

MESTRADO EM ECONOMETRIA APLICADA E PREVISÃO

MICROECONOMETRIA I

EXAME ÉPOCA DE RECURSO – 05/07/2017

Duração 2h30

1. Seja a equação,

$$y_{it} = \beta_0 + \mathbf{X}_{it}\boldsymbol{\beta} + \mathbf{Z}_i\boldsymbol{\gamma} + \lambda_t + c_i + u_{it}, \quad i = 1, \dots, N; t = 1, \dots, T \quad (1)$$

onde $\mathbf{X}_{it} = [\mathbf{X}_{1,it} : \mathbf{X}_{2,it}]$ é vetor $1 \times k$ de variáveis explicativas que variam no tempo; $\mathbf{Z}_i = [\mathbf{Z}_{1,i} : \mathbf{Z}_{2,i}]$ é vetor $1 \times g$ de variáveis explicativas que não variam no tempo; $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ são vetores $k \times 1$ e $g \times 1$, respetivamente, de parâmetros desconhecidos; β_0 é o termo constante escalar; λ_t é efeito fixo específico do tempo; c_i é a heterogeneidade não observada específica ao individuo e constante no tempo; u_{it} é variável aleatória independente e não observável. Sabe-se ainda que $\mathbf{X}_{1,it}$ e $\mathbf{Z}_{1,i}$ são vetores $1 \times k_1$ e $1 \times g_1$, respetivamente, estritamente exógenos relativamente ao erro composto e que $\mathbf{X}_{2,it}$ e $\mathbf{Z}_{2,i}$ são vetores $1 \times k_2$ e $1 \times g_2$, respetivamente, que se supõe que sejam endógenos em relação ao termo de heterogeneidade individual.

- a) Apresente sucintamente o teste ao qual iria recorrer para testar a hipótese de endogeneidade conjunta de $\mathbf{X}_{2,it}$ e de $\mathbf{Z}_{2,i}$ relativamente a c_i ? Neste contexto, a estatística de teste seria identificada? Justifique. (10)
- b) Supondo que $\mathbf{X}_{2,it}$ e $\mathbf{Z}_{2,i}$ são efetivamente variáveis endógenas, explicitamente sumariamente a metodologia que iria utilizar para estimar de forma consistente os coeficientes $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$, indicando a condição necessária de identificação para poder aplicar essa metodologia. (20)
- c) Assuma que $\boldsymbol{\beta} = \mathbf{0}$, o vetor nulo. Comente a seguinte afirmação: “considerando um modelo sem regressores que variem no tempo, no contexto do estimador de Efeitos Fixos, a melhor previsão para a variável dependente é zero, para todo o t ”. Justifique, apresentando todos os raciocínios que considerar úteis. (10)

2. Utilizando dados da *World Health Organization* (WHO) pretende-se estudar o indicador da qualidade dos cuidados de saúde em vários países do Mundo, entre 2010 e 2014. A seguinte equação foi estimada:

$$\begin{aligned} \log(\text{health}_{it}) = & \alpha_0 + \beta_1 \log(\text{hexp}_{it}) + \beta_2 \text{educ}_{it} + \beta_3 \text{disab}_{it} + \\ & + \beta_4 \log(\text{pop}_i) + \beta_5 \text{public}_i + \beta_6 \text{gini}_i + \beta_7 \text{tropics}_i + \beta_7 \text{oced}_i + v_{it} \end{aligned} \quad (2)$$

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

As variáveis têm o seguinte significado:

- health_{it} – indicador que mede a qualidade dos cuidados de saúde no País i , no ano t ;
- hexp_{it} – despesa em saúde, *per capita*, do País i , no ano t ;
- educ_{it} – indicador que mede a qualidade da educacional no País i , no ano t ;
- disab_{it} – esperança média de vida de pessoas com deficiência, no País i , no ano t ;
- pop_i – densidade populacional média do País i ;
- public_i – proporção média da despesa em saúde paga pelo Estado do País i ;
- gini_i – índice de Gini médio do País i ;
- tropics_i – igual a 1 se o País i se situa em regiões tropicais;
- oced_i – igual a 1 se o País i é membro da OCDE.

Adicionalmente foram consideradas as variáveis *idcountry* e *year*, identificadores do País e do ano, respetivamente. Os resultados da estimação deste modelo constam do **anexo 1**.

- a) Identifique os estimadores associados a **est1**, **est2**, **est3** e **est4**, fazendo referência à hipótese de exogeneidade subjacente para que o respetivo estimador seja válido. Adicionalmente, explique porque existem estimativas apresentadas como **omitted**. (20)
- b) Justifique a utilização da opção **vce(cluster idcountry)** em todas as estimações, exceto na segunda estimação. (10)
- c) Interprete o valor dos coeficientes estimados e comente os resultados. Elabore uma comparação sucinta entre as estimativas e o esperado pela teoria económica. (20)
- d) Identifique os estimadores associados a **est5** e **est6**, fazendo referência ao método de estimação utilizado. Qual o motivo do recurso à opção **twostep**? E à opção **robust**? (15)
- e) Tendo em conta o conjunto de regressores incluídos em **est5** e **est6**, identifique aqueles que foram considerados estritamente exógenos, pré-determinados e endógenos nas respetivas estimações. Comente a validade destas condições. (10)

f) Esclareça o significado da subopção **lag(2 .)** para **L.1health**. Formalize a respetiva condição de momentos e comente a sua validade recorrendo a um teste adequado. (10)

g) Comente a seguinte afirmação: “a rejeição do teste de Hansen em **est5** deve-se exclusivamente à estimativa da persistência da variável dependente ser muito elevada, tendo em conta a estimação consistente em **est6**”. (15)

3. Com o intuito de se modelar a probabilidade de um indivíduo em idade ativa estar no mercado trabalho, estimou-se o seguinte modelo:

$$P(job_{it} = 1 | \mathbf{X}_{it}, c_i) = G(\gamma_0 + \phi_1 age_{it} + \phi_2 age_{it}^2 + \phi_3 educ_{it} + \phi_4 mar_{it} + \phi_5 kids_{it} + \phi_6 hinc_{it} + c_i + \lambda_t) \quad (3)$$

onde $G(\cdot)$ é uma função de distribuição conhecida. As variáveis têm o seguinte significado:

- job_{it} – igual a 1 se o indivíduo i está empregado, no ano t ;
- age_{it} – idade do indivíduo i , no ano t ;
- $educ_{it}$ – grau de escolaridade do indivíduo i , no ano t ;
- mar_{it} – igual a 1 se o indivíduo i está casado, no ano t ;
- $kids_{it}$ – número de filhos com menos de 18 anos do indivíduo i , no ano t ;
- $hinc_{it}$ – rendimento mensal familiar do indivíduo i , no ano t (em milhares de euros).

Adicionalmente foram consideradas as variáveis id e $year$, identificadores do indivíduo e do ano, respetivamente. Os resultados da estimação deste modelo constam do **anexo 2**.

a) Identifique os estimadores associados a **nlest1** e a **nlest2**. Estas estimativas são diretamente comparáveis? Justifique. (10)

b) A que dizem respeito as estimativas **margins1**, **margins2**, **margins3** e **margins4**? É possível que alguma dessas estimativas não seja teoricamente válida? Identifique-a, justificando. (10)

c) Interprete e compare os valores das estimativas anteriores. (20)

d) Considere que a variável $educ_{it}$ é endógena relativamente a c_i . Explique sucintamente o procedimento que permita controlar esta fonte de endogeneidade, no presente contexto. Indique o modelo subjacente e o método de estimação a utilizar. Formalize cuidadosamente todas as hipóteses necessárias para que a estimação seja válida. (20)

ANEXO 1

```
reg lhealth lhexp educ disab lpop public gini tropics oecd t13 t14, vce(cluster idcountry)
```

```
estimates store est1
```

```
xtreg lhealth lhexp educ disab lpop public gini tropics oecd t13 t14, re
```

```
estimates store est2
```

```
xtreg lhealth lhexp educ disab lpop public gini tropics oecd t13 t14, fe vce(cluster idcountry)
```

```
estimates store est3
```

```
reg D.(lhealth lhexp educ disab lpop public gini tropics oecd t13 t14), vce(cluster idcountry)
```

```
estimates store est4
```

```
estimates table est1 est2 est3 est4, star(.1 .05 .01)
```

Variable	est1	est2	est3	est4
lhexp	.03656263***	.00032743	.00015105	.00021295
educ	.00398718*	.00099262	.00091179	.00093036
disab	.00889714***	.01329024***	.01350469***	.0134443***
lpopden	.0013626	-.00036616	(omitted)	(omitted)
public	.00038114	.00055999**	(omitted)	(omitted)
gini	-.14661942**	-.02584943	(omitted)	(omitted)
tropics	.01040339	.01286539	(omitted)	(omitted)
oecd	-.00880826	.02070149	(omitted)	(omitted)
t13	-.0015454**	-.00006577	-.00011712	-.00007271***
t14	-.00292362***	-.00019438	-.00027512	-.00021119***
_cons	3.5904221***	3.4954757***	3.5152395***	-.00001163

legend: * p<.1; ** p<.05; *** p<.01

```

xtabond2 lhealth L.lhealth lhexp educ disab lpop public gini tropics oecd t13 t14,
iv(t13 t14 tropics oecd) gmm(L.lhealth, lag(2 .)) gmm(lhexp educ disab lpop public
gini) nolevel eq twostep robust

```

```

Hansen test of overid. restrictions: chi2(30) = 50.39 Prob > chi2 = 0.011
gmm(L.lhealth, lag(2 .))
Difference (null H = ): chi2(3) = 6.41 Prob > chi2 = 0.093
iv(t13 t14 tropics oecd)
Difference (null H = ): chi2(2) = 1.70 Prob > chi2 = 0.426

```

```
estimates store est5
```

```

xtabond2 lhealth L.lhealth lhexp educ disab lpop public gini tropics oecd t13 t14,
iv(t13 t14 tropics oecd) gmm(L.lhealth, lag(2 .)) gmm(lhexp educ disab lpop public
gini) twostep robust

```

```

Hansen test of overid. restrictions: chi2(46) = 43.41 Prob > chi2 = 0.582
GMM instruments for levels
Difference (null H = ): chi2(14) = 15.99 Prob > chi2 = 0.314
gmm(L.lhealth, lag(2 .))
Difference (null H = ): chi2(5) = 9.21 Prob > chi2 = 0.101
iv(t13 t14 tropics oecd)
Difference (null H = ): chi2(4) = 2.60 Prob > chi2 = 0.626

```

```
estimates store est6
```

```
estimates table est5 est6, star(.1 .05 .01)
```

Variable	est5	est6
lhealth		
L1.	.35581164**	.91160768***
lhexp	-.00051834	.00306324
educ	-.00022595	-.00004154
disab	.00812448***	.00100798***
lpopden	0	.00301238**
public	0	.00007136
gini	0	-.07866121*
tropics	0	.00849541
oecd	0	-.00565756
t13	.00005573	-.00003729
t14	.00007054	-.00011224
_cons		.31624354***

```

legend: * p<.1; ** p<.05; *** p<.01

```

ANEXO 2

```
probit working c.age##c.age educ married kids houseinc i.year
estimates store nlest1
```

```
xtprobit working c.age##c.age educ married kids houseinc i.year, re
estimates store nlest2
```

```
estimates table nlest1 nlest2, star(.1 .05 .01)
```

Variable	nlest1	nlest2
working		
age	.16654449***	.93494081***
c.age#c.age	-.0022819***	-.01209971***
educ	.10820556***	.40160931***
married	-.07622638*	-.12961089
kids	-.28497207***	-.50054204***
houseinc	1.8037211***	3.3133195***
year		
1985	.05046647	.24068711***
1986	.01794635	.17968511*
1987	.07998483	.49493921***
1988	.08544649*	.56606856***
_cons	-3.695359***	-19.301979***
lnsig2u		
_cons		2.7931435***

legend: * p<.1; ** p<.05; *** p<.01

```

quiet xtreg working c.age##c.age educ married kids houseinc i.year, fe
vce(cluster id)
margins, dydx(age educ married kids houseinc) post
estimates store margins1

```

```

quiet probit working c.age##c.age educ married kids houseinc i.year
margins, dydx(age educ married kids houseinc) post
estimates store margins2

```

```

quiet xtprobit working c.age##c.age educ married kids houseinc i.year, re
margins, dydx(age educ married kids houseinc) post
estimates store margins3

```

```

quiet xtlogit working c.age##c.age educ married kids houseinc i.year, fe
margins, dydx(age educ married kids houseinc) post
estimates store margins4

```

```

estimates table margins1 margins2 margins3 margins4, star(.1 .05 .01)

```

Variable	margins1	margins2	margins3	margins4
age	-.00308024	-.01347609***	-.01350009***	-.00002556
educ	.0841066	.03324387***	.02826837***	3.231e-06
married	-.0086363	-.02341894*	-.00912302	-8.344e-06
kids	-.02887451*	-.08755163***	-.03523202***	-9.528e-06
houseinc	.26167052***	.55415505***	.23321702***	.00013824

legend: * p<.1; ** p<.05; *** p<.01