INTRODUCTORY ECONOMETRICS

Lesson 3b

Dr Javier Fernández

etpfemaj@ehu.es

Dpt. of Econometrics & Statistics

UPV—EHU

3.3 A General Test for Linear Restrictions.

Testing for Linear Restrictions: Example 1

- Recall GLRM subject to q linear restrictions:

$$\underset{(T\times 1)}{Y} = \underset{(T\times K+1)}{X}\ \underset{(K+1\times 1)}{\beta} + \underset{(T\times 1)}{u},$$

$$H_0: \underset{(q\times K+1)}{R}\ \underset{(K+1\times 1)}{\beta} = \underset{(q\times 1)}{r}.$$

- Previous tests $\equiv$ special cases of LRs:
  1. Let's have the GLRM with $q=1$, $R=\begin{bmatrix}0 & 0 & 1 & \dots & 0\end{bmatrix}$ and $r=0$:

$$H_0: R\beta = \begin{bmatrix}0 & 0 & 1 & \dots & 0\end{bmatrix}\begin{pmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \dots\\ \beta_K\end{pmatrix} = \beta_2$$

$$= r \qquad\qquad = 0$$

es decir, $H_0: \beta_2 = 0$;

the test of individual significance of X_2.

Testing for Linear Restrictions: Example 2

- $H_0: \underset{(q\times K+1)}{R}\ \underset{(K+1\times 1)}{\beta} = \underset{(q\times 1)}{r}.$
  3. Let's assume $q=2$ restrictions such that

$$R=\begin{bmatrix}0 & 2 & 3 & 0 & \dots & 0\\ 1 & 0 & 0 & -2 & \dots & 0\end{bmatrix} \text{ and } r=\begin{bmatrix}5\\ 3\end{bmatrix}:$$

-

$$H_0: R\beta = \begin{bmatrix}0 & 2 & 3 & 0 & \dots & 0\\ 1 & 0 & 0 & -2 & \dots & 0\end{bmatrix}\begin{pmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\\ \dots\\ \beta_K\end{pmatrix} = \begin{bmatrix}2\beta_1+3\beta_2\\ \beta_0-2\beta_3\end{bmatrix}$$

$$= r \qquad\qquad = \begin{bmatrix}5\\ 3\end{bmatrix}$$

that is, the GLRM under $H_0: \begin{cases}2\beta_1+3\beta_2=5\\ \beta_0-2\beta_3=3\end{cases}$

Testing for Linear Restrictions: Example 3

- $H_0: \underset{(q\times K+1)}{R}\ \underset{(K+1\times 1)}{\beta} = \underset{(q\times 1)}{r}.$
  2. Let's assume $q=K$ restrictions such that

$$R=\begin{bmatrix}\mathbf{0} & | & \mathbf{I}_K\end{bmatrix} = \left(\begin{array}{c|cccc}0 & 1 & 0 & \dots & 0\\ 0 & 0 & 1 & \dots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \dots & 1\end{array}\right) \text{ and } r=\mathbf{0}=\begin{pmatrix}0\\ 0\\ \vdots\\ 0\end{pmatrix}$$

-

$$H_0: R\beta = \begin{bmatrix}\mathbf{0}_K & | & \mathbf{I}_K\end{bmatrix}\begin{pmatrix}\beta_0\\ \vdots\\ \beta^\star\end{pmatrix} = \beta^\star$$

$$= r \qquad\qquad = \mathbf{0}$$

that is, $H_0: \beta^\star = \mathbf{0}$;

the test of overall significance of the regression.

Testing for Linear Restrictions: dn

- ...so, can have a general test statistic to cover for all hypothesis of the form

$$H_0: \underset{(q\times K+1)}{R} \; \underset{(K+1\times 1)}{\beta} = \underset{(q\times 1)}{r} \quad ?$$

- Given that $\widehat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X'X)^{-1})$, we have that

$$R\widehat{\beta} \sim \mathcal{N}(R\beta, \sigma^2 R(X'X)^{-1}R')$$

- As before, standardise $R\widehat{\beta}$ and construct SS,

$$\frac{(R\widehat{\beta} - R\beta)'[R(X'X)^{-1}R']^{-1}(R\widehat{\beta} - R\beta)}{\sigma^2} \sim \chi^2(q)$$

- Therefore (recall changing $\sigma^2 \to \widehat{\sigma}^2$):

$$\boxed{\frac{(R\widehat{\beta} - R\beta)'[R(X'X)^{-1}R']^{-1}(R\widehat{\beta} - R\beta)/q}{\widehat{\sigma}^2} \sim \mathcal{F}^{q}_{T-K-1}}$$

General Test for Linear Restrictions: rule

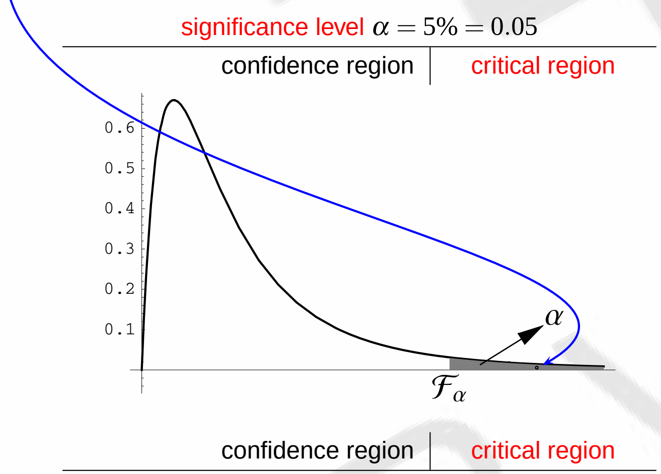
- Which Test? $\{H_0 : R\beta = r$
- **Remember:** Hypothesis ⇝ statistic ⇝ rule...
- Test for linear restrictions:
  - Hypothesis: $H_0 : R\beta = r$ vs. $H_a : R\beta \neq r$
  - Statistic:

$$F = \frac{(R\widehat{\beta} - r)'[R(X'X)^{-1}R']^{-1}(R\widehat{\beta} - r)/q}{\widehat{\sigma}^2} \sim \mathcal{F}^{q}_{T-K-1} \text{ under } H_0:$$

  - Rule: $F > \mathcal{F}_\alpha(q, T-K-1) \quad \Rightarrow$ reject H_0 :
    $\Rightarrow$ linear restrictions aren't (jointly) true.

General Test for Linear Restrictions: rule (cont)

- Rule: $F > \mathcal{F}_\alpha(q, T-K-1) \quad \Rightarrow$ reject H_0 :
-



3.4 Tests based on the Residual Sum of Squares.

General Test for Linear Restrictions: rule 2

- Hypothesis: $H_0 : R\beta = r$ vs. $H_a : R\beta \neq r$
- Statistic:

$$F = \frac{(R\widehat{\beta} - r)'[R(X'X)^{-1}R']^{-1}(R\widehat{\beta} - r)/q}{\widehat{\sigma}^2}$$

- using result on $\widehat{\beta}_R = (I - AR)\widehat{\beta} + Ar$, numerator is difference between SS's:

$$F = \frac{(\text{RSS}_R - \text{RSS})/q}{\text{RSS}/(T\text{–}K\text{–}1)} \sim \mathscr{F}^q_{T\text{–}K\text{–}1} \text{ under } H_0 :$$

- Rule: $F > \mathscr{F}_\alpha(q, T\text{–}K\text{–}1) \quad \Rightarrow$ reject H_0 :
  $\Rightarrow$ linear restrictions aren't (jointly) true.

General Test for Linear Restrictions: Summary

- Hypothesis: $H_0 : R\beta = r$ vs. $H_a : R\beta \neq r$
- Statistic:

$$F = \frac{(R\widehat{\beta} - r)'[R(X'X)^{-1}R']^{-1}(R\widehat{\beta} - r)/q}{\widehat{\sigma}^2} = \frac{(\text{RSS}_R - \text{RSS})/q}{\text{RSS}/(T\text{–}K\text{–}1)} \sim \mathscr{F}^q_{T\text{–}K\text{–}1} \text{ under } H_0 :$$

- Rule: $F > \mathscr{F}_\alpha(q, T\text{–}K\text{–}1) \quad \Rightarrow$ reject H_0 :
  $\Rightarrow$ linear restrictions aren't (jointly) true.
- Note that, SS form needs estimating twice: unrestricted and restricted regressions.
- and, of course, they can also be used to test for individual significance, overall significance, informative restrictions, etc.

Test based on SS: Example Cobb-Douglas

- Hypothesis: $H_0 : \beta_L + \beta_K = 1$ vs. $H_a : \beta_L + \beta_K \neq 1$
- Statistic:
  - $$\widehat{\nu} = \widehat{\beta}_L + \widehat{\beta}_K = 0.67 + 0.27 = 0.89$$
  - $$S_{\widehat{\nu}} = \sqrt{\text{Var}(\widehat{\beta}_L) + \text{Var}(\widehat{\beta}_K) + 2\text{Cov}(\widehat{\beta}_L, \widehat{\beta}_K)} = \widehat{\sigma}\sqrt{a_{11} + a_{22} + 2a_{12}} = 2\sqrt{4 + 7 + 2(-1)} = 2\sqrt{9} = 6$$
  - $$t = \frac{\widehat{\nu} - 1}{S_{\widehat{\nu}}} = \frac{0.89 - 1}{6} = \frac{-0.11}{6} = -0.018.$$
- Rule: $|t| = 0.018 < t_{0.025}(50) = 2.01 \quad \Rightarrow$ don't reject H_0 :
  $\Rightarrow$ the "constant returns to scale" hypothesis is supported by data.

Test based on SS: Example Cobb-Douglas (2)

- Alternatively, use SS form to calculate this t ratio:
  unrestricted:
  $\log Y = \alpha + \beta_L \log L + \beta_K \log K + u, \quad \rightsquigarrow \quad \text{RSS} = 200$
- restricted: $\log Y = \alpha + \beta_L \log L + (1 - \beta_L)\log K + u$
  $\log(Y/K) = \alpha + \beta_L \log(L/K) + u, \quad \rightsquigarrow \quad \text{RSS}_R = 200.001296$
- $$F = \frac{(\text{RSS}_R - \text{RSS})/q}{\text{RSS}/(T\text{–}K\text{–}1)} = \frac{(200.001296 - 200)/1}{200/50} = \frac{.001296}{4} = 0.000324 < \mathscr{F}_{0.05}(1, 50) = 4.04$$
- or (recall $t(m) = \sqrt{\mathscr{F}(1, m)}$)

$$t = \sqrt{F} = \sqrt{0.000324} = 0.018 < t_{0.05}(50) = 2.01$$

General Test: Example 2

- GLRM with $q = 2$, $R = \begin{bmatrix} 0 & 2 & 3 & 0 & \dots & 0 \\ 1 & 0 & 0 & -2 & \dots & 0 \end{bmatrix}$ and $r = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$:

$$R\widehat{\beta} = \begin{bmatrix} d_1'\widehat{\beta} \\ d_2'\widehat{\beta} \end{bmatrix} = \begin{bmatrix} 2\widehat{\beta}_1 + 3\widehat{\beta}_2 \\ \widehat{\beta}_0 - 2\widehat{\beta}_3 \end{bmatrix}$$

-

$$R(X'X)^{-1}R' = \begin{bmatrix} d_1'(X'X)^{-1}d_1 & d_1'(X'X)^{-1}d_2 \\ d_2'(X'X)^{-1}d_1 & d_2'(X'X)^{-1}d_2 \end{bmatrix}$$

$$= \begin{bmatrix} 4a_{11} + 9a_{22} + 12a_{12} & 2a_{10} - 4a_{13} + 3a_{02} - 6a_{23} \\ & a_{00} + 4a_{33} - 4a_{03} \end{bmatrix}$$

- Therefore $F =$

$$\frac{\begin{bmatrix} 2\beta_1 + 3\beta_2 - 5 & \beta_0 - 2\beta_3 - 3 \end{bmatrix} \begin{bmatrix} 4a_{11} + 9a_{22} + 12a_{12} & 2a_{10} - 4a_{13} + 3a_{02} - 6a_{23} \\ & a_{00} + 4a_{33} - 4a_{03} \end{bmatrix}^{-1} \begin{bmatrix} 2\beta_1 + 3\beta_2 - 5 \\ \beta_0 - 2\beta_3 - 3 \end{bmatrix} /2}{\widehat{\sigma}^2}$$

$\sim \mathscr{F}^2_{T\text{–}K\text{–}1}$ under H_0 :

- *es decir*, an "F" statistic for testing two linear restrictions jointly.

General Test: Example 2

- Alternatively (easier), use SS form to calculate this F statistic:

$$H_0 : \begin{cases} 2\beta_1 + 3\beta_2 = 5 \\ \beta_0 - 2\beta_3 = 3 \end{cases}$$

$$\beta_1 = \frac{5 - 3\beta_2}{2}, \qquad \beta_0 = 3 + 2\beta_3$$

- unrestricted:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \cdots + u \rightsquigarrow \text{RSS}$$

- restricted:

$$Y = (3 + 2\beta_3) + (2.5 - 1.5\beta_2)X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \cdots + u$$

$$\underbrace{Y - 3 - 2.5X_1}_{Y^*} = \beta_2 \underbrace{(X_2 - 1.5X_1)}_{X_2^*} + \beta_3 \underbrace{(X_3 + 2)}_{X_3^*} + \beta_4 X_4 \cdots + u$$

$$Y^* = \beta_2 X_2^* + \beta_3 X_3^* + \beta_4 X_4 \cdots + u \rightsquigarrow \text{RSS}_R$$

- and $F = \frac{(\text{RSS}_R - \text{RSS})/q}{\text{RSS}/(T\text{–}K\text{–}1)}$, etc.

General Test: Example 3

- GLRM with $q = K$, $R = \begin{bmatrix} \mathbf{0}_K & | & \mathbf{I}_K \end{bmatrix}$ and $r = \mathbf{0}_K$:

$$R\widehat{\beta} \rightsquigarrow \text{selects } \beta^\star$$

$$R(X'X)^{-1}R' \rightsquigarrow \text{selects } \begin{bmatrix} a_{00} & a_{01} & \dots & a_{0K} \\ a_{10} & a_{11} & \dots & a_{1K} \\ \vdots & \vdots & \dots & \vdots \\ a_{K0} & a_{K1} & \dots & a_{KK} \end{bmatrix} = (x'x)^{-1}$$

- Therefore:

$$F = \frac{(\widehat{\beta}^\star - 0)'[(x'x)^{-1}]^{-1}(\widehat{\beta}^\star - 0)/K}{\widehat{\sigma}^2}$$

$$= \frac{\widehat{\beta}^{\star\prime} x'x \widehat{\beta}^\star / K}{\widehat{\sigma}^2}$$

- *es decir*, the usual "F" statistic for testing the overall significance of the regression.

General Test: Example 3

- Alternatively, use SS form to calculate this F:

unrestricted: $Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_K X_K + u \rightsquigarrow$ RSS

restricted: $Y = \beta_0 + u \rightsquigarrow \text{RSS}_R = \text{TSS}$

- Statistic:

$$F = \frac{(\text{RSS}_R - \text{RSS})/q}{\text{RSS}/(T\text{–}K\text{–}1)} = \frac{(\text{TSS} - \text{RSS})/K}{\text{RSS}/(T\text{–}K\text{–}1)} = \frac{\text{ESS}/K}{\text{RSS}/(T\text{–}K\text{–}1)} = \frac{R^2/K}{(1 - R^2)/(T\text{–}K\text{–}1)}$$

obtaining same formula as before.

3.5 Point Prediction and Prediction Interval.

Prediction

- Previous chapters: **Specification, Estimation and Validation**.
- This chapter: Final stage: **Use = Prediction**.
- Starting point: appropriate model to describe behaviour of variable Y:

Model Specification
⇓
Parameter Estimation
⇓
Validation and Hypothesis Testing

-

⇓

not satisfactory result	satisfactory result
⇓	⇓
respecify model	go to prediction stage

Concept

- Time series: prediction (of future values)
  ⇒ Forecasting
- Cross-section: prediction (of unobserved values)
  ⇒ Simulation
- In general: prediction ⇒ answer to
  "what if. . . ?" questions,
  es decir what value would take Y *if* $X = X_p$?

Basic Elements

- Model or PRF:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \cdots + \beta_K X_{Kt} + u_t$$
$$Y_t = X_t'\beta + u_t, \quad t = 1, \ldots, T.$$

- Estimated model or SRF:

$$\widehat{Y}_t = X_t'\widehat{\beta}, \quad t = 1, \ldots, T. \tag{8}$$

- Prediction observation: with subindex $p =$ (usually $p \notin [1, T]$):

$$Y_p = X_p'\beta + u_p.$$

- Random disturbance u_p:

$$\mathsf{E}(u_p) = 0, \quad \mathsf{E}(u_p^2) = \sigma^2, \quad \mathsf{E}(u_p u_s) = 0 \quad \forall s \neq p.$$

- Known value of vector X_p' .

Point Prediction

- Substituting in SRF (8):

$$\boxed{\widehat{Y}_p = X_p' \widehat{\beta}.}$$

es decir, numeric value as approximation to unknown value.

Prediction Error

- The error made (when taking $\widehat{Y}_p$ instead of the true Y_p) is

$$e_p = Y_p - \widehat{Y}_p,$$

- which can be expressed as:
- a function of the two error sources introduced in the prediction.
- Under normality:

$$(\widehat{\beta} - \beta) \sim \mathcal{N}(0, \sigma^2 (X'X)^{-1}), \quad \text{and} \quad u_p \sim \mathcal{N}(0, \sigma^2),$$

- so that

$$\boxed{e_p \sim \mathcal{N}(0, \sigma_e^2)},$$

- where the prediction error variance is:

$$\sigma_e^2 = X_p' \underbrace{\text{Var}(\widehat{\beta})}_{\sigma^2 (X'X)^{-1}} X_p + \underbrace{\text{Var}(u_p)}_{\sigma^2} + \underbrace{\text{Cov}(\widehat{\beta}, u_p)}_{0}$$
$$= \sigma^2 (1 + X_p' (X'X)^{-1} X_p).$$

Interval Prediction

- Standardised prediction error:

$$\frac{e_p - 0}{\sigma_e} = \frac{e_p}{\sigma \sqrt{1 + X_p'(X'X)^{-1} X_p}} \sim \mathcal{N}(0,1),$$

- Recall how changing $\sigma \to \widehat{\sigma} \quad \Rightarrow \quad \mathcal{N}(0,1) \to t$!!, then

$$\frac{e_p}{\widehat{\sigma}_e} = \frac{e_p}{\widehat{\sigma} \sqrt{1 + X_p'(X'X)^{-1} X_p}} \sim t(T{-}K{-}1).$$

- Therefore:

$$Pr(-t_{\alpha/2} \leq \frac{e_p}{\widehat{\sigma}_e} \leq t_{\alpha/2}) = 1 - \alpha,$$

- and solving for Y_p :

$$Pr(\widehat{Y}_p - \widehat{\sigma}_e t_{\alpha/2} \leq Y_p \leq \widehat{Y}_p + \widehat{\sigma}_e t_{\alpha/2}) = 1 - \alpha.$$

- Then, the $(1-\alpha)$ confidence interval for the unknown Y_p is:

$$CI(Y_p)_{(1-\alpha)} = \left[\widehat{Y}_p \pm \widehat{\sigma}_e t_{\alpha/2}\right],$$

which measures the precision of the point prediction.

Prediction: Example

- In the previous example (linearised Cobb-Douglas fn:)

$$\widehat{\log Y_t} = \widehat{\alpha} + \widehat{\beta}_L \log L_t + \widehat{\beta}_K \log K_t, \quad T = 53;$$
$$\widehat{\log Y_t} = 2.10 + 0.67 \log L_t + 0.32 \log K_t, \quad \widehat{\sigma}^2 = 4$$

- "What value would Y_p take if $\log L_p = 2.5; \log K_p = 2.0$?":
- $X_p' = \begin{bmatrix} 1 & 2.5 & 2.0 \end{bmatrix}$
-

$$\widehat{\log Y_p} = X_p' \widehat{\beta} = \begin{bmatrix} 1 & 2.5 & 2.0 \end{bmatrix} \begin{bmatrix} 2.10 \\ 0.67 \\ 0.32 \end{bmatrix}$$
$$= 2.10 + 0.67 \cdot 2.5 + 0.32 \cdot 2.0 = 4.42$$

Prediction: Example

- Construct a 95% CI for the true Y_p:

$$
\begin{aligned}
\widehat{\sigma}_e^2 &= \sigma^2 (1 + X_p'(X'X)^{-1}X_p) \\
&= 4\left(1 + \begin{bmatrix}1 & 2.5 & 2.0\end{bmatrix}\begin{pmatrix}2 & 0 & 0\\ 0 & 4 & -1 \\ 0 & -1 & 7\end{pmatrix}\begin{bmatrix}1\\2.5\\2.0\end{bmatrix}\right) \\
&= 4\left(1 + \begin{bmatrix}2 & 8 & 11.5\end{bmatrix}\begin{bmatrix}1\\2.5\\2.0\end{bmatrix}\right) \\
&= 4\,(1+45) = 4\cdot 46 = 184
\end{aligned}
$$

-

$$
\begin{aligned}
CI(\log Y_p)_{0.95} &= \left[\widehat{\log Y_p} \pm \widehat{\sigma}_e\, t_{0.025}(50)\right] \\
&= \left[4.42 \pm \sqrt{184}\cdot 2.01\right] \\
&= [4.42 \pm 27.25] \\
&= [-22.84 \quad ; \quad 31.68]
\end{aligned}
$$