

ECONOMETRIA I

ENDOGENEIDADE

Victor Oliveira

Núcleo de Economia Internacional e Desenvolvimento Econômico

PPGDE – 2023

- 1 Endogeneidade
- 2 Mínimos Quadrados em Dois Estágios
- 3 Consistência do Estimador de MQ2E
- 4 Normalidade Assintótica de MQ2E
- 5 Eficiência Assintótica do MQ2E

- Encontrar instrumentos válidos para os nossos modelos é uma tarefa árdua! Na verdade muito difícil!
- Uma das razões dessa dificuldade reside no fato de que as **duas condições** requeridas de um instrumento são muitas vezes **conflitantes**.
- **Exemplo:** considere a estimação da equação de salários em função da educação.
 - **Variável omitida:** habilidade do indivíduo – vies a o coeficiente da educação;
 - **Possível instrumento:** educação da mãe – correlacionada com a educação do indivíduo, isto é, $\text{Cov}(z_1, \text{educ}) \neq 0$ e $\theta_1 \neq 0$.
 - **Atenção:** educação da mãe também deve ser correlacionada com a habilidade do indivíduo presente no erro: $\text{Cov}(z_1, u) \neq 0$.

- A condição $\text{cov}(z_1, educ) \neq 0$ pode ser satisfeita com uma correlação entre z e $educ$ diferente de zero, porém **baixa**. Esse é o caso de um **instrumento fraco**. De fato:

$$\text{var}(\hat{\beta}_1^{VI}) = \frac{\text{var}(u)}{n \text{ var}(educ)[\text{corr}(educ, z_1)]^2} \quad (1)$$

- Como uma consequência, a **variância (assintótica) do estimador de IV aumenta**. Ou seja, o estimador perde precisão. Por isso, **devemos procurar um instrumento que tenha a mais alta correlação possível com $educ$** .

- David Card (1995) publicou um paper em que utiliza a distância entre a residência e a faculdade como um instrumento para educação.
- Para ele, se um aluno em potencial mora perto de uma faculdade, isso reduz o custo de frequência e, portanto, aumenta a probabilidade de o aluno frequentar a faculdade.
- A proximidade da faculdade não afeta diretamente as habilidades ou aptidões de um aluno, portanto, não deve ter um efeito direto em seu salário de mercado.
- Estas considerações sugerem que a proximidade da faculdade pode ser utilizada como um instrumento para educação numa regressão de salários.

- A equação de regressão estimada é dada por:

$$\begin{aligned}\log(\text{salario})_i = & \beta_0 + \beta_1 \text{educ}_i + \beta_2 \text{exp}_i + \beta_3 \text{exp}^2/100_i + \beta_4 \text{black}_i \\ & + \beta_5 \text{sul}_i + \beta_6 \text{urbano}_i + u_i\end{aligned}\quad (2)$$

- **Dados:** National longitudinal Survey of Young Men (NLSYM) para o ano de 1976.
- **Variáveis:** log do salário hora da semana; *educação em anos de escolaridade*; experiência no mercado de trabalho: idade em anos menos a educação mais 6; experiência ao quadrado/100, black; região e urbano.

Tabela 1: Equação de salários com variáveis instrumentais

Variáveis	MQO	VI(a)	VI(b)
Educ	0,074 (0.004)	0,132 (0.049)	0,133 (0.051)
Exp	0,084 (0.007)	0,107 (0.021)	0,056 (0.026)
$Exp^2/100$	-0,224 (0.032)	-0,228 (0.035)	-0,080 (0.133)
Black	-0,190 (0.017)	-0,131 (0.051)	-0,103 (0.075)
Sul	-0,125 (0.015)	-0,105 (0.023)	-0,098 (0.028)
Urbano	0,161 (0.015)	0,131 (0.030)	0,108 (0.049)

- 1 VI (a) usa “college” como instrumento para educação.
- 2 VI (b) usa “college”, idade e $idade^2/100$ como instrumentos para educação, experiência e $experiencia^2/100$.

- Considere um modelo linear populacional

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + u \quad (3)$$

em que x_k pode ser correlacionado com u . Vamos assumir que temos mais de uma variável instrumental para x_k .

- Seja z_1, z_2, \dots, z_M variáveis tal que

$$\text{cov}(z_h, u) = 0, \quad h = 1, 2, \dots, M \quad (4)$$

em que cada z_h é exógena na equação (4). **Se cada um destes possui alguma correlação parcial com x_k , temos M diferentes estimadores de VI.**

- Vamos definir um vetor de variáveis exógenas por

$$\mathbf{z} \equiv (1, x_1, x_2, \dots, x_{k-1}, z_1, z_2, \dots, z_M)$$

como um vetor com dimensão $1 \times L$, em que $L = k + M$.

- Dentre todas as possíveis combinações de \mathbf{z} que podem ser usadas como instrumentos para x_k , o método de **mínimos quadrados em dois estágios (MQ2E)** escolhe aquela que é mais correlacionada com x_k .
- A combinação linear de \mathbf{z} mais correlacionada com x_k é dada pela projeção linear de x_k em \mathbf{z} .

- A forma reduzida é dada por:

$$x_k = \delta_0 + \delta_1 x_1 + \dots + \delta_{k-1} x_{k-1} + \theta_1 z_1 + \dots + \theta_M z_M + r_k \quad (5)$$

- Por definição, $\mathbb{E}(r_k) = 0$ e $\mathbb{E}(r_k | x_1, \dots, x_{k-1}, z_1, \dots, z_M) = 0$. Como qualquer combinação linear de \mathbf{z} é não correlacionada com u ,

$$x_k^* = \delta_0 + \delta_1 x_1 + \dots + \delta_{k-1} x_{k-1} + \theta_1 z_1 + \dots + \theta_M z_M \quad (6)$$

é não correlacionado com u . Na verdade, x_k^* é frequentemente interpretado como a parte de x_k que é não correlacionada com u . Se x_k é endógeno, é porque r_k é correlacionado com u .

- Se observarmos x_k^* , usaríamos ele como um instrumento para x_k . Como o δ_j e θ_j são parâmetros populacionais, x_k^* não um instrumento útil.
- Se mantivermos o pressuposto padrão de que não há uma dependência linear exata entre as variáveis exógenas, podemos estimar de forma consistente os parâmetros da equação (6) por Mínimos Quadrados.
- O análogo amostral do x_k^* para cada observação i são simplesmente os valores ajustados dos mínimos quadrados:

$$\hat{x}_{ik} = \hat{\delta}_0 + \hat{\delta}_1 x_{i1} + \dots + \hat{\delta}_{k-1} x_{ik-1} + \hat{\theta}_1 z_{i1} + \dots + \hat{\theta}_M z_{iM} \quad (7)$$

- Para cada observação i , definimos o vetor $\hat{\mathbf{x}}_i \equiv (1, x_{i1}, \dots, x_{i,k-1}, \hat{x}_{ik})$, $i = 1, 2, \dots, N$. Usando $\hat{\mathbf{x}}_i$ como um instrumento para \mathbf{x}_i gera o estimador de VI

$$\hat{\boldsymbol{\beta}} = \left(\sum_{i=1}^N \hat{\mathbf{x}}_i' \mathbf{x}_i \right)^{-1} \left(\sum_{i=1}^N \hat{\mathbf{x}}_i' y_i \right) = (\widehat{\mathbf{X}}' \mathbf{X})^{-1} \widehat{\mathbf{X}}' \mathbf{Y} \quad (8)$$

- O estimador de VI na eq. (8) vem a ser um estimador de mínimos quadrados.
- Para vermos isso, note que a matriz $\widehat{\mathbf{X}}$ de dimensão $N \times (k+1)$ pode ser expressa como $\widehat{\mathbf{X}} = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X} = \mathbf{P}_Z\mathbf{X}$.
- A matriz de projeção $\mathbf{P}_Z = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$ é idempotente e simétrica.

- Portanto,

$$\widehat{\mathbf{X}}' \mathbf{X} = \mathbf{X}' \mathbf{P}_Z \mathbf{X} = (\mathbf{P}_Z \mathbf{X})' \mathbf{P}_Z \mathbf{X} = \widehat{\mathbf{X}}' \widehat{\mathbf{X}} \quad (9)$$

- Plugando essa expressão na eq. (8) mostra que o estimador de VI que usa instrumentos $\widehat{\mathbf{x}}_i$ pode ser escrito como

$$\widehat{\boldsymbol{\beta}} = (\widehat{\mathbf{X}}' \widehat{\mathbf{X}})^{-1} \widehat{\mathbf{X}}' \mathbf{Y} \quad (10)$$

- O nome de dois estágios vem desse procedimento.

- Para resumirmos, $\hat{\beta}$ pode ser obtido da seguinte forma:
 - 1 **Regressão de primeiro estágio:** obter os valores ajustados \hat{x}_k da regressão x_k em $1, x_1, \dots, x_{k-1}, z_1, \dots, z_M$.
 - 2 **Regressão de segundo estágio:** rodar uma regressão de mínimos quadrados de y em $1, x_1, \dots, x_{k-1}, \hat{x}_k$. Nesta etapa é gerado o $\hat{\beta}_j$.

Tabela 2: Equação de salários com variáveis instrumentais

Variáveis	MQO	VI(a)	VI(b)	MQ2E(a)	MQ2E(b)
Educ	0,074 (0.004)	0,132 (0.049)	0,133 (0.051)	0,161 (0.040)	0,160 (0.041)
Exp	0,084 (0.007)	0,107 (0.021)	0,056 (0.026)	0,119 (0.018)	0,047 (0.025)
$Exp^2/100$	-0,224 (0.032)	-0,228 (0.035)	-0,080 (0.133)	-0,231 (0.037)	-0,032 (0.127)
Black	-0,190 (0.017)	-0,131 (0.051)	-0,103 (0.075)	-0,102 (0.044)	-0,064 (0.061)
Sul	-0,125 (0.015)	-0,105 (0.023)	-0,098 (0.028)	-0,095 (0.022)	-0,086 (0.026)
Urbano	0,161 (0.015)	0,131 (0.030)	0,108 (0.049)	0,116 (0.026)	0,083 (0.041)
Sargan				0.82	0.52
Valor-P				0.37	0.47

- 1 MQ2E (a) usa público e privado como instrumento para educação.
- 2 MQ2E (b) usa público, privado, idade e $idade^2/100$ como instrumentos para educação, experiência e $experiencia^2/100$.

- **Nota:** *O estimador de MQ2E e o estimador de VI são idênticos quando há apenas um instrumento para x_k .*
- Testar a **condição de rank** com uma variável explicativa endógena e múltiplos instrumentos é simples. Basta realizar um teste de hipótese na equação (7)

$$H_0 : \theta_1 = 0, \theta_2 = 0, \dots, \theta_M = 0 \quad (11)$$

contra uma hipótese alternativa que no mínimo um $\theta_j \neq 0$. Este teste é uma razão convincente para rodarmos o primeiro estágio.

- **Sobreidentificação:** é dito que o modelo com uma única variável explicativa endógena é **sobreidentificado** quando $M > 1$ e existe $M - 1$ restrições de sobreidentificação.
- Isto significa que temos $M - 1$ mais variáveis exógenas do que necessário para identificar os parâmetros. Por exemplo, se $M = 2$, poderíamos descartar um dos instrumentos e ainda obter a identificação.

Consistência do Estimador de MQ2E

- Considere o modelo populacional dado por:

$$y = \mathbf{x}\beta + u \quad (12)$$

em que \mathbf{x} é um vetor de dimensão $1 \times k$ e inclui a constante. Alguns elementos de \mathbf{x} podem ser correlacionados com u . Como usual, assumimos que uma amostra aleatória é observada da população.

MQ2E.1 Para algum vetor \mathbf{z} de dimensão $1 \times L$, $\mathbb{E}(\mathbf{z}'u) = \mathbf{0}$

- Esse primeiro pressuposto é derivado da média condicional zero, $\mathbb{E}(u|\mathbf{z}) = 0$.

- O segundo pressuposto traz uma condição geral de *rank*.

$$\text{MQ2E.2a } \text{rank } \mathbb{E}(\mathbf{z}'\mathbf{z}) = L$$

$$\text{MQ2E.2b } \text{rank } \mathbb{E}(\mathbf{z}'\mathbf{x}) = K$$

- A parte (a) é apenas necessária uma vez que as variáveis exógenas serão linearmente independentes na população.
- A parte (b) é **condição de *rank*** crucial para identificação: \mathbf{z} é linear relacionado a \mathbf{x} de forma que o *rank* $\mathbb{E}(\mathbf{z}'\mathbf{x})$ possui posto cheio.
- O que é necessário para a condição de *rank* é a **condição de ordem**, $L \geq k$.

- O estimador **MQ2E** pode ser escrito como a eq. (8):

$$\hat{\beta} = \left(\sum_{i=1}^N \hat{\mathbf{x}}_i' \hat{\mathbf{x}}_i \right)^{-1} \left(\sum_{i=1}^N \hat{\mathbf{x}}_i' y_i \right) = (\widehat{\mathbf{X}}' \widehat{\mathbf{X}})^{-1} \widehat{\mathbf{X}}' \mathbf{Y} \quad (13)$$

ou na forma:

$$\hat{\beta} = \left[\left(\sum_{i=1}^N \mathbf{x}_i' \mathbf{z}_i \right) \left(\sum_{i=1}^N \mathbf{z}_i' \mathbf{z}_i \right)^{-1} \left(\sum_{i=1}^N \mathbf{z}_i' \mathbf{x}_i \right) \right]^{-1} \times \left(\sum_{i=1}^N \mathbf{x}_i' \mathbf{z}_i \right) \left(\sum_{i=1}^N \mathbf{z}_i' \mathbf{z}_i \right)^{-1} \left(\sum_{i=1}^N \mathbf{z}_i' y_i \right) \quad (14)$$

Teorema (Teorema da Consistência do MQ2E)

Sob os pressupostos MQ2E.1 e MQ2E.2, o estimador de MQ2E obtido de uma amostra aleatória é consistente para β : $\hat{\beta} \xrightarrow{p} \beta$ quando $n \rightarrow \infty$.

Prova: escreva

$$\hat{\beta} = \beta + \left[\left(N^{-1} \sum_{i=1}^N \mathbf{x}'_i \mathbf{z}_i \right) \left(N^{-1} \sum_{i=1}^N \mathbf{z}'_i \mathbf{z}_i \right)^{-1} \left(N^{-1} \sum_{i=1}^N \mathbf{z}'_i \mathbf{x}_i \right) \right]^{-1} \times \\ \left(N^{-1} \sum_{i=1}^N \mathbf{x}'_i \mathbf{z}_i \right) \left(N^{-1} \sum_{i=1}^N \mathbf{z}'_i \mathbf{z}_i \right)^{-1} \left(N^{-1} \sum_{i=1}^N \mathbf{z}'_i \mathbf{u}_i \right) \quad (15)$$

e, usando pressuposto MQ2E.1 e MQ2E.2, aplicando a LGN para cada termo junto com o Teorema de Slutsky (ver *Lemma 3.4* do cap. 3 do Wooldridge, 2010).

- Para ver esse resultado no formato matricial, é necessário assumirmos alguns pressupostos:

- 1 As variáveis (Y_{1i}, X_i, Z_i) , $i = 1, 2, \dots, n$, são *i.i.d.*
- 2 $\mathbb{E}(Y_1^2) < \infty$
- 3 $\mathbb{E}\|\mathbf{X}\|^2 < \infty$
- 4 $\mathbb{E}\|\mathbf{Z}\|^2 < \infty$
- 5 $\mathbb{E}(\mathbf{Z}\mathbf{Z}')$ é positiva definida
- 6 $\mathbb{E}(\mathbf{Z}\mathbf{X}')$ tem posto cheio k
- 7 $\mathbb{E}(\mathbf{Z}\mathbf{u}) = \mathbf{0}$

- Na forma matricial, a prova desta consistência é similar a feita para mínimos quadrados.
- Considere a equação estrutural $Y = \mathbf{X}\beta + \mathbf{u}$ e substitua ela dentro da expressão do estimador. Assim, temos:

$$\begin{aligned}\hat{\beta} &= \left[(\mathbf{X}'\mathbf{Z}) (\mathbf{Z}'\mathbf{Z})^{-1} (\mathbf{Z}'\mathbf{X}) \right]^{-1} (\mathbf{X}'\mathbf{Z}) (\mathbf{Z}'\mathbf{Z})^{-1} (\mathbf{Z}'(\mathbf{X}\beta + \mathbf{u})) \\ &= \beta + \left[(\mathbf{X}'\mathbf{Z}) (\mathbf{Z}'\mathbf{Z})^{-1} (\mathbf{Z}'\mathbf{X}) \right]^{-1} (\mathbf{X}'\mathbf{Z}) (\mathbf{Z}'\mathbf{Z})^{-1} (\mathbf{Z}'\mathbf{u})\end{aligned}\tag{16}$$

- Separando o componente estocástico, reescrevendo e aplicando a LFGN e o TCL:

$$\hat{\beta} - \beta = \left[\left(\frac{1}{n} \mathbf{X}' \mathbf{Z} \right) \left(\frac{1}{n} \mathbf{Z}' \mathbf{Z} \right)^{-1} \left(\frac{1}{n} \mathbf{Z}' \mathbf{X} \right) \right]^{-1} \times \quad (17)$$

$$\left(\frac{1}{n} \mathbf{X}' \mathbf{Z} \right) \left(\frac{1}{n} \mathbf{Z}' \mathbf{Z} \right)^{-1} \left(\frac{1}{n} \mathbf{Z}' \mathbf{u} \right) \quad (18)$$

tal que

$$\hat{\beta} - \beta \xrightarrow{p} (\mathbf{Q}_{XZ} \mathbf{Q}_{ZZ} \mathbf{Q}_{ZX})^{-1} \mathbf{Q}_{XZ} \mathbf{Q}_{ZZ}^{-1} \mathbb{E}(\mathbf{Z} \mathbf{u}) = \mathbf{0} \quad (19)$$

em que $\mathbf{Q}_{XZ} = \mathbb{E}(\mathbf{X} \mathbf{Z}')$, $\mathbf{Q}_{ZZ} = \mathbb{E}(\mathbf{Z} \mathbf{Z}')$ e $\mathbf{Q}_{ZX} = \mathbb{E}(\mathbf{Z} \mathbf{X}')$.

- A LFGN se mantém sob o **pressuposto 1** e os pressupostos de segundo momento finito, **pressupostos 2-4**.
- O TMC se aplica se as matrizes Q_{ZZ} e $Q_{XZ}Q_{ZZ}^{-1}Q_{ZX}$ tem inversa, a qual se mantém sob os **pressupostos de identificação 5-6**.
- A igualdade final usa o **pressuposto 7**.

Normalidade Assintótica do Estimador de MQ



- A normalidade assintótica de $\sqrt{N}(\hat{\beta} - \beta)$ vem da normalidade assintótica de $\sqrt{N} \sum_{i=1}^N \mathbf{z}'_i u_i$, que vem do TCL sob o pressuposto MQ2E.1 e do pressuposto de segundo momentos finitos.
- A variância assintótica é mais simples sob o pressuposto de homocedasticidade:

$$\text{MQ2E.3 } \mathbb{E}(u^2 \mathbf{z}' \mathbf{z}) = \sigma^2 \mathbb{E}(\mathbf{z}' \mathbf{z}), \text{ em que } \sigma^2 = \mathbb{E}(u^2)$$

- Pela lei das expectativas iteradas, para o pressuposto MQ2E.3 é suficiente que:

$$\mathbb{E}(u^2 | \mathbf{z}) = \sigma^2 \quad (20)$$

que é o mesmo quando $\text{Var}(u | \mathbf{z}) = \sigma^2$ se $\mathbb{E}(u | \mathbf{z}) = 0$.

Teorema (Teorema normalidade assintótica do MQ2E)

Sob os pressupostos MQ2E.1-MQ2E.3, $\sqrt{N}(\hat{\beta} - \beta)$ é assintoticamente normalmente distribuída com média zero e matriz de variância

$$\sigma^2 \left(\mathbb{E}(\mathbf{x}'\mathbf{z})[\mathbb{E}(\mathbf{z}'\mathbf{z})]^{-1}\mathbb{E}(\mathbf{z}'\mathbf{x}) \right) = \sigma^2[\mathbb{E}(\mathbf{x}^{*'}\mathbf{x}^*)]^{-1} \quad (21)$$

em que $\mathbf{x}^* = \mathbf{z}\Pi$ é um vetor de projeção linear de dimensão $1 \times k$.

- A matriz em (21) é facilmente estimada usando médias amostrais.
- Para estimar σ^2 necessitamos de estimativas apropriadas de u_i . Vamos definir os resíduos de MQ2E como:

$$\hat{u}_i = y_i - \mathbf{x}_i\hat{\beta}, \quad i = 1, 2, \dots, N \quad (22)$$

- Dados os resíduos do MQ2E, um estimador consistente de σ^2 sob os pressupostos MQ2E.1-MQ2E.3 é dado por:

$$\sigma^2 \equiv (N - k)^{-1} \sum_{i=1}^N \hat{u}_i^2 \quad (23)$$

- A matriz $k \times k$

$$\hat{\sigma}^2 \left(\sum_{i=1}^N \hat{\mathbf{x}}_i' \hat{\mathbf{x}}_i \right) = \hat{\sigma}^2 (\widehat{\mathbf{X}}' \widehat{\mathbf{X}})^{-1} \quad (24)$$

é um estimador válido da variância assintótica de $\hat{\beta}$ sob os pressupostos MQ2E.1-MQ2E.3 O desvio-padrão assintótico de $\hat{\beta}_j$ é a raiz quadrada do j -ésimo elemento da diagonal principal da matriz (24).

- O maior incentivo para o uso de MQ2E vem da sua eficiência na classes dos estimadores de VI.

Teorema (Teorema da Relativa Eficiência do MQ2E)

Sob os pressupostos MQ2E.1-MQ2E.3, o estimador de MQ2E é eficiente na classe de todos estimadores de variáveis instrumentais usando instrumentos lineares em \mathbf{z} .

- **Prova:** considere $\hat{\beta}$ o estimador de MQ2E e seja $\tilde{\beta}$ algum estimador de VI usando instrumentos \mathbf{z} .
- Seja $\tilde{\beta}$ tal que $\tilde{\mathbf{x}} \equiv \mathbf{z}\mathbf{\Gamma}$, em que $\mathbf{\Gamma}$ é uma matriz de dimensão $L \times k$.

- Para MQ2E, a escolha da VI é efetivamente $\mathbf{x}^* = \mathbf{z}\Pi$, onde $\Pi = [\mathbb{E}(\mathbf{z}'\mathbf{z})]^{-1}\mathbb{E}(\mathbf{z}'\mathbf{x})$.
- Sob os **pressupostos** $\mathbb{E}(\mathbf{z}'\mathbf{u}) = 0$ [MQ2E.1], $\text{rank } \mathbb{E}(\mathbf{z}'\mathbf{x}) = k$ [MQ2E.2] e $\mathbb{E}(u^2\mathbf{z}'\mathbf{z}) = \sigma^2\mathbb{E}(\mathbf{z}'\mathbf{z})$ [MQ2E.3], sabemos que:
 - 1 $\text{avar}\sqrt{N}(\hat{\beta} - \beta) = \sigma^2\mathbb{E}[(\mathbf{x}^{*'}\mathbf{x}^*)]^{-1}$;
 - 2 $\text{avar}\sqrt{N}(\tilde{\beta} - \beta) = \sigma^2[\mathbb{E}(\tilde{\mathbf{x}}'\mathbf{x})]^{-1}[\mathbb{E}(\tilde{\mathbf{x}}'\tilde{\mathbf{x}})][\mathbb{E}(\mathbf{x}'\tilde{\mathbf{x}})]^{-1}$.
- Queremos mostrar que $\text{avar}\sqrt{N}(\tilde{\beta} - \beta) - \text{avar}\sqrt{N}(\hat{\beta} - \beta)$ é **positiva semi-definida**.
- *Para isso, é suficiente mostrar que $\mathbb{E}(\mathbf{x}^{*'}\mathbf{x}^*) - \mathbb{E}(\mathbf{x}'\tilde{\mathbf{x}})[\mathbb{E}(\tilde{\mathbf{x}}'\tilde{\mathbf{x}})]^{-1}\mathbb{E}(\tilde{\mathbf{x}}'\mathbf{x})$ é positiva semi-definida.*

- Mas $\mathbf{x} = \mathbf{x}^* + \mathbf{r}$, em que $\mathbb{E}(\mathbf{z}'\mathbf{r}) = \mathbf{0}$ e assim $\mathbb{E}(\tilde{\mathbf{x}}'\mathbf{r}) = \mathbf{0}$. Segue que $\mathbb{E}(\tilde{\mathbf{x}}'\mathbf{x}) = \mathbb{E}(\tilde{\mathbf{x}}'\mathbf{x}^*)$ e assim,

$$\begin{aligned}\mathbb{E}(\mathbf{x}^{*'}\mathbf{x}^*) - \mathbb{E}(\mathbf{x}'\tilde{\mathbf{x}})[\mathbb{E}(\tilde{\mathbf{x}}'\tilde{\mathbf{x}})]^{-1}\mathbb{E}(\tilde{\mathbf{x}}'\mathbf{x}) \\ = \mathbb{E}(\mathbf{x}^{*'}\mathbf{x}^*) - \mathbb{E}(\mathbf{x}^{*'}\tilde{\mathbf{x}})[\mathbb{E}(\tilde{\mathbf{x}}'\tilde{\mathbf{x}})]^{-1}\mathbb{E}(\tilde{\mathbf{x}}'\mathbf{x}^*) = \mathbb{E}(\mathbf{s}^{*'}\mathbf{s}^*)\end{aligned}\tag{25}$$

em que $\mathbf{s}^* = \mathbf{x}^* - P(\mathbf{x}^*|\tilde{\mathbf{x}})$ é resíduo populacional da projeção linear de \mathbf{x}^* em $\tilde{\mathbf{x}}$. Como $\mathbb{E}(\mathbf{s}^{*'}\mathbf{s}^*)$ é positiva semi definida, a prova está completa.

- Uma implicação importante desse teorema é que, assintoticamente, teremos melhor resultados sempre que usarmos todos os instrumentos que tivermos disponíveis.

ECONOMETRIA I

ENDOGENEIDADE

Victor Oliveira

Núcleo de Economia Internacional e Desenvolvimento Econômico

PPGDE – 2023