Introduction to Microeconometric Evaluation

SS 2011

Albert-Ludwigs-Universität Freiburg

Alexander Spermann

UNI FREIBURG

1. Endogeneity
2. Simultaneity
3. Missing Variables

- Börsch-Supan, Axel und Jens Köke (2002), An Applied Econometricians‘ View of Empirical Corporate Governance Studies, *German Economic Review,* 3 (3), S. 295-326

1. Endogeneity
Condition of Exogeneity

- True Model (1) $z = \beta x + \varepsilon$

  (2) $x = \gamma z + \eta$

whereas

$$\mathrm{E}(\varepsilon) = 0 \quad \mathrm{E}(\eta) = 0 \quad Cov(\varepsilon, \eta) = \sigma_{\varepsilon\eta}$$

OLS estimation of (1):

$$\hat{\beta} = \frac{Cov(x,z)}{Var(x)}$$

z from (1)

$$= \frac{Cov(x, \beta x + \varepsilon)}{Var(x)}$$

1. Endogeneity
Condition of Exogeneity

I $\quad Cov(x,x) = Var(x)$

II $\quad Cov(z, ax + by) = a\,Cov(x,z) + b\,Cov(y,z)$

Example:

$$\hat{\beta} = \frac{Cov(x, \beta x + \varepsilon)}{Var(x)} = \frac{\beta Cov(x,x) + 1 Cov(\varepsilon, x)}{Var(x)}$$

Rule II

$$= \frac{\beta Var(x)}{Var(x)} + \frac{Cov(\varepsilon, x)}{Var(x)}$$

Rule I

$$= \beta + \frac{Cov(x, \varepsilon)}{Var(x)}$$

(3)

1. Endogeneity

$$Cov(x,\varepsilon)$$

$$= Cov(\gamma z + \eta, \varepsilon)$$

$$= Cov(\gamma(\beta x + \varepsilon) + \eta, \varepsilon)$$

$$= \gamma Cov(\varepsilon, (\beta x + \varepsilon) + 1 Cov(\varepsilon, \eta)$$ Rule II

$$= \gamma[\beta Cov(x,\varepsilon) + 1 \underbrace{Cov(\varepsilon,\varepsilon)]+\sigma_{\varepsilon\eta}}$$ Rule II

$$= Var(\varepsilon)$$ Rule I

$$= \gamma\beta Cov(x,\varepsilon) + \gamma\sigma_{\varepsilon}^{2} + \sigma_{\varepsilon\eta} \qquad (4)$$

1. Endogeneity

Solving for Cov(x,ε):

$$Cov(x,\varepsilon) - \beta\gamma Cov(x,\varepsilon) = \gamma\sigma_{\varepsilon}^{2} + \sigma_{\varepsilon\eta}$$

$$(1-\beta\gamma)Cov(x,\varepsilon) = \gamma\sigma_{\varepsilon}^{2} + \sigma_{\varepsilon\eta}$$

$$Cov(x,\varepsilon) = \frac{\gamma\sigma_{\varepsilon}^{2} + \sigma_{\varepsilon\eta}}{1-\beta\gamma}$$

$$\hat{\beta} = \beta + \frac{Cov(x,\varepsilon)}{Var(x)}$$

1. Endogeneity
Problem of Endogeneity, if

$$Cov(x,\varepsilon) \neq 0$$

Case 1

$$\gamma \neq 0$$

Case 2

$$\sigma_{\varepsilon\eta} \neq 0$$

➔ OLS estimators are biased and inconsistent

1. Endogeneity
Notation Analysis of Endogeneity

1. Covariance

$$Cov(x,\varepsilon)=0$$

$$Cov(u_i,x_i)=0$$

2. Expected Value

$$E(u_i x_i)=0$$

3. Conditional Expected Value

$$E(\varepsilon|x)=0$$

1. Endogeneity
Relation between Covariance and Expected Value

$$Cov(u_i, x_i)$$

$$= E\left[\left(u_i - \underbrace{E[u_i]}_{=0}\right)(x_i - E[x_i])\right]$$

$$= E[u_i(x_i - E[x_i])]$$

$$= E[u_i x_i] - \underbrace{E[u_i]}_{=0} E[x_i]$$

$$= E[u_i x_i]$$

1. Endogeneity
Conditional Expected Value

$$E(\varepsilon|x)=0$$

In case x and y are independent, then:

$$E(y|x)=E(y)$$

conditional expected value	unconditional expected value

$$\Rightarrow E(\varepsilon|x)=E(\varepsilon)=0$$

if ε and x are independent

(assumption of exogeneity).

1. Endogeneity
Condition of Exogeneity

Emphasis of this lesson is the assumption of exogeneity:

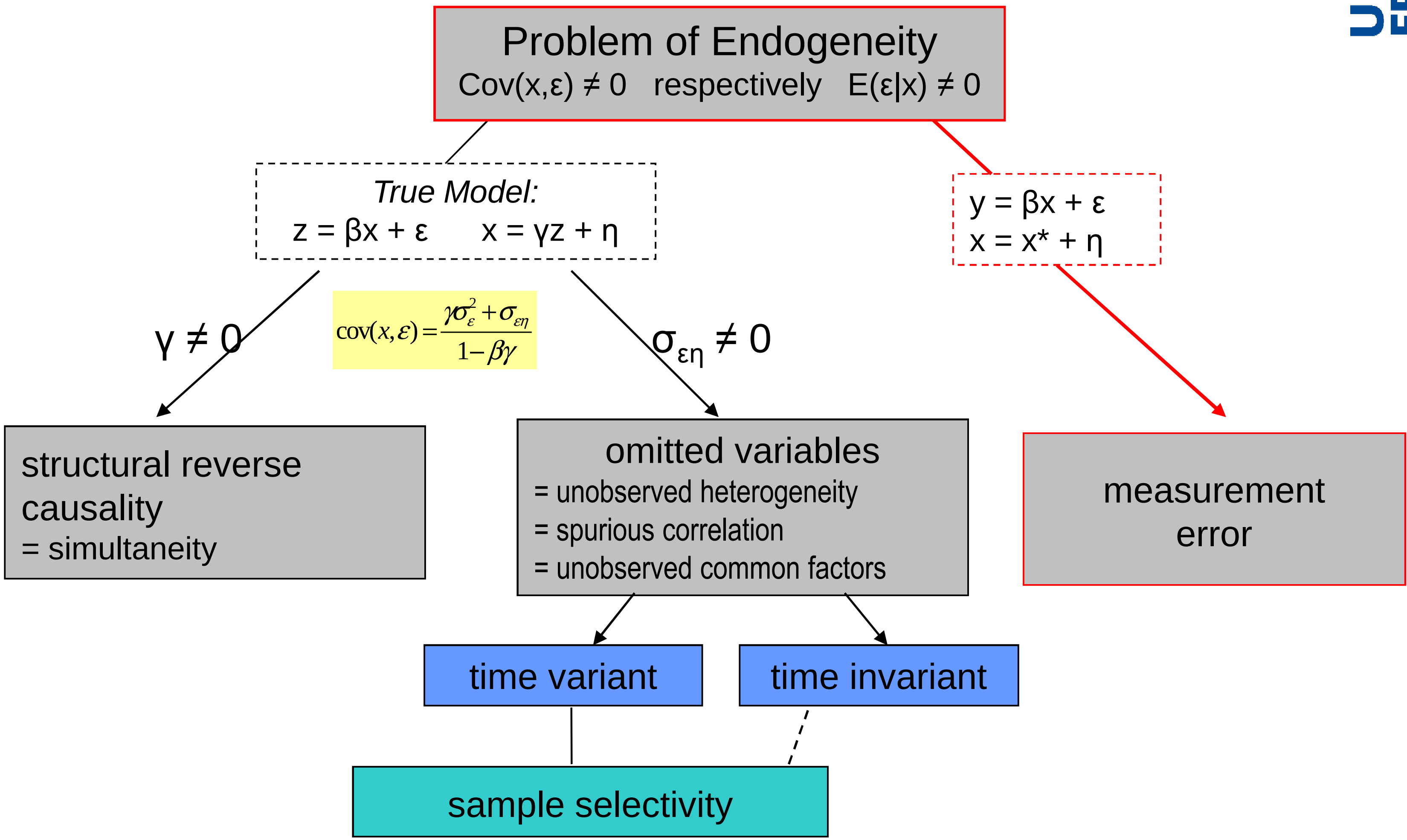
- Independence of residual and explaining variables
- All missing variables are captured by a <u>disturbance term</u>

1. Endogeneity
Condition of Exogeneity

If the assumption of exogeneity is violated then OLS is

- biased
- inconsistent

1. Endogeneity



Problem of Endogeneity

$Cov(x,\varepsilon) \neq 0$ respectively $E(\varepsilon|x) \neq 0$

True Model:

$z = \beta x + \varepsilon$ $x = \gamma z + \eta$

$\gamma \neq 0$

$$\text{cov}(x,\varepsilon) = \frac{\gamma\sigma_\varepsilon^2 + \sigma_{\varepsilon\eta}}{1-\beta\gamma}$$

structural reverse causality
= simultaneity

2. Simultaneity

Basic Problem:

Direction of causal effects between variables is ambiguous.

Example: y ⟷ x

Rate of criminality	⟷	No. of policemen
GDP	⟷	Consumption
Unemployment	⟷	Application of active labour market policy

(1) $y = \beta x + \varepsilon$

(2) $x = \gamma y + \eta$

Estimation of (1) with OLS→ Estimated coefficients biased, if $\gamma \neq 0$, as $Cov(x,\varepsilon) \neq 0$.

3. Omitted Variables Classification

Problem of Endogeneity

Cov(x,ε) ≠ 0 respectively E(ε|x) ≠ 0

True Model:

z = βx + ε x = γz + η

γ ≠ 0

$$\mathrm{cov}(x,\varepsilon) = \frac{\gamma\sigma_\varepsilon^2 + \sigma_{\varepsilon\eta}}{1-\beta\gamma}$$

$\sigma_{\varepsilon\eta} \neq 0$

structural reverse causality
=simultaneity

omitted variables
= unobserved heterogeneity
= spurious correlation[1]
= unobserved common factors

[1] The denomination is deceptive; a better denomination would be *„spurious causality"*

3. Omitted Variables

- **Classification:** Omission of variables leads to an endogeneity bias and thus to misleading regression results
- In case the incorrect specification $y = x_1\beta_1 + \varepsilon$ is assumed instead of $y = x_1\beta_1 + x_2\beta_2 + \varepsilon$, then the effect of the omitted variable is captured in the residual
- If either $Cov(x_1,x_2)=0$ or $\beta_2=0$ is violated, then the disturbance term is correlated with x_1 ➔ endogeneity bias

3. Omitted Variables = unobserved heterogeneity

1) Unobserved Heterogeneity

= unobservable individual effect

Examples:

- Motivation
- Intelligence
- Management Skills

2) Spurious Correlation / Spurious Causality

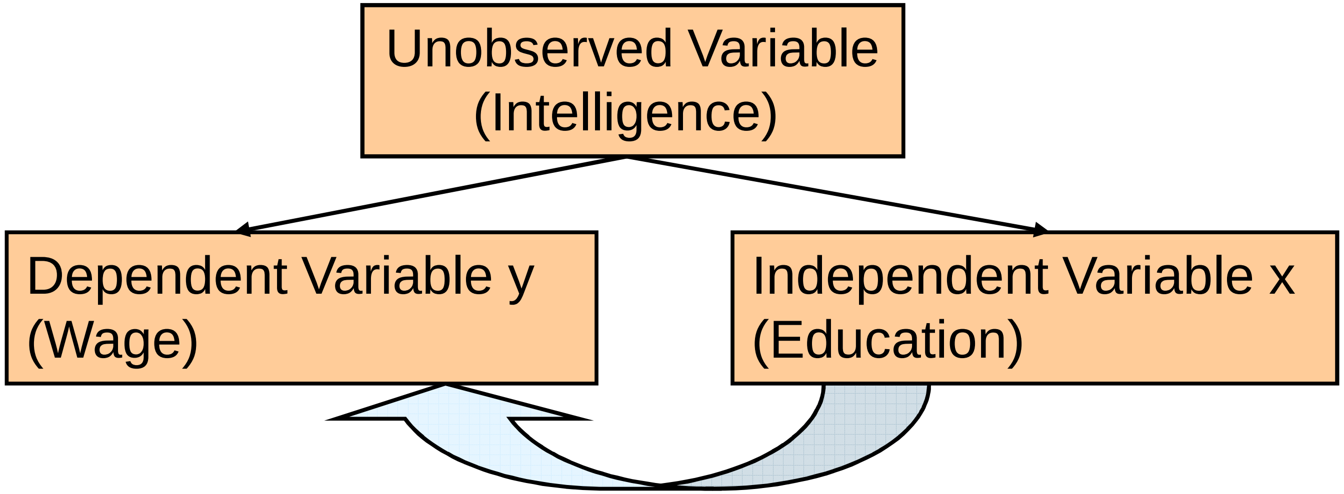
Due to an omitted variable, a pseudo-correlation between regressor x and regressand y emerges

Example: Estimation of the effect of education on wages: Individuals A and B differ in regard to their intelligence

→ Due to higher intelligence, A has more years of education

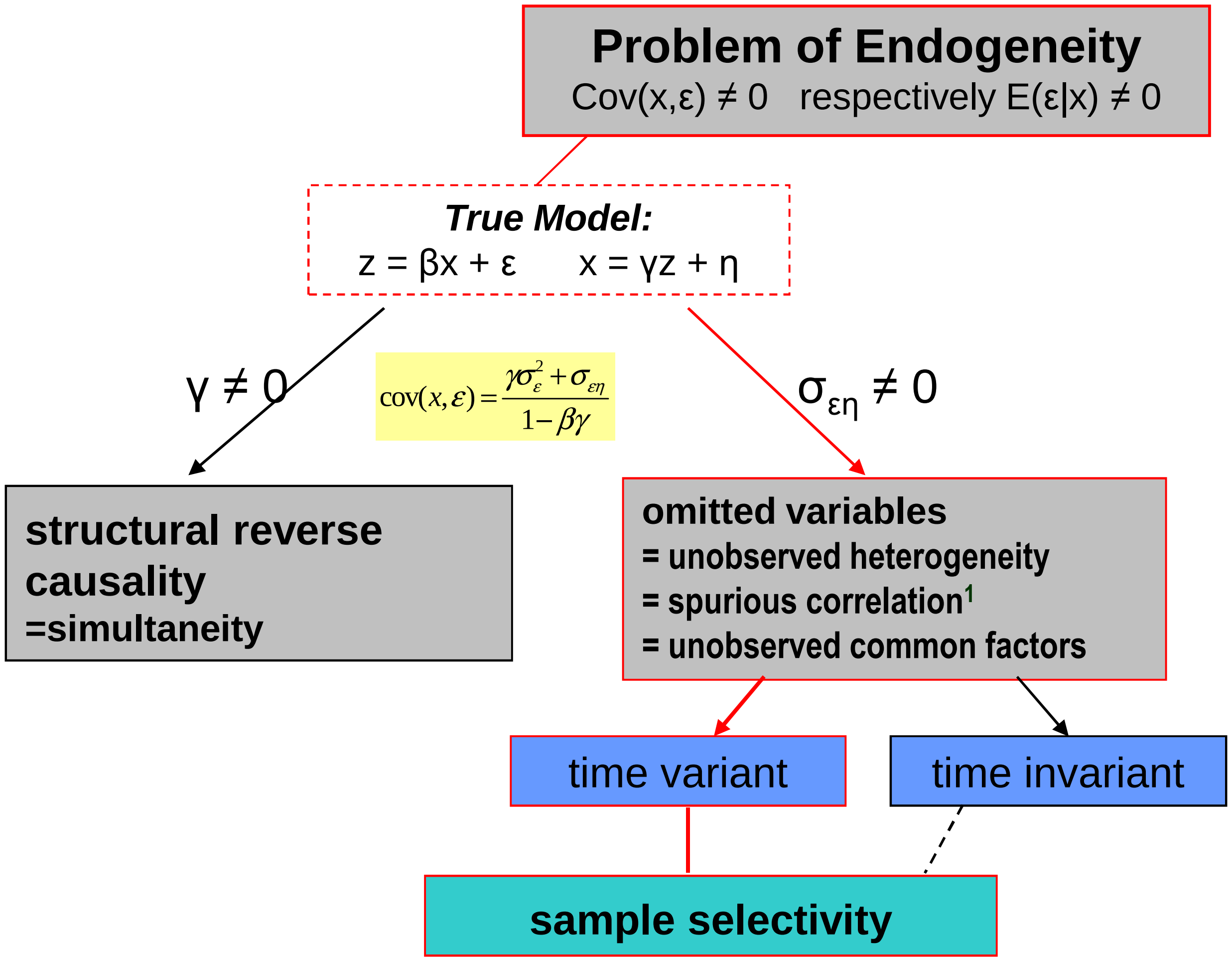
→ Due to higher intelligence, A receives higher wages

3) Unobserved Common Factors

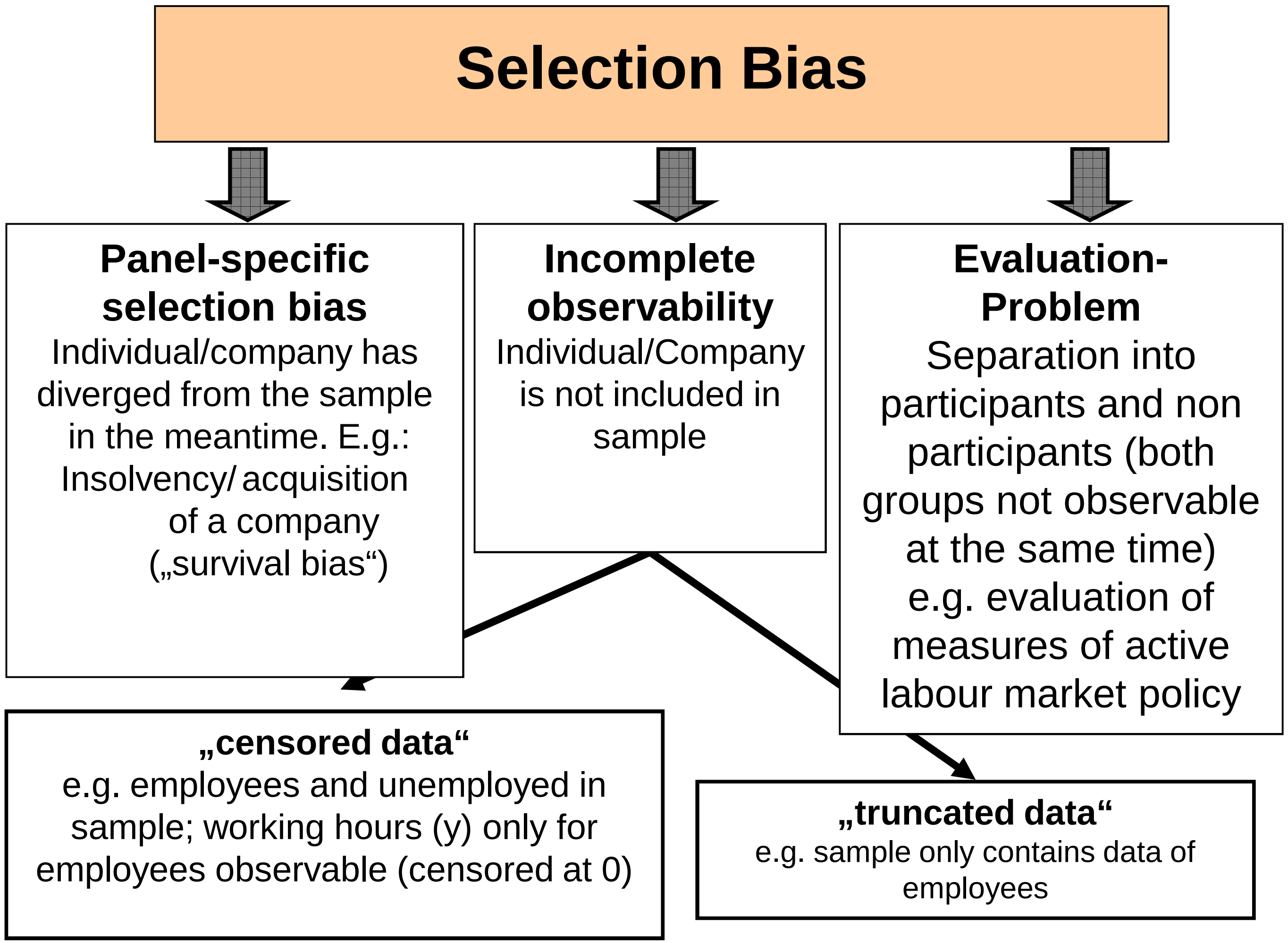


In case intelligence is not specified within the model: Regression overestimates the real effect of education on wages because of a positive correlation between intelligence and education.

3. Omitted Variables
Classification of Selection Bias



3. Omitted Variables
Selection Bias – Criteria for Differentiation I



Selection Bias – Criteria for Differentiation II

Selection Bias

„positive sample selection“	**„negative sample selection“**
e.g. particular motivation in case of placement vouchers	e.g. small companies do not appear in DAX investigations

3. Omitted Variables
Selection Bias – Criteria for Differentation III

Selection Bias

Self Selection

Individuals select themselves into a sample.

e.g. Individual applies for a career advancement

External Selection

Individuals are selected into a sample.

e.g. Individual is registered for a career advancement by a referee's decision

Solution Approaches