

UNIVERSIDADE FEDERAL DE VIÇOSA
CENTRO DE CIÊNCIAS AGRÁRIAS
DEPARTAMENTO DE ECONOMIA RURAL

ERU 626 - ECONOMETRIA I
Segundo Semestre/2010

AULA PRÁTICA Nº 3- “Dados em Painel”

Ana Carolina Campana Nascimento
Fernanda Maria de Almeida

1. DADOS EM PAINEL

Dados em painel consistem na combinação de série temporal e seção cruzada, isto é, têm-se dados de várias unidades medidas ao longo do tempo. Considerando um conjunto de dados com $i = 1, 2, \dots, N$ unidades e $t = 1, 2, \dots, T$ períodos de tempo, o modelo geral será:

$$Y_{it} = \alpha_i + X_{it}\beta + \varepsilon_{it} \quad (1)$$

em que α_i representa os efeitos específicos, ou características, das unidades que não variam ao longo do tempo e ε_{it} o termo de erro.

Este modelo gera dois modelos típicos que são estimados de acordo com as pressuposições que fazemos a respeito da possível correlação entre o termo de erro e as variáveis explicativas X_{it} : modelo de efeitos fixos e modelo de efeitos aleatórios.

a) Modelo de Efeitos Fixos:

$$Y_{it} = \alpha_i + \beta X_{it} + \varepsilon_{it} \quad (2)$$

A principal característica deste modelo é tratar os α_i 's como variáveis aleatórias não observadas e correlacionadas com algum X_{it} .

b) Modelo de Efeitos Aleatórios:

$$Y_{it} = \alpha_i + \beta X_{it} + u_{it} \quad (3)$$

O estimador de efeitos aleatórios considera o erro combinado, isto é, $u_{it} = v_i + \varepsilon_{it}$ e pressupõe que v_i é iid com variância σ_v^2 e que ε_{it} é iid com variância σ_ε^2 . Pode-se mostrar que $V(u_{it}) = \sigma_v^2 + \sigma_\varepsilon^2$ e que $Cov(u_{it}, u_{is}) = \sigma_v^2$, $t \neq s$. Logo,

$\rho_u = Cor(u_{it}, u_{is}) = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_\varepsilon^2}$, para todo $t \neq s$. Assim, o modelo de EA tem como

pressuposição correlação serial no erro (correlação igual em todos lags). O estimador de efeitos aleatórios é um estimador de MQG que considera a correlação entre os erros de cada unidade.

Exemplo:

Considere os dados do trabalho de Y. Grunfeld, cujo objetivo era verificar como o investimento real bruto (Y) depende do valor real da empresa (x2) e do estoque real do capital (x3). Os dados são do período de 1935 a 1954 e correspondem a informações de quatro empresas (GE, US, GM e WEST), ou seja, tem-se 4 unidades de corte transversal e 20 períodos. Este exemplo está na página 514 do Gujarati.

Tabela 1 – Investimento de quatro empresas, 1935-1954 (p.515, Gujarati)

ano	id	empresa	y	x2	x3	ano	id	empresa	y	x2	x3
1935	1	GE	33.1	1170.6	97.8	1935	3	US	209.9	1362.4	53.8
1936	1	GE	45	2015.8	104.4	1936	3	US	355.3	1807.1	50.5
1937	1	GE	77.2	2803.3	118	1937	3	US	469.9	2673.3	118.1
1938	1	GE	44.6	2039.7	156.2	1938	3	US	262.3	1801.9	260.2
1939	1	GE	48.1	2256.2	172.6	1939	3	US	230.4	1957.3	312.7
1940	1	GE	74.4	2132.2	186.6	1940	3	US	361.6	2202.9	254.2
1941	1	GE	113	1834.1	220.9	1941	3	US	472.8	2380.5	261.4
1942	1	GE	91.9	1588	287.8	1942	3	US	445.6	2168.6	298.7
1943	1	GE	61.3	1749.4	319.9	1943	3	US	361.6	1985.1	301.8
1944	1	GE	56.8	1687.2	321.3	1944	3	US	288.2	1813.9	279.1
1945	1	GE	93.6	2007.7	319.6	1945	3	US	258.7	1850.2	213.8
1946	1	GE	159.9	2208.3	346	1946	3	US	420.3	2067.7	232.6
1947	1	GE	147.2	1656.7	456.4	1947	3	US	420.5	1796.7	264.8
1948	1	GE	146.3	1604.4	543.4	1948	3	US	494.5	1625.8	306.9
1949	1	GE	98.3	1431.8	618.3	1949	3	US	405.1	1667	351.1
1950	1	GE	93.5	1610.5	647.4	1950	3	US	418.8	1677.4	357.8
1951	1	GE	135.2	1819.4	671.3	1951	3	US	588.2	2289.5	341.1
1952	1	GE	157.3	2079.7	726.1	1952	3	US	645.2	2159.4	444.2
1953	1	GE	179.5	2371.6	800.3	1953	3	US	641	2031.3	623.6
1954	1	GE	189.6	2759.9	888.9	1954	3	US	459.3	2115.5	669.7
1935	2	GM	317.6	3078.5	2.8	1935	4	WEST	12.93	191.5	1.8
1936	2	GM	391.8	4661.7	52.6	1936	4	WEST	25.9	516	0.8
1937	2	GM	410.6	5387.1	156.9	1937	4	WEST	35.05	729	7.4
1938	2	GM	257.7	2792.2	209.2	1938	4	WEST	22.89	560.4	18.1
1939	2	GM	330.8	4313.2	203.4	1939	4	WEST	18.84	519.9	23.5
1940	2	GM	461.2	4643.9	207.2	1940	4	WEST	28.57	628.5	26.5
1941	2	GM	512	4551.2	255.2	1941	4	WEST	48.51	537.1	36.2
1942	2	GM	448	3244.1	303.7	1942	4	WEST	43.34	561.2	60.8
1943	2	GM	499.6	4053.7	264.1	1943	4	WEST	37.02	617.2	84.4
1944	2	GM	547.5	4379.3	201.6	1944	4	WEST	37.81	626.7	91.2
1945	2	GM	561.2	4840.9	265	1945	4	WEST	39.27	737.2	92.4
1946	2	GM	688.1	4900	402.2	1946	4	WEST	53.46	760.5	86
1947	2	GM	568.9	3526.5	761.5	1947	4	WEST	55.56	581.4	111.1
1948	2	GM	529.2	3245.7	922.4	1948	4	WEST	49.56	662.3	130.6
1949	2	GM	555.1	3700.2	1020.1	1949	4	WEST	32.04	583.8	141.8
1950	2	GM	642.9	3755.6	1099	1950	4	WEST	32.24	635.2	136.7
1951	2	GM	755.9	4833	1207.7	1951	4	WEST	54.38	732.8	129.7
1952	2	GM	891.2	4924.9	1430.5	1952	4	WEST	71.78	864.1	145.5
1953	2	GM	1304.4	6241.7	1777.3	1953	4	WEST	90.08	1193.5	174.8
1954	2	GM	1486.7	5593.6	2226.3	1954	4	WEST	68.6	1188.9	213.5

PASSOS PARA ESTIMAÇÃO NO STATA:

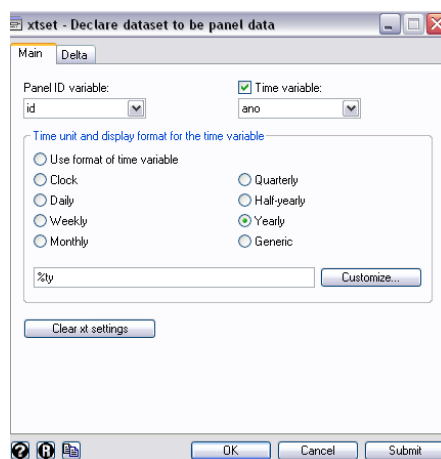
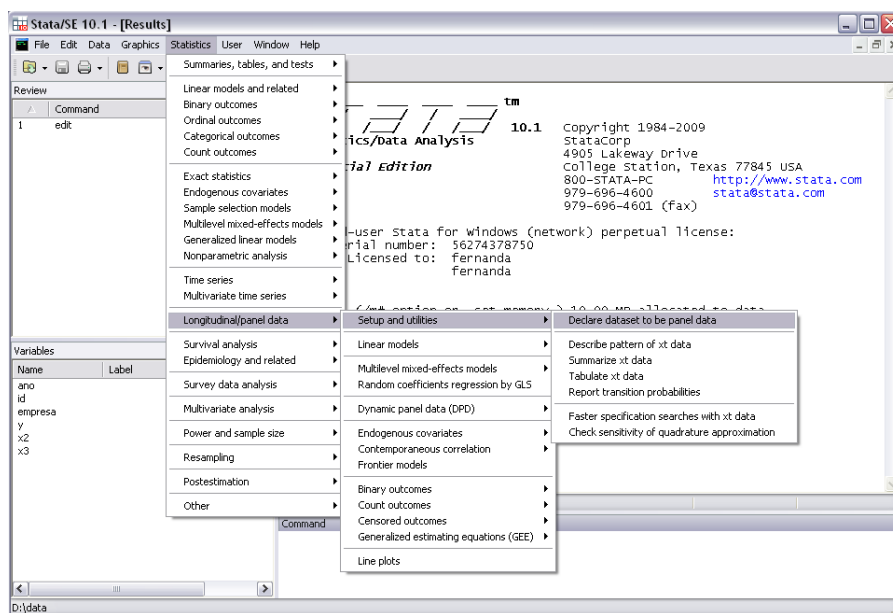
1º) Organização dos dados

No Excel, os dados devem seguir a estrutura da Tabela 1, isto é, ordena-se os dados de acordo com a série temporal para cada unidade de seção cruzada. Note que as unidades de seção cruzada (empresas) são enumeradas, uma vez que o Stata não reconhece textos (nome das unidades).

2º) Declaração dos dados no programa

Após a organização dos dados e inserção dos mesmos no programa, deve-se declarar no programa a variável referente à série de tempo e a referente às unidades. No caso do exemplo em questão, a variável tempo é *ano* e a variável unidade é *id*. Para isso, temos três opções:

i) Barra de ferramentas:



ii) Comando 1:

xtset id ano, yearly

O termo yearly é pra indicar que a série é anual.

```
. xtset id ano, yearly
      panel variable: id (strongly balanced)
      time variable: ano, 1935 to 1954
             delta: 1 year
```

iii) Comando 2:

```
. tis ano
```

```
. iis id
```

3º) Modelo pool

Nesse modelo, todos os coeficientes são constantes ao longo do tempo e entre indivíduos e a forma de estimação é o habitual MQO.

```
. reg y x2 x3
```

Source	SS	df	MS
Model	4849457. 37	2	2424728. 69
Residual	1560689. 67	77	20268. 697
Total	6410147. 04	79	81141. 1018

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
x2	.1100955	.0137297	8.02	0.000	.0827563 .1374348
x3	.3033932	.0492957	6.15	0.000	.2052328 .4015535
_cons	-63.30413	29.6142	-2.14	0.036	-122.2735 -4.334734

4º) Modelo de regressão de efeitos fixos ou de variáveis binárias de mínimos quadrados

i) Caso em que coeficientes angulares são constantes, mas o intercepto varia entre as unidades, isto é:

$$Y_{it} = \beta_{1i} + \beta_2 X_{2it} + \beta_3 X_{3it} + \varepsilon_{it} \quad (4)$$

No Stata, o passo inicial é a criação de *dummies* para cada uma das unidades:

```
. tabulate id, gen (d)
```

id	Freq.	Percent	Cum.
1	20	25.00	25.00
2	20	25.00	50.00
3	20	25.00	75.00
4	20	25.00	100.00
Total	80	100.00	

Note que com o comando **tabulate** foram criadas quatro novas variáveis, uma *dummy* para cada empresa: d1, d2, d3 e d4.

Variables	
Name	Label
ano	
id	
empresa	
y	
x2	
x3	
d1	id== 1.0000
d2	id== 2.0000
d3	id== 3.0000
d4	id== 4.0000

Agora é só rodar o modelo:

```
. reg y x2 x3 d2 d3 d4
```

Source	SS	df	MS			
Model	5990684. 14	5	1198136. 83		Number of obs =	80
Resi dual	419462. 898	74	5668. 41754		F(5, 74) =	211. 37
					Prob > F =	0. 0000
					R-squared =	0. 9346
					Adj R-squared =	0. 9301
Total	6410147. 04	79	81141. 1018		Root MSE =	75. 289

y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
x2	. 1079481	. 0175089	6. 17	0. 000	. 0730608 . 1428354
x3	. 3461617	. 0266645	12. 98	0. 000	. 2930315 . 3992918
d2	161. 5722	46. 45639	3. 48	0. 001	69. 00583 254. 1386
d3	339. 6328	23. 98633	14. 16	0. 000	291. 839 387. 4266
d4	186. 5665	31. 50681	5. 92	0. 000	123. 7879 249. 3452
_cons	-245. 7924	35. 81112	-6. 86	0. 000	-317. 1476 -174. 4371

Obs: a variável d1 foi deixada como referência (evitar problema da multicolinearidade perfeita). Caso sejam utilizadas as 4 *dummies*, o Stata dropará uma automaticamente.

ii) Caso em que coeficientes angulares são constantes, mas o intercepto varia ao longo do tempo (uma *dummy* para cada ano), isto é:

$$Y_{it} = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + \sum_{t=1}^{20} \alpha_t ano_t + \varepsilon_{it} \quad (5)$$

```
. tabulate ano, gen (ano)
```

ano	Freq.	Percent	Cum.
1935	4	5.00	5.00
1936	4	5.00	10.00
1937	4	5.00	15.00
1938	4	5.00	20.00
1939	4	5.00	25.00
1940	4	5.00	30.00
1941	4	5.00	35.00
1942	4	5.00	40.00
1943	4	5.00	45.00
1944	4	5.00	50.00
1945	4	5.00	55.00
1946	4	5.00	60.00
1947	4	5.00	65.00
1948	4	5.00	70.00
1949	4	5.00	75.00
1950	4	5.00	80.00
1951	4	5.00	85.00
1952	4	5.00	90.00
1953	4	5.00	95.00
1954	4	5.00	100.00
Total	80	100.00	


```
. reg y x2 x3 ano1-ano19
```

Source	SS	df	MS			
Model	4938658. 06	21	235174. 193		Number of obs =	80
Resi dual	1471488. 98	58	25370. 4997		F(21, 58) =	9. 27
					Prob > F =	0. 0000
					R-squared =	0. 7704
					Adj R-squared =	0. 6873
Total	6410147. 04	79	81141. 1018		Root MSE =	159. 28

y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
x2	. 1159174	. 01817	6. 38	0. 000	. 0795462 . 1522886
x3	. 2696593	. 0833411	3. 24	0. 002	. 1028339 . 4364847
ano1	21. 02495	129. 7675	0. 16	0. 872	-238. 7329 280. 7828
ano2	-14. 03424	133. 1953	-0. 11	0. 916	-280. 6535 252. 5851
ano3	-58. 41449	135. 1823	-0. 43	0. 667	-329. 0113 212. 1823
ano4	-48. 66584	126. 5326	-0. 38	0. 702	-301. 9483 204. 6167
ano5	-96. 80262	128. 0009	-0. 76	0. 453	-353. 0243 159. 4191
ano6	-36. 10807	129. 116	-0. 28	0. 781	-294. 5618 222. 3457
ano7	21. 16647	127. 7266	0. 17	0. 869	-234. 5062 276. 8391
ano8	30. 29937	124. 1757	0. 24	0. 808	-218. 2653 278. 864
ano9	-12. 76907	124. 8506	-0. 10	0. 919	-262. 6847 237. 1466
ano10	-17. 82782	125. 6498	-0. 14	0. 888	-269. 3431 233. 6875
ano11	-36. 96994	126. 769	-0. 31	0. 760	-292. 7257 214. 7858
ano12	26. 90836	125. 7493	0. 21	0. 831	-224. 8062 278. 623
ano13	27. 81252	119. 3515	0. 23	0. 817	-211. 0954 266. 7204
ano14	26. 05879	117. 3271	0. 22	0. 825	-208. 7969 260. 9145
ano15	-28. 65514	116. 2965	-0. 25	0. 806	-261. 4479 204. 1376
ano16	-20. 3938	115. 8973	-0. 18	0. 861	-252. 3874 211. 5998
ano17	. 9819593	116. 2576	0. 01	0. 993	-231. 733 233. 6969
ano18	21. 96068	114. 6821	0. 19	0. 849	-207. 6005 251. 5219
ano19	39. 43192	113. 441	0. 35	0. 729	-187. 6448 266. 5087
_cons	-56. 33982	99. 75287	-0. 56	0. 574	-256. 0169 143. 3373

ii) Caso em que coeficientes angulares são constantes, mas o intercepto varia com os indivíduos e com o tempo (uma *dummy* para cada empresa e para cada ano).

```
. reg y x2 x3 d2-d4 ano1-ano19
```

Source	SS	df	MS		Number of obs =	80
Model	6082748.11	24	253447.838		F(24, 55) =	42.58
Residual	327398.928	55	5952.70778		Prob > F =	0.0000
					R-squared =	0.9489
					Adj R-squared =	0.9266
Total	6410147.04	79	81141.1018		Root MSE =	77.154

y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
x2	.129307	.0274237	4.72	0.000	.0743487 .1842853
x3	.3672492	.0416591	8.82	0.000	.2837625 .450736
d2	105.2457	67.68668	1.55	0.128	-30.40141 240.8929
d3	341.1008	24.8116	13.75	0.000	291.3773 390.8244
d4	220.324	41.16781	5.35	0.000	137.8219 302.8262
ano1	134.3636	72.00957	1.87	0.067	-9.946798 278.674
ano2	87.3297	66.46131	1.31	0.194	-45.86174 220.5211
ano3	29.58592	66.10801	0.45	0.656	-102.8975 162.0893
ano4	48.12215	66.82507	0.72	0.474	-85.79829 182.0426
ano5	-7.886558	63.9173	-0.12	0.902	-135.9797 120.2066
ano6	51.85022	63.78342	0.81	0.420	-75.9746 179.6751
ano7	107.7241	63.45228	1.70	0.095	-19.43706 234.8854
ano8	118.3592	64.92635	1.82	0.074	-11.75613 248.4745
ano9	71.99879	63.47345	1.13	0.262	-55.20486 199.2024
ano10	68.47821	63.65855	1.08	0.287	-59.09638 196.0528
ano11	44.28525	62.83216	0.70	0.484	-81.63321 170.2037
ano12	104.1942	61.82815	1.69	0.098	-19.71216 228.1006
ano13	100.1916	62.87224	1.59	0.117	-25.80715 226.1904
ano14	92.30316	63.16828	1.46	0.150	-34.2889 218.8952
ano15	31.20783	62.10854	0.50	0.617	-93.26046 155.6761
ano16	35.80472	61.1856	0.59	0.561	-86.81396 158.4234
ano17	47.8422	57.53413	0.83	0.409	-67.45878 163.1432
ano18	57.96435	56.39349	1.03	0.309	-55.05073 170.9794
ano19	54.01372	54.99911	0.98	0.330	-58.20895 164.2344
_cons	-359.5819	82.63602	-4.35	0.000	-525.1882 -193.9756

Outras opções de modelos seria considerar que todos os coeficientes variam entre os indivíduos, isto é, utilizar *dummies* de inclinação. Todavia, deve-se ter cautela quanto ao uso de *dummies*, uma vez que um grande número delas reduz os graus de liberdade, além de aumentar a possibilidade de multicolinearidade.

Quanto à escolha entre o modelo da equação pool e cada uma das especificações apresentadas anteriormente, utiliza-se o teste F restrito.

5º) Modelo de efeitos fixos

```
. xtreg y x2 x3, fe
```

Fixed-effects (within) regression	Number of obs =	80
Group variable: id	Number of groups =	4
R-sq: within = 0.8068	Obs per group: min =	20
between = 0.7304	avg =	20.0
overall = 0.7554	max =	20
corr(u_i, Xb) = -0.1001	F(2, 74) =	154.53
	Prob > F =	0.0000

y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
x2	.1079481	.0175089	6.17	0.000	.0730608 .1428354
x3	.3461617	.0266645	12.98	0.000	.2930315 .3992918
_cons	-73.84946	37.52291	-1.97	0.053	-148.6155 .9165759
sigma_u	139.05116				
sigma_e	75.288894				
rho	.77329633				(fraction of variance due to u_i)

F test that all u_i=0:	F(3, 74) =	67.11	Prob > F =	0.0000
------------------------	------------	-------	------------	--------

Os diferentes valores de R^2 indicam como o modelo se ajusta dentro das unidades (R^2_{within}), entre unidades ($R^2_{between}$) e no geral ($R^2_{overall}$). O termo *sigma_u* é o

erro padrão de α_i e σ_e é o erro padrão de ε_{it} (σ_e). A expressão ρ é

$$\text{uma estimativa da relação } \text{Corr}(\varepsilon_{it}, \varepsilon_{is}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\varepsilon^2}, \text{ ou seja, a razão da variância de } \alpha_i$$

para a variância do erro composto. O teste F na última linha é o teste utilizado para verificar se o modelo pool é mais adequado que o modelo de efeitos fixos (Ho: modelo pool é preferível ao modelo de efeitos fixos).

Efeitos fixos para unidades:

```
. predict fe_id, u
. list fe_id
```

	fe_id		
1.	-171.9429	41.	167.6899
2.	-171.9429	42.	167.6899
3.	-171.9429	43.	167.6899
4.	-171.9429	44.	167.6899
5.	-171.9429	45.	167.6899
6.	-171.9429	46.	167.6899
7.	-171.9429	47.	167.6899
8.	-171.9429	48.	167.6899
9.	-171.9429	49.	167.6899
10.	-171.9429	50.	167.6899
11.	-171.9429	51.	167.6899
12.	-171.9429	52.	167.6899
13.	-171.9429	53.	167.6899
14.	-171.9429	54.	167.6899
15.	-171.9429	55.	167.6899
16.	-171.9429	56.	167.6899
17.	-171.9429	57.	167.6899
18.	-171.9429	58.	167.6899
19.	-171.9429	59.	167.6899
20.	-171.9429	60.	167.6899
21.	-10.37068	61.	14.62365
22.	-10.37068	62.	14.62365
23.	-10.37068	63.	14.62365
24.	-10.37068	64.	14.62365
25.	-10.37068	65.	14.62365
26.	-10.37068	66.	14.62365
27.	-10.37068	67.	14.62365
28.	-10.37068	68.	14.62365
29.	-10.37068	69.	14.62365
30.	-10.37068	70.	14.62365
31.	-10.37068	71.	14.62365
32.	-10.37068	72.	14.62365
33.	-10.37068	73.	14.62365
34.	-10.37068	74.	14.62365
35.	-10.37068	75.	14.62365
36.	-10.37068	76.	14.62365
37.	-10.37068	77.	14.62365
38.	-10.37068	78.	14.62365
39.	-10.37068	79.	14.62365
40.	-10.37068	80.	14.62365

Efeitos fixos para tempo:

Para que o Stata reconheça que ele deve calcular o efeito fixo na variável de tempo, deve-se setar: **iis** para a variável de tempo e **tis** para a variável cross-section. Ou seja:

```

. tis id
. iis ano
. qui xtreg y x2 x3, fe
. predict ano_fe, u
. list ano_fe

```

	ano_fe		
1.	28. 87476	41.	28. 87476
2.	-6. 184439	42.	-6. 184439
3.	-50. 56469	43.	-50. 56469
4.	-40. 81604	44.	-40. 81604
5.	-88. 95282	45.	-88. 95282
6.	-28. 25827	46.	-28. 25827
7.	29. 01627	47.	29. 01627
8.	38. 14917	48.	38. 14917
9.	-4. 919267	49.	-4. 919267
10.	-9. 978021	50.	-9. 978021
11.	-31. 12014	51.	-31. 12014
12.	34. 75816	52.	34. 75816
13.	35. 66232	53.	35. 66232
14.	33. 90859	54.	33. 90859
15.	-20. 80534	55.	-20. 80534
16.	-12. 544	56.	-12. 544
17.	8. 831759	57.	8. 831759
18.	29. 81048	58.	29. 81048
19.	47. 28172	59.	47. 28172
20.	7. 849801	60.	7. 849801
21.	28. 87476	61.	28. 87476
22.	-6. 184439	62.	-6. 184439
23.	-50. 56469	63.	-50. 56469
24.	-40. 81604	64.	-40. 81604
25.	-88. 95282	65.	-88. 95282
26.	-28. 25827	66.	-28. 25827
27.	29. 01627	67.	29. 01627
28.	38. 14917	68.	38. 14917
29.	-4. 919267	69.	-4. 919267
30.	-9. 978021	70.	-9. 978021
31.	-31. 12014	71.	-31. 12014
32.	34. 75816	72.	34. 75816
33.	35. 66232	73.	35. 66232
34.	33. 90859	74.	33. 90859
35.	-20. 80534	75.	-20. 80534
36.	-12. 544	76.	-12. 544
37.	8. 831759	77.	8. 831759
38.	29. 81048	78.	29. 81048
39.	47. 28172	79.	47. 28172
40.	7. 849801	80.	7. 849801

6º) Modelo de efeitos aleatórios

```

. tis ano
. iis id
. xtreg y x2 x3, re

```

```

Random-effects GLS regression              Number of obs   =      80
Group variable:  id                       Number of groups =       4

R-sq:  within = 0. 8068                    Obs per group:  min =      20
          between = 0. 7303                  avg =      20.0
          overall = 0. 7554                  max =      20

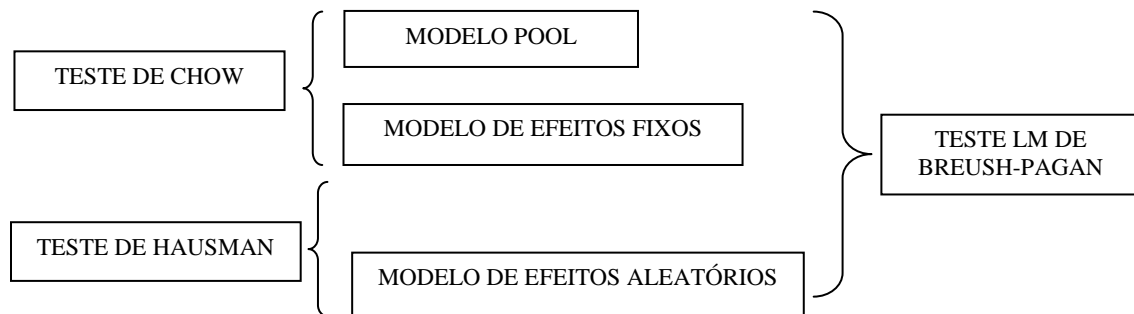
Random effects u_i ~ Gaussian              Wald chi2(2)     =    317. 79
corr(u_i, X)      = 0 (assumed)            Prob > chi2      =     0. 0000

```

y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
x2	. 1076555	. 0168169	6. 40	0. 000	. 0746949	. 140616
x3	. 3457104	. 0265451	13. 02	0. 000	. 2936829	. 3977378
_cons	-73. 03529	83. 94957	-0. 87	0. 384	-237. 5734	91. 50284
sigma_u	152. 15823					
sigma_e	75. 288894					
rho	. 80332024	(fraction of variance due to u_i)				

7º) Escolha entre pool, efeitos fixos e efeitos aleatórios

Para realizar a escolha entre os modelos, utiliza-se os seguintes testes:



i) Teste de Chow

H0: modelo restrito (pooled)

H1: modelo irrestrito (efeitos fixos)

A estatística do teste F da linha inferior da estimativa de efeitos fixos, bem como seu respectivo p-valor indica que o modelo de efeitos fixos é melhor que o pool.

```

. xtreg y x2 x3, fe
Fixed-effects (within) regression
Group variable: id
R-sq:  within = 0.8068
       between = 0.7304
       overall = 0.7554
corr(u_i, Xb) = -0.1001
Number of obs   = 80
Number of groups = 4
Obs per group:  min = 20
                avg = 20.0
                max = 20
F(2, 74) = 154.53
Prob > F = 0.0000

```

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
x2	.1079481	.0175089	6.17	0.000	.0730608 .1428354
x3	.3461617	.0266845	12.98	0.000	.2930315 .3992918
_cons	-73.84946	37.52291	-1.97	0.053	-148.6155 .9165759
sigma_u	139.05116				
sigma_e	75.288894				
rho	.77329633				(fraction of variance due to u_i)

F test that all u_i=0: F(3, 74) = 67.11 Prob > F = 0.0000

ii) Teste de Hausman

H0: modelo de efeitos aleatórios

H1: modelo de efeitos fixos

```

. qui xtreg y x2 x3, fe
. estimates store fe
. qui xtreg y x2 x3, re
. estimates store re
. hausman fe re

```

	Coefficients		(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
	(b) fe	(B) re		
x2	.1079481	.1076555	.0002926	.0048738
x3	.3461617	.3457104	.0004513	.0025204

b = consistent under H0 and Ha; obtained from xtreg
B = inconsistent under Ha, efficient under H0; obtained from xtreg

Test: H0: difference in coefficients not systematic

chi2(2) = (b-B)' [(V_b-V_B)^(-1)] (b-B)
= 0.07
Prob>chi2 = 0.9678

Pela estatística do teste de hausman, tem-se que o modelo de efeitos aleatórios é melhor que o de efeitos fixos.

iii) Teste LM de Breusch-Pagan

H0: modelo pooled

H1: modelo de efeitos aleatórios

```
. qui xtreg y x2 x3, re
. xttest0

Breusch and Pagan Lagrangian multiplier test for random effects

y[id, t] = Xb + u[id] + e[id, t]

Estimated results:

```

	Var	sd = sqrt(Var)
y	81141.1	284.8528
e	5668.418	75.28889
u	23152.13	152.1582

```
Test: Var(u) = 0
chi2(1) = 379.08
Prob > chi2 = 0.0000
```

O resultado do teste indica que efeitos aleatórios são preferíveis ao modelo pool.

8º) Detecção de autocorrelação e heterocedasticidade em painel

i) autocorrelação (teste de Wooldridge)

Instalação: *findit xtserial*, clicar em *st0039* e depois *click here to install*

```
. xtserial y x2 x3, output
Linear regression
```

		Number of obs =		76	
		F(2, 3) =		45.61	
		Prob > F =		0.0057	
		R-squared =		0.4578	
		Root MSE =		65.499	

(Std. Err. adjusted for **4** clusters in id)

D. y	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
x2					
D1.	.0901393	.0168937	5.34	0.013	.0363761 .1439025
x3					
D1.	.3229579	.1240902	2.60	0.080	-.0719524 .7178683

```
Wooldridge test for autocorrelation in panel data
H0: no first-order autocorrelation
F( 1, 3) = 1300.479
Prob > F = 0.0000
```

Rejeita-se a hipótese nula de ausência de autocorrelação.

ii) Teste de Wald para heterocedasticidade em grupo (efeitos fixos)

Instalação: *findit xttest3*, clicar em *st0004* e depois *click here to install*

```
. qui xtreg y x2 x3, fe
. xttest3

Modified Wald test for groupwise heteroskedasticity
in fixed effect regression model

H0: sigma(1)^2 = sigma^2 for all i

chi2 (4) = 240.33
Prob>chi2 = 0.0000
```

Rejeita-se a hipótese nula de ausência de heterocedasticidade.

A correção desses problemas pode ser feita por estimações considerando erros padrão robustos ou por bootstrap.