Econometrics - Lecture 6

GMM-Estimator and Econometric Models

Contents

Assumption (A7): $E\{x_t \varepsilon_t\} = 0$ for all t

Linear model for y_t

$$y_t = x_t'\beta + \varepsilon_t,\ t = 1, \ldots, T \quad (\text{or } y = X\beta + \varepsilon)$$

given observations x_{tk}, $k = 1, \ldots, K$, of the regressor variables, error term ε_t

(A7) $E\{x_t \varepsilon_t\} = 0$ for all t, i.e., no contemporary correlation

- Guaranties unbiasedness and consistency of the OLS estimator
- In reality, (A7) not always fulfilled
- $E\{x_t \varepsilon_t\} \neq 0$: biased, inconsistent OLS estimator
- Examples of situations with $E\{x_t \varepsilon_t\} \neq 0$
  - Regressors with measurement errors
  - Regression on the lagged dependent variable with autocorrelated error terms
  - Endogeneity of regressors
  - Simultaneity

Instrumental Variables

The model is

$$y_t = x_t'\beta + \varepsilon_t$$

with $V\{\varepsilon_i\} = \sigma_\varepsilon^2$ and

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

- Instrumental variables z_t
  1. Exogenous: $E\{\varepsilon_t\, z_t\} = 0$: z_t uncorrelated with error term
  2. Relevant: $Cov\{x_t\, ,\, z_t\} \neq 0$: z_t correlated with endogenous regressors

IV Estimator

- Based on the moment conditions

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

- Solution of corresponding sample moment conditions

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

IV estimator based on the instruments z_t

$$\hat{\beta}_{IV} = \left(\sum\nolimits_t z_t x_t'\right)^{-1} \left(\sum\nolimits_t z_t y_t\right) = \left(Z'X\right)^{-1} Z'y$$

Identification requires that the KxK matrix $\Sigma_t z_t x_t' = Z'X$ is finite and invertible; instruments z_t are relevant when this is fulfilled

IV Estimator: Properties

IV estimator is

- Consistent
- (Asymptotic) covariance matrix

$$V\{\hat{\beta}_{IV}\} = \sigma^2 \left[\left(\sum_t x_t z_t'\right)\left(\sum_t z_t z_t'\right)^{-1}\left(\sum_t z_t x_t'\right)\right]^{-1}$$

- Estimated covariance matrix: σ² is substituted by

$$\hat{\sigma}^2 = \tfrac{1}{T}\sum_t \left(y_t - x_t'\hat{\beta}_{IV}\right)^2$$

- The asymptotic distribution of IV estimators, given IID(0, σ_ε^2) error terms, leads to the approximate distribution

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

with the estimated covariance matrix

Contents

The General Case

R: number of instrument variables and of components of z_i

The *R* moment equations are

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

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: Z'X* has not full rank, is not invertible; infinite many solutions fulfill moment equations, but no consistent estimator; under-identified or not identified model
3. *R* > *K*: more instruments than necessary for identification; over-identified model; a unique solution cannot be obtained such that all *R* sample moment conditions are fulfilled; strategy for choosing the estimator among all possible estimators

The GIV Estimator

For $R > K$, in general, a unique solution of all R sample moment conditions cannot be obtained; instead:

Generalized instrumental variable (GIV) estimator

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

uses best approximations $\hat{X} = Z(Z'Z)^{-1}Z'X$ for columns of X

- The GIV estimator can be written as

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

- 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($0,\sigma_\varepsilon^2$) error terms ε_t, leads to the approximate distribution

  $$N\left(\beta, \hat{V}(\hat{\beta}_{IV})\right)$$
- The (asymptotic) covariance matrix of is given by

  $$V\left\{\hat{\beta}_{IV}\right\} = \sigma^2 \left[\left(\sum_t x_t z_t'\right)\left(\sum_t z_t z_t'\right)^{-1}\left(\sum_t z_t x_t'\right)\right]^{-1}$$
- Estimated covariance matrix: σ^2 is substituted by

  $$\hat{\sigma}^2 = \tfrac{1}{T}\sum_t \left(y_t - x_t'\hat{\beta}_{IV}\right)^2$$

Contents

The Generalized IV Estimator

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 IV estimator

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

- For each positive definite weighting matrix W_N, the generalized IV estimator is consistent
- GIV estimator: special case with W_N^{opt} (see below)

For $R = K$, the matrix $Z'X$ is square and invertible; the IV estimator is $(Z'X)^{-1}Z'y$ for any W_N

Most Efficient IV Estimator

Weighting matrix W_N

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

  $W_N^{opt} = [1/N(Z'Z)]^{-1}$

  - Corresponds to the most efficient IV estimator

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

    with $\hat{X} = Z(Z'Z)^{-1}Z'X$
  - Coincides with the GIV (or TSLS) estimator

Consistency of the Generalized IV Estimator

With a *RxR* positive definite weighting matrix W_N, minimizing the weighted quadratic form in the sample moments

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

results in a consistent estimator for β

- Sample moments converge asymptotically to the corresponding population moments
- The population moments are zero for the true parameters
- Minimizing the quadratic loss function in the sample moments results in solutions which asymptotically coincide with the true parameters

This idea is basis of the generalized method of moments estimator

Generalized Method of Moments (GMM) Estimator

- GMM generalizes the IV estimation concept
- Estimates of model parameters are derived from moment conditions which are not necessarily linear
- Number of moment conditions at least as large as number of unknown parameters

Generalized Method of Moments (GMM) Estimator

The model is characterized by R moment conditions

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

[generalization of $E\{(y_i - x_i'\beta) z_i\} = 0$]

- $f(.)$: 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 moment conditions

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

1. $R = K$: unique solution for θ; if $f(.)$ is nonlinear in θ, numerical solution might be derived
2. $R < K$: parameters not identified

GMM Estimator

3. *R* > *K*: minimization, wrt θ, of the objective function, i.e., the quadratic form

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

W_N: symmetric, positive definite weighting matrix

GMM estimator corresponds to the optimal weighting matrix

$$W_N^{opt} = \left(E\{f(w_i, z_i, \hat{\theta})\, f(w_i, z_i, \hat{\theta})'\} \right)^{-1}$$

the inverse of the covariance matrix of the sample moments, and is the most efficient estimator

For nonlinear *f*(.)

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

Example: The Linear Model

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

- Moment or orthogonality conditions:

  $$E\{\varepsilon_t\, x_t\} = E\{(y_t - x_t\text{‘}\beta)x_t\} = 0$$

  $f(.) = (y_i - x_i\text{‘}\beta)x_i$, $\theta = \beta$, instrument variables: x_i; moment conditions are exogeneity conditions for x_i
- Sample moment conditions:

  $$1/N\ \Sigma_i\ (y_i - x_i\text{‘}b)\ x_i = 1/N\ \Sigma_i\ e_i\ x_i = g_N(b) = 0$$
- With $W = I$, $Q_N(\beta) = [1/N\ \Sigma_i\ e_i\ x_i\ ]^2$
- OLS and GMM estimators coincide, but for the estimators
  - OLS: residual sum of squares $S_N(b) = 1/N\ \Sigma_i\ e_i^2$ has its minimum
  - GMM: $Q_N(b) = 0$

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

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

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

- 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: Labor 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

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

Labor Demand Function: OLS Estimation

In logarithmic transforms: Output from GRETL

Dependent variable : l_LABOR
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

Specification of GMM Estimation

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

Labor 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_LABOR - 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)

Linear Model: MM Estimator

Model

$y_i = x_i'\beta + \varepsilon_i$

with $V\{\varepsilon_i\} = \sigma_\varepsilon^2$ and $E\{\varepsilon_i\, x_i\} \neq 0$ and R instrumental variables z_i

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

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

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

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

method of moments estimator

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

b_{MM} coincides with the IV estimator if $R = K$

Contents

GMM Estimator

Model with *R* moment conditions

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

Sample moment conditions

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

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

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

W_N: symmetric, positive definite weighting matrix

- The GMM estimator is consistent for any choice of W_N
- Optimal weighting matrix

$$W_N^{opt} = \left(E\{f(w_i, z_i, \hat{\theta})\, f(w_i, z_i, \hat{\theta})'\} \right)^{-1}$$

the inverse of the covariance matrix of the sample moments, gives the most efficient estimator

For nonlinear *f*(.)

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

GMM Estimator: Properties

Under weak regularity conditions, the GMM estimator is

- consistent (for any *W*)
- most efficient if *W* = $\mathrm{W}^{\mathrm{opt}} = \left(\mathrm{E}\{ \mathrm{f}(\mathrm{w}_i, \mathrm{z}_i, \hat{\theta})\ \mathrm{f}(\mathrm{w}_i, \mathrm{z}_i, \hat{\theta})' \} \right)^{-1}$
- asymptotically normal: $\sqrt{N}(\hat{\theta} - \theta) \to N\left(0, V^{-1}\right)$

  where $V = D\, W^{opt} D'$ with the *K*x*R* 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 in *D* and W^{opt} the population moments by sample equivalents evaluated at the GMM estimates

GMM Estimator: Calculation

1. One-step GMM estimator: Choose a positive definite *W, e.g.,* $W = I$, 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 $W = V^{-1}$; 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, 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

Contents

Klein‘s Model 1

$$C_t = \alpha_1 + \alpha_2 P_t + \alpha_3 P_{t-1} + a_4(W_t^p + W_t^g) + \varepsilon_{t1} \quad \text{(consumption)}$$
$$I_t = \beta_1 + \beta_2 P_t + \beta_3 P_{t-1} + \beta_4 K_{t-1} + \varepsilon_{t2} \quad \text{(investments)}$$
$$W_t^p = \gamma_1 + \gamma_2 X_t + \gamma_3 X_{t-1} + \gamma_4 t + \varepsilon_{t3} \quad \text{(private wages and salaries)}$$
$$X_t = C_t + I_t + G_t$$
$$K_t = I_t + K_{t-1}$$
$$P_t = X_t - W_t^p - T_t$$

C (consumption), P (profits), W^p (private wages and salaries), W^g (public wages and salaries), I (investments), K_{-1} (capital stock, lagged), X (production), G (governmental expenditures without wages and salaries), T (taxes) and t [time (trend)]

Endogenous: C, I, W^p, X, P, K; exogeneous: 1, 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

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 to 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 demand and prices of the previous period determine the demand (dynamic 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 seems always to be 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; modeling 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 $\beta_i = \lambda_0\lambda^i$
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 partial adjustment model

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

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

    Example: Investments determined by expected profit X^e:

    $$X^e_{t+1} = \lambda X^e_t + (1 - \lambda) X_t$$

Contents

Multi-equation Models

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

Multi-equation model: the number of equations determines the number of dependent variables which describe the model

Characteristics of multi-equation models:

- Types of equations
- Types of variables
- Identifiability

Types of Equations

- Behavioral or structural equations: describe the behavior 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 I: $X_t = C_t + I_t + G_t$
- Equilibrium conditions: assume a certain relationship, which can be interpreted as an equilibrium

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)
- Variables which are in addition 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*)
- predetermined variables: uncorrelated with current and future error terms (ε_{t+i}, $i \geq 0$)

Error terms:

- Uncorrelated over time
- Contemporaneous correlation of error terms of different equations possible

Identifiability: An Example

(1) Both demand and supply function are

$$Q = \alpha_1 + \alpha_2 P + \varepsilon$$

Fitted to data gives for both functions the same relationship: not distinguishable whether the coefficients of the demand or the supply function was estimated!

(2) Demand and supply function, respectively, are

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

$$Q = \beta_1 + \beta_2 P + \varepsilon_2$$

Endogenous: Q, P; exogenous: Y

Reduced forms for Q and P are

$$Q = \pi_{11} + \pi_{12} Y + v_1$$

$$P = \pi_{21} + \pi_{22} Y + v_2$$

with parameters π_{ij}

Identifiability: An Example, cont'd

The coefficients of the supply function can uniquely be derived from the parameters π_{ij}:

$$\beta_2 = \pi_{12}/\pi_{22}$$

$$\beta_1 = \pi_{11} - \beta_2\, \pi_{21}$$

consistent estimates of π_{ij} result in consistent estimates for β_i

For the coefficients of the demand function, such unique relations of the π_{ij} can not be found

The supply function is identifiable, the demand function is not identifiable or under-identified

The conditions for identifiability of the coefficients of a model equation are crucial for the applicability of the various estimation procedures

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)