

## MESTRADO EM ECONOMETRIA APLICADA E PREVISÃO

### MICROECONOMETRIA I EXAME EN – 12/06/2017

Duração 2h30

1. Seja a equação,

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \lambda_t + c_i + u_{it} \quad i = 1, \dots, N \quad t = 1, \dots, T \quad (1)$$

com  $\mathbf{x}_{it}$  vetor  $1 \times k$  de variáveis explicativas que variam no tempo,  $\boldsymbol{\beta}$  vetor  $k \times 1$  de parâmetros desconhecidos,  $\lambda_t$  efeito fixo específico do tempo,  $c_i$  a heterogeneidade não observada específica ao indivíduo, constante no tempo, e  $u_{it}$  variável aleatória não observável. Suponha que  $Var(c_i) = \sigma_c^2$  e  $Var(u_{it}) = \sigma_u^2$ . Suponha ainda que, se nada for dito em contrário, são válidas as hipóteses clássicas.

- a) Deduza as variâncias e covariâncias do erro do modelo (1) e mostre qual o estimador mais eficiente para estimar  $\boldsymbol{\beta}$ , assinalando as hipóteses necessárias para que seja válido. **(20)**
- b) Suponha agora que reespecificava o modelo (1) para  $y_{it} = \alpha y_{it-1} + \mathbf{x}_{it}\boldsymbol{\beta} + \lambda_t + c_i + u_{it}$ . Explique porque não deve estimar este modelo pelo *pooled* OLS. Poderá utilizar o estimador de efeitos fixos? Justifique. **(20)**
2. Considere a seguinte equação que explica as importações de Portugal do país  $i$ , num dado momento de tempo,  $t$ :

$$lm_{it} = \alpha_0 + \beta_1 lypcpt_t + \beta_2 lypcx_{it} + \beta_3 ldist_i + \beta_4 border_i + \beta_5 brics_i + \beta_6 palops_i + \beta_7 eu15_i + v_{it} \quad (2)$$

$$v_{it} = c_i + \lambda_t + u_{it}$$

As variáveis têm o seguinte significado:

$lm_{it}$  - logaritmo das importações de Portugal do país  $i$  no ano  $t$ ;

$lypcpt_t$  - logaritmo do PIB real *per capita* de Portugal no ano  $t$ ;

$ldist_i$  - logaritmo da distância entre Lisboa e a capital do país  $i$ ;

$border_i$  - variável Dummy igual a 1 se Portugal tem fronteira com o país  $i$ ;

$brics_i$  - variável Dummy igual a 1 se o país  $i$  pertence ao grupo dos BRICS;

$palops_i$  - variável Dummy igual a 1 se o país  $i$  pertence ao grupo dos PALOPS;

$EU15_i$  - variável Dummy igual a 1 se o país  $i$  pertence ao grupo dos EU15;

Foram usadas ainda as variáveis:

$idpais$  - identificador do País;

$ano$  - identificador do ano.

A estimação deste modelo conduziu aos resultados que constam do Anexo 1.

- a) Identifique o estimador utilizado na equação 1 para os coeficientes do modelo e para a matriz de covariâncias e explique porque se utilizou este estimador da matriz de covariâncias. **(10)**
- b) Interprete o valor dos coeficientes e comente os resultados comparando com o esperado de acordo com a teoria económica. **(20)**
- c) Identifique o estimador utilizado na equação 2 e explique quais as vantagens e os inconvenientes deste estimador face ao da equação 1, comparando com os resultados obtidos nessa equação. **(20)**
- d) Escreva o modelo estimado na equação 3 e identifique o respetivo estimador (e método de estimação). **(10)**
- e) Identifique as hipóteses de exogeneidade assumidas na equação 3 sobre as variáveis explicativas. Concorda com estas hipóteses? Justifique. **(10)**
- f) Comente a validade dos resultados da equação 3 fazendo todos os testes que se justifiquem adequados. **(20)**
3. Pretende-se modelar a probabilidade de um indivíduo ter uma consulta médica em função da variável dummy igual a 1 se o indivíduo tem uma saúde excelente, *hlthe*, do logaritmo do seu rendimento, *linc*, do número de doenças crónicas, *ndisease*, a idade, *age*, e o quadrado da idade. O Anexo 2 contém várias estimações relacionadas com este modelo. Destas estimações consta ainda a variável binária *dmdu* igual a 1 se o indivíduo teve uma consulta médica.
- a) Qual o método de estimação usado na equação 1? Escreva a função maximizada/minimizada que deu origem às estimativas. **(15)**
- b) Porque foi obtido o quadro 1? Interprete os valores que constam deste quadro e comente-os. **(20)**
- c) Exponha os argumentos que levaram à estimação da equação 2. Verifique se a evidência empírica confirma estes argumentos efetuando um teste de hipóteses. **(15)**
- d) O que se pretende com os comandos **1** a **4** no final da página 6? Interprete o resultado obtido, comparando eventualmente com a estimação obtida com base na equação 1. **(10)**
- e) Por qual das estimações iria optar? Justifique. **(10)**

## ANEXO 1

### Equação 1

```
. reg lm lypcpt lypcx ldist border brics palops eu15 D4 D5 D6 D7, vce(cluster idpais)
```

```
Linear regression                Number of obs    =        254
                                F(10, 36)         =            .
                                Prob > F             =            .
                                R-squared            =        0.6980
                                Root MSE         =        1.3551
```

(Std. Err. adjusted for 37 clusters in idpais)

lm	Coef.	Robust Std. Err.	t	P> t
lypcpt	-.5585286	2.122606	-0.26	0.794
lypcx	.8906026	.3701358	2.41	0.021
ldist	1.409772	1.942768	0.73	0.473
border	3.805592	1.350679	2.82	0.008
brics	2.576852	.7303909	3.53	0.001
palops	-.2542976	.6669226	-0.38	0.705
eu15	2.536895	.7167895	3.54	0.001
D4	-.1537553	.1389402	-1.11	0.276
D5	-.145187	.1785407	-0.81	0.421
D6	-.0725742	.2058224	-0.35	0.726
D7	-.0192061	.3693267	-0.05	0.959
_cons	9.717505	21.02132	0.46	0.647

### Equação 2

```
. xtreg lm lypcpt lypcx ldist border brics palops eu15 D4 D5 D6 D7, fe vce(cluster idpais)
```

```
Fixed-effects (within) regression                Number of obs    =        254
Group variable: idpais                          Number of groups =         37
```

```
R-sq:                                           Obs per group:
within = 0.0137                                min =           5
between = 0.3867                               avg =          6.9
overall = 0.3917                               max =           7
```

```
corr(u_i, Xb) = 0.3343                          F(6,36)         =        0.82
                                                Prob > F        =        0.5611
```

Continua na página seguinte

(Std. Err. adjusted for 37 clusters in idpais)

lm	Coef.	Robust Std. Err.	t	P> t
lypcpt	-.4186021	2.418789	-0.17	0.864
lypcx	.8137824	.4078262	2.00	0.046
ldist	0	(omitted)		
border	0	(omitted)		
brics	0	(omitted)		
palops	0	(omitted)		
eu15	0	(omitted)		
D4	-.1495873	.133444	-1.12	0.270
D5	-.1403124	.1736379	-0.81	0.424
D6	-.1596678	.1501272	-1.06	0.295
D7	-.2403936	.2753223	-0.87	0.388
_cons	15.23558	21.79365	0.70	0.489
sigma_u	1.9959335			
sigma_e	.53888782			
rho	.93205666	(fraction of variance due to u_i)		

### Equação 3

```
. xtabond2 lm l.lm lypcpt lypcx ldist border brics palops eu15 D4 D5 D6 D7,  
iv(ldist border brics palops eu15 D4 D5 D6 D7) gmm(l.lm lypcpt lypcx, lag(. 2))  
twostep robust
```

Dynamic panel-data estimation, two-step system GMM

Group variable: idpais	Number of obs	=	254
Time variable : ano	Number of groups	=	37
Number of instruments = 52	Obs per group: min	=	5
Wald chi2(12) = 1292.59	avg	=	6.86
Prob > chi2 = 0.000	max	=	7

lm	Coef.	Corrected Std. Err.	z	P> z
lm				
L1.	.7157344	.0956788	7.48	0.000
lypcpt	-6.231841	2.595663	-2.40	0.016
lypcx	.2809003	.1340138	2.10	0.036
ldist	.1871284	.0701791	2.67	0.008
border	1.029762	1.067832	0.96	0.335
brics	.8232471	.3615546	2.28	0.023
palops	-.1103787	.4282355	-0.26	0.797
eu15	.5463922	.3892631	1.40	0.160
D4	-.0372656	.1232048	-0.30	0.762
D5	.2524895	.138092	1.83	0.067
D6	.3936083	.2723566	1.45	0.148
D7	.6303051	.3546093	1.78	0.075
_cons	63.34348	26.71783	2.37	0.018

Continua na página seguinte

Instruments for first differences equation  
 Standard  
 D.(ldist border brics palops eu15 D4 D5 D6 D7)  
 GMM-type (missing=0, separate instruments for each period unless collapsed)  
 L(1/2).(L.lm lypcpt lypcx)

Instruments for levels equation  
 Standard  
 ldist border brics palops eu15 D4 D5 D6 D7  
 \_cons  
 GMM-type (missing=0, separate instruments for each period unless collapsed)  
 D.(L.lm lypcpt lypcx)

-----  
 Arellano-Bond test for AR(1) in first differences: z = -1.82 Pr > z = 0.069  
 Arellano-Bond test for AR(2) in first differences: z = -0.81 Pr > z = 0.417  
 -----

Sargan test of overid. restrictions: chi2(39) = 107.24 Prob > chi2 = 0.000  
 (Not robust, but not weakened by many instruments.)  
 Hansen test of overid. restrictions: chi2(39) = 31.21 Prob > chi2 = 0.808  
 (Robust, but weakened by many instruments.)

Difference-in-Hansen tests of exogeneity of instrument subsets:

GMM instruments for levels  
 Hansen test excluding group: chi2(25) = 31.44 Prob > chi2 = 0.175  
 Difference (null H = exogenous): chi2(14) = 5.23 Prob > chi2 = 0.982  
 iv(ldist border brics palops eu15 D4 D5 D6 D7)  
 Hansen test excluding group: chi2(32) = 25.48 Prob > chi2 = 0.786  
 Difference (null H = exogenous): chi2(7) = 5.73 Prob > chi2 = 0.572

## ANEXO 2

### Equação 1

. probit dmdu hlthe linc ndisease age c.age#c.age

Probit regression  
 Number of obs = 5,574  
 LR chi2(5) = 261.13  
 Prob > chi2 = 0.0000  
 Pseudo R2 = 0.0370  
 Log likelihood = -3399.0706

dmdu	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
hlthe	.037126	.0372477	1.00	0.319	-.0358783	.1101302
linc	.1201946	.0143733	8.36	0.000	.0920235	.1483657
ndisease	.0330512	.0029935	11.04	0.000	.0271841	.0389183
age	-.0097199	.0040043	-2.43	0.015	-.0175683	-.0018716
c.age#c.age	.0002105	.0000674	3.12	0.002	.0000784	.0003427
_cons	-.9208892	.1365992	-6.74	0.000	-1.188619	-.6531597

Continua na página seguinte

Quadro 1

. margins, dydx(\*)

Number of obs = 5,574

Model VCE : OIM

Expression : Pr(dmdu), predict()  
 dy/dx w.r.t. : hlthe linc ndisease age

	dy/dx	Delta-method Std. Err.	z	P> z	[95% Conf. Interval]	
hlthe	.0128911	.0129303	1.00	0.319	-.0124518	.038234
linc	.0417346	.004898	8.52	0.000	.0321347	.0513345
ndisease	.0114762	.0010083	11.38	0.000	.0095	.0134524
age	.0002308	.0004123	0.56	0.576	-.0005774	.001039

. biprobit (dmdu=hlthe linc ndisease age c.age#c.age) (hlthe=linc ndisease age c.age#c.age)

Seemingly unrelated bivariate probit Number of obs = 5,574

Wald chi2(9) = 977.98

Log likelihood = -6944.4124 Prob > chi2 = 0.0000

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
dmdu						
hlthe	-.7445952	.3294783	-2.26	0.024	-1.390361	-.0988295
linc	.149197	.0164018	9.10	0.000	.1170501	.1813439
ndisease	.0209781	.0067993	3.09	0.002	.0076517	.0343044
age	-.0188231	.0054205	-3.47	0.001	-.0294471	-.008199
c.age#c.age	.0002749	.0000696	3.95	0.000	.0001385	.0004113
_cons	-.4742128	.254963	-1.86	0.063	-.973931	.0255054
hlthe						
linc	.1326688	.0147463	9.00	0.000	.1037666	.161571
ndisease	-.0348019	.0028082	-12.39	0.000	-.0403059	-.029298
age	-.0325646	.0039181	-8.31	0.000	-.040244	-.0248852
c.age#c.age	.0002551	.0000652	3.91	0.000	.0001273	.0003829
_cons	-.0639944	.1376237	-0.46	0.642	-.3337318	.2057431
/athrho	.5274994	.2698634	1.95	0.051	-.0014232	1.056422
rho	.4834673	.2067854			-.0014232	.7842906

LR test of rho=0: chi2(1) = 2.5951 Prob > chi2 = 0.1072

1. g P1 = normal(-.7445952 + .149197\*linc + .0209781 \*ndisease -.0188231 \*age +.0002749\*age^2)
2. g P0 = normal( .149197\*linc + .0209781 \*ndisease -.0188231 \*age +.0002749\*age^2)
3. g PE = P1 - P0
4. sum PE

Variable	Obs	Mean	Std. Dev.	Min	Max
PE	5,574	-.1886958	.0345815	-.2903283	-.0581061