Econometrics - Lecture 6

GMM-Estimator and Econometric Models

Contents

- The GIV Estimator
- GMM Estimation
- Econometric Models
- Dynamic Models
- Multi-equation Models
- Time Series Models
- Models for Limited Dependent Variables
- Panel Data Models
- Econometrics II

From OLS to IV Estimation

Linear model $y_i = x_i'\beta + \varepsilon_i$ with K-vector of regressors

- OLS estimator: solution of the K normal equations

  $$1/N\ \Sigma_i(y_i - x_i'b)\ x_i = 0$$

- Corresponding moment conditions

  $$E\{\varepsilon_i\ x_i\} = E\{(y_i - x_i'\beta)\ x_i\} = 0$$

- IV estimator given R instrumental variables z_i which may overlap with x_i: based on the R moment conditions

  $$E\{\varepsilon_i\ z_i\} = E\{(y_i - x_i'\beta)\ z_i\} = 0$$

- IV estimator: solution of corresponding sample moment conditions

  $$\tfrac{1}{N}\sum_i (y_i - x_i'\hat{\beta}_{IV})\, z_i = 0$$

Number of Instruments

Moment conditions

$$E\{\varepsilon_i z_i\} = E\{(y_i - x_i'\beta) z_i\} = 0$$

one equation for each component of z_i

- z_i possibly overlapping with x_i

General case: R moment conditions

Substitution of expectations by sample averages gives R equations

$$\frac{1}{N}\sum_i (y_i - x_i'\hat{\beta}_{IV}) z_i = 0$$

1. $R = K$: one unique solution, the IV estimator; identified model

$$\hat{\beta}_{IV} = \left(\sum_i z_i x_i'\right)^{-1} \sum_t z_i y_i = (Z'X)^{-1} Z'y$$

2. $R < K$: infinite number of solutions, not enough instruments for a unique solution; under-identified or not identified model

The GIV Estimator

3. *R* > *K*: more instruments than necessary for identification; over-identified model

For *R* > *K*, in general, no unique solution of all *R* sample moment conditions can be obtained; instead:

- the weighted quadratic form in the sample moments

$$Q_N(\beta) = \left[\tfrac{1}{N} \sum_i (y_i - x_i'\beta) z_i \right]' W_N \left[\tfrac{1}{N} \sum_i (y_i - x_i'\beta) z_i \right]$$

  with a *RxR* positive definite weighting matrix W_N is minimized
- gives the generalized instrumental variable (GIV) estimator

$$\hat{\beta}_{IV} = (X'ZW_N Z'X)^{-1} X'ZW_N Z'y$$

The weighting matrix W_N

W_N: positive definite, order *RxR*

- Different weighting matrices result in different consistent GIV estimators with different covariance matrices
- Optimal choice for W_N?
- For $R = K$, the matrix $Z'X$ is square and invertible; the IV estimator is $(Z'X)^{-1}Z'y$ for any W_N

GIV and TSLS Estimator

Optimal weighting matrix: $W_N^{opt} = [1/N(Z'Z)]^{-1}$; corresponds to the most efficient IV estimator

$$\hat{\beta}_{IV} = (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y$$

- If the error terms are heteroskedastic or autocorrelated, the optimal weighting matrix has to be adapted
- Regression of each regressor, i.e., each column of *X*, on *Z*, i.e., on the *R* column of *Z*, results in $\hat{X} = Z(Z'Z)^{-1}Z'X$ and

$$\hat{\beta}_{IV} = (\hat{X}'\hat{X})^{-1}\hat{X}'y$$

- This is why the GIV estimator is also called "two stage least squares" (TSLS) estimator:
  1. First step: regress each column of *X* on *Z*
  2. Second step: regress *y* on predictions of *X*

GIV Estimator and Properties

- GIV estimator is consistent
- The asymptotic distribution of the GIV estimator, given IID(0, σ_ε^2) error terms, leads to

  $$N\left(\beta, \hat{V}\{\hat{\beta}_{IV}\}\right)$$

  which is used as approximate distribution in case of finite *N*
- The (asymptotic) covariance matrix of the GIV estimator is given by

  $$V\left\{\hat{\beta}_{IV}\right\} = \sigma^2 \left[\left(\sum_i x_i z_i'\right)\left(\sum_i z_i z_i'\right)^{-1}\left(\sum_i z_i x_i'\right)\right]^{-1}$$

- In the estimated covariance matrix, σ^2 is substituted by

  $$\hat{\sigma}^2 = \tfrac{1}{N}\sum_i \left(y_i - x_i'\hat{\beta}_{IV}\right)^2$$

  the estimate based on the IV residuals $y_i - x_i'\hat{\beta}_{\mathrm{IV}}$

Contents

Moment Conditions of OLS and IV Estimation

Linear model $y_i = x_i'\beta + \varepsilon_i$

- OLS estimator: solution of the *K* normal equations

  $$1/N\ \Sigma_i(y_i - x_i'b)\ x_i = 0$$

- Corresponding moment conditions

  $$E\{\varepsilon_i\ x_i\} = E\{(y_i - x_i'\beta)\ x_i\} = 0$$

- IV estimator given *R* instrumental variables z_i (which may overlap with x_i) is based on the *R* moment conditions

  $$E\{\varepsilon_i\ z_i\} = E\{(y_i - x_i'\beta)\ z_i\} = 0$$

- IV estimator: solution of corresponding sample moment conditions

  $$\tfrac{1}{N}\sum_i (y_i - x_i'\hat{\beta}_{IV})\, z_i = 0$$

Generalized Method of Moments (GMM) Estimation

The model is characterized by R moment conditions and the corresponding equations

$$E\{f(w_i, z_i, \theta)\} = 0$$

[cf. $E\{(y_i - x_i'\beta)\, z_i\} = 0$]

- $f(.)$: R-vector function
- w_i: vector of observable variables, exogenous or endogenous
- z_i: vector of instrumental variables
- θ: K-vector of unknown parameters

Sample equivalents $g_N(\theta)$ of moment conditions should fulfil

$$g_N(\theta) = \frac{1}{N}\sum_i f(w_i, z_i, \theta) = 0$$

Estimates $\hat{\theta}$ are chosen such that the sample moment conditions are fulfilled

GMM Estimation

$R \geq K$ is a necessary condition for GMM estimation

- $R = K$: unique solution, the K-vector $\hat{\theta}$, of

  $$g_N(\theta) = 0$$

  if $f(.)$ is nonlinear in θ, numerical solution might be derived
- $R > K$: in general, no choice $\hat{\theta}$ for the K-vector θ will result in $g_N(\hat{\theta})$ = 0 for all R equations; for a good choice $\hat{\theta}$, $g_N(\hat{\theta}) \sim 0$, i.e., all components of $g_N(\hat{\theta})$ are close to zero

  estimate $\hat{\theta}$ is obtained through minimization with respect to θ of the quadratic form

  $$Q_N(\theta) = g_N(\theta)' W_N g_N(\theta)$$

  W_N: symmetric, positive definite weighting matrix

The GMM Estimator

Weighting matrix W_N

- Different weighting matrices result in different consistent estimators with different covariance matrices
- Optimal weighting matrix

  $$W_N^{opt} = [E\{f(w_i, z_i, \theta)\, f(w_i, z_i, \theta)'\}]^{-1}$$

  i.e., the inverse of the covariance matrix of the sample moments
- For $R = K$: $W_N = I_N$ with unit matrix I_N

Minimization of $Q_N(\theta) = g_N(\theta)' \, W_N \, g_N(\theta)$: For nonlinear $f(.)$

- Numerical optimization algorithms
- W_N depends on θ; iterative optimization

Example: The Linear Model

Model: $y_i = x_i'\beta + \varepsilon_i$ with $E\{\varepsilon_i x_i\} = 0$ and $V\{\varepsilon_i\} = \sigma_\varepsilon^2$

- Moment or orthogonality conditions:

  $$E\{\varepsilon_t x_t\} = E\{(y_t - x_t'\beta)x_t\} = 0$$

  $f(.) = (y_i - x_i'\beta)x_i$, $\theta = \beta$, instrumental variables: x_i; moment conditions are exogeneity conditions for x_i
- Sample moment conditions:

  $$1/N\, \Sigma_i\, (y_i - x_i'b)\, x_i = 1/N\, \Sigma_i\, e_i\, x_i = g_N(b) = 0$$
- With $W_N = I_N$, $Q_N(\beta) = [1/N]^2\, (\Sigma_i\, \varepsilon_i\, x_i)'(\Sigma_i\, \varepsilon_i\, x_i) = [1/N]^2\, X'\varepsilon\varepsilon'X$
- OLS and GMM estimators coincide, give the estimator b, but
  - OLS: residual sum of squares $S_N(b) = 1/N\, \Sigma_i\, e_i^2$ has its minimum
  - GMM: $Q_N(b) = [1/N]^2\, (\Sigma_i\, e_i\, x_i)'(\Sigma_i\, e_i\, x_i) = 0$

Linear Model, $E\{\varepsilon_t x_t\} \neq 0$

Model $y_i = x_i'\beta + \varepsilon_i$ with $V\{\varepsilon_i\} = \sigma_\varepsilon^2$, $E\{\varepsilon_i x_i\} \neq 0$ and *R* instrumental variables z_i

- Moment conditions:

  $$E\{\varepsilon_i z_i\} = E\{(y_i - x_i'\beta)z_i\} = 0$$

- Sample moment conditions:

  $$1/N \, \Sigma_i \, (y_i - x_i'b) \, z_i = g_N(b) = 0$$

- Identified case ($R = K$): the single solution is the IV estimator

  $$b_{IV} = (Z'X)^{-1} Z'y$$

- Over-identified case ($R > K$): GMM estimator from

  $$\min_\beta Q_N(\beta) = \min_\beta g_N(\beta)' W_N \, g_N(\beta)$$

Linear Model: GMM Estimator

Minimization of $Q_N(\beta)= \min_\beta g_N(\beta)' W_N g_N(\beta)$ wrt β:

- For $W_N = I$, the first order conditions are

$$\frac{\partial Q_N(\beta)}{\partial \beta} = 2\left(\frac{\partial g_N(\beta)}{\partial \beta}\right)' g_N(\beta) = 2\left(\tfrac{1}{N}X'Z\right)\left(\tfrac{1}{N}Z'y - \tfrac{1}{N}Z'X\beta\right) = 0$$

  resulting in the estimator

$$b = [(X'Z)(Z'X)]^{-1} (X'Z)Z'y$$

  b coincides with the IV estimator if $R = K$

- The optimal weighting matrix $W_N^{opt} = (E\{\varepsilon_i^2 z_i z_i'\})^{-1}$ is estimated by

$$W_N^{opt} = \left(\tfrac{1}{N}\sum_i e_i^2 z_i z_i'\right)^{-1}$$

  generalizes the covariance matrix of the GIV estimator to White's heteroskedasticity-consistent covariance matrix estimator (HCCME)

Example: Labour Demand

Verbeek's data set "labour2": Sample of 569 Belgian companies (data from 1996)

- Variables
  - *labour*: total employment (number of employees)
  - *capital*: total fixed assets
  - *wage*: total wage costs per employee (in 1000 EUR)
  - *output*: value added (in million EUR)
- Labour demand function

$$labour = \beta_1 + \beta_2 * output + \beta_3 * capital$$

Labour Demand Function: OLS Estimation

In logarithmic transforms: Output from GRETL

Dependent variable : l_LABOUR
Heteroskedastic-robust standard errors, variant HC0,

	coefficient	std. error	t-ratio	p-value
const	3,01483	0,0566474	53,22	1,81e-222 ***
l_ OUTPUT	0,878061	0,0512008	17,15	2,12e-053 ***
l_CAPITAL	0,003699	0,0429567	0,08610	0,9314

Mean dependent var	4,488665	S.D. dependent var	1,171166
Sum squared resid	158,8931	S.E. of regression	0,529839
R- squared	0,796052	Adjusted R-squared	0,795331
F(2, 129)	768,7963	P-value (F)	4,5e-162
Log-likelihood	-444,4539	Akaike criterion	894,9078
Schwarz criterion	907,9395	Hannan-Quinn	899,9928

GMM Estimation in GRETL

Specification of function and orthogonality conditions for labour demand model

```
# initializations go here
matrix X = {const , l_OUTPUT, l_CAPITAL}
series e = 0
scalar b1 = 0
scalar b2 = 0
scalar b3 = 0
matrix V = I(3)

gmm e = l_LABOuR – b1*const – b2*l_OUTPUT – b3*l_CAPITAL
  orthog e; X
  weights V
  params b1 b2 b3
end gmm
```

Labour Demand Function: GMM Estimation

In logarithmic transforms: Output from GRETL

```
Using numerical derivatives
Tolerance = 1,81899e-012
Function evaluations: 44
Evaluations of gradient: 8

Model 8: 1-step GMM, using observations 1-569
e = l_LABOUR – b1*const – b2*l_OUTPUT – b3*l_CAPITAL

           estimate      std. error   t-ratio     p-value
  ----------------------------------------------------------------------
  b1       3,01483       0,0566474    53,22       0,0000     ***
  b2       0,878061      0,0512008    17,15       6,36e-066 ***
  b3       0,00369851    0,0429567     0,08610    0,9314

  GMM criterion: Q = 1,1394e-031 (TQ = 6,48321e-029)
```

Labour Demand Functions: Comparison of Estimates

OLS and GMM estimates coincide

	OLS	GMM
const	3,015	3,015
	0,057	0,057
L_OUTPUT	0,878	0,878
	0,051	0,051
l_CAPITAL	0,0037	0,0037
	0,0430	0,0430

GMM Estimator: Properties

Under weak regularity conditions, the GMM estimator is

- consistent (for any W_N)
- most efficient if $W_N = W_N^{opt} = [E\{f(w_i, z_i, \theta)\, f(w_i, z_i, \theta)'\}]^{-1}$
- asymptotically normal: $\sqrt{N}(\hat{\theta} - \theta) \rightarrow N\left(0, V^{-1}\right)$

  where $V = D\, W_N^{opt}\, D'$ with the KxR matrix of derivatives

$$D = E\left\{\frac{\partial f(w_i, z_i, \theta)}{\partial \theta'}\right\}$$

The covariance matrix V^{-1} can be estimated by substituting the population parameters θ by sample equivalents $\hat{\theta}$ evaluated at the GMM estimates in D and W_N^{opt}

GMM Estimator: Calculation

1. One-step GMM estimator: Choose a positive definite W_N, *e.g.*, $W_N = I_N$, optimization gives $\hat{\theta}_1$ (consistent, but not efficient)
2. Two-step GMM estimator: use the one-step estimator $\hat{\theta}_1$ to estimate $V = D\, W_N^{opt}\, D'$, repeat optimization with $W_N = V^{-1}$; this gives $\hat{\theta}_2$
3. Iterated GMM estimator: Repeat step 2 until convergence

If $R = K$, the GMM estimator is the same for any W_N, only step 1 is needed; the objective function $Q_N(\theta)$ is zero at the minimum

If $R > K$, step 2 is needed to achieve efficiency

GMM and Other Estimation Methods

- GMM estimation generalizes the method of moments estimation
- Allows for a general concept of moment conditions
- Moment conditions are not necessarily linear in the parameters to be estimated
- Encompasses various estimation concepts such as OLS, GLS, IV, GIV, ML

	moment conditions
OLS	$E\{(y_i - x_i'\beta)\, x_i\} = 0$
GLS	$E\{(y_i - x_i'\beta)\, x_i / \sigma^2 (x_i)\} = 0$
IV	$E\{(y_i - x_i'\beta)\, z_i\} = 0$
ML	$E\{\partial/\partial\beta\, f[\varepsilon_i(\beta)]\} = 0$

Contents

Klein's Model 1

$C_t = \alpha_1 + \alpha_2 P_t + \alpha_3 P_{t-1} + a_4(W_t^p + W_t^g) + \varepsilon_{t1}$ (consumption)

$I_t = \beta_1 + \beta_2 P_t + \beta_3 P_{t-1} + \beta_4 K_{t-1} + \varepsilon_{t2}$ (investments)

$W_t^p = \gamma_1 + \gamma_2 X_t + \gamma_3 X_{t-1} + \gamma_4 t + \varepsilon_{t3}$ (private wages and salaries)

$X_t = C_t + I_t + G_t$

$K_t = I_t + K_{t-1}$

$P_t = X_t - W_t^p - T_t$

C (consumption), *P* (profits), W^p (private wages and salaries), W^g (public wages and salaries), *I* (investments), *K* (capital stock), *X* (production), *G* (governmental expenditures without wages and salaries), *T* (taxes) and *t* [time (trend)]

Endogenous: *C*, *I*, W^p, *X*, *P*, *K*; exogeneous: W^g, *G*, *T*, *t*, P_{-1}, K_{-1}, X_{-1}

Early Econometric Models

Klein's Model

- Aims:
  - to forecast the development of business fluctuations and
  - to study the effects of government economic-political policy
- Successful forecasts of
  - economic upturn rather than a depression after World War II
  - mild recession at the end of the Korean War

Model	year	eq's
Tinbergen	1936	24
Klein	1950	6
Klein & Goldberger	1955	20
Brookings	1965	160
Brookings Mark II	1972	~200

Econometric Models

Basis: the multiple linear regression model

- Adaptations of the model
  - Dynamic models
  - Systems of regression models
  - Time series models
- Further developments
  - Models for panel data
  - Models for spatial data
  - Models for limited dependent variables

Contents

- The GIV Estimator
- GMM Estimation
- Econometric Models
- Dynamic Models
- Multi-equation Models
- Time Series Models
- Models for Limited Dependent Variables
- Panel Data Models
- Econometrics II

Dynamic Models: Examples

Demand model: describes the quantity *Q* demanded of a product as a function of its price *P* and consumers' income *Y*

(a) Current price and current income determine the demand (static model):

$$Q_t = \beta_1 + \beta_2 P_t + \beta_3 Y_t + \varepsilon_t$$

(b) Current price and income of the previous period determine the demand (dynamic model):

$$Q_t = \beta_1 + \beta_2 P_t + \beta_3 Y_{t-1} + \varepsilon_t$$

(c) Current price and demand of the previous period determine the demand (autoregressive model):

$$Q_t = \beta_1 + \beta_2 P_t + \beta_3 Q_{t-1} + \varepsilon_t$$

Dynamic of Processes

Static processes: independent variables have a direct effect, the adjustment of the dependent variable on the realized values of the independent variables is completed within the current period, the process is assumed to be always in equilibrium

Static models may be unsuitable:

(a) Some activities are determined by the past, such as: energy consumption depends on past investments into energy-consuming systems and equipment

(b) Actors of the economic processes often respond with delay, e.g., due to the duration of decision-making and procurement processes

(c) Expectations: e.g., consumption depends not only on current income but also on income expectations in future; modelling of income expectation based on past income development

Elements of Dynamic Models

1. Lag-structures, distributed lags: describe the delayed effects of one or more regressors on the dependent variable; e.g., the lag-structure of order *s* or DL(*s*) model (DL: distributed lag)
   $$Y_t = \alpha + \Sigma^s_{i=0}\beta_i X_{t-i} + \varepsilon_t$$
2. Geometric lag-structure, Koyck's model: infinite lag-structure with $\beta_i = \lambda_0\lambda^i$ $(0 < \lambda < 1)$
3. ADL-model: autoregressive model with lag-structure, e.g., the ADL(1,1)-model
   $$Y_t = \alpha + \varphi Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t$$
4. Error-correction model
   $$\Delta Y_t = -(1-\varphi)(Y_{t-1} - \mu_0 - \mu_1 X_{t-1}) + \beta_0 \Delta X_t + \varepsilon_t$$
   obtained from the ADL(1,1)-model with $\mu_0 = \alpha/(1-\varphi)$ und $\mu_1 = (\beta_0+\beta_1)/(1-\varphi)$

The Koyck Transformation

Transforms the model

$$Y_t = \lambda_0 \Sigma_i \lambda^i X_{t-i} + \varepsilon_t$$

into an autoregressive model ($v_t = \varepsilon_t - \lambda\varepsilon_{t-1}$):

$$Y_t = \lambda Y_{t-1} + \lambda_0 X_t + v_t$$

- The model with infinite lag-structure in X becomes a model
  - with an autoregressive component λY_{t-1}
  - with a single regressor X_t and
  - with autocorrelated error terms
- Econometric applications
  - The adaptive expectations model

    Example: Investments determined by expected profit X^e:

    $$X^e_{t+1} = \lambda X^e_t + (1 - \lambda) X_t \quad \text{(with } 0 < \lambda < 1\text{)}$$
  - The partial adjustment model

    Example: K^p_t: planned stock for t; strategy for adapting K_t on K^p_t

    $$K_t - K_{t-1} = \delta(K^p_t - K_{t-1})$$

Contents

- The GIV Estimator
- GMM Estimation
- Econometric Models
- Dynamic Models
- Multi-equation Models
- Time Series Models
- Models for Limited Dependent Variables
- Panel Data Models
- Econometrics II

Example: Income and Consumption

Consumption C_t and disposable income Y_t are simultaneously determined by

$$C_t = \beta_1 + \beta_2 Y_t + \varepsilon_t \qquad \text{(A)}$$

$$Y_t = C_t + I_t \qquad \text{(B)}$$

- The disposable income Y_t is determined by the consumption C_t
- Equations (A) and (B) are the structural equations or the structural form of the simultaneous equation model that describes both C_t and Y_t
- The coefficients β_1 and β_2 are behavioural parameters
- In equation (A), Y_t is endogenous: The OLS estimates b_1 and b_2 are biased and not consistent

Multi-equation Models

Economic phenomena are usually characterized by the behaviour of more than one dependent variable

Multi-equation model: the number of equations determines the number of dependent variables which are described by the model

Characteristics of multi-equation models:

- Types of equations
- Types of variables
- Identifiability

Klein's Model 1

$C_t = \alpha_1 + \alpha_2 P_t + \alpha_3 P_{t-1} + a_4(W_t^p + W_t^g) + \varepsilon_{t1}$ (consumption)

$I_t = \beta_1 + \beta_2 P_t + \beta_3 P_{t-1} + \beta_4 K_{t-1} + \varepsilon_{t2}$ (investments)

$W_t^p = \gamma_1 + \gamma_2 X_t + \gamma_3 X_{t-1} + \gamma_4 t + \varepsilon_{t3}$ (private wages and salaries)

$X_t = C_t + I_t + G_t$

$K_t = I_t + K_{t-1}$

$P_t = X_t - W_t^p - T_t$

C (consumption), *P* (profits), W^p (private wages and salaries), W^g (public wages and salaries), *I* (investments), *K* (capital stock), *X* (production), *G* (governmental expenditures without wages and salaries), *T* (taxes) and *t* [time (trend)]

Endogenous: *C*, *I*, W^p, *X*, *K*, *P*; exogeneous: W^g, *G*, *T*, *t*, P_{-1}, K_{-1}, X_{-1}

Types of Equations

- Behavioural or structural equations: describe the behaviour of a dependent variable as a function of explanatory variables
- Definitional identities: define how a variable is defined as the sum of other variables, e.g., decomposition of gross domestic product as the sum of its consumption components

  Example: Klein's model 1: $X_t = C_t + I_t + G_t$
- Equilibrium conditions: assume a certain relationship, which can be interpreted as an equilibrium

  Example: equality of demand (Q^d) and supply (Q^s) in a market model: $Q_t^d = Q_t^s$

Definitional identities and equilibrium conditions have no error terms

Types of Variables

Specification of a multi-equation model: definition of

- variables which are explained by the model (endogenous variables)
- other variables which are used in the model

Number of equations needed in the model: same number as that of the endogenous variables in the model

Explanatory or exogenous variables: uncorrelated with error terms

- strictly exogenous variables: uncorrelated with error terms ε_{t+i} (for any $i \neq 0$)
- predetermined variables: uncorrelated with current and future error terms (ε_{t+i}, $i \geq 0$); lagged explanatory variables

Error terms:

- Uncorrelated over time
- Error terms from different equations and same observation period typically correlated, contemporaneous correlation

Systems of Regression Equations

Economic processes encompass the simultaneous developments as well as interrelations of a set of dependent variables

- For modelling economic processes: system of relations, typically in the form of regression equations: multi-equation model

Example: Two dependent variables y_{t1} and y_{t2} are modelled as

$$y_{t1} = x'_{t1}\beta_1 + \varepsilon_{t1}$$

$$y_{t2} = x'_{t2}\beta_2 + \varepsilon_{t2}$$

with $V\{\varepsilon_{ti}\} = \sigma_i^2$ for $i = 1, 2$, $Cov\{\varepsilon_{t1}, \varepsilon_{t2}\} = \sigma_{12} \neq 0$

Typical situations:

1. The set of regressors x_{t1} and x_{t2} coincide
2. The set of regressors x_{t1} and x_{t2} differ, may overlap
3. Regressors contain one or both dependent variables
4. Regressors contain lagged variables

Capital Asset Pricing Model

Capital asset pricing (CAP) model: describes the return R_i of asset i

$$R_i - R_f = \beta_i(E\{R_m\} - R_f) + \varepsilon_i$$

with

- R_f: return of a risk-free asset
- R_m: return of the market's optimal portfolio

- β_i: indicates how strong fluctuations of the returns of asset i are determined by fluctuations of the market as a whole
- Knowledge of the return difference $R_i - R_f$ will give information on the return difference $R_j - R_f$ of asset j, at least for some assets
- Analysis of a set of assets $i = 1, \ldots, s$
  - The error terms ε_i, $i = 1, \ldots, s$, represent common factors, e.g., inflation rate, have a common dependence structure
  - Efficient use of information: simultaneous analysis

A Model for Investment

Grunfeld investment data (Grunfeld & Griliches, 1960): Panel data set on gross investments I_{it} of firms i = 1, ..., 6 over 20 years and related data

- Investment decisions are assumed to be determined by

  $$I_{it} = \beta_{i1} + \beta_{i2}F_{it} + \beta_{i3}C_{it} + \varepsilon_{it}$$

  with
  - F_{it}: market value of firm i at the end of year t-1
  - C_{it}: value of stock of plant and equipment at the end of year t-1
- Simultaneous analysis of equations for the various firms i: efficient use of information
  - Error terms for the firms include common factors such as economic climate
  - Coefficients may be the same for the firms

The Hog Market

Model equations:

$$Q^d = \alpha_1 + \alpha_2 P + \alpha_3 Y + \varepsilon_1 \quad \text{(demand equation)}$$
$$Q^s = \beta_1 + \beta_2 P + \beta_3 Z + \varepsilon_2 \quad \text{(supply equation)}$$
$$Q^d = Q^s \text{ (equilibrium condition)}$$

with Q^d: demanded quantity, Q^s: supplied quantity, P: price, Y: income, and Z: costs of production, or

$$Q = \alpha_1 + \alpha_2 P + \alpha_3 Y + \varepsilon_1 \quad \text{(demand equation)}$$
$$Q = \beta_1 + \beta_2 P + \beta_3 Z + \varepsilon_2 \quad \text{(supply equation)}$$

- Model describes quantity and price of the equilibrium transactions
- Model determines simultaneously Q and P, given Y and Z
- Error terms
  - May be correlated: $\text{Cov}\{\varepsilon_1, \varepsilon_2\} \neq 0$
- Simultaneous analysis necessary for efficient use of information

Types of Multi-equation Models

Multivariate regression or multivariate multi-equation model

- A set of regression equations, each explaining one of the dependent variables
  - Possibly common explanatory variables
  - Seemingly unrelated regression (SUR) model: each equation is a valid specification of a linear regression, related to other equations only by the error terms
  - See cases 1 and 2 of “typical situations” on slide 40

Simultaneous equations models

- Describe the relations within the system of economic variables
  - in form of model equations
  - See cases 3 and 4 of “typical situations” on slide 40

Error terms: dependence structure is specified by means of second moments or as joint probability distribution

Examples of Multi-equation Models

Multivariate regression models

- Capital asset pricing (CAP) model: for all assets, return R_i (or risk premium $R_i - R_f$) is a function of $E\{R_m\} - R_f$; dependence structure of the error terms caused by common variables
- Model for investment: firm-specific regressors, dependence structure of the error terms like in CAP model
- Seemingly unrelated regression (SUR) models

Simultaneous equations models

- Hog market model: endogenous regressors, dependence structure of error terms
- Klein's model I: endogenous regressors, dynamic model, dependence of error terms from different equations and possibly over time

Single- vs. Multi-equation Models

Complications for estimation of parameters of multi-equation models:

- Dependence structure of error terms
- Violation of exogeneity of regressors

Example: Hog market model, demand equation

$$Q = \alpha_1 + \alpha_2 P + \alpha_3 Y + \varepsilon_1$$

- Covariance matrix of $\varepsilon = (\varepsilon_1, \varepsilon_2)'$

$$\text{Cov}\{\varepsilon\} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

- P is not exogenous: $\text{Cov}\{P,\varepsilon_1\} = (\sigma_1^2 - \sigma_{12})/(\beta_2 - \alpha_2) \neq 0$

Statistical analysis of multi-equation models requires methods adapted to these features

Multi-equation Models: Estimation of Parameters

Estimation procedures

- Multivariate regression models
  - FGLS , GLS, ML
- Simultaneous equations models
  - Single equation methods: indirect least squares (ILS), two stage least squares (TSLS), limited information ML (LIML)
  - System methods of estimation: three stage least squares (3SLS), full information ML (FIML)
  - Dynamic models: estimation methods for vector autoregressive (VAR) and vector error correction (VEC) models

Contents

Types of Trend

Trend: The expected value of a process $\{Y_t, t = 1, 2, ...\}$ increases or decreases with time

- Deterministic trend: a function $f(t)$ of the time t, describing the evolution of $E\{Y_t\}$ over time

  $$Y_t = f(t) + \varepsilon_t, \ \varepsilon_t\text{: white noise}$$

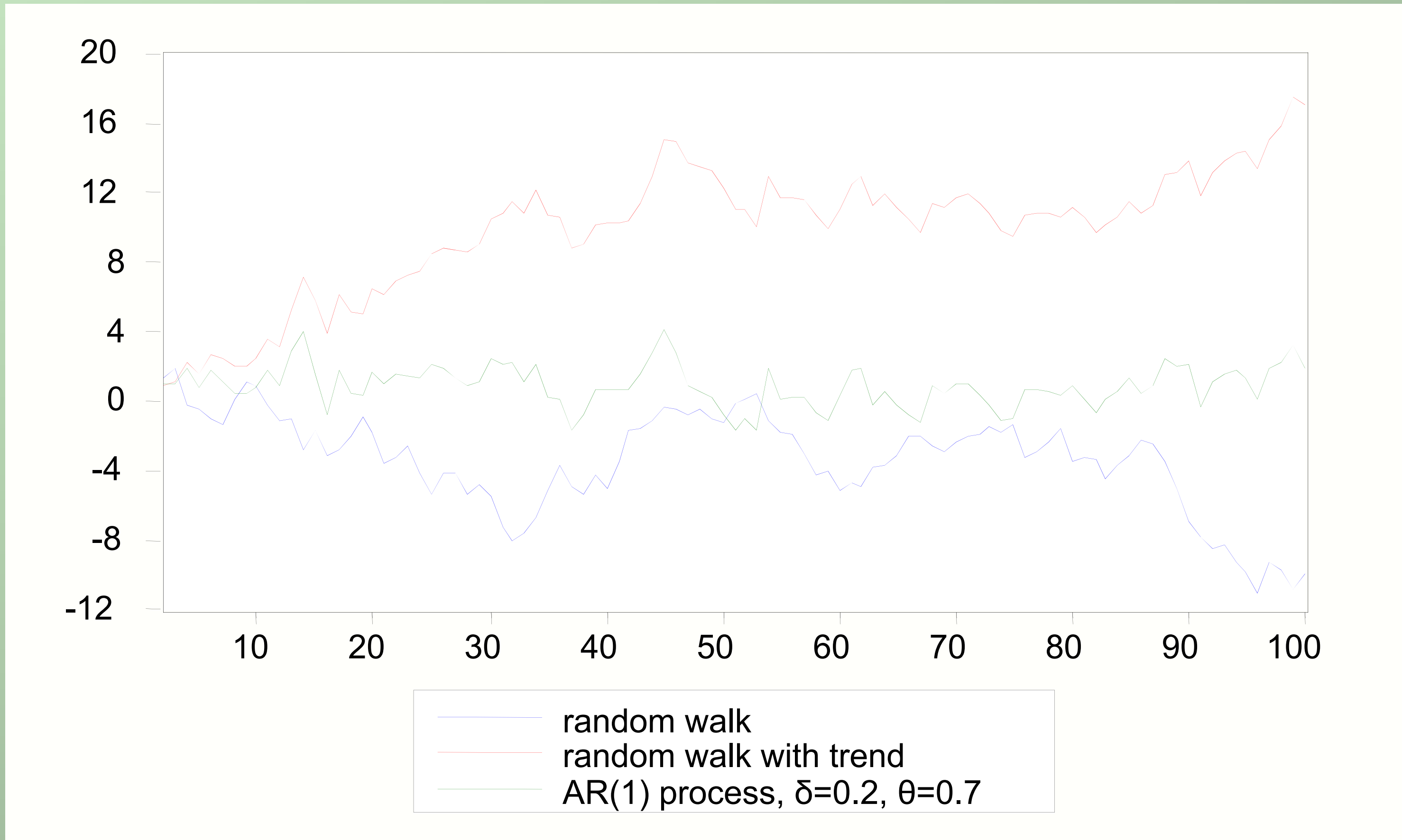
  Example: $Y_t = \alpha + \beta t + \varepsilon_t$ describes a linear trend of Y; an increasing trend corresponds to $\beta > 0$
- Stochastic trend: $Y_t = \delta + Y_{t-1} + \varepsilon_t$ or

  $$\Delta Y_t = Y_t - Y_{t-1} = \delta + \varepsilon_t, \ \varepsilon_t\text{: white noise}$$

  - describes an irregular or random fluctuation of the differences ΔY_t around the expected value δ
  - AR(1) – or AR(p) – process with unit root
  - "random walk with trend"

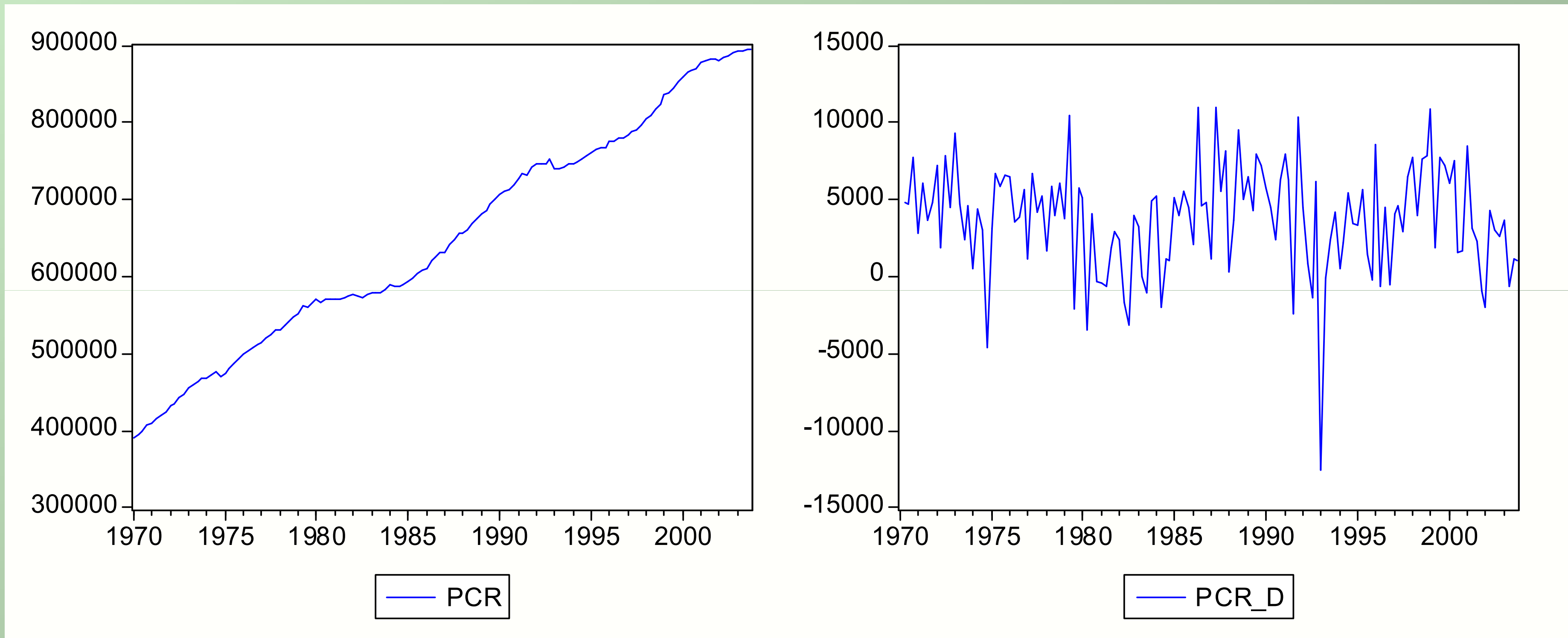
Trends: Random Walk and AR Process

Random walk: $Y_t = Y_{t-1} + \varepsilon_t$; random walk with trend: $Y_t = 0.1 + Y_{t-1} + \varepsilon_t$; AR(1) process: $Y_t = 0.2 + 0.7 Y_{t-1} + \varepsilon_t$; ε_t simulated from $N(0,1)$



Example: Private Consumption

Private consumption, AWM database; level values (PCR) and first differences (PCR_D); random walk?



Mean of PCD_D: 3740

How to Model Trends?

Specification of a

- deterministic trend, e.g., $Y_t = \alpha + \beta t + \varepsilon_t$: risk of spurious regression, wrong decisions
- stochastic trend: analysis of differences ΔY_t if a random walk, i.e., a unit root, is suspected

Spurious Regression: An Illustration

Independent random walks: $Y_t = Y_{t-1} + \varepsilon_{yt}$, $X_t = X_{t-1} + \varepsilon_{xt}$

ε_{yt}, ε_{xt}: independent white noises with variances $\sigma_y^2 = 2$, $\sigma_x^2 = 1$

Fitting the model

$Y_t = \alpha + \beta X_t + \varepsilon_t$

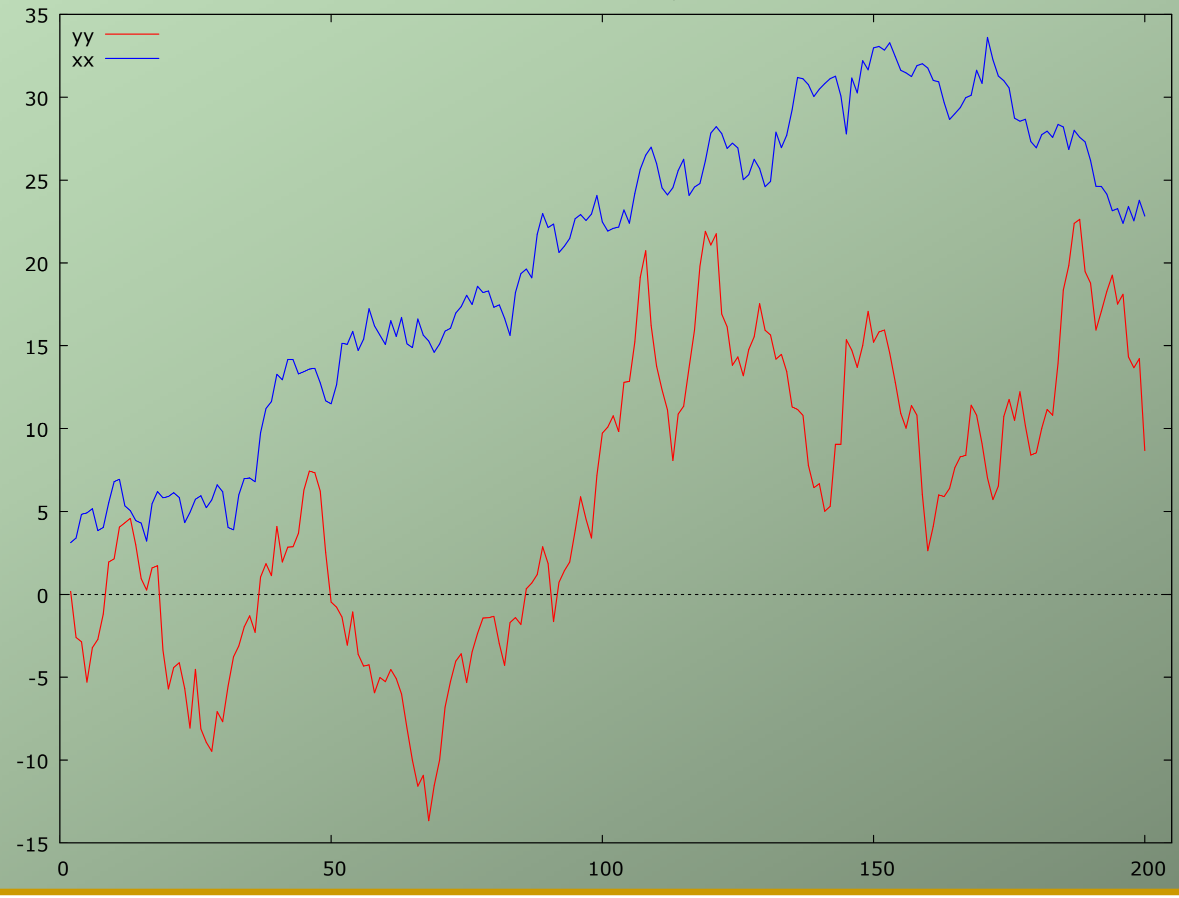
gives

$\hat{Y}_t = - 8.18 + 0.68 X_t$

t-statistic for *X*: $t = 17.1$

p-value = 1.2 E-40

$R^2 = 0.50$, DW = 0.11



Models in Non-stationary Time Series

Let $X_t \sim I(1)$, $Y_t \sim I(1)$ be integrated of order 1 and the model be

$$Y_t = \alpha + \beta X_t + \varepsilon_t$$

it follows in general that $\varepsilon_t \sim I(1)$, i.e., the error terms are non-stationary

Consequences for OLS estimation of α and β

- (Asymptotic) distributions of *t*- and *F*-statistics are not the *t*- and *F*-distribution
- *t*-statistic, R^2 indicate explanatory potential
- Highly autocorrelated residuals, DW statistic converges for growing length of time series to zero

Nonsense or spurious regression (Granger & Newbold, 1974)

- Non-stationary time series are trended; non-stationarity causes an apparent relationship

Avoiding Spurious Regression

- Identification of non-stationarity: unit-root tests
- Models for non-stationary variables
  - Elimination of stochastic trends: specifying the model for differences
  - Inclusion of lagged variables may result in stationary error terms
  - Explained and explanatory variables may have a common stochastic trend, are cointegrated: equilibrium relation, error-correction models

Unit Root Tests

AR(1) process $Y_t = \delta + \theta Y_{t-1} + \varepsilon_t$ with white noise ε_t

- Dickey-Fuller or DF test (Dickey & Fuller, 1979)

  Test of H_0: $\theta = 1$ against H_1: $\theta < 1$
- KPSS test (Kwiatkowski, Phillips, Schmidt & Shin, 1992)

  Test of H_0: $\theta < 1$ against H_1: $\theta = 1$
- Augmented Dickey-Fuller or ADF test

  extension of DF test
- Various modifications like Phillips-Perron test, Dickey-Fuller GLS test, etc.

The Error-correction Model

ADL(1,1) model with $Y_t \sim I(1)$, $X_t \sim I(1)$

$$Y_t = \delta + \theta Y_{t-1} + \varphi_0 X_t + \varphi_1 X_{t-1} + \varepsilon_t$$

- Common trend implies an equilibrium relation, i.e.,

  $$Y_{t-1} - \beta X_{t-1} \sim I(0)$$

  error-correction form of the ADL(1,1) model

  $$\Delta Y_t = \varphi_0 \Delta X_t - (1 - \theta)(Y_{t-1} - \alpha - \beta X_{t-1}) + \varepsilon_t$$

Error-correction model describes

- the short-run behaviour
- consistently with the long-run equilibrium $Y_t = \alpha + \beta X_t$

Cointegration

Non-stationary variables $X_t \sim I(1)$, $Y_t \sim I(1)$

$$Y_t = \alpha + \beta X_t + \varepsilon_t$$

- X_t and Y_t are cointegrated: $\varepsilon_t \sim I(0)$
- X_t and Y_t are not cointegrated: $\varepsilon_t \sim I(1)$

Tests for cointegration:

- If β is known, unit root test based on differences $Y_t - \beta X_t$
- Test procedures
  - Unit root test (DF or ADF) based on residuals e_t
  - Cointegrating regression Durbin-Watson (CRDW) test: DW statistic
  - Johansen technique: extends the cointegration technique to the multivariate case

Vector Error-Correction Model

Y_t: *k*-vector, each component *I*(1)

VAR(*p*) model for the *k*-vector Y_t

$$Y_t = \delta + \Theta_1 Y_{t-1} + \ldots + \Theta_p Y_{t-p} + \varepsilon_t$$

transformed into

$$\Delta Y_t = \delta + \Gamma_1 \Delta Y_{t-1} + \ldots + \Gamma_{p-1} \Delta Y_{t-p+1} + \Pi Y_{t-1} + \varepsilon_t$$

with r{Π} = *r* and Π = γβ' gives

$$\Delta Y_t = \delta + \Gamma_1 \Delta Y_{t-1} + \ldots + \Gamma_{p-1} \Delta Y_{t-p+1} + \gamma\beta' Y_{t-1} + \varepsilon_t \qquad \text{(B)}$$

- *r* cointegrating relations $\beta' Y_{t-1}$
- Adaptation parameters γ measure the portion or speed of adaptation of Y_t in compensation of the equilibrium error $Z_{t-1} = \beta' Y_{t-1}$
- Equation (B) is called the vector error-correction (VEC) form of the VAR(*p*) model

Contents

Example

To be explained whether a household owns a car: explanatory power have

- income
- household size
- etc.

Regression for describing car-ownership is not suitable!

- Owning a car has two manifestations: yes/no
- Indicator for owning a car is a binary variable

Models are needed that allow describing a binary dependent variable or a, more generally, limited dependent variable

Cases of Limited Dependent Variable

Typical situations: functions of explanatory variables are used to describe or explain

- Dichotomous dependent variable, e.g., ownership of a car (yes/no), employment status (employed/unemployed), etc.
- Ordered response, e.g., qualitative assessment (good/average/bad), working status (full-time/part-time/not working), etc.
- Multinomial response, e.g., trading destinations (Europe/Asia/Africa), transportation means (train/bus/car), etc.
- Count data, e.g., number of orders a company receives in a week, number of patents granted to a company in a year
- Censored data, e.g., expenditures for durable goods, duration of study with drop outs

Contents

- The GIV Estimator
- GMM Estimation
- Econometric Models
- Dynamic Models
- Multi-equation Models
- Time Series Models
- Models for Limited Dependent Variables
- Panel Data Models
- Econometrics II

Panel Data

Population of interest: individuals, households, companies, countries

Types of observations

- Cross-sectional data: Observations of all units of a population, or of a (representative) subset, at one specific point in time; e.g., wages in 2015
- Time series data: Series of observations on units of the population over a period of time; e.g., wages of a worker in 2009 through 2015
- Panel data: Repeated observations of (the same) population units collected over a number of periods; data set with both a cross-sectional and a time series aspect; multi-dimensional data

Cross-sectional and time series data are one-dimensional, special cases of panel data

Pooling independent cross-sections: (only) similar to panel data

Panel Data: Three Types

Typically data at micro-economic level (individuals, households, firms), but also at macro-economic level (e.g., countries)

Notation:

- N: Number of cross-sectional units
- T: Number of time periods

Types of panel data:

- Large T, small N: "long and narrow"
- Small T, large N: "short and wide"
- Large T, large N: "long and wide"

Some Examples

Verbeek's data set "males": Wages and related variables

- short and wide panel (N = 545, T = 8)
- rich in information (~40 variables)

Grunfeld investment data: Investments in plant and equipment by

- N = 10 firms
- for each of T = 20 yearly observations for 1935-1954

Penn World Table: Purchasing power parity and national income accounts for

- N = 189 countries/territories
- for some or all of the years 1950-2009 ($T \leq 60$)

Example: Individual Wages

Verbeek's data set "males"

- Sample of
  - 545 full-time working males, end of schooling in 1980
  - from each person: yearly data collection from 1980 till 1987
- Variables
  - *wage*: log of hourly wage (in USD)
  - *school*: years of schooling
  - *exper*: age – 6 – *school*
  - dummies for union membership, married, black, Hispanic, public sector
  - others

Use of Panel Data

Econometric models for describing the behaviour of cross-sectional units over time

Panel data models

- Allow controlling individual differences, comparing behaviour, analysing dynamic adjustment, measuring effects of policy changes
- More realistic models than cross-sectional and time-series models
- Allow more detailed or sophisticated research questions

Methodological implications

- Dependence of sample units in time-dimension
- Some variables might be time-constant (e.g., variable *school* in “males”, population size in the Penn World Table dataset)
- Missing values

Examples for Fixed- and Random-effects

Grunfeld investment data: Investment model

$$I_{it} = \alpha_i + \beta_{i1}F_{it} + \beta_{i2}C_{it} + u_{it}$$

with F_{it}: market value, C_{it}: value of stock of plant and equipment, both of firm *i* at the end of year *t*-1

- N = 10 firms, T = 20 yearly observations
- Fixed effects α_i allow for firm-specific, time-constant factors

Wage equation

$$wage_{it} = \beta_1 + \beta_2\, exper_{it} + \beta_3\, exper2_{it} + \beta_4\, school_{it} + \beta_5\, union_{it} + \beta_6\, mar_{it} + \beta_7\, black_{it} + \beta_8\, rural_{it} + \alpha_i + u_{it}$$

with composite error $\varepsilon_{it} = \alpha_i + u_{it}$

- α_i: unit-specific parameter for each of 545 units
- Time-constant factors α_i: stochastic variables with identical distribution
- Regressors are uncorrelated with u_{it}

Models for Panel Data

Model for y, based on panel data from N cross-sectional units and T periods

$$y_{it} = \beta_0 + x_{it}'\beta_1 + \varepsilon_{it}$$

$i = 1, ..., N$: sample unit

$t = 1, ..., T$: time period of sample

x_{it} and β_1: K-vectors

- β_0 and β_1: represent intercept and K regression coefficients; are assumed to be identical for all units and all time periods
- ε_{it}: represents unobserved factors that may affect y_{it}
  - Assumption that ε_{it} are uncorrelated over time not realistic; refer to the same unit or individual
  - Standard errors of OLS estimates misleading, OLS estimation not efficient (does not exploit dependence structure over time)

Fixed Effects Model

The general model

$$y_{it} = \beta_0 + x_{it}'\beta_1 + \varepsilon_{it}$$

- Specification for the error terms: two components

  $$\varepsilon_{it} = \alpha_i + u_{it}$$

  - α_i fixed, unit-specific, time-constant factors, also called unobserved (individual) heterogeneity; may be correlated with x_{it}
  - $u_{it} \sim IID(0, \sigma_u^2)$; homoskedastic, uncorrelated over time; represents unobserved factors that change over time, also called idiosyncratic or time-varying error
  - ε_{it} : also called composite error
- Fixed effects (FE) model

  $$y_{it} = \Sigma_j \alpha_i d_{ij} + x_{it}'\beta_1 + u_{it}$$

  d_{ij}: dummy variable for unit i: $d_{ij} = 1$ if $i = j$, otherwise $d_{ij} = 0$
- Overall intercept β_0 omitted; unit-specific intercepts α_i

Fixed Effects Estimator

"Within transformation": transforms y_{it} into time-demeaned $\ddot{y}_{it}$ by subtracting the average $\bar{y}_i = (\Sigma_t y_{it})/T$: $\ddot{y}_{it} = y_{it} - \bar{y}_i$; analogously $\ddot{x}_{it}$ and $\ddot{u}_{it}$, for all i and t

$$b_{FE} = (\Sigma_i \Sigma_t \ddot{x}_{it} \ddot{x}_{it}')^{-1} \Sigma_i \Sigma_t \ddot{x}_{it} \ddot{y}_{it}$$

- Unbiased if all x_{it} are independent of all u_{it}
- Consistent (for $N \rightarrow \infty$) if x_{it} are strictly exogenous, i.e., $E\{x_{it} u_{is}\} = 0$ for all s, t
- Asymptotically normally distributed
- Covariance matrix

$$V\{b_{FE}\} = \sigma_u^2 (\Sigma_i \Sigma_t \ddot{x}_{it} \ddot{x}_{it}')^{-1}$$

Random Effects Model

Starting point is again the model

$$y_{it} = \beta_0 + x_{it}'\beta_1 + \varepsilon_{it}$$

with composite error $\varepsilon_{it} = \alpha_i + u_{it}$

- Specification for the error terms:
  - $u_{it} \sim IID(0, \sigma_u^2)$; homoskedastic, uncorrelated over time
  - $\alpha_i \sim IID(0, \sigma_a^2)$; represents all unit-specific, time-constant factors; correlation of error terms over time only via the α_i
  - α_i and u_{it} are assumed to be mutually independent and independent of x_{js} for all j and s
- Random effects (RE) model

$$y_{it} = \beta_0 + x_{it}'\beta_1 + \alpha_i + u_{it}$$

- Unbiased and consistent ($N \rightarrow \infty$) estimation of β_0 and β_1
- Efficient estimation of β_0 and β_1: takes error covariance structure into account; GLS estimation

Contents

Econometrics II

1. ML Estimation and Specification Tests (MV, Ch.6)
2. Models with Limited Dependent Variables (MV, Ch.7)
3. Univariate time series models (MV, Ch.8)
4. Multivariate time series models, part 1 (MV, Ch.9)
5. Multivariate time series models, part 2 (MV, Ch.9)
6. Models Based on Panel Data (MV, Ch.10)

Univariate Time Series Models

- Time Series
- Stochastic Processes
- Stationary Processes
- The ARMA Process
- Deterministic and Stochastic Trends
- Models with Trend
- Unit Root Tests

Multivariate Time Series Models

- Dynamic Models
- Lag Structures, ADL Models
- Models with Non-stationary Variables
- Cointegration, Tests for Cointegration
- Error-correction Model
- Systems of Equations
- VAR Models
- Simultaneous Equations and VAR Models
- VAR Models and Cointegration
- VEC Models