

# Objetivos

Ao final desse grupo de *slides* os alunos deverão ser capazes de:

- ✓ **Conduzir** testes de hipóteses adequados para:
  - ✓ verificar se um regressor é endógeno;
  - ✓ verificar a validade dos instrumentos.

# TESTE DE ENDOGENEIDADE

**Aula 17c**

Leitura DETALHADA: Wooldridge, 2010 (4ª edição) – Capítulo 15

# Teste de Endogeneidade

O estimador de 2SLS é menos eficiente que o de MQO, quando as variáveis explicativas são exógenas.

Assim sendo, é razoável conduzir um teste de endogeneidade para saber se existe a necessidade de implementar o método 2SLS ou não.

# Teste de Endogeneidade

Considere o modelo

$$y_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon_1 \quad (1)$$

em que

$x_1$  e  $x_2$  – variáveis exógenas

$x_3$  – variável endógena

Ainda, considere que  $z_1$  seja um instrumento para  $x_3$ .

# Teste de Endogeneidade

## Fatos

1. Se  $x_3$  for não correlacionada com  $\varepsilon_1$ , então, devemos estimar os parâmetros do modelo por MQO (mais eficiente).
2. MQO e 2SLS fornecem estimadores consistentes se a condição de exogeneidade estiver satisfeita.

# Teste de Endogeneidade

**HAUSMAN (1978)**, sugeriu fazer uma comparação direta das estimativas de MQO e 2SLS e determinar se as diferenças são estatisticamente significantes.

Se as estimativas geradas por MQO e 2SLS diferirem de forma significativa, concluimos que  $x_3$  deve ser endógena (supondo  $x_1$  e  $x_2$  exógenas).

# Teste de Endogeneidade

## Procedimento para aplicação do Teste de Hausman:

- i. Encontre a forma reduzida de  $x_3$  ( $x_3$  em função de todas as variáveis exógenas contidas em  $Z$  – matriz de instrumentos).
- ii. Estime os parâmetros de (i), por MQO.
- iii. Obtenha os resíduos da estimação feita em (ii).
- iv. Estime os parâmetros da equação estrutural (equação de interesse), por MQO, incluindo os resíduos, que foram obtidos em (iii), como variável explicativa.
- v. Se o parâmetro associado ao resíduo for estatisticamente significativo, concluiremos que  $x_3$  é endógena.

# Aplicação

## Exercício sobre IV – MROZ (1987)

Considere o seguinte modelo de regressão linear múltipla:

$$\ln(wage) = \beta_0 + \beta_1 \text{exper} + \beta_2 \text{exper}^2 + \beta_3 \text{educ} + \varepsilon$$

(para as mulheres que trabalham)

Verifique se a variável *educ* é endógena, tendo como instrumentos *educação do pai*, *educação da mãe* e *educação do marido*. Com base no resultado obtido, qual método de estimação você indicaria para estimar os parâmetros da equação de interesse: MQO ou 2SLS? Adote 10% de significância. Para tanto, utilize o arquivo *MROZ.xls*.



# TESTE DE SARGAN

## VERIFICAÇÃO DA VALIDADE DOS INSTRUMENTOS

# TESTE DE SARGAN

Qual a validade do instrumento, ou seja, como sabemos se os instrumentos escolhidos são independentes do termo de erro?

Para responder à pergunta anterior, Sargan (1964) desenvolveu um teste estatístico, chamado de SARG, para testar a validade dos instrumentos.

# TESTE DE SARGAN

O procedimento é o seguinte:

- 1) Divida os regressores da equação estrutural em dois conjuntos:
  - (a) conjunto dos regressores exógenos; e
  - (b) conjunto dos regressores endógenos;
- 2) Estime os parâmetros da equação estrutural, instrumentalizando adequadamente os regressores endógenos.
- 3) Obtenha os resíduos de (2).
- 4) Escreva a forma reduzida para os resíduos que foram obtidos em (3) e estime os parâmetros da equação auxiliar por MQO.
- 5) Calcule a estatística  $SARG = (n - (k+1))R^2 \sim \chi^2_{(p-q)}$ , em que  $p$  é o número de instrumentos e  $q$  é o número de regressores endógenos.
- 6) Rejeite  $H_0$  (instrumentos válidos), se  $SARG > \chi^2_{(crítico)}$ .

# Exercício 1

Admita que em seu modelo de regressão exista apenas uma variável explicativa endógena. Ainda, suponha que você tenha uma variável instrumental disponível. Do exposto, seria possível realizar o teste de SARGAN? Justifique a sua resposta.

# Exercício 2

Considere o seguinte modelo de regressão linear múltipla:

$$\ln(wage) = \beta_0 + \beta_1 \text{exper} + \beta_2 \text{exper}^2 + \beta_3 \text{educ} + \varepsilon$$

(para as mulheres que trabalham)

Utilizando o arquivo *MROZ.xls*, estime os parâmetros do modelo de regressão de interesse, via 2SLS, usando *educação do pai*, *educação da mãe* e *educação do marido* como instrumentos para *educ*. Ainda, conduza o teste de SARG. Adote  $\alpha = 10\%$ . Comente o resultado obtido.

Insper

# LEITURA COMPLEMENTAR

**TESTE**

**DE**

**RESTRICÇÕES SOBREIDENTIFICADORAS**

**(análogo ao Teste de SARGAN – validade dos instrumentos)**

# Teste de Restrições Sobreidentificadoras

Suponha que na equação estrutural de interesse apareça somente uma variável explicativa endógena.

Nesse caso:

- Se tivermos somente uma única IV, não teremos restrições sobreidentificadoras. Ou seja, não haverá nada que possa ser testado.
- Se tivermos duas IVs, teremos uma restrição sobreidentificadora. Se tivermos três IVs, teremos duas restrições sobreidentificadoras, e assim por diante.



# Teste de Restrições Sobreidentificadoras

## Procedimento para aplicação do teste:

- i. Estime os parâmetros da equação estrutural por 2SLS.
- ii. Obtenha os resíduos.
- iii. Regrida os resíduos em função de todas as variáveis exógenas.
- iv. Obtenha o  $R^2$  (coeficiente de explicação).
- v. Sob a hipótese nula de que todas as IVs são não correlacionadas com o erro da equação estrutural,

$$nR^2 \sim \chi^2_q$$

*em que*

$q$  – é o número de IVs menos o número de regressores endógenos presentes no modelo.

# Teste de Restrições Sobreidentificadoras

Rejeitar a hipótese nula significa que pelo menos uma das IVs não é exógena.

# Aplicação

## Exercício sobre IV – MROZ (1987)

Considere o seguinte modelo de regressão linear múltipla:

$$\ln(wage) = \beta_0 + \beta_1 \text{exper} + \beta_2 \text{exper}^2 + \beta_3 \text{educ} + \varepsilon$$

(para as mulheres que trabalham)

Para responder o item, a seguir, utilize o arquivo **MROZ.xls**:

- Usando *educação do pai, educação da mãe e educação do marido* como instrumentos para *educ*, estime os parâmetros do modelo por 2SLS.
- Teste se o uso de três instrumentos (duas restrições sobreidentificadoras) gera viés no estimador de 2SLS (se algum dos instrumentos é correlacionado com o erro do modelo).

# Aplicação (cont.)

## Exercício sobre IV – MROZ (1987)

Considere o seguinte modelo de regressão linear múltipla:

$$\ln(wage) = \beta_0 + \beta_1 \text{exper} + \beta_2 \text{exper}^2 + \beta_3 \text{educ} + \varepsilon$$

(para as mulheres que trabalham)

Para responder o item, a seguir, utilize o arquivo *MROZ.xls*:

- c) Seria possível realizar o teste anteriormente proposto se tivéssemos apenas um instrumento? Justifique a sua resposta.

## Referência da Aplicação

MROZ, T. A. (1987). The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions, *Econometrica*, 55, 765-799.