1 e_{t-2}) = \alpha_1\sigma_e^2$
$Cov(y_t, y_{t-2}) = 0$ no overlap

▸ MA(3): $Cov(y_t, y_{t-1}) = (\alpha_1 + \alpha_1\alpha_2 + \alpha_2\alpha_3)\sigma_e^2$
เขียนไล่ขึ้นก่อน
$y_t = e_t + \alpha_1 e_{t-1} + \alpha_2 e_{t-2} + \alpha_3 e_{t-3}$
$y_{t-1} = e_{t-1} + \alpha_1 e_{t-2} + \alpha_2 e_{t-3} + \alpha_3 e_{t-4}$

* Autocorrelation function (ACF): tells lengths strengh of memory of process

$corr(y_t, y_{t-k}) \Rightarrow \gamma_k = \frac{Cov(y_t, y_{t-k})}{\sqrt{Var(y_t)Var(y_{t-k})}} = \frac{Cov(y_t, y_{t-k})}{Var(y_t)} = \frac{Cov(y_t, y_{t-k})}{Var(y_{t-k})}$ → function of k not t
($\in [-1,1]$; gap k)

→ AR(1): $Cov(y_t, y_{t-k}) = \rho^k\sigma_y^2 \Rightarrow ACF(k) = \rho^k$
if $|\rho|$ is low → ACF diminishes quickly
if $|\rho|$ close to 1 → ACF decreases very slow
AR(1) has strong memory & highly persistent

→ ACF of MA(q) is zero after q lags

" ACF & model selection "
→ Test whether ACF(k) = 0 ⇒ $y_t = \beta_0 + \gamma_k y_{t-k} + u_t$ (t = (k+1), …, T)
$H_0: \gamma_k = 0$ vs. $H_1: \gamma_k \neq 0$
$\sqrt{T}(\hat{\gamma}_k - \gamma_k) \xrightarrow{d} N(0, v_k)$
→ t-stat $= \frac{\sqrt{T}\hat{\gamma}_k}{\sqrt{1 + 2\hat{\gamma}_1^2 + 2\hat{\gamma}_2^2 + \dots + 2\hat{\gamma}_{k-1}^2}} \sim N(0,1)$ (# obs; special variance v_k)

STATA e.g.
reg y_t y_{t-1} → $\hat{\gamma}_1^{OLS} \neq 0$
reg y_t y_{t-2} → $\hat{\gamma}_2^{OLS} \neq 0$
reg y_t y_{t-10} → $\hat{\gamma}_{10}^{OLS} = 0$
} sample ≠ population
then, do hypothesis testing
if fail to reject → choose MA(9)

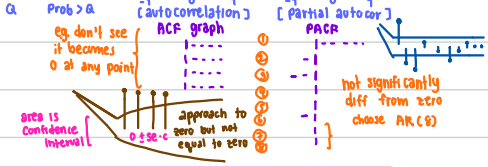
* Partial autocorrelation coefficient function (PACF) → for AR(p)

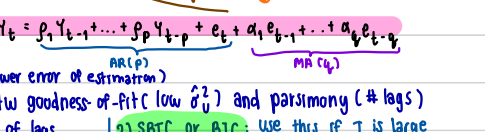
PACF(1) is ρ_{11} in AR(1): $Y_t = \rho_{11} Y_{t-1} + u_t$
PACF(2) is ρ_{22} in AR(2): $Y_t = \rho_{21} Y_{t-1} + \rho_{22} Y_{t-2} + u_t$ (if this is actually AR(1))
PACF(k) is ρ_{kk} in AR(k): $Y_t = \rho_{k1} Y_{t-1} + \rho_{k2} Y_{t-2} + \dots + \rho_{kk} Y_{t-k} + u_t$

▸ If AR(p) is the right model, then coefficient of terms Y_{t-k} for k>p should be zero ⇒ PACF become roughly 0 after p lags.

" PACF & model Selection "
→ Test whether PACF(k) is zero ⇒ $Y_t = \rho_{k1} Y_{t-1} + \dots + \rho_{kk} Y_{t-k} + u_t$
$H_0: \rho_{kk} = 0$ vs. $H_1: \rho_{kk} \neq 0$
→ Test-stat: $\sqrt{T}\hat{\rho}_{kk} \xrightarrow{d} N(0,1)$ (# obs; nn result; PACF(k))
If fail to reject H_0 from k on, then use AR(k−1)

STATA: LAG | AC | PAC | Q | Prob > Q
need to do hypothe whether it is significantly diff from 0
AC: estimated ACF; PAC: estimated PACF





can be non-stationary ARMA(1,0)

③ ARMA process: ARMA(p,q): $Y_t = \rho_1 Y_{t-1} + \dots + \rho_p Y_{t-p} + e_t + \alpha_1 e_{t-1} + \dots + \alpha_q e_{t-q}$
AR(p) | MA(q)
→ more lag → lose obs but more regressors → better fit (lower error of estimation)

* Information criteria: trade off btw goodness-of-fit (low $\hat{\sigma}_u^2$) and parsimony (# lags)

1) AIC: tends to suggest large # of lags
$AIC = \log\hat{\sigma}_u^2 + \frac{2(p+q)}{T}$ (# obs) — when T is large this approaches 0

2) SBIC or BIC: use this if T is large
$BIC = \log\hat{\sigma}_u^2 + \frac{(p+q)}{T}\log(T)$ (estimated var for u_t)

→ model with lowest value of information criterion is preferred

STATA
• tsset t ← tell STATA that this is time-series
• arima name of r.v., arima (p, d, q) (integration = 0 for stationary process; e.g. arima (1,0,1) is ARMA(1,1))
(estimating model)
• estat ic → then look for AIC & BIC

not OLS but MLE; consistent but biased → need large data set

run regression:
$y_t = \beta_0 + \beta_1 y_{t-1} + u_t \Rightarrow$ reg y l.y
$\Delta y_t = \gamma_0 + \gamma_1 \Delta y_{t-1} + u_t \Rightarrow$ reg dy l.dy

$\frac{\partial}{\partial x}\left(\frac{u}{v}\right) = \frac{vu' - uv'}{v^2}$ (ระวังเครื่องหมาย)

$\frac{\partial}{\partial inc}(1 + e^{x\beta}) = \frac{\partial}{\partial inc}e^{x\beta} = e^{x\beta} \cdot \frac{\partial}{\partial inc}(x\beta)$

If $|\rho| \geq 1 \Rightarrow$ non-stationary

▸ Non-Stationary process → best guess for next period $E[Y_{t+h} | Y_t] = Y_t$, $var(Y_t) = t \cdot \sigma_e^2$; $Y_{t+h} = Y_t + e_{t+1} + \dots + e_{t+h}$

① Random walk process : AR(1) with $\rho = 1$: $Y_t = Y_{t-1} + e_t$

② Unit Root Process : $Y_t = Y_{t-1} + \eta_t$ (η_t is allowed to be any stationary process e.g. MA(q)) → unit root is highly persistent but not necessarily trending.
($\rho = 1$)

* Random walk with drift : $Y_t = \beta_0 + Y_{t-1} + e_t$ (drift term)
$Y_t = \beta t + e_1 + e_2 + \dots + e_t$ (slope → trend) — unlike $|\rho|$ that capture persistency

$\rho \uparrow$ → more persistent, not fluctuate, smooth
↑ tells correlation of itself in the past

★ Dickey-Fuller Test for a Unit Root : <reject = stationary>

From the process : $Y_t = \rho Y_{t-1} + \eta_t$

▸ Hypothesis testing : $H_0 : \rho = 1$ vs. $H_1 : |\rho| < 1 \Rightarrow$ hard to do, don't know dist

transform to: $\Delta Y_t = (\rho - 1) Y_{t-1} + e_t \equiv \theta Y_{t-1} + e_t$ ($Y_t - Y_{t-1}$)

▸ Hypothesis : $H_0 : \theta = 0$ vs. $H_1 : \theta < 0$

▸ t-stat : $t_{\hat\theta} = \frac{\hat\theta^{OLS}}{SE(\hat\theta^{OLS}_\theta)}$ ~ Dickey-Fuller dist ; we reject H_0 if $t_{\hat\theta} < C \Rightarrow$ implies $|\rho| < 0$
เป็นตารางค่า t ไม่ใช่ t-stat!!! or there's evidence that $\{Y_t\}$ is not a unit root process

STATA • dfuller Y_t , nocon → may get non-stationary (fail to reject)
• dfuller ΔY_t , nocon (I(1))

Result : Test-stat $t_{\hat\theta}$ < interpolated dickey-fuller critical value 1% 5% 10% ⇒ reject
z(t)

Integration : $Y_t - Y_{t-1} = \eta_t$ becomes stationary ; integration of order ρ or I(ρ)

d.y $\{y_t\}$ → Dickey fuller → reject ✓ → Stationary process by itself = I(0)

$\Delta y_t = \{y_t - y_{t-1}\}$ → " ——— " reject ✓ → I(1) (t-tail)

$\Delta y_t - \Delta y_{t-1}$ (d2 y) → " ——— " reject ✓ → I(2) (t-tail)

If causation of vol_t on tax_t → inconsistent OLS → 2SLS

* Relationship across Stationary variables :
• Static time-series model : $Y_t = \beta_0 + \beta_1 X_{1t} + \dots + \beta_k X_{kt} + U_t$

• Finite distributed lag (FDL) model : e.g. $Vol_t = \beta_0 + \gamma_0 tax_t + \gamma_1 tax_{t-1} + \gamma_2 tax_{t-2} + U_t$

temporary change : consider if govt ↑ tax 1 unit from c to c+1 on period t and revert back

$vol_{t-1} = \beta_0 + \gamma_0 C + \gamma_1 C + \gamma_2 C$
* $vol_t = \beta_0 + \gamma_0 (C+1) + \gamma_1 C + \gamma_2 C$
$vol_{t+1} = \beta_0 + \gamma_0 C + \gamma_1 (C+1) + \gamma_2 C$
$vol_{t+2} = \beta_0 + \gamma_0 C + \gamma_1 C + \gamma_2 (C+1)$
$vol_{t+3} = \beta_0 + \gamma_0 C + \gamma_1 C + \gamma_2 C$

t	x_t	x_{t-1}	x_{t-2}	⇒ y_t
-2	0	0	0	β_0
-1	0	0	0	β_0
* 0	1	0	0	$\beta_0 + \gamma_0$
1	0	1	0	$\beta_0 + \gamma_1$
2	0	0	1	$\beta_0 + \gamma_2$
3	0	0	0	β_0

impact propensity | multiplier
↳ immediate effect = $Vol_t - Vol_{t-1} = \gamma_0$
one-period lagged effect = γ_1
two " = γ_2

Permanent change : เพิ่ม 1 หน่วย period
$vol_{t-1} = \beta_0 + \gamma_0 C + \gamma_1 C + \gamma_2 C$
$vol_t = \beta_0 + \gamma_0 (C+1) + \gamma_1 C + \gamma_2 C$ → immediate = γ_0
$vol_{t+1} = \beta_0 + \gamma_0 (C+1) + \gamma_1 (C+1) + \gamma_2 C$ → 1 lagged = $\gamma_0 + \gamma_1$
$vol_{t+2} = \beta_0 + \gamma_0 (C+1) + \gamma_1 (C+1) + \gamma_2 (C+1)$ → 2 lag : $\gamma_0 + \gamma_1 + \gamma_2$
$vol_{t+3} = \beta_0 + \gamma_0 (C+1) + \gamma_1 (C+1) + \gamma_2 (C+1) = \gamma_0 + \gamma_1 + \gamma_2$ → long run multiplier

* Cointegration : if two process $\{X_t\}$ & $\{Y_t\}$ both non-stationary (don't move together in SR)
↳ may be long run relationship
↳ $\{X_t\}$ & $\{Y_t\}$ are cointegrated if both are I(1) processes and there's a constant γ such that $Y_t - \gamma X_t$ is I(0)
e.g. aggregate consumption & income

IV Serial correlation & time-varying volatility models :

▸ $cov(U_t, U_{t-k})$ is not zero in time series → $Var[U|X]$ is not diagonal → "autocorrelation"

▸ Use Newey-West Standard Error for both homo|heteroscedastic with serial correlation. (not change $\hat\beta^{OLS}$ (se))
↳ 1) assume fading serial correlation 2) not affect consistency & biasedness of $\hat\beta$ → change only inference

→ not mean that $cov(X, U) \neq 0$ → OLS still consistent

* Breusch-Godfrey test : AR tests for serial correlation check whether error term correlate with itself in past
Run OLS : $Y_t = \beta X_t + U_t$ and keep residuals $\{\hat U_t\}$
AR(p) process : $\hat U_t = \rho_1 \hat U_{t-1} + \rho_2 \hat U_{t-2} + \dots + \rho_p \hat U_{t-p} + \epsilon_t$ (or) $\hat U_t = \delta X_t + \rho_1 \hat U_{t-1} + \rho_2 \hat U_{t-2} + \dots + \rho_p \hat U_{t-p} + \epsilon_t$

▸ Hypothesis : $H_0 : \rho_1 = \rho_2 = \dots = \rho_p = 0$ vs. $H_1 : \rho_1 \neq 0$ or ... or $\rho_p \neq 0$ (no serial correlation)

▸ t-stat (F or LM) : $LM = (n - p) R_0^2 \sim \chi^2_p$ (n = #obs, p = # lags → lose p obs)

STATA • estat bgodfrey, lags(p) — chi 2, df, LM stat, prob>chi2
if < α% reject H_0 then, there's serial correlation
p-value

* Durbin-Watson test : not recommend, too old & based on 1 lag & if dL < DW < dU cannot conclude

* Newey-West Standard error (HAC) (new period) (cov across obs) (capture serial correlation)

$$\sum_{t=1}^{T} \hat U_t^2 X_t X_t' + \sum_{j=1}^{H} \left\{ \left(1 - \frac{j}{H+1}\right) \sum_{s=j+1}^{T} \hat U_s \hat U_{s-j} [X_s X_{s-j}' + X_{s-j} X_s'] \right\}$$

capture hetero — assume there're serial correlations up to H lags — autocorrelation of j periods apart diminishes when j increases

STATA • reg y x1 x2
• newey y x1 x2, lag (H)

RESULT : y | coef. (same as OLS) | Newey se | t | P>|t|

* Volatility → can be time-varying, $Var[U_t | X_t]$ is not constant across obs t ⇒ return data could be heteroscedastic

① ARCH (q) model :

1st mean equation : $Y_t = X_t \beta + U_t$ (parameter) — $Var(U_t)$ is not constant <hetero> (shocks) — EX $R_t = \beta_0 + U_t$ (total shock) — use MLE
* β & α run together bcs we cannot observe U
$R_t = \beta_0 + \beta_1 R_t^m + \beta_2 SMB_t + \beta_3 HML_t + U_t$ (idiosyncratic shock)

2nd volatility equation : $\sigma_t^2 = \alpha_0 + \sum_{j=1}^{q} \alpha_j U_{t-j}^2$
where $U_t = \sigma_t z_t$ and $z_t \sim N(0,1) \Rightarrow \sigma_t^2$ is square of volatility
functional forms of volatility : ARCH(q), GARCH(p,q), EGARCH, TGARCH, GJR, etc. (heteroskedasticity models)

GARCH e.g. GARCH (1,1) (ARCH)
$\sigma_t^2 = \alpha_0 + \gamma \sigma_{t-1}^2 + \alpha_1 U_{t-1}^2$ (allow effect of shocks for many periods)

3rd Determine # lags by using AIC, BIC

Drawbacks of ARCH(3) : $\sigma_t^2 = \alpha_0 + \alpha_1 U_{t-1}^2 + \alpha_2 U_{t-2}^2 + \alpha_3 U_{t-3}^2$
1. restrict $\alpha_0, \alpha_1, \dots, \alpha_q > 0$ 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 U_{t-j} affect volatility in same way ✗ (capture asymmetric shock) — EGARCH, TGARCH, GJR-GARCH
↳ people normally response this more

STATA : LM test for ARCH (evidence for ARCH structure)
H_0 : no ARCH effects vs. H_1 : ARCH(p) disturbance
• estat archlm, lag (q) (lags(p)) — chi 2, df, Prob>chi2 (p-value < α) — reject H_0

STATA GARCH
• arch mean eq, arch (1) garch (1) (#lags)
RESULT : Y coef. — X_1 $\hat\beta_1$ } mean equation; cons $\hat\beta_0$ — ARCH $\hat\alpha_1$, $\hat\gamma_1$ } volatility equation
• arch mean eq, arch(1|3) nolog (volatility eq)

V MLE & Limited Dependent Variable Models :

▸ Binary responses model : e.g. prob that individual has brokerage acct.
$Y \sim$ Bernoulli : $E[y] = 1 \cdot p + 0 \cdot (1-p) = p$
$E[y|X] = E[\beta X + U | X]$ (assume 0) ⇒ OLS unbiased & consistent
$Pr\{y = 1 | X\} = \beta X + E[U|X]$

Linear probability model (LPM)
$Pr\{Y = 1 | X\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$
when X_1 ↑ 1 unit → $Pr\{Y=1|X\}$ ↑ β_1 unit
* $\hat\beta X$ may not in [0,1] as there's no restriction for $\hat Y_i$ to be [0,1]
↳ function $G(\cdot) \in [0,1]$
$P\{Y_i = 1 | X_i\} = G(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_{ki}) = G(X_i \beta)$

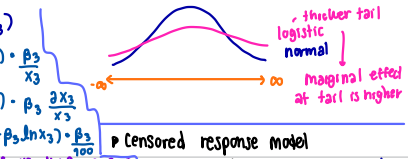
① Probit model : assume $U_i \sim$ iid $N(0,1)$
$Y_i^* = X_i \beta + U_i$; $Y_i = 1$ if $Y_i^* > 0$
$P\{Y_i = 1 | X_i\} = \Phi(X_i \beta)$ (CDF standard normal ⇒ use z table to find prob)
e.g. $X\hat\beta = 2.31$ → $Pr\{y=1|X\} = 98.96\%$

② Logit Model : G(·) is log function
$P\{Y = 1 | X\} = \Lambda(X\beta) = \frac{\exp(X\beta)}{1 + \exp(X\beta)}$
e.g. $X\hat\beta = 2.31 \Rightarrow Pr\{y=1|X\} = \frac{e^{2.31}}{1+e^{2.31}} = 90.97\%$

* Marginal effect and interpretation of β's :
$Pr\{y=1|X\} = G(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 \ln X_3)$ (CDF)
$\frac{\partial}{\partial X_1} Pr\{y=1|X\} = PDF(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 \ln X_3) \times \beta_1$ (difference by diff people) (direction depends on β_1)

Formula of PDF :
• Probit : $g(X\beta) = \frac{1}{\sqrt{2\pi}} \exp(-(X\beta)^2)$ (higher effect & higher likelihood)
• Logit : $g(X\beta) = \frac{\exp(X\beta)}{(1+\exp(X\beta))^2}$

$Pr\{y=1|X\} = G(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 \ln X_3)$
$\frac{\partial Pr\{y=1|X\}}{\partial X_3} = PDF(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 \ln X_3) \cdot \frac{\beta_3}{X_3}$
$Pr\{y=1|X\} = PDF(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 \ln X_3) \cdot \beta_3 \frac{\partial X_3}{X_3}$
∴ Marginal effect of $X \uparrow 1\% = PDF(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 \ln X_3) \cdot \frac{\beta_3}{100}$



STATA LR chi2 (k) → t-stat ← H_0 : all β = 0 vs. H_1 : Some β ≠ 0
* estimating marginal effect :
• mfx, predict(p) at (age = 30, edu = 21) dy/dx std z P>|z| [95%C.I] x
variable age edu — marginal effect = pdf · β — 30 21 (s: year) (column von studies)

▸ Censored response model

① Tobit Model : strongly assume that error term is iid normally distributed and homoskedastic. * If U|X not normal, Tobit is not good estimator
$Y_i^* = X_i \beta + U_i$ (income), where $U_i | X_i \sim$ iid $N(0, \sigma^2)$ for β
$Y_i = \max(c, Y_i^*)$ (real spending) (condition for MLE to be consistent)
↳ if Y_i^* isn't high enough

• β is not effect of X on Y but on Y^* (willingness to invest) (actual level of investment in stock mkt)

↳ cannot use OLS (biased & inconsistent)
↳ sensor from below "lower level of zero"

STATA : • tobit y x, ll(0) (ul(0) sensor from above)
y | coef. | Std er. — use to do inference

• Apply MLE technique
• Likelihood of observing $Y_i = c$ is
$P\{Y_i = c | X_i\} = P\{U_i < c - X_i\beta | X_i\} = \Phi\left(\frac{c - X_i\beta}{\sigma}\right)$
$E[Y|X] = \left\{ X\beta + \sigma \frac{\phi(\frac{X\beta}{\sigma})}{\Phi(\frac{X\beta}{\sigma})} \right\} \cdot \Phi\left(\frac{X\beta}{\sigma}\right)$ not linear

STATA * Marginal effect : expected value
• mfx, predict (e(0,.)) $E[y|y>0]$
$y = E[y|y>0]$ for whole sample
mean of y before change
dy/dx marginal effect
• mfx, predict (p(0,.)) prob (y>0)
$y = Pr(y>0)$ = prob before change
change that will invest increased by — dy/dx (x ? whose effect)

VI Intro to panel data analysis :

▸ Use long format → i t Y X_1 X_2 ⇒ # obs = # of individual × T (time)

many cross-sectional individual with many periods ≠ repeated cross-sectional data (cannot link indiv) survey 1st time 2nd time
doesn't need consecutive periods

* 3 linear regression models :

① Pooled cross-section : treat each data point (individual i at time t) as one observation & apply usual cross-sectional methods <can use OLS, IV, 2SLS, P|GLS> ⇒ as no endogeneity (assume iid)
$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_k X_{kit} + U_{it}$ โดยมาก estimator
• Structural break : when relationship change at some point → use dummy variable
$Y_{it} = \beta_0 + \beta_1 X_{1it} + \dots + \beta_k X_{kit} + \gamma_0 D_{it} + \gamma_1 D_{it} X_{1it} + \dots + \gamma_k D_{it} X_{kit} + U_{it}$
↳ Test whether two groups have different relationship "Chow test" ← a kind of F-test
$H_0 : \gamma_0 = \gamma_1 = \dots = \gamma_k = 0$ vs. $H_1 : \gamma_0 \neq 0$, or $\gamma_1 \neq 0$, or ..., or $\gamma_k \neq 0$

② Fixed Effects Model : there's unobserved heterogeneity is correlated with regressors → $cov(X, \alpha) \neq 0$ (left in error term) — cannot use OLS (endogeneity)
$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + (\alpha_i + U_{it})$ & $cov(X, \alpha_i) \neq 0$
α_i is time invariant, specific to each individual, unobserved or unquantifiable

Solutions 1) Use IV or 2SLS estimator to deal with endogeneity ← need to find more data to run with Z (instrument) (hard!)
โดยมาก FD, FE, 2SLS|IV estimators
2) take α_i out from error term by having dummy for each individual
↳ $Y_{it} = \beta_0 + \alpha^A D_i^A + \alpha^B D_i^B + \beta_1 X_{1it} + \beta_2 X_{2it} + U_{it}$
3) Kill α_i from equation

▸ First difference estimator : <reg dy dx> (no endogeneity ☺)
$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \alpha_i + U_{it}$
$Y_{it-1} = \beta_0 + \beta_1 X_{1it-1} + \beta_2 X_{2it-1} + \alpha_i + U_{it-1}$
$\Delta y_{it} = \beta_1 \Delta X_{1it} + \beta_2 \Delta X_{2it} + \Delta U_{it}$
lose obs นั่นคือ N×(T-1)
df = N(T-1) - # beta

/within <reg y-ȳ x-x̄>
▸ Fixed Effect Estimator : e.g. $\bar y_i$ = average income across time for each individual
$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \alpha_i + U_{it}$
$\bar Y_i = \beta_0 + \beta_1 \bar X_{1i} + \beta_2 \bar X_{2i} + \alpha_i + \bar U_i$
$Y_{it} - \bar Y_i = \beta_1 (X_{1it} - \bar X_{1i}) + \beta_2 (X_{2it} - \bar X_{2i}) + error$
"time-demeaned method"
obs = NT obs
df = N(T-1) - # beta
bcs we fix mean
~t n(T-1) - # of regress
≈ N(0,1) ถ้า df ↑

FD vs. FE : • both have same properties regarding unbiasedness & consistency.
• If idiosyncratic error U_{it} is serially uncorrelated → FE ✓
• If T large relative to n → like time series & require stationary assumption → FD ✓ (Integrated series, can turn non-sta to be weak sta in some cases)
• If there's measurement error in X → FE ✓ <bias declines at rate 1/T>

▸ Between Effects Estimator : Cross-sectional of OLS regression of Y on $\bar X$ ⇒ $\bar Y_i = \beta \bar X_i + (\alpha_i + \bar U_i)$
↳ cannot eliminate fixed effects α_i ; inconsistent under Fixed Effect Model assumption that $cov(\alpha_i, X_{it}) \neq 0$ (still has endogeneity)
↳ helpful for case of measurement error in regressor if E[ε] is zero → $\bar X_i + \bar\epsilon_i$ → E[x] + E[ε] = E[x]
reg $\bar Y_i$ $\bar X_i$; df : NT - # beta
e.g. $Y_{it} = \beta X_{it} + U_{it}$ (unobserved : $\bar X_{it} = \bar X_{it} + \bar\epsilon_{it}$)
reg $y\bar x \to \bar y = \beta \bar x + U - \beta\bar\epsilon$ (no endogeneity problem) → can use OLS and get consistent & unbiased estimator for β
↳ not efficient bcs $v_{it} = (\alpha_i + U_{it})$ are correlated across obs (serial correlation)
FGLS is better estimator

③ Random Effects Model : $Y_{it} = \beta_0 + \beta X_{it} + (\alpha_i + U_{it})$ & $cov(X_{it}, \alpha_i) = 0$ (โดยมาก estimator) (hard to justify)
↳ apply GLS by transforming model
$Y_{it} - \theta \bar Y_i = \beta_0 (1 - \theta) + \beta (X_{it} - \theta \bar X_i) + v_{it} - \theta \bar v_i$
$\theta = 1 - \left(\frac{\sigma_\alpha^2}{\sigma_\alpha^2 + T\sigma_u^2}\right)^{\frac{1}{2}}$ — if this small → serial correlation very small → OLS ✓
reg $Y_{it} - \theta \bar Y_i$ $X_{it} - \theta \bar X_i$

FGLS "Random Effect Estimator"

FE vs. RE : • if no variation in X_i across time ($X_{it} - \bar X_i = 0$) → use RE or pooled OLS

* Hausman Test : hypothesis testing to see if data fit RE vs. FE model
$H_0 : \beta_{RE} = \beta_{FE}$ vs. H_1 $\beta_{RE} \neq \beta_{FE}$ — t-stat ~ χ^2_k (# of regressors)
fail to reject → RE model — reject → FE model

STATA • xtset r.v. i r.v. t
• xtreg y x ..., be (dummy between effect)
α unobserved ↳ u_i
* Number of group is # of individual
FE • xtreg, fe
↳ some rows of regressors are omitted bcs value doesn't change across time
↳ cons doesn't have any meaning
RE • xtreg, re
corr (u_i, x) = 0

Hausman test • hausman FE RE → if prob>chi2 = p-value < 10%, 5%, 1% → reject H_0 → fixed effect model (cannot use - RE estimator - Between - OLS)

MIDTERM

▸ Rule to reject H_0 (two-tailed) (ใช้ตารางค่า t)
t-stat 1. if two-tailed : |t-stat| > Critical value $_{\frac{\alpha}{2}}$
H_1 : > : t-stat > C_α
H_1 : < : t-stat < $-C_\alpha$
p-value < α 1%, 5%, 10%

F-dist — reject H_0 if t-stat > C_α

Chi-dist

* (1-α)% CI = estimate ± $C_{\frac{\alpha}{2}}$ · se($\hat\beta$) (two-tailed t-dist)
upper H_1 > ∈ (estimate - C_α · se($\hat\beta$), ∞)
lower H_1 < ∈ (-∞, estimate + C_α · se($\hat\beta$))

$\log X^\alpha = \alpha \log X$, $\log(X \cdot Y) = \log X + \log Y$
• $\ln(XY) = \ln(X) + \ln(Y)$, $\ln\left(\frac{X}{Y}\right) = \ln(X) - \ln(Y)$
• wage = $\beta_0 + 100 \exp + U$ → exp↑ 1 yr → wage↑ 100 บาท
• lnwage = $\beta_0 + 0.1 \exp + U$
↳ exp↑ 1 yr → wage↑ 10%
• wage = $\beta_0 + 10000 \ln(\exp) + U$
↳ exp↑ 1% → wage↑ $\frac{10000}{100}$ บาท
• ln wage : $\beta_0 + \ln(\exp) + U$ → exp↑ 1%
↳ wage↑ 1%

* Rescale :
→ rescaling to cY_1 & cY = $c\beta_0 + c\beta_1 X_1 + \dots + c\beta_k X_k + cU$; $c\hat\beta_j$; $|c|$ se($\hat\beta_j$)
cX_1 & Y = $\beta_0 + \left(\frac{\beta_1}{c}\right)(cX_1) + \dots + \beta_k X_k + U$
↳ $\frac{\hat\beta_1}{c}$; $\left|\frac{1}{c}\right|$ se($\hat\beta_1$)

* R^2 remain the same