Econometria II

Lista de exercícios – Painel (inclui MQOE, EA, EF e 1ªDif)

1. Considere o modelo:

$$lny_{it} = \beta_0 + \beta_1 x_{it} + \varepsilon_{it}$$
$$\varepsilon_{it} = a_i + u_{it}$$

Os dados consistem de N observações de y (salário) e x (escolaridade) para $T = 2$ períodos de tempo.

a. Interprete possíveis variáveis omitidas que possam fazer parte do efeito fixo (a). Você acredita que "a" possa ser não-correlacionado com x? Que implicações isso tem para a escolha de sua estratégia de estimação de β?

b. Um pesquisador propõe eliminar o efeito fixo, a_i, estimando por MQO um modelo em primeira diferenças do tipo:

$$y_{i1} - y_{i0} = \beta_1(x_{i1} - x_{i0}) + (u_{i1} - u_{i0})$$

Outro pesquisador propõe eliminar o efeito fixo estimando por MQO um modelo de diferenças em relação à média:

$$y_{it} - \overline{y_i} = \beta_1(x_{it} - \overline{x_i}) + (u_{it} - \overline{u_i})$$

Onde $\overline{y_i} = \frac{1}{N}\sum_t y_{it}$ é a média da variável y para o indivíduo i.

Um terceiro pesquisador propõe ainda a inclusão de variáveis dummy (d) indicando se a observação pertence ao indivíduo i, estimando assim um MQO do modelo:

$$\ln(y_{it}) = \beta_0 + \beta_1 x_{it} + \sum_i a_i d_{it} + u_{it}$$

Qual destes estimadores é melhor? Por quê?

c. Supondo que $var(a\,|X) = \sigma_a^2$ e $var(u\,|a, X) = \sigma_u^2$, qual o formato da matriz de covariância de ε?

d. Suponha que E(ε|X) = 0. Qual o estimador de efeitos aleatórios neste caso? Qual sua distribuição?

e. Uma alternativa ao estimador proposto no item anterior é empilhar os dados e tratá-los como um corte transversal. Qual a distribuição deste estimador?

---

2. Em uma amostra aleatória com 10 mil observações de fazendas de soja com informações sobre terra cultivada e número de máquinas, você pretende estimar a seguinte função de produção: $y_{it} = AK_{it}^{b_1}T_i^{b_2}e^{\varepsilon_{it}}$, supondo que $\varepsilon_{it} = u_i + v_{it}$ e que:

$$v_{it} \perp\!\!\!\perp u_i, K_{it}, T_i, v_{js}\,; para\ todo\ i, t, j \neq i, s \neq t$$

Além disso, supomos que $E(u) = E(v) = 0$; $var(u\,|K,T) = \sigma_u^2$ e $var(v\,|\,u,K,T) = \sigma_v^2$.

a. Calcule var(ε_{it}), cov(ε_{it}, ε_{is}) e cov(ε_{it}, ε_{js}).

b. Suponha que você estimou o modelo e obteve os seguintes resultados:

Estimativa	MQE	EA	1D
$\hat{b}_1$	5,5	7	6,2
Var($\hat{b}_1$)	3	1	4
Hausman (z) (H$_0$:$\hat{b}_1 = \hat{b}_1^{1D}$)	11,3	4,2	-

Qual sua melhor estimativa do impacto de lnK sobre lnY? Por que?

c. Neste caso, que dificuldades teríamos para estimar b_2?

---

3. Considere o modelo:

$$y_{it} = \beta_0 + \beta_1 x_{it} + \varepsilon_{it}$$
$$\varepsilon_{it} = a_i + u_{it}$$

Os dados consistem de N observações de y (salário) e x (escolaridade) para $T = 2$ períodos de tempo. Você tem em mente um modelo onde, em equilíbrio, salários refletem a produtividade marginal do trabalho e a escolaridade é portanto interpretada como parte do estoque de capital humano do trabalhador que determina sua produtividade.

a. Qual o efeito marginal que um ano adicional de escolaridade tem sobre a produtividade do trabalhador médio da economia neste modelo?

b. Interprete possíveis variáveis omitidas que possam fazer parte do efeito fixo (a). Você acredita que "a" possa ser não-correlacionado com x? Que implicações isso tem para a escolha de sua estratégia de estimação de β?

c. Um pesquisador propõe eliminar o efeito fixo, a_i, estimando por MQO um modelo em primeira diferenças do tipo:

$$y_{i1} - y_{i0} = \beta_1(x_{i1} - x_{i0}) + (u_{i1} - u_{i0})$$

Outro pesquisador propõe eliminar o efeito fixo estimando por MQO um modelo de diferenças em relação à média:

$$y_{it} - \overline{y_\iota} = \beta_1(x_{it} - \overline{x_\iota}) + (u_{it} - \overline{u_\iota})$$

Onde $\overline{y_\iota} = \frac{1}{N}\sum_t y_{it}$ é a média da variável y para o indivíduo i.

Um terceiro pesquisador propõe ainda a inclusão de variáveis dummy (d) indicando se a observação pertence ao indivíduo i, estimando assim um MQO do modelo:

$$\ln(y_{it}) = \beta_0 + \beta_1 x_{it} + \sum_i a_i d_{it} + u_{it}$$

Qual destes estimadores é melhor? Por quê?

d. Suponha que E(ε|X) = 0. Qual o estimador de efeitos aleatórios neste caso? Qual sua distribuição?

4. Considere o modelo:

$$\begin{aligned}
&\ln(salário)_{it} = b_0 + b_1 escolaridade_{it} + b_2 sexo_i + b_3 idade_{it} + \varepsilon_{it} \\
&\varepsilon_{it} = u_i + v_{it} \\
&v_{it} \perp v_{js}, u_i, X_{it}; \forall i,t,j,s : j \neq i, t \neq s \\
&\mathrm{var}(u \mid X) = \sigma_u^2; \mathrm{var}(v \mid X, u) = \sigma_v^2
\end{aligned}$$

a. Qual o formato da matriz de covariância de $\varepsilon = [\varepsilon_{11}, \dots, \varepsilon_{1T}, \varepsilon_{21}, \dots, \varepsilon_{it}, \dots, \varepsilon_{N1}, \dots, \varepsilon_{NT}]'$?

b. Proponha duas maneiras de transformar o modelo acima de modo a poder estimar de forma consistente o impacto de escolaridade sobre o logaritmo do salário no caso em que u é correlacionado com escolaridade. Se u não fosse correlacionado com a escolaridade você seria capaz de propor um estimador melhor? Justifique.

c. Se o objetivo do exercício fosse verificar se há discriminação por sexo no mercado de trabalho, como você procederia no caso em que (i) E(u | X) = 0?

d. Suponha que a poupança renda 8% ao ano, e que um ano de estudos custe R$10000. Eu gostaria de com meu modelo descobrir se a melhor forma de usar meus R$10 mil é aplicar na poupança ou estudar por um ano a mais. Mostre que na equação acima b_1 pode ser interpretado como a taxa de retorno ao investimento em educação e proponha uma forma de testar se vale a pena investir em educação.

---

5. Considere o modelo:

$$\begin{aligned}
&y_{it} = b_0 + b_1 x_{1it} + b_2 x_{2i} + b_3 x_{3it} + \varepsilon_{it} \\
&\varepsilon_{it} = u_i + v_{it} \\
&v_{it} \perp v_{js}, u_i, X_{it}; \forall i,t,j,s : j \neq i, t \neq s \\
&\mathrm{var}(u \mid X) = \sigma_u^2; \mathrm{var}(v \mid X, u) = \sigma_v^2 \\
&E[v \mid u, X] = 0
\end{aligned}$$

a. Suponha que E[u|X] = 0. Derive a distribuição do estimador de Mínimos Quadrados Empilhados de β = [b0 b1 b2 b3]'.

b. Proponha uma maneira de estimar b_1 caso $E[u|X] \neq 0$. Obtenha a distribuição do seu estimador.

c. Discuta as vantagens e desvantagens do estimador de primeiras diferenças, se comparado ao estimador de efeitos aleatórios.