

INTRODUÇÃO AO PAINEL ESPACIAL

VANTAGENS AO USAR UM PAINEL DE DADOS

Um painel de dados, onde várias localidades ($i \dots n$) são acompanhadas ao longo do tempo ($t > 1$), possui certas vantagens, em relação à opção *cross-section* (cujo foco reside em um único período, $t = 1$). Dentre estas: a) possibilita a inclusão de mais observações ($i * t$), fortalecendo as propriedades assintóticas dos estimadores e das inferências (testes t e F mais robustos); b) admite ajustamentos dinâmicos e análises intertemporais c) permite utilizar certas características locais, não-observadas e invariantes no tempo (e.g.: preferências, cultura, relevo, clima, etc.), para reduzir a variância residual (i.e.: modelo de efeitos aleatórios - EA) ou eliminar vieses nas estimativas (i.e.: modelo de efeitos fixos - EF) (Wooldridge, 2010).

AGRUPANDO OS DADOS

COD_IBGE	t (período)	i (local)	y	x1	x2	Dica: se estiver usando o Excel, use o comando "PROCV" (baseado em "COD_IBGE") para incluir novas variáveis (ex. x3, x4)
3127701	2010	GV	10	2	5	
3136702	2010	JF	8	3	2.5	
3106200	2010	BH	15	9	6	
3127701	2020	GV	7	1	3.3	
3136702	2020	JF	12	4	8	
3106200	2020	BH	16	13	11.8	

METODOLOGIA EM PAINEL SIMPLES

A **Equação 1**, baseada em i regiões e t períodos, possui dados na forma de painel e pode ser estimada via **Pooled Ordinary Least Squares (POLS)**, usando o teste de Breusch-Pagan (1980) para verificar se existe algum efeito não observado, c_i , constante no tempo (e.g.: cultura, preferências, clima, etc.), afetando os resíduos do modelo. Caso $H_0: \sigma_c^2 = 0$ prevaleça, o POLS é o mais indicado.

$$y_{it} = X_{it}\beta_k + \varepsilon_{it} \quad (1)$$

Onde: $\beta_{i\dots k}$ são os coeficientes de impacto que acompanham as k variáveis explicativas incluídas na matriz X_{it} e ε_{it} é um termo de erro.

Caso contrário ($\sigma_c^2 \neq 0$), deve-se estimar os modelos de **efeitos fixos (EF)** e de **efeitos aleatórios (EA)**, usando o teste de Hausman (1978) para definir se c_i causa viés em alguma variável explicativa (x_{it}). Formalmente (Wooldridge, 2010):

- EF:** assume que c_i causa viés. Logo, se $y_{it} = X_{it}\beta_k + c_i + \varepsilon_{it}$, elimina-se c_i ao usar os desvios das variáveis em relação à média (e.g.: $\dot{y}_{it} = y_{it} - \bar{y}_i$). Neste caso, estima-se: $\dot{y}_{it} = \dot{X}_{it}\beta_k + \dot{\varepsilon}_{it}$.
- EA:** assume que c_i é aleatório e faz parte do termo de erro (i.e.: $y_{it} = X_{it}\beta_k + v_{it}$, onde: $v_{it} = c_i + \varepsilon_{it}$). Logo, usa-se a correlação serial de c_i em v_{it} para tornar as estimativas mais eficientes.

Aceitando-se $H_0: E(c_i|x_{it}) = 0$, EF e EA serão consistentes, porém EA será mais eficiente. Do contrário, apenas EF será consistente (Wooldridge, 2010). Em todos os casos, recomenda-se o uso da matriz robusta de White (1980), a fim de minimizar problemas de heterocedasticidade.

METODOLOGIA EM PAINEL-ESPACIAL

Como é possível que uma região i influencie seus j vizinhos e vice-versa (Gallo e Ertur, 2003) e ignorar esses efeitos espaciais poderia gerar estimativas viesadas e/ou ineficientes (Lesage e Pace, 2009), deve-se controlar a possível dependência espacial associada à Equação 1.

“There are a number of factors - trade between regions, technology and knowledge diffusion and more generally regional spillovers - that lead to geographically dependent regions. Because of spatial interactions between regions, geographical location is important in accounting for the economic performances of regions.” Gallo e Ertur (2003, p.176).

Para tanto, é comum usar a matriz de contiguidade espacial “rainha” (W),¹ para definir os j vizinhos das i regiões analisadas. Deste modo, é possível controlar a dependência espacial associada à variável dependente (Wy_{it}), às variáveis explicativas (WX_{it}) e ao termo de erro ($W\xi_{it}$), conforme exposto na **Equação 2**.

$$y_{it} = \rho W y_{it} + X_{it} \beta_k + W X_{it} \tau_k + \xi_{it}, \text{ sendo: } \xi_{it} = \lambda W \xi_{it} + \varepsilon_{it} \quad (2)$$

A equação 2, pode ser estimada via *Quasi-Máxima Verossimilhança* (QMV), com base no método de **Efeitos Fixos Espacial (EFE)**, de Lee e Yu (2010), e no modelo de **Efeitos Aleatórios Espacial (EAE)**, proposto por Kapoor, Kelejian e Prucha (2007). Novamente, deve-se usar o teste de Hausman (1978), para definir entre EFE ou EAE.

Cabe destacar que, em modelos espaciais, o efeito total de x_k (*i.e.*: $\partial y / \partial x_k$), depende do seu impacto inicial (β_k) e do efeito multiplicador gerado pela vizinhança. Assim, com base na matriz de efeitos parciais $|(I_n - \rho W)^{-1} I_n \beta_k|$, onde I_n é uma matriz identidade de dimensão n , pode-se obter o efeito direto - ED (média da diagonal principal), indireto - EI (média dos elementos fora da diagonal principal) e total, $ET = ED + EI$, proveniente de x_k (Lesage e Pace, 2009).

REFERÊNCIAS

- BREUSCH, T.S; PAGAN A.R. The LM Test and Its Applications to Model Specification in Econometrics, *Review of Economic Studies* 47, 239–254, 1980.
- GALLO, J.L.; ERTUR, C. Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980–1995. *Papers in Regional Science*. v.82, p.175-201, 2003.
- HAUSMAN, J.A. Specification Tests in Econometrics. *Econometrica* 46, p.1251–1271, 1978.
- KAPOOR, M.; KELEJIAN, H.H.; PRUCHA, I.R. Panel data models with spatially correlated error components. *Journal of Econometrics* 140: 97–130, 2007.
- LEE, L.; YU, J. A spatial dynamic panel data model with both time and individual fixed effects. *Econometric Theory* 26.2, pp. 564–597. 2010.
- LESAGE, J.P.; PACE, R.K. *Introduction to spatial econometrics*. 1ª ed. Taylor-Francis, 2009.
- LESAGE, J.P.; PACE, R.K. The Biggest Myth in Spatial Econometrics. *Econometrics*, n.2, p.217-249, 2014.
- WHITE, H.A. Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity. *Econometrica* 48: 817-838, 1980.
- WOOLDRIDGE, J.M. *Econometric Analysis of Cross Section and Panel Data*. 2nd Ed. The MIT Press, 2010, 1096p.

¹ Embora existam outras opções para W , LeSage e Pace (2014) afirmam que os resultados econométricos-espaciais seriam pouco sensíveis à escolha de matrizes baseadas na proximidade e/ou contiguidade.

**COMANDOS STATA.**

- 1) Crie uma pasta na área de trabalho com o nome “Espacial_STATA”;
- 2) Insira seu Shape File (Shape_BRmun) e o arquivo em Excel, com os dados, na Pasta;
- 3) Certifique-se de que os dados já estão na forma de painel ($i * t$ linhas por $k + 1$ colunas);
- 4) Crie uma matriz de contiguidade “rainha” no GEODA e chame-a “WQ”

*DICA: mantenha todos os arquivos na mesma pasta e não use espaços no diretório (ex.: use “base de dados” e não “base de dados”)

***a) instalando toolbox espaciais no STATA**

```
ssc install spmap
ssc install sppack
ssc install shp2dta
ssc install mif2dta
ssc install spgrid
ssc install geodist
ssc install mergepoly
ssc install geoinpoly
ssc install xsmle
```

***b) copie e cole os dados do excel no STATA e salve-os como “Dados_Espaciais.dta”**

```
save "C:\Users\Asus\Desktop\Espacial_STATA\Dados_Espaciais.dta", replace
```

***c) Defina o diretório e crie um arquivo com os resultados que serão gerados no STATA**

```
set more off
cd "C:\Users\Asus\Desktop\Espacial_STATA"
cap log using "resultados_STATA", replace
```

***d) transforme “Shape_BRmun” (GEODA) em .dta (STATA)**

```
shp2dta using Shape_BRmun, database(Shape_BRmun) coordinates(X_COORD Y_COORD) genid (_ID)
```

***e) Salve as coordenadas X e Y em .dta (STATA) – para geo-referenciar os dados**

```
use Shape_BRmun.dta
keep _ID X_COORD Y_COORD MUNIC_PIO
save cord.dta, replace
```

***f) Insira as coordenadas do “Shape_BRmun” na base de dados e chame-a “base.dta” (STATA)**

```
use Dados_Espaciais.dta
rename codigo MUNIC_PIO
merge m:m MUNIC_PIO using cord.dta
save base.dta, replace
clear all
```

***g) Usando “base.dta”, importe a matriz “rainha” (WQ) para o STATA**

```
use base.dta
spmat import WQ using rainha.gal, geoda
spmat save WQ using "WQ.dta", replace
spmat export WQ using "WQ.txt", replace
spmatrix import WQ using C:\Users\Asus\Desktop\Espacial_STATA\WQ.txt, replace
```

***h) Defina o Painel (i = MUNIC_PIO; t = ano) e atribua coordenadas espaciais às localidades**

```
xtset MUNIC_PIO ano
spset MUNIC_PIO, coord(x_coord y_coord) coordsys(latlong)
```

***i) Estimções em Painel Simples e Painel Espacial (EA=RE e EF=FE)**

```
xtreg y x1 x2 x3, re rob
estimates store PS_re
xtreg y x1 x2 x3, fe rob
estimates store PS_fe

spxtregress y x1 x2 x3, re sarpanel dvarlag(WQ) errorlag(WQ) ivarlag(WQ: x1 x2 x3)
estimates store PE_re2
spxtregress y x1 x2 x3, fe dvarlag(WQ) errorlag(WQ) ivarlag(WQ: x1 x2 x3)
estimates store PE_fe
```

***j) Teste de Hausman e tabelamento dos resultados**

```
hausman PS_fe PS_re, sigmamore
hausman PE_fe PE_re

estimates table PS_re PS_fe, b star(0.1 0.05 0.01) stats(N r2 aic bic rank) b(%7.2f)stfmt(%8.2f)
estimates table PE_re PE_fe, b star(0.1 0.05 0.01) stats(N r2 aic bic rank) b(%7.2f)stfmt(%8.2f)
```

***k) Estime os efeitos diretos, indiretos e totais do melhor modelo (Ex.: PE_fe = melhor)**

```
spxtregress y x1 x2 x3, fe dvarlag(WQ) errorlag(WQ) ivarlag(WQ: x1 x2 x3)
estat impact
```

FONTE: <https://www.slideshare.net/rcschuch/tutorial-de-econometria-espacial-utilizando-o-stata>