INTRODUCTORY ECONOMETRICS

Lesson 2c

Dr Javier Fernández

etpfemaj@ehu.es

Dpt. of Econometrics & Statistics

UPV—EHU

2.7a Omission of relevant variables.

Omission of relevant variables

- true relationship:

$$Y = X\beta + u = \left[X_I \;\middle|\; X_{II} \right] \begin{pmatrix} \beta_I \\ \hline \beta_{II} \end{pmatrix} + u$$

$$X = \left[\begin{array}{cccc|ccc} 1 & x_{11} & \dots & x_{K_1,1} & x_{K_1+1,1} & \dots & x_{K1} \\ 1 & x_{12} & \dots & x_{K_1,2} & x_{K_1+1,2} & \dots & x_{K2} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1T} & \dots & x_{K_1,T} & x_{K_1+1,T} & \dots & x_{KT} \end{array} \right], \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{K_1} \\ \hline \beta_{K_1+1} \\ \vdots \\ \beta_K \end{pmatrix}.$$

$$Y = X_I \beta_I + X_{II}\beta_{II} + u$$

- estimated relationship:

$$Y = X_I \beta_I + v \qquad \text{where} \quad v = X_{II}\beta_{II} + u,$$

$$\text{then} \quad \mathrm{E}(v) \neq 0 \quad \rightsquigarrow \quad \mathrm{E}(\widehat{\beta}) \neq \beta.$$

i.e. $\widehat{\beta}$ is biased.

Omission of relevant variables: consequences

Summary:

- OLS estimator of coefficients is *biased* (except if $x_I' x_{II} = 0$).
- OLS estimator of intercept is *always biased*.
- Estimator of Error variance is *always biased*.

2.7b Multicollinearity

Perfect Multicollinearity

Extreme case:

- exact linear combination:
  - $\sum_{k=0}^{K} \lambda_k X_{kt} = 0, \quad \lambda \neq 0, \quad X_{0t} = 1,$
  - $\exists X_i \mid X_i = \lambda_0^* + \sum_{\substack{k=1 \\ k \neq i}}^{K} \lambda_k^* X_{kt},$
  - $\exists X_i, X_j \mid$ Corr$(X_i, X_j) = 1,$
  - $\exists X_i \mid$ aux regres X_i on $\{X_k\}_{\substack{k=1 \\ k \neq i}}^{K} \rightsquigarrow R_i^2 = 1.$
- Problem:
  - $\text{rk}X < K+1$, (X isn't of full rank)
  - $\rightsquigarrow \quad \det(X) = 0$
  - $\rightsquigarrow \quad \nexists (X'X)^{-1}$
  - $\rightsquigarrow \quad \widehat{\beta}$?

Perfect Multicollinearity: example

- Let $X_{4t} = 2X_{1t} \quad \forall t$:

$$X_{4t} = 0 + 2X_{1t} + 0 \cdot X_{2t} + 0 \cdot X_{3t} + 0 \cdot X_{5t} + \cdots + 0 \cdot X_{Kt},$$

- no error? $\Rightarrow$ aux regres X_4 on $\{X_k\}_{\substack{k=1 \\ k \neq 4}}^{K} \rightsquigarrow R_4^2 = 1!!$
- Model specification:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \cdots + u_t, t = 1, 2 \ldots, T,$$
$$X_{4t} = 2X_{1t},$$

- and substituting in model:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 (2X_{1t}) + \cdots + u_t,$$
$$= \beta_0 + \underbrace{(\beta_1 + 2\beta_4)}_{\beta_1^\star} X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \cdots + u_t$$

- now we have one less parameter to estimate.

Multicollinearity: counterexample

$$Y_t = \beta_0 + \beta_1^\star X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \cdots + u_t$$

- Just K parameters remain to be estimated,
  but β_1 and β_4 cannot be estimated separately:
  - we can just estimate a linear combination of them:
    $\beta_1^\star = \beta_1 + 2\beta_4$,
  - *i.e.* combined effect of X_{1t} and X_{4t} on Y_t!!
- (Exercise: Try it yourself with $X_{2t} - 3X_{3t} = 10, \quad \forall t$.)
- multicollinearity = *linear relationships*
  but. . . what if relationship *isn't linear*? *e.g.:*

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{1t}^2 + u_t$$

  - X is of full column rank $\rightsquigarrow$ no problem.

Perfect Multicollinearity: consequences

- some parameters cannot be estimated separately.
- some estimates are just l.c. of parameters.
- R^2 is correct:
  correctly picks up proportion of (variance of) Y_t explained by the regression.
- Predictions of Y are still valid.

2.7c Imperfect Multicollinearity

Imperfect Multicollinearity

- Problem:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \cdots + u_t, t = 1, 2 \ldots, T,$$
$$X_{4t} = 2X_{1t} + v_t,$$

v_t = gap between X_{4t} and $2X_{1t}$,

- approximate relationship:
- auxiliary regression X_{4t} on rest $\rightsquigarrow$ $R^2 \approx 1$.
- it's a matter of degree ($x'x$ not diagonal $\rightsquigarrow$ correlated variables)
- Note: whenever perfect/imperfect is not specified we mean imperfect mc.

Multicollinearity: Symptoms

- Typical symptom:
  - high R^2 (relevant group of regressors)
  - but they appear to be not relevant individually (inability to separate effects of regressors).
- more formally:

$$\text{Var}(\widehat{\beta}^\star) = \sigma^2 (x'x)^{-1} = \frac{\sigma^2}{T} \text{Var}(X^\star)^{-1}$$

$$\Rightarrow \quad \text{Var}(\widehat{\beta}_k) = \frac{\sigma^2}{T\,\text{Var}(X_k)(1 - R_k^2)},$$

- so that, in the previous example $X_{4t} \approx 2X_{1t}$:
  - Corr$(X_4, X_1) \uparrow$
  - R_4^2 and R_1^2 $\upuparrows$
  - denominator $\downarrow$
  - variances $\upuparrows$

Multicollinearity: Consequences

- Some coefficients aren't significant, even if their variables have an important effect on dependent variable.
- Nevertheless, Gauss-Markov $\Rightarrow$ linear, unbiased and of minimum variance estimators, then *it isn't possible to find a Better LUE*.
- R^2 is correct: correctly picks up proportion of (variance of) Y_t explained by the regression.
- Predictions of Y are still valid.

Multicollinearity: How to detect

- Small changes in data
  $\Rightarrow$ important changes in estimates
  (they can even affect their signs).
- Coefficient estimations
  not individually significant. . .
- . . . but they are jointly significant.
- High coefficient of determination R^2.
- Auxiliary regressions among regressors
  $\Rightarrow$ high R_k^2.

Multicollinearity: Some solutions

Multicollinearity is not an easy problem to solve.
Nevertheless, from

$$\text{Var}(\widehat{\beta}_k) = \frac{\sigma^2}{T\text{Var}(X_k)(1-R_k^2)},$$

it turns out that to lower the variance we may:

- $\mathbf{T}\uparrow$: Increase number of observations T .
  Also, differences among regressors may increase.
- $\text{Var}(\mathbf{X})\uparrow$: Increase data dispersion; *e.g.* study about consumption function:
  sample of families $\leftrightsquigarrow$ all possible incomes.
- $\text{Var}(\mathbf{X})\uparrow$:Include additional information.
  e.g. impose restrictions suggested by Ec. Th.
- $\sigma^2\downarrow$: Add new relevant regressor not yet included.
  It would also avoid serious bias problems.
- $\mathbf{R_k^2}\downarrow$: Eliminate variables that may produce multicollinearity.
  (Take care of omitting some relevant regressor though).

2.8 The OLS Estimator under Restrictions.

GLRM under linear restrictions (1)

- previous chapter objectives:
  - Econometric model (GLRM), characteristics and basic assumptions. . .
  - but. . . no knowledge about model parameters.
  - Least Squares Method for parameter estimation (OLS).
  - Properties of resulting estimators.
- present chapter objectives:
  - *a priori* information about parameter values (or l.c.) . . .
  - given by
    - economic theory,
    - other empirical work,
    - own experience, etc.
  - Non-Restricted Model $\Rightarrow$ Ordinary LS.
  - Restricted Model $\Rightarrow$ Restricted LS.
  - Check, given the estimated model, if the information is compatible with available data.

GLRM under linear restrictions: examples

- production function with constant returns to scale: $\beta_K + \beta_L = 1$.
- product demands as function of price: $\beta = -1$ (say).
- in GLRM: let us assume that $\beta_2 = 0$ and $2\beta_3 = \beta_4 - 1$:
  - Full model:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \cdots + \beta_{Kt} X_{Kt} + u_t, \text{ with } \beta_2 = 0 \text{ and } 2\beta_3 + 1 = \beta_4;$$

  - Alternative transformed model:

$$Y_t = \beta_0 + \beta_1 X_{1t} + 0 X_{2t} + \beta_3 X_{3t} + (2\beta_3 + 1) X_{4t} + \cdots + \beta_{Kt} X_{Kt} + u_t$$
$$Y_t - X_{4t} = \beta_0 + \beta_1 X_{1t} + \beta_3 (X_{3t} + 2X_{4t}) + \cdots + \beta_K X_{Kt} + u_t$$
$$Y_t^* = \beta_0 + \beta_1 X_{1t} + \beta_3 Z_t + \cdots + \beta_K X_{Kt} + u_t$$

    where $Y_t^* = Y_t - X_{4t}$ and $Z_t = X_{3t} + 2X_{4t}$.
  - This transformed model:
    - can be estimated by OLS:
      $\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_3, \widehat{\beta}_5, \ldots, \widehat{\beta}_K$, together with $\widehat{\beta}_2 = 0$ and $\widehat{\beta}_4 = 2\widehat{\beta}_3 + 1$.
    - has new endogenous variable Y_t^* (not always so: e.g. if $\beta_2 = 0$ alone) and new explanatory variable Z_t.

GLRM under linear restrictions (2)

- The “transformation” method is good for simple cases only.
- In general, q (nonredundant) linear restrictions among parameters:

$$\begin{matrix} 1 \\ \vdots \\ q \end{matrix} \begin{pmatrix} \diamond & \diamond & \diamond & \ldots & \diamond \\ \vdots & & & & \\ \diamond & \diamond & \diamond & \ldots & \diamond \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \cdots \\ \beta_K \end{pmatrix} = \begin{pmatrix} \diamond \\ \vdots \\ \diamond \end{pmatrix}$$

  - for given matrix R and vector r,

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

  - example of non-valid case (why?):

$$\beta_3 = 0, \quad 2\beta_2 + 3\beta_4 = 1, \quad \beta_1 - 2\beta_4 = 3, \quad 6\beta_4 = 2 - 4\beta_2 + \beta_3$$

GLRM under linear restrictions (2cont)

- Write previous example $\beta_2 = 0$ and $2\beta_3 = \beta_4 - 1$ ($q = 2$ restrictions) as in general formula:

$$\underset{(2 \times K+1)}{\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 2 & -1 & 0 & \ldots & 0 \end{pmatrix}} \underset{(K+1 \times 1)}{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \cdots \\ \beta_K \end{pmatrix}} = \underset{(2 \times 1)}{\begin{pmatrix} 0 \\ -1 \end{pmatrix}}.$$

$R_{(2 \times K+1)}$, $\beta_{(K+1 \times 1)}$, $r_{(2 \times 1)}$

- In general, we write GLRM subject to q linear restrictions as:

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

Estimation: restricted least squares (RLS).

- Typical optimization exercise:

$$\min_{\beta}(u'u) \quad \text{where} \quad u = Y - X\beta,$$
subject to $R\beta = r$.

- Lagrangian:

$$L(\beta, \lambda) = u'u - 2\lambda'(R\beta - r)$$
$$\min_{\beta,\lambda} L(\beta, \lambda).$$

- First derivatives:

$$\frac{\partial L(\beta,\lambda)}{\partial \beta} = -2X'u - 2R'\lambda,$$
$$\frac{\partial L(\beta,\lambda)}{\partial \lambda} = -2(R\beta - r),$$

Estimation: restricted least squares (RLS) (cont).

- 1st.o.c. $\rightsquigarrow$ normal equations:

$$X'\widehat{u}_R + R'\widehat{\lambda} = 0, \qquad (4)$$

$$R\widehat{\beta}_R = r, \qquad (5)$$

where $\widehat{\beta}_R$ and $\widehat{\lambda}$ are values of β, λ that satisfy 1st.o.c. and residuals

$$\widehat{u}_R = Y - X\widehat{\beta}_R. \qquad (6)$$

- Solving for $\widehat{\beta}_R$ and $\widehat{\lambda}$:

$$\widehat{\lambda} = [R(X'X)^{-1}R']^{-1}(r - R\widehat{\beta}),$$

$$\widehat{\beta}_R = \widehat{\beta} + (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(r - R\widehat{\beta})$$

$$= \widehat{\beta} + A(r - R\widehat{\beta}) = (I - AR)\widehat{\beta} + Ar \qquad (7)$$

where $A = (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}$.

RLS estimation: characteristics

- Expression (7): $\widehat{\beta}_R = \widehat{\beta} + A(r - R\widehat{\beta}) \rightsquigarrow$
  - the restricted estimate $\widehat{\beta}_R$ can be obtained as a function of the (not restricted) ordinary estimate: $\widehat{\beta}$
  - $R\widehat{\beta} \simeq r \quad \Rightarrow \quad \widehat{\beta}_R$ (restricted) $\simeq \widehat{\beta}$ (not restricted) .
- Normal equations (4): $X'\widehat{u}_R + R'\widehat{\lambda} = 0 \rightsquigarrow$
  - satisfy the restrictions (obvious).
  - $X'\widehat{u}_R \neq 0$, *i.e.*:
    - sum of restricted residuals not zero,
    - restricted residuals not orthogonal to explanatory variables,
    - then, restricted residuals not orthogonal to fitted $\widehat{Y}_R$.
  - TSS $\neq$ RSS$_R$ + ESS$_R$ (compare with ordinary case and with transformed equation: R^2 ??).

Properties of the RLS estimator (1)

Expression (7) : $\widehat{\beta}_R = (I - AR)\widehat{\beta} + Ar \quad \rightsquigarrow$

1. **Linear:** RLS estimator $\widehat{\beta}_R$ is l.c. of OLS estimator $\widehat{\beta}$, which is linear , then $\widehat{\beta}_R$ is linear also .
2. **Bias:** RLS estimator $\widehat{\beta}_R$ is $\begin{cases} \text{biased,} & \text{if } R\beta \neq r \text{ ,} \\ \text{unbiased,} & \text{if } R\beta = r \text{ true} \end{cases}$

   Demo:

$$\mathrm{E}\left(\widehat{\beta}_R\right) = (I - AR)\,\mathrm{E}\left(\widehat{\beta}\right) + Ar = (I - AR)\,\beta + Ar = \beta + A(r - R\beta).$$

3. **Covariance Matrix:** $\mathrm{Var}\left(\widehat{\beta}_R\right) = (I - AR)\mathrm{Var}\left(\widehat{\beta}\right) = \sigma^2 (I - AR)(X'X)^{-1}$

   Demo:

$$\mathrm{Var}\left(\widehat{\beta}_R\right) = (I - AR)\,\mathrm{Var}\left(\widehat{\beta}\right)(I - AR)' = \sigma^2 (I - AR)\,(X'X)^{-1}(I - AR)'$$

$$= \sigma^2 \left[(X'X)^{-1} + AR(X'X)^{-1}R'A' - AR(X'X)^{-1} - (X'X)^{-1}R'A'\right]$$

where: $AR(X'X)^{-1}R'A' = (X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}R(X'X)^{-1}R'A'$

$$= (X'X)^{-1}R'A'.$$

Properties of the RLS estimator (2)

4. **Smaller variance** than OLS estimators, *even if restrictions aren't true*:

   Demo:

$$\mathrm{Var}\left(\widehat{\beta}_R\right) = \mathrm{Var}\left(\widehat{\beta}\right) - AR\,\mathrm{Var}\left(\widehat{\beta}\right)$$

$$= \mathrm{Var}\left(\widehat{\beta}\right) - (\text{psd matrix}).$$

□

5. surprising result (apparently):
   - less “uncertainty” about parameters $\rightsquigarrow$ greater precision in estimation. . .
   - but. . . towards an erroneous result (biased) if restriction isn't true.

Multicolinearity vs restrictions

Must clearly distinguish two different cases:

- linear relationships among regressors (*i.e.* multicollinearity):

  e.g. $X_{4t} = 2X_{1t}$

  $\Rightarrow$ missing information for individual estimates.

- linear relationships among coefficients:

  e.g. $\beta_4 = 2\,\beta_1$

  $\Rightarrow$ extra information about parameters

  $\rightsquigarrow$ estimators with smaller variance.

- respective models to estimate:

$$Y_t = \beta_0 + (\underbrace{\beta_1 + 2\beta_4}_{\beta_1^*})X_{1t} + \beta_2 X_{2t} + + \cdots + u_t\,,$$

$$\Rightarrow \widehat{\beta}_1^* \quad \text{but} \quad \widehat{\beta}_1, \widehat{\beta}_4\ ?$$

$$Y_t = \beta_0 + \beta_1(\underbrace{X_{1t} + 2X_{4t}}_{X_{1t}^*}) + \beta_2 X_{2t} + + \cdots + u_t\,,$$

$$\Rightarrow \widehat{\beta}_1 \quad \text{and} \quad \widehat{\beta}_4 = 2\,\widehat{\beta}_1$$