

Econometria Aplicada

Variáveis Instrumentais

João Ricardo Costa Filho

"The most important questions of life are, for the most part, really only problems in probability."

Laplace (1812)

"In God we trust. All others must bring data."

William Edwards Deming

O problema da endogeneidade

Motivação (tudo começa com uma pergunta)

Como estimar uma curva de demanda?

- Podemos estimar a equação

$$\ln [\text{Preço}_i] = \beta_0 + \beta_1 \ln [\text{Quantidade}_i] \text{ por MQO?}$$

Oferta e demanda

- Podemos estimar a equação
 $\ln [\text{Preço}_i] = \beta_0 + \beta_1 \ln [\text{Quantidade}_i]$ por MQO?
- Qual é a relação que eu espero encontrar entre o preço e a quantidade?

Oferta e demanda

- Podemos estimar a equação
 $\ln [\text{Preço}_i] = \beta_0 + \beta_1 \ln [\text{Quantidade}_i]$ por MQO?
- Qual é a relação que eu espero encontrar entre o preço e a quantidade?
 - Usualmente, a curva de **demand**a representa uma relação negativa entre o preço e a quantidade vendida.

Oferta e demanda

- Podemos estimar a equação
 $\ln [\text{Preço}_i] = \beta_0 + \beta_1 \ln [\text{Quantidade}_i]$ por MQO?
- Qual é a relação que eu espero encontrar entre o preço e a quantidade?
 - Usualmente, a curva de **demanda** representa uma relação negativa entre o preço e a quantidade vendida.
 - Usualmente, a curva de **oferta** representa uma relação positiva entre o preço e a quantidade vendida.

Oferta e demanda

- Podemos estimar a equação $\ln [\text{Preço}_i] = \beta_0 + \beta_1 \ln [\text{Quantidade}_i]$ por MQO?
- Qual é a relação que eu espero encontrar entre o preço e a quantidade?
 - Usualmente, a curva de **demanda** representa uma relação negativa entre o preço e a quantidade vendida.
 - Usualmente, a curva de **oferta** representa uma relação positiva entre o preço e a quantidade vendida.

Preço e quantidade são variáveis endógenas. Portanto, elas violam uma hipótese importante do estimador de MQO.

Oferta e demanda

- Podemos estimar a equação
 $\ln [\text{Preço}_i] = \beta_0 + \beta_1 \ln [\text{Quantidade}_i]$ por MQO?
- Qual é a relação que eu espero encontrar entre o preço e a quantidade?
 - Usualmente, a curva de **demanda** representa uma relação negativa entre o preço e a quantidade vendida.
 - Usualmente, a curva de **oferta** representa uma relação positiva entre o preço e a quantidade vendida.

Preço e quantidade são variáveis endógenas. Portanto, elas violam uma hipótese importante do estimador de MQO. O que fazer?

- Nosso problema é que a variável X é correlacionada com os erros (i.e. $Cov(X, \varepsilon) \neq 0$).

Instrumento

- Nosso problema é que a variável X é correlacionada com os erros (i.e. $Cov(X, \varepsilon) \neq 0$).
- E se encontrarmos uma variável Z que seja exógena (i.e. $Cov(X, \varepsilon) = 0$)

- Nosso problema é que a variável X é correlacionada com os erros (i.e. $Cov(X, \varepsilon) \neq 0$).
- E se encontrarmos uma variável Z que seja exógena (i.e. $Cov(X, \varepsilon) = 0$) e correlacionada com X (i.e. $Cov(X, Z) \neq 0$)?

Instrumento

- Nosso problema é que a variável X é correlacionada com os erros (i.e. $Cov(X, \varepsilon) \neq 0$).
- E se encontrarmos uma variável Z que seja exógena (i.e. $Cov(X, \varepsilon) = 0$) e correlacionada com X (i.e. $Cov(X, Z) \neq 0$)?
- No caso da curva de demanda, queremos uma (ou mais) variável(is) que não sejam correlacionadas com a quantidade demandada (exógenas), mas que sejam correlacionadas com o preço.

Instrumento

- Quais são as dificuldades?

Instrumento

- Quais são as dificuldades?
- Uma variável Z que seja exógena (i.e. $Cov(X, \varepsilon) = 0$). Como sabemos que ela é exógena? Teoria econômica?

Instrumento

- Quais são as dificuldades?
- Uma variável Z que seja exógena (i.e. $\text{Cov}(X, \varepsilon) = 0$). Como sabemos que ela é exógena? Teoria econômica?
- Uma variável Z que seja correlacionada com X (i.e. $\text{Cov}(X, Z) \neq 0$). Nesse caso, podemos testar:

$$\mathcal{H}_0 : \pi_1 = 0$$

$$\mathcal{H}_a : \pi_1 \neq 0$$

$$X_i = \pi_0 + \pi_1 Z_i + v_i.$$

Instrumento

- Quais são as dificuldades?
- Uma variável Z que seja exógena (i.e. $\text{Cov}(X, \varepsilon) = 0$). Como sabemos que ela é exógena? Teoria econômica?
- Uma variável Z que seja correlacionada com X (i.e. $\text{Cov}(X, Z) \neq 0$). Nesse caso, podemos testar:

$$\mathcal{H}_0 : \pi_1 = 0$$

$$\mathcal{H}_a : \pi_1 \neq 0$$

$$X_i = \pi_0 + \pi_1 Z_i + v_i.$$

Isso é o que chamamos de **primeiro estágio**.

Instrumento

Podemos usar instrumentos em modelos estruturais de regressões múltiplas. Apenas precisamos de um instrumento para cada variável endógena:

Instrumento

Podemos usar instrumentos em modelos estruturais de regressões múltiplas. Apenas precisamos de um instrumento para cada variável endógena:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_i + v_i.$$

onde X é endógena e W é exógena.

Instrumento

Podemos usar instrumentos em modelos estruturais de regressões múltiplas. Apenas precisamos de um instrumento para cada variável endógena:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_i + v_i.$$

onde X é endógena e W é exógena. Seja Z um instrumento:

$$\mathcal{H}_0 : \pi_1 = 0$$

$$\mathcal{H}_a : \pi_1 \neq 0$$

$$X_i = \pi_0 + \pi_1 Z_i + \pi_2 W_i + u_i.$$

Podemos utilizar mais de um instrumento para cada variável endógena:

Podemos utilizar mais de um instrumento para cada variável endógena:

$$X_i = \pi_0 + \pi_1 Z_{1,i} + \pi_2 Z_{2,i} + \pi_3 W_i u_i.$$

onde Z_1 e Z_2 são os instrumentos (variáveis exógenas).

Regressão com variáveis instrumentais

- 1) Primeiro estágio: estimar a regressão da(s) variável(is) exógena(s) sobre o(s) instrumento(s).

Regressão com variáveis instrumentais

- 1) Primeiro estágio: estimar a regressão da(s) variável(is) exógena(s) sobre o(s) instrumento(s).
- 2) Calcular o valor esperado para a variável exógena à partir da regressão no primeiro estágio.

Regressão com variáveis instrumentais

- 1) Primeiro estágio: estimar a regressão da(s) variável(is) exógena(s) sobre o(s) instrumento(s).
- 2) Calcular o valor esperado para a variável exógena à partir da regressão no primeiro estágio.
- 3) Estimar o modelo estrutural com o valor esperado para a variável exógena (e outros regressores/controles).

Vamos aos dados!

Vamos para a atividade em grupo!

