

MESTRADO EM ECONOMETRIA APLICADA E PREVISÃO

MICROECONOMETRIA E AMOSTRAGEM

EXAME EN – 21/06/2016

Duração 2h30

1. Seja a equação,

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

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

a) Indique a componente do modelo (1) que não é estimada (residual) quando se utiliza o estimador de efeitos fixos e a componente do modelo (1) que não é estimada (residual) quando se usa o estimador de efeitos aleatórios e compare as variâncias de ambas. Comente. **(20)**

b) Supondo que $T = 4$, indique os regressores que permitem a estimação dos parâmetros λ_t $t = 1, 2, \dots, T$. **(15)**

c) Aplique a transformação de efeitos fixos ao modelo (1) e mostre que se \mathbf{x}_i for estritamente exógena (relativamente a u_{it}) então o modelo transformado não tem problemas de endogeneidade. **(15)**

2. Considere a seguinte equação que explica o grau de abertura real de um País i , num dado momento de tempo t :

$$lropen_{it} = \alpha_0 + \delta_1 lropen_{i,t-1} + \beta_2 lrgdppc_{it} + \beta_3 lrgdppc_{i,t-1} + \beta_4 lconsu_{it} + \beta_5 lconsu_{i,t-1} + \beta_6 linv_{it} + \beta_7 linv_{i,t-1} + u_{it} \quad (1)$$

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

As variáveis têm o seguinte significado:

$lropen$ - logaritmo do grau de abertura real sendo este igual a (Exportações + Importações)/PIB real;

$lrgdppc$ - logaritmo do PIB real per capita dos E.U.A.;

$lconsu$ - logaritmo do peso do consumo no PIB real per capita;

$linv$ - logaritmo do peso do investimento no PIB real per capita;

Foram usadas ainda as variáveis:

$country$ - identificador do País;

$year$ - identificador do ano.

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

- a) Identifique o estimador utilizado na equação 1 e explique porque se utiliza a opção **robust**. (10)
- b) Proporia o estimador de efeitos fixos para estimar a equação (1)? Justifique. (20)
- c) Identifique o estimador utilizado na equação 2 e explique quais as vantagens deste estimador face ao da equação 1. Exemplifique com evidência empírica. (20)
- d) Escreva os momentos adicionais usados na estimação da equação 2 relativamente à estimação da equação 1 e teste a sua validade. (15)
3. Pretende-se modelar a satisfação de um indivíduo ao visitar o Parque Nacional do Gerês. Considere a variável $SATIS^*$ que representa a utilidade que o indivíduo retirou da visita. No entanto a variável $SATIS^*$ não foi observada, visto que no inquérito feito aos visitantes do Parque se perguntou apenas se estavam satisfeitos com a visita ou não. Deste modo criou-se a variável binária $SATIS$ que é igual a 1 se o indivíduo ficou satisfeito com a visita. No Anexo 2 encontra resultados da modelação desta variável onde as variáveis ID, ED e REND representam respetivamente a idade do indivíduo, o seu grau de escolaridade e o seu rendimento anual em milhares de euros.
- a) Deduza o modelo estimado na equação 1 do Anexo 2 a partir da especificação da variável latente $SATIS^*$, identificando as hipóteses necessárias. (15)
- b) Interprete os valores obtidos no quadro 1 do Anexo 2. (15)
- c) Suponha que a variância da variável latente é heterocedástica, sendo igual a $\exp(\delta REND)$.
- i) Explique como poderia testar esta hipótese. (10)
- ii) Caso encontrasse evidência de heterocedasticidade comente a validade da estimação da equação 1. (10)
- d) Suponha que a variável REND é endógena sendo Z uma variável instrumental para REND.
- i) Explique como se estimou o modelo para $SATIS$ nesta circunstância sem esquecer de identificar o método usado, comparar as estimativas obtidas para os coeficientes com as obtidas na equação 1 (quanto ao sinal e à grandeza), referindo porque se utilizou a opção **vce(boot)**. (25)
- ii) Será que a evidência empírica aponta para a presença de endogeneidade? Justifique através de um teste estatístico. (10)

ANEXO 1

Equação 1

```
. xtabond2 lropen L.lropen lconsu L.lconsu linv L.linv lrgdppc L.lrgdppc y2002 y2003
y2004 y2005 y2006 y2007, iv( i.year ) gmm(L.lropen linv lconsu lrgdppc) nolevelq
twostep robust
```

```
-----
Group variable: country                Number of obs   =    1113
Time variable : year                  Number of groups =    186
Number of instruments = 108           Obs per group: min =     3
Wald chi2(13) =    370.18              avg =    5.98
Prob > chi2   =     0.000              max =     6
-----
```

	Coef.	Corrected Std. Err.	z	P> z	[95% Conf. Interval]	

lropen						
L1.	.3849266	.0690946	5.57	0.000	.2495037	.5203495

lconsu						
--.	.3973899	.1228614	3.23	0.001	.156586	.6381937
L1.	-.1402155	.0702591	-2.00	0.046	-.2779208	-.0025102

linv						
--.	.185715	.0488057	3.81	0.000	.0900576	.2813725
L1.	-.0102445	.0232781	-0.44	0.660	-.0558688	.0353798

lrgdppc						
--.	-.1793486	.0989354	-1.81	0.070	-.3732584	.0145612
L1.	.0748486	.0692424	1.08	0.280	-.060864	.2105612

y2002	.0134708	.0074273	1.81	0.070	-.0010864	.028028
y2003	.0245295	.0074111	3.31	0.001	.010004	.039055
y2004	.0598014	.0088735	6.74	0.000	.0424096	.0771932
y2005	.0630943	.0100982	6.25	0.000	.0433021	.0828864
y2006	.0740122	.0120347	6.15	0.000	.0504246	.0975998
y2007	.0819756	.0127615	6.42	0.000	.0569635	.1069878

Instruments for first differences equation

Standard

D.(2000b.year 2001.year 2002.year 2003.year 2004.year 2005.year 2006.year 2007.year)

GMM-type (missing=0, separate instruments for each period unless collapsed)

L(1/7).(L.lropen linv lconsu lrgdppc)

```
-----
Arellano-Bond test for AR(1) in first differences: z = -2.79 Pr > z = 0.005
Arellano-Bond test for AR(2) in first differences: z = -1.45 Pr > z = 0.147
-----
```

```
Sargan test of overid. restrictions: chi2(95) = 257.49 Prob > chi2 = 0.000
(Not robust, but not weakened by many instruments.)
```

```
Hansen test of overid. restrictions: chi2(95) = 97.41 Prob > chi2 = 0.412
(Robust, but weakened by many instruments.)
```

Equação 2

```
. xtabond2 lropen L.lropen lconsu L.lconsu linv L.linv lrgdppc L.lrgdppc y2002 y2003
y2004 y2005 y2006 y2007, iv( i.year ) gmm(L.lropen linv lconsu lrgdppc) twostep robust
```

Dynamic panel-data estimation, two-step system GMM

```
-----
Group variable: country                Number of obs      =      1299
Time variable : year                  Number of groups   =       186
Number of instruments = 136           Obs per group: min =         4
Wald chi2(13) = 3348.53                avg                =       6.98
Prob > chi2 = 0.000                    max                =         7
-----
```

	Coef.	Corrected Std. Err.	z	P> z	[95% Conf. Interval]	

lropen						
L1.	.9330298	.0241795	38.59	0.000	.8856389	.9804207
lconsu						
--.	.2911327	.0718818	4.05	0.000	.150247	.4320184
L1.	-.2749576	.0736542	-3.73	0.000	-.4193172	-.130598
linv						
--.	.1412309	.0348011	4.06	0.000	.073022	.2094398
L1.	-.1209515	.0316766	-3.82	0.000	-.1830365	-.0588665
lrgdppc						
--.	-.0130064	.0960063	-0.14	0.892	-.2011754	.1751626
L1.	.0265936	.0955276	0.28	0.781	-.160637	.2138242
y2002	.0090222	.0098896	0.91	0.362	-.0103611	.0284055
y2003	.0179934	.0084314	2.13	0.033	.0014682	.0345186
y2004	.0451898	.0099666	4.53	0.000	.0256556	.0647241
y2005	.0244378	.0076941	3.18	0.001	.0093577	.0395178
y2006	.0210252	.0091113	2.31	0.021	.0031674	.0388831
y2007	.0191465	.0082784	2.31	0.021	.0029212	.0353719
_cons	.1263147	.149778	0.84	0.399	-.1672447	.4198742

Instruments for first differences equation

Standard

D.(2000b.year 2001.year 2002.year 2003.year 2004.year 2005.year 2006.year 2007.year)

GMM-type (missing=0, separate instruments for each period unless collapsed)

L(1/7).(L.lropen linv lconsu lrgdppc)

Instruments for levels equation

Standard

2000b.year 2001.year 2002.year 2003.year 2004.year 2005.year 2006.year 2007.year

_cons

GMM-type (missing=0, separate instruments for each period unless collapsed)

D.(L.lropen linv lconsu lrgdppc)

```
-----
Arellano-Bond test for AR(1) in first differences: z = -4.03 Pr > z = 0.000
Arellano-Bond test for AR(2) in first differences: z = -1.05 Pr > z = 0.295
-----
```

Sargan test of overid. restrictions: chi2(122) = 263.18 Prob > chi2 = 0.000
(Not robust, but not weakened by many instruments.)

Hansen test of overid. restrictions: chi2(122) = 134.38 Prob > chi2 = 0.209
(Robust, but weakened by many instruments.)

ANEXO 2

Equação 1

. probit SATIS ID ED REND

```

Probit regression                               Number of obs   =          243
                                                LR chi2(3)      =          12.45
                                                Prob > chi2     =          0.0060

Log likelihood = -152.8517
    
```

SATIS	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
ID	-.0161422	.0076581	-2.11	0.035	-.0311519	-.0011325
ED	.0639385	.0384337	1.66	0.096	-.0113901	.1392672
REND	.341764	.1924707	1.78	0.076	-.0354715	.7189996
_cons	.1479399	.3834095	0.39	0.700	-.6035289	.8994086

Quadro 1

. margins, dydx(*)

Number of obs = 243

	Delta-method					
	dy/dx	Std. Err.	z	P> z	[95% Conf. Interval]	
ID	-.0057929	.0026699	-2.17	0.030	-.0110257	-.00056
ED	.0229453	.0135613	1.69	0.091	-.0036344	.049525
REND	.1226472	.0677623	1.81	0.070	-.0101644	.2554588

Equação 2

reg REND ID ED Z

Source	SS	df	MS	Number of obs	=	243
Model	4.12671781	3	1.3755726	F(3, 239)	=	6.29
Residual	52.3012198	239	.218833556	Prob > F	=	0.0004
				R-squared	=	0.0731
				Adj R-squared	=	0.0615
Total	56.4279376	242	.233173296	Root MSE	=	.4678

REND	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ID	.0044864	.0027925	1.61	0.109	-.0010146	.0099874
ED	.0470991	.0137303	3.43	0.001	.0200511	.074147
Z	.0037847	.0020248	1.87	0.063	-.0002039	.0077734
_cons	.2469058	.1422015	1.74	0.084	-.0332225	.5270342

. predict vhat, residuals

Equação 3

```
. probit SATIS ID ED REND vhat, vce(boot)
(running probit on estimation sample)
```

Bootstrap replications (50)

```
-----+----- 1 -----+----- 2 -----+----- 3 -----+----- 4 -----+----- 5
..... 50
```

```
Probit regression                Number of obs    =          243
                                Replications        =           50
                                Wald chi2(4)         =          27.76
                                Prob > chi2         =          0.0000
Log likelihood = -136.71062      Pseudo R2       =          0.1406
```

	Observed	Bootstrap	Normal-based			
SATIS	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
ID	-.0607786	.0137323	-4.43	0.000	-.0876934	-.0338637
ED	-.4383894	.1355892	-3.23	0.001	-.7041393	-.1726394
REND	10.30543	2.518026	4.09	0.000	5.370192	15.24067
vhat	-10.05107	2.51904	-3.99	0.000	-14.9883	-5.113848
_cons	-2.86651	.7801534	-3.67	0.000	-4.395583	-1.337438