⑨

2SLS
3SLS

STOCHASTIC REGRESSORS [x's not fixed]

$\exists$ 3 cases: **SUMMARY**

1) X & ϵ independent;
   OLS "OK", usual $\hat{\beta}$ unbiased, t-ratios OK, ...
2) X & ϵ contemporaneously uncorrelated;
   OLS $\hat{\beta}$ biased, but consistent – be careful.
3) X & ϵ correlated;
   OLS not OK. Big problem!

Case 1): X & ϵ independent.

i.e. x_{ti} is independent of ϵ_s, $\forall$ $t, s = 1, \ldots, T$
and $\forall$ $i = 1, \ldots, K$.

i.e. Allowing X to be random,
but uncorrelated with any disturbance.

Now X is a random variable,
with its own distribution and properties.

Now we must talk about all the other r.v.'s
[ϵ, estimates of β_i, σ^2, ...]
conditional on X.

i.e. given X, we can discuss these.

e.g.

The independence implies,

$$f(\epsilon | X) = f(\epsilon) = N(0, \sigma^2 I) \text{ density}$$

i.e. the conditional density of ϵ (given X)
= the <u>un</u>conditional density of ϵ.

This means that <u>conditional on X</u>,
all the previous OLS properties hold.

→ Conditional on X, $\hat{\beta}$ is nice,
s^2 is nice,
t & F statistics are nice

What if we want to talk about
the distribution of $\hat{\beta}$ <u>not</u> <u>conditional</u> <u>on</u> X?

Consider $\hat{\beta} = \beta + (X'X)^{-1}X'\epsilon$.
If X is random,
the distribution of this is weird
and not generally Normal.

<u>Still, some properties hold.</u> ⇒

~~Hence, again,~~
~~we must talk about estimates~~
~~conditional on X.~~
~~[given X & its [illegible]~~ ϵ $\hat{\beta}$ s^2]

xero

[To discuss these properties, need a Lemma.]

Lemma: Let x & y be random variables and g be a function ∋ $E[g(x,y)]$ exists. Then

$$E[g(x,y)] = E_y[\underbrace{E_{x|y}[g(x,y)]}_{\substack{E[x,\text{ conditional on } y] \\ \text{— a function of } y.}}]$$

Proof: Suppose $f(x,y)$ = joint density.

Then $$E[g(x,y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, f(x,y)\; dx\; dy$$

$$= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} \frac{g(x,y)\, f(x,y)}{f(y)}\, dx\right] f(y)\,dy$$

(where $f(y)$ = marginal density)

$$= E_y\left[E_{x|y}[g(x,y)]\right]$$

QED

Thm: If $E(\hat{\beta})$ exists, then $E(\hat{\beta}) = \beta$.

Proof: $\hat{\beta} = \beta + (X'X)^{-1}X'\epsilon$

$$E(\hat{\beta}) = \beta + \underbrace{E_X\left[E_{\epsilon|X}\left[(X'X)^{-1}X'\epsilon\right]\right]}_{\text{(If } X \text{ \& } \epsilon \text{ independent, this } = 0.)}$$

$$= \beta$$

* The second step is just an applic. of the Lemma with $x = \epsilon$, $y = X$, and $g(x, y) = (X'X)^{-1}X'\epsilon$.

Alternative Proof:

$$E(\hat{\beta}) = \beta + E\left[(X'X)^{-1}X'\epsilon\right]$$

$$= \beta + E\left[(X'X)^{-1}X'\right]E(\epsilon) \qquad \left(\text{since } X \text{ \& } \epsilon \text{ independent}\right)$$

$$= \beta$$

Problem: What is $E\left[(X'X)^{-1}X'\right]$?

It depends on the distribution of X!

— It may not exist (eg. Cauchy).

If it exists, $\hat{\beta}$ unbiased.

<u>Thm.:</u> The Covariance Matrix of $\hat{\beta}$ is $\sigma^2 E[(X'X)^{-1}]$.

(again, if it exists)

Proof:

$$\begin{aligned}
Cov(\hat{\beta}) &= E(\hat{\beta}-\beta)(\hat{\beta}-\beta)' \\
&= E[(X'X)^{-1}X'\epsilon\epsilon' X(X'X)^{-1}] \\
&= E_X\left[E_{\epsilon/X}[(X'X)^{-1}X'\epsilon\epsilon' X(X'X)^{-1}]\right]
\end{aligned}$$

(by Lemma)

$$= E_X[(X'X)^{-1}X' \underset{\epsilon/X}{E}(\epsilon\epsilon') X(X'X)^{-1}]$$

(since X & ϵ independent)

$$\begin{aligned}
&= E_X[(X'X)^{-1}X'(\sigma^2 I)X(X'X)^{-1}] \\
&= \sigma^2 E[(X'X)^{-1}]
\end{aligned}$$

QED

<u>Claim:</u> The usual tests are valid.

$f(t\text{-ratio} \mid X)$ is distributed t
& does not depend on X.

i.e. Conditional on X, t-ratio $\sim t$.

Likewise, F is OK.

<u>Bottom Line</u> – X & ϵ being independent is very similar to X being nonstochastic, for all practical purposes

Case 2): X & ϵ contemporaneously uncorrelated.

Defn 1: $Cov(\epsilon_t, X_{ti}) = 0 \quad \forall\ t = 1, \ldots, T$ and $\forall\ i = 1, \ldots, K$.

Says X's in period t are uncorrelated with ϵ's in period t.
$\therefore$ There may be correlation between X's & ϵ's in *different* time periods, but *not* in the *same* time period.

Defn 2: $\text{plim}\ \frac{X'\epsilon}{T} = 0$.

This is like an asymptotic correlation coeff. ($\therefore$ like Defn 1, for large samples)

Consider $\frac{1}{T}\underset{(K \times 1)}{[X'\epsilon]_i} = \frac{1}{T}\sum_{t=1}^{T} X_{ti}\epsilon_t \;;\; i = 1, \ldots, K$

If X_{ti} & ϵ_t are asymptotically uncorrelated, then these will have plim $= 0$.

contemporaneously

Defn 1 & Defn 2 are consistent with Case 2).
They are not mathematically equivalent, but very similar.

[1 - uncorrel. in small samples as well as large;
2 - asymptotically uncorrel.]

Consider the OLS estimates.

Thm: $\hat{\beta}$ is biased.

Proof: $\hat{\beta} = \beta + (X'X)^{-1}X'\epsilon$

$$E(\hat{\beta}) = \beta + E[(X'X)^{-1}X'\epsilon]$$
$$= \beta + E_X[E_{\epsilon|X}(X'X)^{-1}X'\epsilon] \quad \text{(by Lemma)}$$
$$= \beta + E_X[(X'X)^{-1}X' \; E_{\epsilon|X}(\epsilon)]$$

This is not generally 0; some ϵ's can be correl. with some X's in different periods

For $E_{\epsilon|X}(\epsilon)$ to be zero, all of X must be independent of all of ϵ.

Alternative Proof: $E(\hat{\beta}) = \beta + E[(X'X)^{-1}X'\epsilon]$

$$\neq \beta + E(X'X)^{-1}X' \; E(\epsilon),$$

since X & ϵ not entirely independent.

$\therefore$ $\exists$ small sample problems with OLS.

$\hat{\beta}$ [illegible]

Consider Large Sample Properties of OLS $\hat{\beta}$ & s^2 in Case 2.

Thm: If $\text{plim}\left(\frac{X'X}{T}\right)$ is finite & nonsingular, then $\hat{\beta}$ & s^2 are consistent.

Proof: $$\text{plim}\,\hat{\beta} = \beta + \text{plim}\,(X'X)^{-1}X'\epsilon$$
$$= \beta + \underbrace{\text{plim}\left(\frac{X'X}{T}\right)^{-1}}_{\text{(finite)}} \underbrace{\text{plim}\,\frac{X'\epsilon}{T}}_{=0 \text{ by Defn 2 of Case 2.}}$$
$$= \beta$$

QED

No!

? !

s^2, similar.

Did this

SKIP

Thm: If $\frac{X'\epsilon}{\sqrt{T}} \longrightarrow N\left[0, \sigma^2\, \text{plim}\left(\frac{X'X}{T}\right)\right]$,

then $\sqrt{T}(\hat{\beta}-\beta) \longrightarrow N\left[0, \sigma^2\, \text{plim}\left(\frac{X'X}{T}\right)^{-1}\right]$.

Proof: Recall Useful Fact; If $\xi \to N(0, \Omega)$, then $A\xi \to N(0, \underset{\text{plim}}{A}\Omega \underset{\text{plim}}{A'})$

Consider $$\sqrt{T}(\hat{\beta}-\beta) = \sqrt{T}\,(X'X)^{-1}X'\epsilon$$
$$= \left(\frac{X'X}{T}\right)^{-1} \frac{X'\epsilon}{\sqrt{T}}.$$

let $A = \left(\frac{X'X}{T}\right)^{-1}$, $\xi = \frac{X'\epsilon}{\sqrt{T}}$; $A\xi = \sqrt{T}(\hat{\beta}-\beta)$.

Then $\sqrt{T}(\hat{\beta}-\beta) \longrightarrow N\left[0, \text{plim}\left(\frac{X'X}{T}\right)^{-1}\left(\sigma^2\, \text{plim}\,\frac{X'X}{T}\right) \text{plim}\left(\frac{X'X}{T}\right)\right.$

Note: These 2 Thm's $\Rightarrow$

$\hat{\beta}$ & s^2 are consistent, and

$\hat{\beta}$ has the same asymptotic distribution as it did under the full ideal conditions!

(when X is non stochastic.)

$\Longrightarrow$ Asymptotically nice! (eff.)

$\Longrightarrow$ Even though $\hat{\beta}$ is biased in small samples, it becomes nice as T increases.

* How big must T be, for OLS to be nice?

Answer: It depends on the distribution of the X's.

$\rightarrow$ How weird are the X's?

If the distributions of estimates become "Normal" quickly (as T reaches $30 - 50$), then you're OK with that many observations.

Example: Autoregressive Model

— y_t depends on lagged y's.

$$y_t = \underbrace{\beta_1 x_{t1} + \cdots + \beta_K x_{tK}}_{\text{usual regressors}} + \underbrace{\gamma_1 y_{t-1} + \cdots \gamma_p y_{t-p}}_{\text{lagged } y\text{'s}} + \epsilon_t$$

This is a Case 2 problem.

The lagged y's are stochastic regressors!
They depend on past ϵ's,
but are uncorrelated with the current ϵ_t.

→ Contemporaneously uncorrelated.

The X matrix (of regressors) looks like;

t^{th} row -- $x_t = [x_{t1} \ldots x_{tK}\; y_{t-1} \cdots y_{t-p}]$

And $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix} \to T \times (K+p)$.

Case 2 implications apply;
$\hat{\beta}$ biased but consistent: nice with large samples

Note: If you have Autocorrelated ϵ's in this example,
then $\epsilon_t = \rho \epsilon_{t-1} + u_t$,
and ϵ_t is correlated with y_{t-1}!

i.e. you do have contemporaneous correlation.

— See Case 3, example 3, following.

Case 3): X & ϵ are correlated.

$\operatorname{Cov}(x_{ti}, \epsilon_t) \neq 0$ for some i and some t;

equivalently, $\operatorname{plim} \frac{1}{T} X'\epsilon \neq 0$.

Problem: $\hat{\beta}$ is biased & inconsistent.

a) Already shown $\hat{\beta}$ is biased [Case 2].

b) $\operatorname{plim} \hat{\beta} = \beta + \operatorname{plim} (X'X)^{-1} X'\epsilon$

$$= \beta + \underbrace{\operatorname{plim} \left(\frac{X'X}{T}\right)^{-1}}_{(\text{finite})} \underbrace{\operatorname{plim} \frac{X'\epsilon}{T}}_{(\neq 0)}$$

$\neq \beta$

Thus t's, F's are not good.

Same magnitude of problem
as in Specification Error by Omitted Variables

$\Longrightarrow$ This is a big problem!

EXAMPLES

Example 1 : Simultaneous Equation Problem

[more next term!]

$$(1) \quad C_t = \alpha + \beta(y_t) + \epsilon_t$$

$$(C_t + I_t)$$

$$(2) \quad C_t + I_t = y_t$$

C_t = Consumption
y_t = Income
I_t = Investment

Endogenous Variables ; C_t, y_t
[determined in the model]
Exogenous Variable ; I_t

Intuitively, $y_t \to C_t$ and $C_t \to y_t$;
they are determined simultaneously!

If this is the case,
then y_t is correlated with ϵ_t,
and OLS yields biased and inconsistent estim[ates]
$\rightarrow$ Called Simultaneous Equation Bias.

To show that y_t & ϵ_t are correlated, solve this <u>structural system</u> for the <u>reduced form</u>.
— [for the equilibrium y_t]

Substituting (1) into (2);

$$\alpha + \beta y_t + \epsilon_t + I_t = y_t$$

$$\Longrightarrow \quad y_t = \frac{\alpha}{1-\beta} + \frac{1}{1-\beta} I_t + \frac{1}{1-\beta} \epsilon_t$$

From this it is obvious that y_t & ϵ_t are correlated.

Again, OLS $\hat{\beta}$ is bad.

Example 2: Measurement Error

The model; $y_t = \alpha + \beta x_t + \epsilon_t$

Suppose we don't observe x_t;
but some $\boxed{\text{proxy}}$ of x_t, x_t^*.

Assume $x_t^* = x_t + v_t$, with v_t iid $N(0, \sigma_v^2)$
and v's indep. of ϵ's.

v_t is the measurement error.

We can add & subtract βx_t^*;

$$y_t = \alpha + \beta x_t + \epsilon_t + \beta x_t^* - \beta x_t^*$$

OR $$y_t = \alpha + \beta x_t^* + \epsilon_t + \beta(x_t - x_t^*)$$

$$= \alpha + \beta x_t^* + (\epsilon_t - \beta v_t)$$

This is the model that is run.

Problem: $\mathrm{Cov}(x_t^*, \epsilon_t - \beta v_t) \neq 0$.

(has v_t in it)

Result: OLS yields biased & inconsistent estimates.

The extent of the inconsistency depends on the extent of the measurement error, v_t, and its variance, σ_v^2.

Here, $\text{plim}\ \hat{\beta} = \beta - \beta \frac{\sigma_v^2}{\sigma_v^2 + \sigma_\epsilon^2}$

$$= \beta \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + \sigma_v^2}$$

[See Kmenta, Johnston for algebra.]

Note: This biases $\hat{\beta}$ toward zero under this assumption. "attenuating bias"

* Note: Measurement Error in y_t is not a problem

$$y_t = \alpha + \beta x_t + \epsilon_t \ ; \quad \text{observe } y_t^* = y_t + v_t$$

$$\Rightarrow y_t^* = \alpha + \beta x_t + (\epsilon_t + v_t)$$

$$\text{Cov}(x_t, \epsilon_t + v_t) = 0.$$

This simply results in more randomness in y_t. OLS $\hat{\beta}$ is still good.

What can be done about this? Generally, nothing, unless you can eliminate the measurement error.

If you knew σ_v^2,
then you could correct $\hat{\beta}$
to get consistent estimates.

$\Longrightarrow$ Multiply it by $\left(\frac{\sigma_\epsilon^2 + \sigma_v^2}{\sigma_\epsilon^2}\right)$.

BUT in economics,
things like σ_v^2 are rarely known!

In something like Physics,
we may know something like this ...

— Johnston shows how to manipulate $\hat{\beta}$
to get consistent estimates.

Schmidt argues that you can't do this
given the assumptions as to v_t,
the parameters of the model
are not identified.

Identification -- (next term)

Example 3: Autoregressive Model with Autocorrelated Errors

$$y_t = \alpha + \beta x_t + \gamma y_{t-1} + \epsilon_t$$

$$\text{where } \epsilon_t = \rho \epsilon_{t-1} + u_t$$

Note: y_{t-1} is correlated with ϵ_{t-1}
and $\therefore$ with ϵ_t.

All estimates are biased and inconsistent.

This is a very complicated case;
$\exists$ Other Problems:

(i) Usual tests for Autocorrelation don't work.
e.g. D-W statistic is biased towards 2
(against finding Autocorrelation).

Intuition; OLS $\hat{\beta}$, $\hat{\gamma}$ are bad;
$\therefore$ OLS e_t are bad;
$\therefore$ $\hat{\rho}$ & DW are bad!

See Durbin, ECON, 1970 & Johnston
for other tests that do work in this case.

$\Downarrow$

e.g. —

e.g. Regress e_t on e_{t-1} plus other rhs variables.
Then use t-test on coefficient of e_{t-1}. $\to (\hat{\rho}$
This is asymptotically OK.

Note: this is like the C-O method (# 2) for estimating $\hat{\rho}$.

Given this problem, what to do?
GLS? Yes, but ∃ problems with GLS.

(ii) Estimates of $\hat{\rho}$ are inconsistent.

Again, Intuition: OLS $\hat{\beta}$, $\hat{\gamma}$ are bad;
$\therefore$ OLS e_t are bad;
$\therefore$ $\hat{\rho}$ & DW are bad!

Solutions:

a) Use MLE (Hildreth-Lu) procedure.

b) estimate $\hat{\rho}$ like before, but use Instrumental Variables residual instead of OLS residuals.

—(more in a minute!)

(iii) —

(iii) GLS estimates [of β & γ] based on a consistent estimate of ρ need not be asymptotically efficient in this case (not as good as MLE).

Solution c) $\exists$ an asymptotically efficient 2-step procedure;

Define: $y_t^* = y_t - \hat{\rho}\, y_{t-1}$; $x_t^* = x_t - \hat{\rho}\, x_{t-1}$

Summary of Stochastic Regressors

$\exists$ 3 cases: $Y = X\beta + \epsilon$

Case (1) X & ϵ independent.

OLS "OK"; usual $\hat{\beta}$ unbiased, t-ratios OK,

Case (2) X & ϵ contemporaneously uncorrelated.

OLS $\hat{\beta}$ biased but consistent;
Be careful!

Case (3) X & ϵ correlated.

OLS not OK;
$\hat{\beta}$ biased & inconsistent;
big problem.

INSTRUMENTAL Variables

A way to deal with Case 3.

– Given X & ϵ correlated, this is a method for getting consistent estimates.

INTUITION: $Y = X\beta + \epsilon$

X is highly correlated with Y; this is the relationship we wish to measure.

X is correlated with ϵ; this inhibits our attempt to measure the rel.

If we can transform X into a matrix of instruments, Z, that are highly correlated with X (and ∴ still highly correlated with Y), but not correlated with ϵ;

Highly correlated? What if initially X & Y are not highly correlated?

then we can get consistent estimates.

Thm.: Suppose $\exists$ a $(T\times K)$ matrix, Z, $\ni$

(1) $\text{plim}\,\frac{Z'X}{T}$ is finite and nonsingular,

(2) $\text{plim}\,\frac{Z'\epsilon}{T} = 0$.

Then the Instrumental Variables Estimator,

$$\tilde{\beta} = (Z'X)^{-1}Z'Y \quad \text{is } \underline{\text{consistent}}.$$

(3) Furthermore, if $\frac{1}{\sqrt{T}}Z'\epsilon \rightarrow N(0, \sigma^2 \text{plim}\,\frac{Z'Z}{T})$, then

$$\sqrt{T}(\tilde{\beta}-\beta) \longrightarrow N\left[0,\ \sigma^2\,\text{plim}\left(\frac{Z'X}{T}\right)^{-1}\left(\frac{Z'Z}{T}\right)\left(\frac{X'Z}{T}\right)^{-1}\right]$$

Proof:

$$\begin{aligned}\tilde{\beta} &= (Z'X)^{-1}Z'Y \\ &= (Z'X)^{-1}Z'(X\beta+\epsilon) \\ &= \beta + (Z'X)^{-1}Z'\epsilon\end{aligned}$$

$$\begin{aligned}\text{plim}\,\tilde{\beta} &= \beta + \underbrace{\text{plim}\left(\frac{Z'X}{T}\right)^{-1}}_{(\text{finite})}\ \underbrace{\text{plim}\,\frac{Z'\epsilon}{T}}_{(0)} \\ &= \beta\end{aligned}$$

[Assumptions (1) & (2)]

Then, $\sqrt{T}(\tilde{\beta}-\beta) = \left(\frac{Z'X}{T}\right)^{-1}\frac{Z'\epsilon}{\sqrt{T}}$;

and given assumption (3),
and the useful fact, [if $\xi \rightarrow N(0,\Omega)$ then $A\xi \rightarrow N[0, (\text{plim}\,A)\,\Omega\,(\text{p}$

$$\sqrt{T}(\tilde{\beta}-\beta) \longrightarrow N\left[0,\ \sigma^2\,\text{plim}\left(\frac{Z'X}{T}\right)^{-1}\left(\frac{Z'Z}{T}\right)\left(\frac{X'Z}{T}\right)^{-1}\right].$$

QED

Discussion:

Z is a matrix of Instruments.

Assumption (1) says that Z must be correlated with X.
[ie we want the instruments to be correlated with the variables of interest.]

Assumption (2) says that Z must not be correlated with ϵ.

If we can find such a Z, we've solved the problem.

* If we want the asymptotic distribution, we need assumption (3).

[but no one worries about that...]

<u>Example</u>:

$$y_t = \alpha + \beta x_t + \gamma y_{t-1} + \epsilon_t$$

where $\epsilon_t = \rho \epsilon_{t-1} + u_t$

$\Rightarrow$ [Case 3 problem!]

— y_{t-1} correlated with ϵ_t.

— we need an <u>instrument</u> for y_{t-1};

use x_{t-1}!
- highly correlated with y_{t-1}, uncorrelated with ϵ_t.

Thus let $Z_t = [1, \; x_t, \; x_{t-1}]$

correl. with α uncorrel. with ϵ_t	correl. with x_t uncorrel. with ϵ_t	correl. with y_{t-1} uncorrel. with ϵ_t

* We want something as highly correlated with y_{t-1} as possible; this will get us better estimates.

$\rightarrow$ Then $\text{plim} \frac{Z'X}{T}$ is finite, $\text{plim} \frac{Z'\epsilon}{T} = 0$...

Another possible instrument: $Z = [1 \;\; x \;\; y]$