ECON 61001: *Lecture 7*

Alastair R. Hall

The University of Manchester

Outline of this lecture

- Time series regression models with non-spherical errors
  - OLS-based inference
  - GLS-based inference
  - Testing for serial correlation
- Instrumental Variables estimation
  - Models with endogenous regressors
  - OLS as MoM $\rightarrow$ IV
  - Examples of instruments

In Lecture 5 we show that conventional OLS inference framework goes through under the following assumptions.

- Assumption TS1: $y_t = x_t'\beta_0 + u_t, \; t = 1, 2, \ldots T$
- Assumption TS2: (y_t, h_t') is a weakly stationary, weakly dependent time series.

- Assumption TS3: $E[x_t x_t'] = Q$, a finite, positive definite matrix.

- Assumption TS4: $E[u_t \,|\, x_t] = 0$ for all $t = 1, 2, \ldots, T$.
- Assumption TS5: $Var[u_t \,|\, x_t] = \sigma_0^2$.
- Assumption TS6: For all $t \neq s$, $E[u_t u_s \,|\, x_t, x_s] = 0$.

OLS - with conditionally heteroscedastic errors

In video on "OLS-based inference in time series regression models with conditionally heteroscedastic errors" we consider OLS based inference in models satisfying:

- Assumption TS1: $y_t = x_t'\beta_0 + u_t, \ t = 1, 2, \ldots T$
- Assumption TS2-SS: (y_t, h_t') is a strongly stationary, weakly dependent time series.

- Assumption TS3: $E[x_t x_t'] = Q$, a finite, positive definite matrix.

- Assumption TS4: $E[u_t \,|\, x_t] = 0$ for all $t = 1, 2, \ldots, T$.
- Assumption TS5-H: $Var[y_t \,|\, \mathcal{I}_t] = \sigma_t^2$.
- Assumption TS6: For all $t \neq s$, $E[u_t u_s \,|\, x_t, x_s] = 0$

Dynamic completeness

Alternatively, we can impose Assumptions TS1, TS2, TS3 and TS5 (or TS1,TS2-SS, TS3 and (TS5-H) and

- **Assumption** TS7 $E[y_t \mid \mathcal{I}_t] = x_t'\beta_0$ where $\mathcal{I}_t$ is information set at time t.

- if Assumption TS7 holds then the model is said to be *dynamically complete*.

Serial correlation in errors

The framework adopted affects how we view the consequences of serial correlation in the errors and so how we proceed.

If Assumptions TS1, TS2-SS, TS3, TS4, and TS5-H hold but TS6 does not (so $E[u_t u_s \mid x_t, x_s] \neq 0$) then:

- OLS is still consistent as Assumption TS4 holds.

- Scope for serial correlation robust inference based on OLS

If Assumptions TS1, TS2, TS3 & TS5 (or TS1, TS2-SS, TS3 & TS5-H) hold but Assumption TS7 does not (so $E[y_t \mid \mathcal{I}_t] \neq x_t'\beta_0$) then:

- OLS is likely inconsistent as *TS4* may not hold.
- Need to reconsider specification.

If Assumptions TS1, TS2-SS, TS3, TS4, & TS5-H hold then: $\hat{\beta}_T$ is consistent and

$$T^{1/2}(\hat{\beta}_T - \beta_0) \xrightarrow{d} N(0, V_{sc})$$

where

- $V_{sc} = Q^{-1}\Omega Q^{-1}$,
- $\Omega = \Gamma_0 + \sum_{i=1}^{\infty}(\Gamma_i + \Gamma_i')$, for $\Gamma_i = E[u_t u_{t-i} x_t x_{t-i}']$.

Estimation of long run variance Ω?

Use heteroscedasticity and autocorrelation covariance (HAC) matrix estimator:

$$\hat{\Omega}_{HAC} = \hat{\Gamma}_0 + \sum_{i=1}^{T-1} \omega(i, T)(\hat{\Gamma}_i + \hat{\Gamma}_i')$$

where

- $\hat{\Gamma}_i = T^{-1} \sum_{t=i+1}^{T} e_t e_{t-i} x_t x_{t-i}'$

- $\omega(\cdot, \cdot)$ known as kernel
- $\omega(\cdot, \cdot)$ chosen so that $\hat{\Omega}_{HAC}$ is both consistent and psd.

Popular choice of kernel:

- Bartlett (Newey & West, 1987): putting $a_i = i/(b_T + 1)$

$$
\begin{aligned}
\omega(i, T) &= 1 - a_i, \ \text{for } a_i < 1 \\
&= 0, \ \text{else}
\end{aligned}
$$

- This is an example of a truncated kernel estimator:

$$
\hat{\Omega}_{HAC} = \hat{\Gamma}_0 + \sum_{i=1}^{b_T} \omega(i, T) \left\{ \hat{\Gamma}_i + \hat{\Gamma}_i' \right\}
$$

OLS-based inference

- Let $\hat{Q} = T^{-1}X'X$.
- Set $\hat{V}_{sc} = \hat{Q}^{-1}\hat{\Omega}_{HAC}\hat{Q}^{-1}$.

- Then can show $\hat{V}_{sc} \xrightarrow{p} V_{sc}$.

So can perform inference based on OLS provided substitute $\hat{V}_{sc}$ for $\hat{\sigma}^2_T\hat{Q}^{-1}$ in statistics discussed in Lecture 5. (See Tutorial 7)

OLS if Assumption TS7 fails

If justify standard OLS-based inference framework via Assumptions TS1-TS3, TS5 & TS7 then serial correlation arises through failure of Assumption TS7 and so:

$$E[y_t \,|\, \mathcal{I}_t] \neq x_t' \beta_0$$

This means model is misspecified and so either:

- model is linear but have omitted variables

- model is nonlinear

So need to re-consider specification.

Recall need model for Σ.

Here consider model with AR(1) errors.

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

$$u_t = \rho_0 u_{t-1} + \varepsilon_t$$

where

- $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$.
- Regularity conditions hold (esp $E[x_t u_t] = 0$) $\Rightarrow \hat{\beta}_T \xrightarrow{p} \beta_0$.

Then GLS is OLS applied to:

$$\ddot{y}_t = \ddot{x}_t'\beta_0 + \ddot{u}_t$$

where

- $\ddot{y}_t = y_t - \rho_0 y_{t-1}$ *etc.* $\sim$ "quasi-difference".

Consistency of GLS requires

$$E[\ddot{x}_t \ddot{u}_t] = E[(x_t - \rho_0 x_{t-1})(u_t - \rho_0 u_{t-1})] = 0,$$

$\rightarrow$ (in general) need $E[x_s u_t] = 0$ for $s = t-1, t, t+1$.

So conditions for consistency of GLS are stronger than for OLS. Thus, the conditions for the consistency of GLS are stronger than those for the consistency of OLS.

If sub $u_t = y_t - x_t'\beta_0$ in regression model then:

$$y_t = \rho_0 y_{t-1} + x_t'\beta_0 - x_{t-1}'(\rho_0\beta_0) + \varepsilon_t,$$

which is same as (re-written) transformed model in GLS.

Also special case of the the linear regression model

$$y_t = \delta_0 y_{t-1} + x_t'\delta_1 + x_{t-1}'\delta_2 + \varepsilon_t,$$

in which the parameters satisfy:

$$\delta_2 + \delta_0 * \delta_1 = 0,$$

known as the common factor or COMFAC restrictions.

(c.f. response to violation of *TS7*.)

Breusch-Godfrey test for serial correlation

Assume errors follow AR(p) process that is,

$$u_t = \rho_{0,1}u_{t-1} + \rho_{0,2}u_{t-2} + \ldots + \rho_{0,p}u_{t-p} + \varepsilon_t, \text{ where } \varepsilon_t \sim iid(0, \sigma_\varepsilon^2).$$

$$\begin{aligned} H_0: \quad \rho_{0,i} &= 0, \text{ for all } i = 1, 2, \ldots, p \\ H_A: \quad \rho_{0,i} &\neq 0, \text{ for at least one } i = 1, 2, \ldots, p \end{aligned}$$

Test stat is:

$$LM_p = TR^2$$

where R^2 is from regression of e_t on $x_t, e_{t-1}, e_{t-2}, \ldots, e_{t-p}$.

Under H_0: $LM_p \xrightarrow{d} \chi_p^2$.

Choice of p?

- trade-off
- may reflect sampling frequency of data

Interpretation of significant test:

- can be caused by AR(p) errors
- can also be caused by misspecified model

Empirical example: See Section 4.4.5 of Lecture Notes

Consequences of $E[x_t u_t] \neq 0$ (violation of *CS4* or *TS4*).

Recall

$$\hat{\beta}_T - \beta_0 = \left(T^{-1} \sum_{t=1}^{T} x_t x_t' \right)^{-1} T^{-1} \sum_{t=1}^{T} x_t u_t,$$

and using the WLLN, we have

$$T^{-1} \sum_{t=1}^{T} x_t x_t' \xrightarrow{p} E[x_t x_t'] = Q,$$

$$T^{-1} \sum_{t=1}^{T} x_t u_t \xrightarrow{p} E[x_t u_t] = \mu \neq 0.$$

So using Slutsky's Theorem: $\hat{\beta}_T \xrightarrow{p} \beta_0 + Q^{-1}\mu \neq \beta_0$.

So OLS is an inconsistent estimator of β_0.

Source of problem is endogenous regressor(s).

This can arise in three main ways in econometric models:

- reverse causality
- omitted variables
- measurement error

Now consider an example of each.

Acemoglu *et al* (2001):

$$ln[y_i] = \beta_0 + r_i\beta_1 + \text{controls} + u_i$$

where

- y_i is income per capita in developing country i
- r_i is quality of institutions in developing country i

r is likely correlated with u_i due to reverse causality.

Angrist & Krueger (1991):

$$ln[w_i] = \theta_1 + \theta_e ed_i + \text{controls} + u_i$$

where

- w_i equals the weekly wage of individual i;
- ed_i is the number of years of education of individual i;

ed_i likely correlated with u_i due to omitted variables such as "ability" that affect w_i and ed_i.

Clarida, Gali & Gertler (2000):

$$r_t = c + \beta_\pi E[\pi_{t+1}|\Omega_t] + \beta_y E[y_{t+1}|\Omega_t] + \beta_1 r_{t-1} + \beta_2 r_{t-2} + u_t,$$

- r_t = Federal Funds rate in year.quarter t

- π_{t+1} = inflation in $t+1$
- Ω_t information set at time t

- y_{t+1} = output gap in $t+1$
- u_t is error satisfying $E[u_t \mid \Omega_t] = 0$

Expectations not observed if replace by actual values then introduce measurement error because:

$$r_t = c + \beta_\pi \pi_{t+1} + \beta_y y_{t+1} + \beta_1 r_{t-1} + \beta_2 r_{t-2} + \underbrace{\left(u_t \; - \; \beta_\pi v_{\pi,t+1} \; - \; \beta_y v_{y,t+1}\right)}_{= e_t}$$

where

- $v_{(\cdot),t+1} = (\cdot)_{t+1} - E[(\cdot)_{t+1}|\Omega_t]$,
- $E[(\cdot)_{t+1} e_t] = \beta_{(\cdot)} Var[v_{(\cdot),t+1}] \neq 0$.

To motivate IV estimation, we reinterpret OLS as a Method of Moments (MoM) estimator.

From normal equations: OLS is MoM based on the population moment condition:

$$E[x_t u_t(\beta_0)] = 0 \qquad (\star)$$

where $u_t(\beta) = y_t - x_t'\beta$ (and so $u_t(\beta_0) = u_t$).

From this perspective, estimation is based on information in $(\star)$ about β_0. So if

- $E[x_t u_t(\beta_0)] = 0 \rightarrow$ consistent estimator.
- $E[x_t u_t(\beta_0)] \neq 0 \rightarrow$ inconsistent estimator.

IV estimation

Find a $q \times 1$ vector z_t such that:

$$E[z_t u_t(\beta_0)] = 0 \qquad (\sharp)$$

- need $q \geq k$ and z_t "sufficiently related to" x_t + certain other conditions discussed in the next lecture
- z_t known as an instrument

Use the population moment equation in ($\sharp$) as basis for estimation $\rightarrow$ Instrumental Variables (IV) estimator of β_0.

Discuss next time how to calculate IV estimator and what its properties are. Conclude by looking at choices of instruments in our three examples.

Acemoglu *et al* (2001):

$$ln[y_i] \;=\; \beta_0 + r_i\beta_1 \;+\; \text{controls} + \;u_i$$

where

- y_i is income per capita in developing country i
- r_i is quality of institutions in developing country i

r is likely correlated with u_i due to reverse causality.

Base estimation on the population moment condition:

$$E[z_i u_i] \;=\; 0$$

where $z_i = [1, \text{controls}, M_i]'$ and M_i is settler mortality in country i.

Example 2: returns to education

Angrist & Krueger (1991):

$$ln[w_i] = \theta_1 + \theta_e ed_i + \text{controls} + u_i$$

where

- w_i equals the weekly wage of individual i;
- ed_i is the number of years of education of individual i;

ed_i likely correlated with u_i due to omitted variables such as "ability" that affect w_i and ed_i.

Base estimation on

$$E[z_i u_i] = 0$$

where $z_i = [1, \text{controls}, Q_i]'$ and Q_i is quarter of birth of i.

Clarida, Gali & Gertler (2000):

$$r_t = c + \beta_\pi E[\pi_{t+1}|\Omega_t] + \beta_y E[y_{t+1}|\Omega_t] + \beta_1 r_{t-1} + \beta_2 r_{t-2} + u_t,$$

- r_t = Federal Funds rate in year.quarter t
- π_{t+1} = inflation in $t+1$
- Ω_t information set at time t
- y_{t+1} = output gap in $t+1$
- u_t is error satisfying $E[u_t \mid \Omega_t] = 0$

Expectations not observed if replace by actual values then introduce measurement error because:

$$r_t = c + \beta_\pi \pi_{t+1} + \beta_y y_{t+1} + \beta_1 r_{t-1} + \beta_2 r_{t-2} + \underbrace{(u_t - \beta_\pi v_{\pi,t+1} - \beta_y v_{y,t+1})}_{=e_t}$$

where

- $v_{(\cdot),t+1} = (\cdot)_{t+1} - E[(\cdot)_{t+1}|\Omega_t]$,
- $E[(\cdot)_{t+1}e_t] = -\beta_{(\cdot)}Var[v_{(\cdot),t+1}] \neq 0$.

However: $E[z_t e_t] = 0$ for any $z_t \in \Omega_t$ (*e.g.* macro variables).

Further reading

- Notes: Sections 4.4, 5.1, 5.2
- Greene:
  - Time series regression models with serial correlation,
    - OLS 20.4 (but 20.4.1 and 20.4.2 go into more detail on "other conditions" in limit theorems than needed for the course), 20.5
    - GLS/FGLS 20.8, 20.9
    - Testing for serial correlation 20.7
  - IV 8.1, 8.2 and 8.3