Esta aula

- Plano
  - Revisão: Variáveis Instrumentais
  - Equações Simultâneas

- Bibliografia
  - Wooldridge, J. M. Introductory Econometrics: A modern Approach, 6th Ed.

Variáveis Instrumentais

Variáveis Instrumentais

- Se há uma variável explicativa endógena no modelo, podemos utilizar variáveis instrumentais
- Ou seja se Cov(x,u) $\neq$ 0, então podemos corrigir o viés na estimação dos parâmetros através do emprego de variáveis instrumentais

Variáveis Instrumentais

- Para que a variável z seja um bom instrumento para a variável x endógena, deve satisfazer as seguintes condições:
  1. A variável z deve ser exógena:

     $\text{Cov}(z,u) = 0$
  2. A variável z deve ser correlacionada com x:

     $\text{Cov}(z,x) \neq 0$

Variáveis Instrumentais

- A condição $Cov(z,u) = 0$ não é passível de ser testada, e deve ser guiada pela teoria de RI
- Mas podemos testar $Cov(z,x) \neq 0$, através da regressão
- $x = \pi_0 + \pi_1 z + v$
- onde testamos a hipótese H_0: $\pi_1 = 0$
- Essa regressão é chamada de primeiro estágio do método de mínimos quadrados de dois estágios.

Variáveis Instrumentais

- Considere $y = \beta_0 + \beta_1 x + u$, nesse casso temos:
- $Cov(z,y) = \beta_1 Cov(z,x) + Cov(z,u)$, so
- $\beta_1 = Cov(z,y) / Cov(z,x)$
- E o estimador de variáveis instrumentais (IV) é dado por:

$$\hat{\beta}_1 = \frac{\sum (z_i - \bar{z})(y_i - \bar{y})}{\sum (z_i - \bar{z})(x_i - \bar{x})}$$

Variáveis Instrumentais

- O pressuposto de homocedasticidade é dado por $E(u^2|z) = \sigma^2 = Var(u)$
- O erro padrão estimado é dado por

$$se\left(\hat{\beta}_1\right) = \frac{\hat{\sigma}^2}{SST_x R^2_{x,z}}$$

IV versus OLS

- O erro padrão de IV difere daquele de OLS pela presença do termo R^2 da regressão de x em z
- Dado $R^2 < 1$, erro-padrão IV é maior do que aquele de OLS.
- Entretanto, o estimador IV é consistente enquanto que OLS é inconsistente se $\mathrm{Cov}(x,u) \neq 0$
- Quanto maior a correlação entre z e x, mais precisa será a estimação.

Variáveis Instrumentais

- O estimador IV pode ser aplicado em regressão múltipla.
- Nesse caso, precisaremos de pelo menos um instrumento para cada variável explicativa endógena.

Variáveis Instrumentais

- Suponha que o modelo de regressão múltipla seja

  $y_1 = \beta_0 + \beta_1 y_2 + \beta_2 z_1 + u_1,$

  onde y_2 é endógena e z_1 é uma variável explicativa exógena

- Se z_2 for um instrumento para y_2 , então $Cov(z_2, u_1) = 0$
- Nesse caso, devemos fazer a regressão de y_2 contra todas as variáveis exógenas do modelo e testar pela significância do instrumento:
- $y_2 = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + v_2$

  e testar H0: $\pi_2 \neq 0$

Variáveis Instrumentais

- Se houver múltiplos instrumentos, z_2 e z_3, o que fazer?

  Nesse caso, as seguintes condições devem ser satisfeitas:

  1. As variáveis z_2 e z_3 devem ser exógenas:

     $Cov(z_2, u) = 0$ e $Cov(z_3, u) = 0$

  2. As variáveis z_2 e z_3 devem ser correlacionadas com y_2 :

     $Cov(z_2, y_2) \neq 0$ e $Cov(z_3, y_2) \neq 0$

- *Fazemos a regressão* $y_2 = \pi_0 + \pi_1 z_1 + \pi_2 z_2 + \pi_3 z_3 + v_2$

  e testamos H0: $\pi_2 = 0, \pi_3 = 0$

Teste para Endogeneidade

- Se não há endogeneidade, as estimativas de OLS e IV são consistentes
- Porém, se não há endogeneidade, OLS é preferível a IV

- Teste Hausman: testa pelas diferenças entre as estimativas de OLS e IV

Teste para Endogeneidade

- Se y_2 for endogena, então v_2 do primeiro estágio e u_1 do modelo estrutural serão correlacionadas.
- Procedimento para o teste:
- Salvar os residuos do primeiro estágio.
- Incluir esses resíduos na estimação do modelo estrutural
- Se eles forem significantes, rejeitar a hipótese nula de exogeneidade.

Teste para Heterocedasticidade: Breusch-Pagan

- Queremos testar H_0:$Var(u|\mathbf{z_1}, \mathbf{z_2}, x_1, x_2, \ldots, x_k) = \sigma^2$
- Se assumirmos uma relação linear:

$$u^2 = \delta_0 + \delta_1 x_1 + \ldots + \delta_k x_k + \delta_{k+1} z_1 + \delta_{k+2} z_2 + v$$

, podemos testar:

- H_0: $\delta_1 = \delta_2 = \ldots = \delta_k = \delta_{k+1} = \delta_{k+2} = 0$
- Esse é o teste Breusch-Pagan.

Equações Simultâneas

Equações Simultâneas

- Equações Simultâneas ocorrem quando pelo menos uma variável explicativa é conjuntamente determinada com a variável dependente.
- Como qualquer problema de endogeneidade, veremos que esse problema pode ser resolvido com a utilização de variáveis instrumentais.

Equações Simultâneas

- Suponha que você queira entender as exportações de um produto ofertadas pelo país B ao país A:
- $X_s = \alpha_1 P + \beta_1 z + u_1$, onde
- P é o preço do produto e z é um determinante de oferta chamado "supply shifter"
- Essa é uma equaço estrutural visto que é determinada pela teoria de relações internacionais

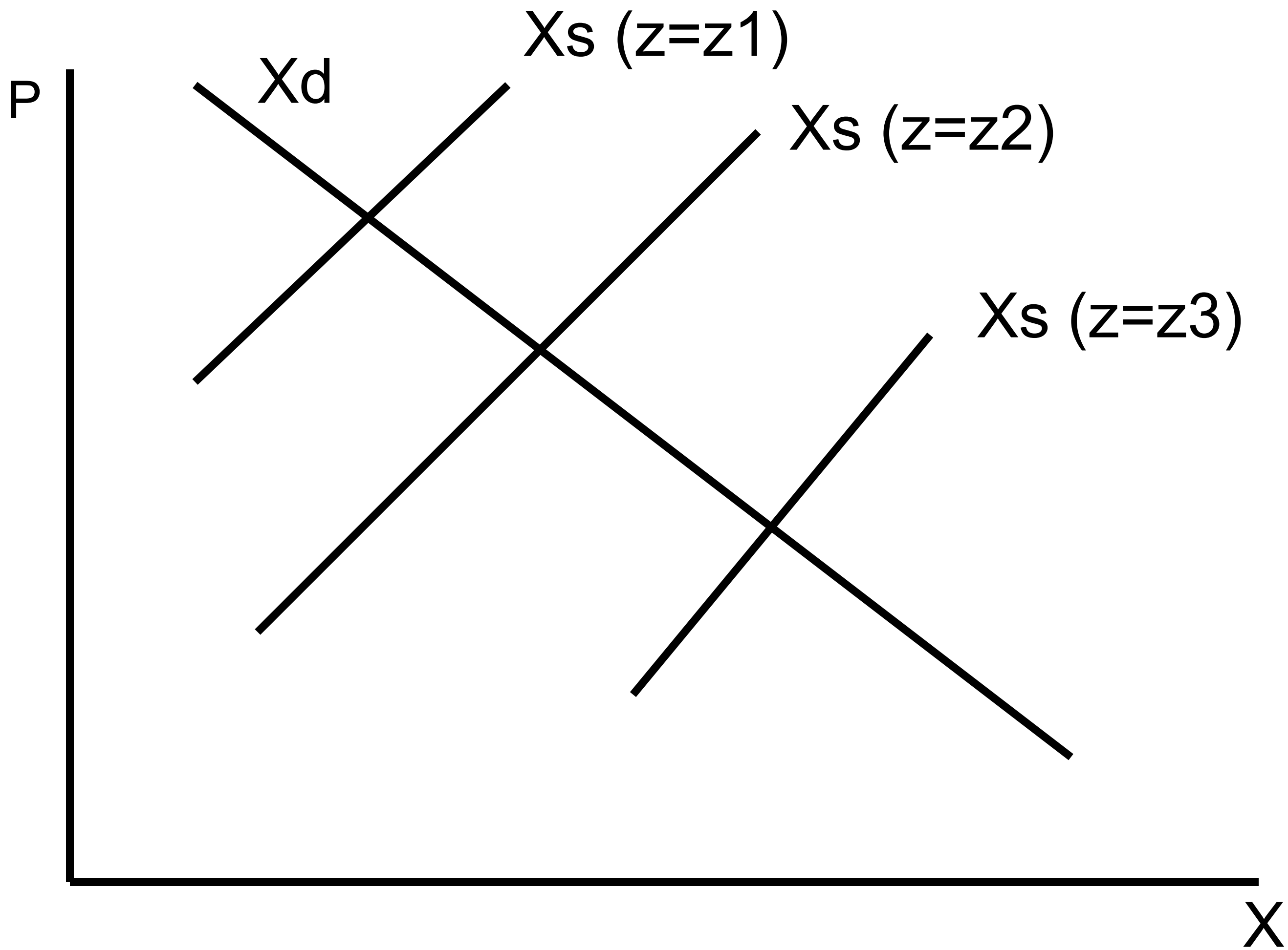
Equações Simultâneas

- Não podemos fazer a regressão das exportações contra o preço visto que as exportações são determinadas também pela demanda por exportações.
- Considere uma segunda equação estrutural, representando a demanda do país pelos produtos de B:
- $X_d = \alpha_2 P + u_2$
- Dessa forma, exportações são determinadas por equações simultâneas.

Equações Simultâneas

- Ambas X e P são endógenas porque são determinadas pelo equilíbrio entre demanda e oferta.
- *O "supply shifter"* z, por ser exógeno, nos permite identificar a equação estrutural de demanda.
- Se não observarmos "demand shifters", a equação de oferta não pode ser estimada.

Equações Simultâneas



Equações Simultâneas

- Podemos estimar a equação estrutural de demanda, utilizando z como instrumento para P.
- A equação de primeiro estágio é $P = \pi_0 + \pi_1 z + v_2$
- E a equação de segundo estágio: $X = \alpha_2 P_{hat} + u_2$
- O estimador 2SLS fornece um estimador consistente de α_2, a inclinação da curva de demanda
- Não podemos estimar α_1, a inclinação da curva de oferta.

Equações Simultâneas

- Suponha que queira estimar: $y_1 = \alpha_1 y_2 + \beta_1 z_1 + u_1$

  onde , $y_2 = \alpha_2 y_1 + \beta_2 z_2 + u_2$
- Portanto, $y_2 = \alpha_2(\alpha_1 y_2 + \beta_1 z_1 + u_1) + \beta_2 z_2 + u_2$
- Então, $(1 - \alpha_2\alpha_1)y_2 = \alpha_2 \beta_1 z_1 + \beta_2 z_2 + \alpha_2 u_1 + u_2$, que pode ser reescrita como
- $y_2 = \pi_1 z_1 + \pi_2 z_2 + v_2$
- Essa é a equação reduzida para y_2 e deverá ser estimada no primeiro estágio.

Equações Simultâneas

- Olhando essa equação reduzida para y_2, vemos que v_2 é uma função linear de u_1, *por isso a estimativa de* α_1 por OLS é enviesada.
- Devemos aplicar 2SLS!

Equações Simultâneas

- De uma forma geral, suponha que z_1 sejam todas as variáveis exógenas na primeira equação, e z_2 sejam todas variáveis exógenas na segunda equação.
- Podem haver variáveis equivalentes em z_1 e z_2, porém:

1. Para identificar a primeira equação, deve haver pelo menos uma variável em z_2 que não está em z_1.
2. Para identificar a segunda equação, deve haver pelo menos uma variável em z_1 que não está em z_2.

Equações Simultâneas

- Note que:

1. Para que possamos identificar a primeira equação, a variável de z_2 que não está em z_1 deve ser significantemente diferente de zero.
2. Para que possamos identificar a segunda equação, a variável de z_1 que não está em z_2 deve ser significantemente diferente de zero.

- Essas são denominadas "rank conditions"

Equações Simultâneas

- Estimação de equações simultâneas por 2SLS é simples:
- O sistema pode ter duas ou mais equações.
- Para as equações identificadas, estime as variáveis dependentes contra todas as variáveis exógenas do sistema.

Obrigada!