Carvalho, Alexandre Xavier Ywata; Albuquerque, Pedro Henrique Melo

Working Paper

Tópicos em econometria espacial para dados cross-section

Texto para Discussão, No. 1508

Provided in Cooperation with:
Institute of Applied Economic Research (ipea), Brasília

Suggested Citation: Carvalho, Alexandre Xavier Ywata; Albuquerque, Pedro Henrique Melo (2010) : Tópicos em econometria espacial para dados cross-section, Texto para Discussão, No. 1508, Instituto de Pesquisa Econômica Aplicada (IPEA), Brasília

This Version is available at:
https://hdl.handle.net/10419/91002

1508
TEXTO PARA DISCUSSÃO

TÓPICOS EM ECONOMETRIA ESPACIAL PARA DADOS *CROSS-SECTION*

Alexandre Xavier Ywata Carvalho
Pedro Henrique Melo Albuquerque

ipea

Instituto de Pesquisa Econômica Aplicada

1508

TEXTO PARA DISCUSSÃO

Brasília, agosto de 2010

TÓPICOS EM ECONOMETRIA ESPACIAL PARA DADOS *CROSS-SECTION*

Alexandre Xavier Ywata Carvalho*
Pedro Henrique Melo Albuquerque**

* Técnico de Planejamento e Pesquisa da Diretoria de Estudos e Políticas Regionais, Urbanas e Ambientais (Dirur) do Ipea. E-mail: alexandre.ywata@ipea.gov.br.

** Pesquisador do Programa de Pesquisa para o Desenvolvimento Nacional (PNPD) na Coordenação de Métodos Quantitativos da Dirur do Ipea e professor do departamento de administração da Universidade de Brasília (UnB).

Governo Federal

Secretaria de Assuntos Estratégicos da Presidência da República
Ministro Samuel Pinheiro Guimarães Neto

ipea **Instituto de Pesquisa Econômica Aplicada**

Fundação pública vinculada à Secretaria de Assuntos Estratégicos da Presidência da República, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiro – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

URL: http://www.ipea.gov.br
Ouvidoria: http://www.ipea.gov.br/ouvidoria

Texto para Discussão

Publicação cujo objetivo é divulgar resultados de estudos direta ou indiretamente desenvolvidos pelo Ipea, os quais, por sua relevância, levam informações para profissionais especializados e estabelecem um espaço para sugestões.

ISSN 1415-4765
JEL: C21, R15

SUMÁRIO

SINOPSE

Este texto apresenta uma discussão sobre diversos modelos econométricos para estimação de modelos paramétricos na presença de dependência espacial, com dados *cross-section*. O foco inicial são modelos de dependência espacial com *lags* espaciais da variável resposta ou *lags* espaciais do resíduo, com estimação dos parâmetros feita via máxima verossimilhança. Uma análise crítica destes modelos é apresentada em seguida, além de se discutirem testes para detectar presença de dependência espacial. Finalmente, discutem-se métodos de estimação mais robustos, os quais permitem a contabilização de endogeneidade em algumas das variáveis explicativas.

ABSTRACT[i]

This paper presents a discussion on several econometric models for estimating parametric models in the presence of spatial dependence with cross-section data. Initially, we cover models for spatial dependence with spatial lags of the response variable and spatial lags of the residues, and estimation is accomplished by maximum likelihood. A critical analysis for these models is also presented, followed by a discussion on tests for spatial dependence. Finally, we present a discussion no more robust estimation methods, allowing for endogeneity in some of the explanatory variables.

---

i. *The versions in English of the abstracts of this series have not been edited by Ipea's editorial department.*
As versões em língua inglesa das sinopses (*abstracts*) desta coleção não são objeto de revisão pelo Editorial do Ipea.

1 INTRODUÇÃO

Nas últimas décadas, um conjunto cada vez maior de ferramentas analíticas para tratamento de dados espaciais tem surgido na literatura especializada. Estas ferramentas têm auxiliado pesquisadores em diferentes campos da ciência a lidar com a crescente disponibilidade de bases de dados georreferenciados. De fato, diferentemente de séries temporais macroeconômicas, por exemplo, uma base de dados totalmente nova e detalhada, com dados *cross-section* espaciais, pode surgir de um ano para o outro. Além disso, o crescente desenvolvimento de dispositivos de coleta e armazenamento de dados geográficos tem contribuído para a construção de inúmeras bases de dados com componentes espaciais.

Apesar de todo o avanço ocorrido nas décadas recentes, ainda há um grande terreno a ser explorado em termos de ferramentas para dados geograficamente localizados. Os avanços esperados para os próximos anos têm a ver tanto com a formalização de resultados matemáticos, quanto com avanços mais conceituais sobre a aplicação dos modelos que vêm sendo utilizados até o presente momento. Uma discussão sobre tópicos de natureza mais conceitual pode ser encontrada, em Pinkse e Slade (2010), Holmes (2010), e McMillen (2010).

Holmes (2010) apresenta uma discussão interessante sobre os três tipos básicos de abordagem para estudos empíricos em análise de dados espaciais. As três abordagens discutidas são: *i*) abordagem estruturalista; *ii*) abordagem experimentalista; e *iii*) abordagem descritiva. Um entendimento destas três abordagens é importante, para que os pesquisadores possam identificar em quais das três um determinado trabalho empírico se situa, de forma que as vantagens e as limitações do trabalho fiquem mais claras.

Na abordagem estruturalista, o exercício empírico parte de um modelo econômico totalmente especificado, com base em uma teoria geralmente microfundamentada. O objetivo do exercício é estimar parâmetros estruturais do modelo (*deep model parameters*), relativos a preferências e/ou tecnologias. A partir do modelo estimado, é possível simular impactos de políticas, inclusive políticas que ainda não foram implementadas. Na literatura de organização industrial mais recente,[1] os modelos

1. Ver Berry, Levinsohn e Pakes (1995; 2004), Nevo (2001), Petrin (2002) e Ackerberg *et al.* (2007)

microfundamentados estimados permitem, por exemplo, avaliar *a priori* o impacto da fusão de duas empresas. Apesar de a abordagem estruturalista estar mais desenvolvida para pesquisas em organização industrial, pesquisadores em economia política (EPPLE e SIEG, 1999) e economia do trabalho (KEANE e WOLPIN, 1997; ECKSTEIN e WOLPIN, 1999) já começaram a utilizá-la.

A abordagem experimentalista surgiu inicialmente na literatura de economia do trabalho. Nesta abordagem, o interesse principal é a identificação do efeito causal de uma determinada política (efeito tratamento). Ao invés de se preocupar com a especificação de um modelo teórico, a ideia básica é encontrar experimentos naturais ou instrumentos válidos para a identificação de causalidade de políticas que já foram implementadas. Para maiores detalhes, o leitor pode recorrer a manuais como Angrist e Pischke (2009) ou Cameron e Trivedi (2005). Nesse contexto, métodos de estimação do tipo mínimos quadrados de dois estágios, ou de forma mais geral, métodos de momentos generalizados, têm um papel muito importante. Outro procedimento comumente empregado é a regressão de descontinuidade (HAHN, TODD e VAN DER KLAAUW, 2001).

Ao contrário das duas abordagens anteriores, a abordagem descritiva não tem por objetivo quantificar o efeito causal de determinadas políticas. Em geral, os artigos que utilizam a abordagem descritiva se iniciam com uma discussão da teoria econômica, que pode estar ou não embasada em modelos matematicamente fundamentados. A partir de regressões e outros indicadores estatísticos, os autores buscam encontrar evidências nas relações entre as variáveis que possam corroborar uma determinada teoria (possivelmente, em detrimento de teorias alternativas). As regressões em geral correspondem a formas reduzidas de equações estruturais mais completas. Uma das limitações desta abordagem é que, além de não permitir inferências causais, ela também está sujeita à crítica de Lucas. Dessa forma, alterações no regime econômico podem incorrer em alterações nos parâmetros do modelo, tornando a utilização dos modelos reduzidos menos críveis do ponto de vista de simulações *a priori* de impactos de políticas.[2]

A maioria dos estudos em economia regional e urbana segue a abordagem descritiva. Nos últimos anos, têm surgido estudos que utilizam a abordagem experimentalista

2. Ver Hendry (1995).

para avaliação de políticas. Por sua vez, a utilização da abordagem estruturalista pode trazer vários benefícios para a economia regional, dada a dificuldade de se encontrar bons instrumentos ou bons experimentos naturais. Uma das dificuldades na utilização da abordagem experimentalista em economia regional é a disponibilidade de dados (comparando-se ao número de observações de estudos em economia do trabalho, por exemplo). Uma sugestão para o uso da abordagem experimentalista em economia regional e urbana é a utilização de dados em nível de firmas, por exemplo, ao invés de dados em nível de municípios.

A utilização da abordagem estruturalista para economia regional ou urbana deve se iniciar com a construção de um modelo teórico (o que pode não ser tão fácil como no caso de modelos de organização industrial). Por seu turno, a utilização de abordagens estruturalistas em economia regional poderia ser interessante para simulações de políticas públicas. No entanto, pouco tem sido feito neste sentido até agora.

Neste trabalho, apresenta-se uma discussão sobre alguns dos modelos econométricos comumente utilizados para modelagem de dados espaciais. De maneira geral, os modelos apresentados estariam mais adequados para estudos empíricos seguindo as abordagens experimentalista e descritiva. De fato, os estimador de mínimos quadrados de dois estágios, de Kelejian e Prucha, e o estimador de método de momentos generalizado, de Conley, permitem a estimação de parâmetros na presença de variáveis endógenas do lado direito da equação, contabilizando e/ou corrigindo para a presença de autocorrelação espacial nos resíduos do modelo. Mesmo não tratando diretamente a abordagem estruturalista, as ideias apresentadas neste texto fornecerão ao leitor uma noção dos procedimentos para estimação com dados com presença de dependência espacial, o que poderá ser útil para a estimação de parâmetros estruturais em modelos microfundamentados.

Dado o grande avanço pelo qual a literatura em métodos estatísticos para dados espaciais tem passado nos últimos anos, não há interesse aqui em ser exaustivo em termos de metodologias discutidas. Pelo contrário, optou-se por apresentar apenas alguns dos métodos mais comumente utilizados, de forma a transmitir ao leitor uma ideia básica, mas elucidativa, sobre os fundamentos da estimação de modelos econométricos com dependência espacial. Nesse sentido, não serão tratados, por exemplo, dados de painel (vejam-se, entre outros, Elhorst, 2003; Druska e Horrace, 2004; Egger *et al.*, 2005), mas apenas dados *cross-section*. Além disso, a abordagem será predominantemente frequentista. Apesar da simpatia em relação aos métodos bayesianos – principalmente no

contexto de dados espaciais –, para não se estenderem demasiado os autores preferiram ater-se aos procedimentos frequentistas. O leitor poderá encontrar boas exposições em Banerjee, Carlin e Gelfand (2004) e Schabenberger e Gotway (2009).

Além desta introdução, este texto contém mais seis seções. Na seção 2, apresenta-se uma discussão sobre os modelos econométricos espaciais para dados *cross-section* provavelmente mais utilizados na literatura. Na seção 3, discutem-se algumas das críticas mais comuns aos modelos espaciais apresentados na seção 2. Na seção 4, são apresentados alguns dos testes mais utilizados para verificação da presença ou não de dependência espacial. As seções 5 e 6 discutem procedimentos de estimação para contabilizar para a presença de variáveis endógenas no lado direito da equação: a seção 5 apresenta o estimador espacial de mínimos quadrados de dois estágios, e a seção 6 apresenta o estimador de método de momentos generalizados, com correção para a presença de autocorrelação espacial. Comentários finais encontram-se na seção 7.

2 MODELOS PARAMÉTRICOS PARA DEPENDÊNCIA ESPACIAL

Nesta seção, será feita uma discussão de alguns dos modelos paramétricos comumente utilizados em econometria espacial. A discussão se limitará a regressões com dados *cross-section*.[3] Para modelos envolvendo dados de painel espacial, o leitor pode recorrer a Elhorst (2003), Druska e Horrace (2004), Egger, Pfaffermayr e Winner (2005).

2.1 MODELOS SAR

Um dos modelos mais comumente utilizados para modelagem de correlação espacial é o modelo autorregressivo espacial (*spatial autorregressive model*), ou simplesmente modelo SAR. A ideia dos modelos SAR é utilizar a mesma ideia dos modelos AR (autorregressivos) em séries temporais, por meio da incorporação de um termo de *lag* entre os regressores da equação. Na sua forma mais simples, o modelo SAR tem expressão:

3. Ver Anselin (1988), Anselin e Florax (2000), Anselin, Florax e Rey (2004), Lesage e Pace (2009), Lesage (1997 e 1999), e Pace e Barry (1997 e 1998).

$$y = \rho W y + \epsilon, \tag{1}$$

onde y é um vetor coluna, contendo n observações na amostra para a variável resposta y_i, o coeficiente escalar ρ corresponde ao parâmetro autorregressivo, o termo ϵ corresponde a um vetor coluna contendo os resíduos ϵ_i da equação. Por enquanto, considera-se que os resíduos ϵ_i são independentes e identicamente distribuídos, com distribuição normal, média zero e variância homogênea σ^2. Um dos componentes presentes em uma grande quantidade de modelos espaciais é a matriz W. Esta matriz é conhecida como matriz de vizinhança, e pode ser definida de diversas formas, o que traz críticas aos modelos espaciais utilizando W (muitos autores consideram as definições para W deveras arbitrárias; a este respeito, ver Pinkse e Slade, 2010).

Uma das formas mais comumente empregadas de definição da matriz W se dá por meio da identificação de vizinhos de primeira ordem. Considere-se que cada observação no vetor y esteja associada a um polígono e um sistema georreferenciado. Por exemplo, o vetor y pode corresponder a observações de uma determinada variável observada para cada município brasileiro, ou corresponder a observações de uma variável para cada setor censitário na cidade de São Paulo. Neste caso, o elemento $W_{i,j}$ da matriz W assume valor $W_{i,j} = \mathbf{1}$, caso os polígonos i e j sejam vizinhos, e $W_{i,j} = \mathbf{0}$, caso i e j não sejam vizinhos. A diagonal principal de W possui todos os elementos iguais a zero, por definição.

Para identificar polígonos (municípios, setores censitários etc.) vizinhos, pode-se considerar uma vizinhança do tipo *queen*, quando os dois polígonos possuem pelo menos um vértice em comum, ou pode-se considerar uma vizinhança do tipo *rook*, quando os polígonos possuem pelos menos um lado inteiro em comum. Note-se que a vizinhança do tipo *queen* é menos restritiva que a vizinhança do tipo *rook*. Além da vizinhança de primeira ordem, podem-se utilizar vizinhanças de ordem maior. Na definição de vizinhança de segunda ordem, por exemplo, os polígonos i e j são vizinhos caso exista um outro polígono k, para o qual i e k sejam vizinhos de primeira ordem, e j e k também sejam vizinhos de primeira ordem. [4]

4. Ver Lesage e Pace (2009).

A matriz W , com elementos 0 ou 1, é conhecida como matriz de vizinhança não normalizada, em contraposição à matriz W^* normalizada. A matriz W^* normalizada é construída a partir da matriz W original (não normalizada), dividindo-se todos os elementos de cada linha de W pela soma da linha. Portanto, a matriz W^* possui todas as linhas com soma igual a 1. Por sua vez, a matriz W original é simétrica, o que não vale para a matriz W^*. O vetor $y_W = Wy$ é conhecido como *lag* espacial. No caso de se utilizar a matriz de contiguidade normalizada, o vetor $y_W = W^*y$ corresponde a um vetor de médias simples das observações para a variável y dos vizinhos. A partir de agora, a matriz de contiguidade será referida simplesmente como W , independentemente de ser uma matriz normalizada ou não normalizada.

O modelo paramétrico em (1) contém, como parâmetros desconhecidos, o coeficiente ρ e a variância σ^2. A estimação do parâmetro ρ permite, por exemplo, inferir o grau de correlação espacial entre as observações y_i. Além disso, testando-se a significância do parâmetro ρ, tem-se um procedimento para inferir a presença ou não de dependência espacial entre as observações. A seguir, se discutirá o processo de inferência dos parâmetros do modelo em (1).

Uma das primeiras sugestões para a estimação do coeficiente ρ é a utilização do estimador de mínimos quadrados ordinários. No entanto, quando o vetor de covariáveis (variáveis do lado direito da equação) é correlacionado com o resíduo da regressão, sabe-se que o estimador de mínimos quadrados ordinários é inconsistente. Esta correlação entre os resíduos e o regressor é observada no modelo em (1).[5] Portanto, estimação via mínimos quadrados ordinários resultaria em uma estimativa inconsistente para o coeficiente ρ.

Como alternativa, o analista pode utilizar estimação via máxima verossimilhança, que não sofre do problema de inconsistência do estimador de mínimos quadrados ordinários, devido à endogeneidade do regressor Wy. Em linhas gerais, a estimação via máxima verossimilhança dos parâmetros ρ e σ^2 parte da distribuição normal multivariada para o vetor de resíduos ϵ. A partir de (1), pode-se escrever

5. Ver Anselin (1988) e Lesage e Pace (2009).

$$y = (I - \rho W)^{-1}\epsilon, \tag{2}$$

onde I é uma matriz identidade com dimensão n. Dado que ϵ possui distribuição normal multivariada, com média nula e covariância $\sigma^2 I$, então o vetor observado y possui distribuição normal multivariada com média nula e covariância $\Sigma_y = \sigma^2(I - \rho W)^{-1}[(I - \rho W)]^{-1}]^T$. A partir desta matriz de covariância, pode-se escrever a função de log-verossimilhança $\mathbf{l}(\boldsymbol{\rho}, \boldsymbol{\sigma^2}) = \log L(\rho, \sigma^2)$. Maximizando-se $\log \mathbf{L}(\rho, \sigma^2)$, obtêm-se os estimadores de máxima verossimilhança dos parâmetros do modelo.

Uma das dificuldades na estimação de modelos SAR (mesmo no caso mais simples, no qual não há covariáveis exógenas) é a necessidade de se realizarem operações com matrizes de grandes dimensões. No processo iterativo para obtenção do máximo da função $\log \mathbf{L}(\rho, \sigma^2)$, é preciso calcular o logaritmo do determinante da matriz $(I - \rho W)$, que possui dimensão n. Se o analista estiver fazendo uma aplicação com observações de setores censitários da cidade de São Paulo, por exemplo, o valor de n está em torno de 18 mil; portanto, a matriz $(I - \rho W)$ possui dimensão 18 mil por 18 mil. Felizmente, pela própria definição da matriz de contiguidade W, pode-se tratá-la como matriz esparsa; ou seja, a grande maioria dos elementos de W são nulos. Para matrizes esparsas, existe uma literatura bem desenvolvida sobre algoritmos que tornam o processo computacional mais eficiente.[6] Portanto, apesar de a codificação do estimador de máxima verossimilhança não ser trivial (é preciso programar algumas rotinas para matrizes esparsas), o esforço computacional pode ser bastante reduzido.

Uma vez dentro do arcabouço de estimação via máxima verossimilhança, pode-se recorrer a vários dos resultados para este tipo de estimador. Pode-se, então, testar a significância do parâmetro ρ, utilizando-se o teste de Wald, o teste da razão de verossimilhança ou o teste dos multiplicadores de Lagrange. Testando-se a significância do parâmetro ρ, se está implicitamente testando a presença de dependência espacial das observações para a variável y_i.

6. Ver Davis (2006).

O modelo SAR em (1) pode ser estendido, para incorporar variáveis exógenas no lado direito da equação, obtendo-se

$$y = \rho W y + X\beta + \epsilon, \tag{3}$$

onde a matriz X é uma matriz contendo as observações das variáveis exógenas. A dimensão de X é $n \times k$, sendo k o número de regressores. Cada linha da matriz X corresponde a uma observação na base de dados (um polígono, em um sistema georreferenciado). No caso de a regressão incluir um intercepto, a primeira coluna da matriz X possui apenas valores 1. O vetor β é um vetor coluna de coeficientes para as variáveis exógenas, e possui dimensão $k \times 1$. O modelo em (3) é conhecido como modelo SAR misto.

Da mesma forma que no SAR simples (equação (1)), a estimação dos parâmetros no modelo SAR misto via mínimos quadrados ordinários também produz estimativas inconsistentes, uma vez que o vetor de *lags* espaciais Wy é correlacionado com o vetor de resíduos ϵ. Novamente, pode-se utilizar máxima verossimilhança, a partir da hipótese de que o vetor de resíduos ϵ possui distribuição normal multivariada com média nula e covariância $\sigma^2 I$. Pode-se então escrever

$$y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1}\epsilon, \tag{4}$$

e o vetor de variáveis observadas y possui distribuição (condicional a X) normal multivariada, com média condicional

$$E[y|X] = (I - \rho W)^{-1} X\beta \text{ ,} \tag{5}$$

e matriz de variância condicional

$$\Sigma_{(y|X)} = \sigma^2 (I - \rho W)^{-1} [[(I - \rho W)]^{-1}]^T \text{ .} \tag{6}$$

A partir da distribuição de y, obtém-se a função de log-verossimilhança condicional $\log L(\rho, \beta, \sigma^2)$. Maximizando-se a função de log-verossimilhança em relação aos parâmetros do modelo, encontram-se as estimativas para os coeficientes e para a variância dos resíduos. Para uma discussão sobre o processo iterativo para estimação dos parâmetros do modelo SAR misto, podem-se consultar Anselin (1988) e Lesage e Pace (2009).

2.2 MODELOS SEM

Da mesma forma que os modelos SAR partem da especificação de modelos AR para séries temporais, uma outra classe de modelos espaciais parte da especificação de modelos MA (médias móveis) para observações no tempo. Estes modelos espaciais são denominados modelos de erros espaciais (*spatial error models*), ou simplesmente SEM. Os modelos SEM possuem a seguinte especificação:

$$y = X\beta + u, \tag{7}$$

No caso, os resíduos da equação observada possuem uma estrutura autorregressiva, da forma

$$u = \lambda W u + \epsilon, \tag{8}$$

O vetor de resíduos ϵ possui distribuição normal multivariada, com média nula e matriz de covariância $\sigma^2 I$. O coeficiente escalar λ indica a intensidade da autocorrelação espacial entre os resíduos da equação observada. Note-se que, ao contrário dos modelos SAR, os modelos SEM não apresentam a variável resposta como uma função direta dos seus *lags* espaciais. A autocorrelação espacial nos modelos SEM aparece nos termos de erro.

Outra diferença dos modelos SEM em relação aos modelos SAR é que os coeficientes no vetor β podem ser estimados consistentemente via mínimos quadrados ordinários. De fato, a regressão em (7) pode ser vista como uma regressão linear com resíduos correlacionados. O estimador de mínimos quadrados ordinários produz estimativas consistentes, mas a matriz de covariância das estimativas $\hat{\beta}_{ols}$ não será mais $\sigma^2[X'X]^{-1}$. Devido aos erros correlacionados, a matriz de covariância de $\hat{\beta}_{ols}$ é dada por[7]

$$Var[\hat{\beta}_{ols}] = [X'X][X'\Omega^{-1}X]^{-1}[X'X], \tag{9}$$

7. Ao longo deste texto, a expressão da forma A' denotará o transposto do elemento em A, onde A é uma matriz, um vetor coluna, um vetor linha, ou mesmo um escalar.

onde $\Omega = Var[u] = \sigma^2(I - \lambda W)^{-1}[(I - \lambda W)]^{-1}]^T$. Note-se que a matriz Ω depende do coeficiente λ e da variância σ^2. A estimativa destes dois parâmetros pode ser obtida consistentemente a partir da estimação de um modelo SAR via máxima verossimilhança, conforme discutido no item anterior, para os resíduos $\hat{u} = y - X\hat{\beta}_{ols}$. Uma vez estimados os escalares λ e σ^2, pode-se obter uma estimativa para a matriz de covariância de $\hat{\beta}_{ols}$

$$\widehat{Var[\hat{\beta}_{ols}]} = [X'X]\left[X'\hat{\Omega}^{-1}X\right]^{-1}[X'X], \tag{10}$$

onde $\hat{\Omega} = \hat{\sigma}^2\left(I - \hat{\lambda}W\right)^{-1}[\left[\left(I - \hat{\lambda}W\right)]^{-1}\right]^T$.

Sabe-se que, no caso de modelos lineares com regressores exógenos (que é o caso nos modelos SEM), com resíduos correlacionados, o estimador de mínimos quadrados ordinários é consistente, mas não é eficiente, havendo outros estimadores lineares que produzem variâncias menores.[8] Especificamente para o modelo SEM, o estimador linear com variância mínima é o estimador de mínimos quadrados generalizados (*generalized least squares* – GLS), dado por

$$\hat{\beta}_{gls} = [X'\Omega^{-1}X]^{-1}[X'\Omega^{-1}y], \tag{11}$$

Na prática, não se conhece a matriz Ω, uma vez que esta depende dos parâmetros desconhecidos λ e σ^2. Utiliza-se então o estimador de mínimos quadrados generalizados executável (*feasible generalized least squares* – FGLS), com expressão

$$\hat{\beta}_{fgls} = \left[X'\hat{\Omega}^{-1}X\right]^{-1}\left[X'\hat{\Omega}^{-1}y\right], \tag{12}$$

onde $\hat{\Omega} = \hat{\sigma}^2\left(I - \hat{\lambda}W\right)^{-1}[\left[\left(I - \hat{\lambda}W\right)]^{-1}\right]^T$, com $\hat{\sigma}^2$ e $\hat{\lambda}$ estimativas via máxima verossimilhança do modelo SAR simples, a partir dos resíduos $\hat{u} = y - X\hat{\beta}_{ols}$. Portanto, uma alternativa para a estimação dos parâmetros do modelo SEM é dada pelos passos:

8. Quando os autores se referem a variâncias menores, na verdade referem-se ao fato de que a diferença $Var[\hat{\beta}_{ols}] - Var[\hat{\beta}]$ é uma matriz positiva definida, onde $\hat{\beta}$ é um estimador linear mais eficiente do que o estimador de mínimos quadrados ordinários.

i) Obter a estimativa de mínimos quadrados ordinários $\hat{\beta}_{ols} = [X'X]^{-1}[X'y]$.

ii) Calcular os resíduos $\hat{u} = y - X\hat{\beta}_{ols}$.

iii) Estimar os parâmetros λ e σ^2, via máxima verossimilhança, para o modelo SAR em $\hat{u}$, $\hat{u} = \lambda W\hat{u} + \epsilon$.

iv) Calcular a estimativa $\hat{\Omega} = \hat{\sigma}^2 \left(I - \hat{\lambda}W\right)^{-1} [\![\left(I - \hat{\lambda}W\right)]\!]^{-1}\Big]^T$.

v) Obter a estimativa $\hat{\beta}_{fgls} = \left[X'\hat{\Omega}^{-1}X\right]^{-1}\left[X'\hat{\Omega}^{-1}y\right]$.

vi) Obter a estimativa para a covariância $\hat{\beta}_{fgls}$, $[\![\widehat{Var}\left[\hat{\beta}\right]_{fgls}\Big] = \left[X'\hat{\Omega}^{-1}X\right]^{-1}$.

Inferência para os coeficientes em β pode ser efetuada a partir da matriz $\left[X'\hat{\Omega}^{-1}X\right]^{-1}$. Note-se que a estimativa final para o vetor β não precisa parar no passo (*v*) acima. De fato, uma vez obtida uma estimativa $\hat{\beta}_{fgls}$, pode-se obter um novo vetor $\hat{u} = y - X\hat{\beta}_{fgls}$. Para este novo vetor $\hat{u}$, estimam-se novamente os parâmetros λ e σ^2, repetindo-se em seguida os passos (*iv*) e (*v*). Este processo pode ser efetuado repetidamente até que os valores no vetor $\hat{\beta}_{fgls}$ atinjam a convergência. Finalizam-se então as estimações com o passo (*vi*).

Além das estimativas via mínimos quadrados ordinários (com correção da matriz de covariância das estimativas dos coeficientes) e das estimativas via mínimos quadrados generalizados efetuáveis (FGLS), a literatura apresenta uma discussão sobre estimação dos parâmetros do modelo SEM via máxima verossimilhança. Combinando as expressões (7) e (8), obtém-se

$$y = X\beta + (I - \lambda W)^{-1}\epsilon, \tag{13}$$

onde ϵ possui distribuição normal multivariada com média nula e covariância $\sigma^2 I$. Portanto, o vetor de variável resposta y possui distribuição normal multivariada com média condicional

$$E[y|X] = X\beta\ , \tag{14}$$

e matriz de variância condicional

$$\Sigma_{(y|X)} = \sigma^2(I - \lambda W)^{-1}[\![(I - \lambda W)]\!]^{-1}\Big]^T. \tag{15}$$

A partir da distribuição de y, obtém-se a função de log-verossimilhança condicional $\log L(\lambda, \beta, \sigma^2)$. Maximizando-se a função de log-verossimilhança em relação aos parâmetros do modelo, encontram-se as estimativas para os coeficientes e para a variância dos resíduos. Para uma discussão sobre o processo iterativo para estimação dos parâmetros do modelo SEM, consultem-se Anselin (1988) e Lesage e Pace (2009). Similarmente às estimações no caso de modelos SAR, a estimação de modelos SEM também envolve operações com matrizes esparsas. Novamente, utilizando-se rotinas mais eficientes para matrizes esparsas, o esforço computacional pode ser bem menor.

2.3 MODELOS SARMA

Finalmente, os modelos SEM e SAR podem ser combinados em uma especificação mais geral, seguindo a ideia nos modelos ARMA (*autorregressive and moving average*) para séries temporais. Os modelos SARMA (*spatial autorregressive and moving average*) têm uma especificação da forma

$$y = \rho W_1 y + X\beta + u, \tag{16}$$

na qual os resíduos da equação observada possuem uma estrutura autorregressiva, da forma

$$u = \lambda W_2 u + \epsilon. \tag{17}$$

As matrizes W_1 e W_2 são matrizes de contiguidade não necessariamente iguais. De fato, quando $W_1 = W_2$, o modelo é não identificado, e as estimativas para os coeficientes λ e ρ podem resultar bastante instáveis,[9] a menos que a matriz de delineamento X contenha pelo menos uma variável exógena além do intercepto. Uma das críticas em relação à utilização dos modelos SARMA é justamente o fato de eles exigirem, em alguns casos, a especificação de duas matrizes de contiguidade diferentes. Em geral, a escolha de uma matriz de contiguidade é arbitrária; a escolha de duas matrizes diferentes implica um grau de arbitrariedade ainda mais criticável.

9. Ver Anselin (1988) e Lesage e Pace (2009).

Estimação dos parâmetros do modelo SARMA pode ser feita via máxima verossimilhança. A partir das expressões (16) e (17), pode-se escrever

$$(I - \rho W_1)y = X\beta + (I - \lambda W_2)^{-1}\epsilon$$

$$\Rightarrow y = (I - \rho W_1)^{-1}X\beta + (I - \rho W_1)^{-1}(I - \lambda W_2)^{-1}\epsilon \ .$$

Assumindo-se que ϵ possui distribuição normal multivariada, com média zero e covariância $\sigma^2 I$, conclui-se que o vetor de observações para a variável resposta y possui distribuição normal multivariada com média condicional

$$E[y|X] = (I - \rho W_1)^{-1}X\beta \ , \tag{18}$$

e matriz de variância condicional

$$\Sigma_{(y|x)} = \sigma^2(I - \rho W_1)^{-1}(I - \lambda W_2)^{-1}\left[(I - \rho W_1)^{-1}(I - \lambda W_2)^{-1}\right]^T . \tag{19}$$

Utilizando-se a fórmula para a distribuição normal multivariada, pode-se chegar à função de log-verossimilhança $\log \mathbf{L}(\lambda, \rho, \beta, \sigma^2)$, como função dos parâmetros desconhecidos do modelo. Similarmente aos modelos SAR e SEM, as estimativas de máxima verossimilhança não possuem fórmula fechada, necessitando de um processo iterativo para maximização da função $\log \mathbf{L}(\lambda, \rho, \beta, \sigma^2)$. Uma discussão sobre os passos no processo iterativo para estimação dos parâmetros no modelo SARMA pode ser encontrada em Anselin (1988) e Lesage e Pace (2009).

3 CRÍTICAS AOS MODELOS DE DEPENDÊNCIA ESPACIAL

Apesar do seu uso bastante disseminado, os modelos paramétricos para tratamento de dependência espacial (exemplos: SAR, SEM e SARMA) vêm recebendo várias críticas na literatura. Estas críticas não necessariamente retiram destes modelos quaisquer utilidades em pesquisas empíricas. No entanto, alguns dos pontos levantados pelos críticos são importantes para: *i*) antecipar aos usuários alguns cuidados e limitações acerca dos quais eles devem estar cientes; *ii*) fornecer um certo balizamento para pesquisas futuras para os modelos espaciais, de maneira a corrigir ou amenizar algumas das limitações. Nesta seção, será feita uma discussão sobre algumas das críticas aos

modelos apresentados na seção 3 (e seus equivalentes para dados de painel). Estas críticas se aplicam mais fortemente ao problema de especificação paramétrica (ou não) para capturar corretamente a dependência espacial. No caso de testes de hipótese para presença ou não de dependência espacial, os testes atualmente disponíveis (conforme seção 4) se comportam de forma bastante satisfatória. Maiores detalhes podem ser encontrados, em Pinkse e Slade (2010).

De maneira geral, o embasamento teórico para a modelagem em econometria espacial ainda se encontra em um estágio inicial. Dessa forma, uma das dificuldades é encontrar um modelo que se adeque a todos os tipos de situação. Nesse sentido, alguns autores defendem que os pesquisadores se concentrem no desenvolvimento de teorias específicas para classes particulares de aplicações, ao invés de seguirem na busca de extensões para técnicas já existentes.

Entre as limitações para os modelos de SAR e outros modelos da forma ARMA espaciais (incluindo extensões para dados de painel), podem-se citar os itens a seguir.

i) Hipótese improvável e desnecessária de normalidade dos resíduos.

ii) O fato de y_i depender dos seus próprios *lags* espaciais pode implicar que y_i também dependa dos *lags* espaciais do vetor de covariáveis x_i, incorrendo no problema de reflexão (*reflexion problem*), apontado por Manski (1993). A consequência prática é que a inclusão de *lags* espaciais de x_i pode ocasionar uma matriz de design X com altíssimo grau de multicolinearidade.

iii) Os modelos SAR e demais modelos ARMA assumem relações lineares entre os regressores e a variável resposta y_i. Este fato nem sempre é verdade na prática, e pode haver a necessidade de especificações não lineares da relação entre o vetor de regressões x_i e a variável y_i.

iv) Os modelos SAR e correlatos não levam em consideração a presença de dependência entre o vetor de regressores x_i e o resíduo u_i, causada pela presença de regressores endógenos em x_i e/ou pela presença de heteroscedasticidade.

v) Há fortes críticas à representação excessivamente simplista de toda a dependência espacial em um único coeficiente ρ.

vi) A matriz de contiguidade W implica um alto grau de arbitrariedade na sua especificação, principalmente levando-se em consideração a irregularidade dos mapas de municípios e de setores censitários.

De maneira geral, os modelos SAR e correlatos foram inicialmente propostos como possíveis extensões dos modelos para dependência em séries temporais. No entanto, há uma série de críticas à analogia dos procedimentos para dependência espacial com os procedimentos para dependência temporal. Algumas destas críticas estão listadas a seguir.

a) A hipótese de estacionariedade, diferentemente de diversas aplicações em séries temporais, não é válida para o caso espacial.

b) Os dados não são igualmente espaçados.

c) A presença de observações ausentes (*missing values*) pode incorrer na presença de endogeneidade, ocasionando vieses nos estimadores de máxima verossimilhança.

d) Observações espaciais, em muitos casos, são agregações de observações (por polígono, por exemplo) do comportamento de vários agentes. Portanto, modelos baseados no comportamento de agentes individuais podem não ser mais válidos.

e) Nos modelos para séries temporais, os procedimentos são teoricamente validados a partir de proposições sobre o comportamento assintótico dos estimadores, quando o número de observações T (intervalo total da série histórica) assume valores cada vez maiores ($T \to \infty$). Para modelos para dados espaciais, não é claro se a expansão assintótica ocorre com o aumento da densidade de observações dentro do mapa (*infill asymptotics*), ocorre com o aumento das fronteiras (*increasing domain asymptotics*), ou ocorre com as suas expansões simultaneamente.

f) O item anterior é particularmente importante, porque não há garantia de que as relações de dependência espacial se alteram quando mais observações são adicionadas aos dados. Por exemplo, no caso de *infill asymptotics*, a adição de novas observações pode ocasionar um aumento da dependência espacial, uma vez que as observações estarão cada vez mais próximas em média.

g) Diferentemente dos modelos para séries temporais, a estimação dos modelos com dados espaciais pode sofrer do grave problema de endogeneidade das decisões locacionais das unidades observadas na amostra. Uma consequência da endogeneidade das localizações é que as distâncias entres os agentes, bem como as estruturas de vizinhança, também são endógenas. Este problema tem se mostrado de difícil solução até o momento, e vem sendo desprezado na maioria das aplicações.

Diversos artigos recentes têm focalizado alguns dos problemas discutidos anteriormente. Para adicionar maior flexibilidade à modelagem da vizinhança, por exemplo, algumas extensões do modelo SAR tradicional consistem em substituir a matriz de contiguidade W por uma expansão de funções base, da forma

$$y = \left[\sum_{k=0}^{\infty} \rho_k W_k y\right] + X\beta + u. \tag{20}$$

Na prática, é necessário truncar o número de elementos no somatório da expressão (20), até um número $K(n)$. Como é típico em estimações com expansões de funções base, faz-se $K(n)$ aumentar para o infinito, quando o tamanho n da amostra aumenta. Neste caso, a expressão torna-se

$$y = \left[\sum_{k=0}^{K(n)} \rho_k W_k y\right] + X\beta + u, \tag{21}$$

e o problema de rigidez em relação à forma funcional da dependência espacial pode ser amenizado (para maiores detalhes, ver Pinkse, Slade e Bret, 2002; Pinkse e Slade, 2004; e Pofahl, 2007).

Boa parte dos problemas de endogeneidade pode ser tratada com a utilização de variáveis instrumentais apropriadas, conforme discutido nas seções 5 e 6. Para o problema de observações ausentes (*missing data*), no qual o processo gerador das observações ausentes é exógeno, podem-se utilizar procedimentos de mínimos quadrados de dois estágios (LEE, 2007). Para situações nas quais a geração das observações ausentes é endógena, não há solução conhecida na literatura. De maneira geral, ainda existe um grande caminho a ser trilhado em termos de procedimentos e tratamentos teóricos, para lidar com os problemas nos modelos para dados espaciais.

4 TESTES PARA DEPENDÊNCIA ESPACIAL

Na seção anterior, foram discutidos alguns modelos mais comumente utilizados para contabilizar para a presença de dependência espacial nos resíduos (ou na própria variável resposta) do modelo de regressão. Nesta seção, será apresentada uma discussão sobre testes para dependência espacial. De maneira geral, os modelos paramétricos apresentados na seção 2 têm sofrido diversas críticas, conforme será visto na seção 4. Por seu turno, os testes para a presença de dependência espacial não sofrem o mesmo ataque, e são relativamente bem aceitos na literatura.

4.1 ESTATÍSTICA DE MORAN

Uma das estatísticas para testes de dependência espacial mais disseminadas é a estatística *I* de Moran. Esta estatística pode ser aplicada à variável y_i diretamente, ou aos resíduos da regressão de y_i *versus* um conjunto de variáveis explicativas. Considere-se então um modelo de regressão linear, da forma

$$y = X\beta + u, \tag{22}$$

onde y é um vetor coluna $(n \times 1)$ de variáveis resposta, X é uma matriz com cada linha contendo as observações para as variáveis explicativas, β é um vetor de coeficientes e u é um vetor coluna contendo os resíduos da regressão. A partir da estimativa de mínimos quadrados ordinários para o vetor de coeficientes, obtém-se a seguinte expressão para os resíduos

$$\hat{u} = y - X[X'X]^{-1}[X'y]. \tag{23}$$

A estatística *I* de *Moran* para a autocorrelação espacial pode ser aplicada nos resíduos do modelo de regressão de maneira direta. Formalmente, a estatística *I* é dada por

$$I = \frac{n}{s}\left[\frac{\hat{u}'W\hat{u}}{\hat{u}'\hat{u}}\right], \tag{24}$$

onde $\hat{u}$ é o vetor de resíduos da regressão por mínimos quadrados ordinários, W é a matriz de contiguidade espacial, n é o número de observações da amostra e s é um fator de padronização igual à soma de todos os elementos da matriz W. A partir da estatística *I*, pode-se construir um teste para a hipótese nula de presença de independência espacial. Por sua vez, a especificação da hipótese alternativa não é tão simples.

A distribuição assintótica para a estatística *I* foi derivada por Cliff e Ord (1972). Dessa forma, considere-se

$$z_I = \frac{I - E(I)}{\sqrt{V(I)}}, \tag{25}$$

onde $E(I)$ e $\sqrt{V(I)}$ são respectivamente a média e a variância assintótica da estatística *I* de *Moran*. Sob a hipótese nula, a distribuição da estatística z_I pode ser estimada via simulações de Monte Carlo. Quando a estatística I é construída a partir dos resíduos $\hat{u}$, a rejeição da hipótese nula implica em evidências de que há autocorrelação espacial no modelo de regressão. A partir daí, o analista pode recorrer a um dos modelos paramétricos discutidos na seção 2, na seção 4 ou na seção 5.

4.2 TESTE DE KELEJIAN-ROBINSON

Kelejian e Robinson (1992) propuseram um teste com o mesmo objetivo do teste *I* de *Moran*. No entanto, diferentemente do teste *I* de *Moran*, o teste de *Kelejian-Robinson* não pressupõe normalidade da variável sendo testada (a variável observada y_i ou os resíduos $\hat{u}_i$ da regressão). Portanto, o teste de Kelejian-Robinson é mais robusto à não normalidade dos resíduos ou da variável observada, sendo mais apropriado quando a hipótese similaridade ao padrão gaussiano seja questionável.

O teste de Kelejian-Robinson tem como pressuposto inicial

$$Cov(\epsilon_i, \epsilon_j) = \sigma_{ij} = Z_{ij}\alpha \ , \tag{26}$$

onde Z_{ij} é um vetor $1 \times q$ de covariáveis, tipicamente tomadas como funções das variáveis explicativas originais para i e j, com i e j sendo localidades "contíguas" em um espaço geral de observações ordenadas. Por exemplo, Z_{ij} pode ser construído a partir de produtos cruzados dos elementos de X_i e X_j. O vetor Z_{ij} não necessariamente possui a mesma dimensão de X_i (ou X_j). O elemento α é um vetor $q \times \mathbf{1}$ de parâmetros, indicando o quanto os componentes de Z_{ij} podem explicar a covariância entre os resíduos. Intuitivamente, a ausência de autocorrelação espacial poderá não produzir relações significativas entre $Cov(\epsilon_i, \epsilon_j)$ e Z_{ij}, resultando em estimativas não significantes para os coeficientes no vetor α.

A hipótese nula é então construída como $H_0: \alpha = \mathbf{0}$ em (24). Dada uma amostra de tamanho n, seja c um vetor de dimensões $h_n \times \mathbf{1}$, contendo as covariâncias σ_{ij}'s

não nulas[10] (por construção) para todo $i < j$. O teste é implementado regredindo-se os h_n produtos cruzados $\hat{c}_{ij} = \hat{u}_i \hat{u}_j$ dos resíduos *versus* os vetores Z_{ij}, para todo $i < j$, com i e j polígonos vizinhos. Seja então a matriz Z, com dimensão $h_n \times q$, construída a partir do empilhamento dos vetores linha Z_{ij}, e seja $\hat{c}$ um vetor coluna, com dimensão $h_n \times 1$, construído a partir do empilhamento dos valores de $\hat{c}_{ij} = \hat{u}_i \hat{u}_j$. Uma estimativa para α pode ser obtida via mínimos quadrados ordinários, resultando em

$$\hat{\alpha} = (Z'Z)^{-1} Z'\hat{c}.$$

A partir da estimativa $\hat{\alpha}$, pode-se construir a estatística teste de Kelejian-Robinson, dada pela expressão

$$KR = \frac{\hat{\alpha}' Z' Z \hat{\alpha}}{\hat{\sigma}^4}, \tag{27}$$

onde $\hat{\sigma}^4$ é um estimador consistente de σ^4, e σ^2 é a variância para o resíduo da regressão de $\hat{c}_{ij} = \hat{u}_i \hat{u}_j$ *versus* Z_{ij}. Uma estimativa para σ^4 pode ser dada, por exemplo, por

$$\hat{\sigma}^4 = \left[\frac{(\hat{c} - Z\hat{\alpha})'(\hat{c} - Z\hat{\alpha})}{h_n}\right]^2.$$

Sob a hipótese nula, tem-se que $\frac{\hat{c}'\hat{c}}{h_n}$ converge em probabilidade para σ^2. Pode-se mostrar então que uma forma alternativa para a estatística teste é dada por

$$KR = h_n^2 \frac{c' Z (Z'Z)^{-1} Z' c}{[c'c]^2}. \tag{28}$$

Sob a hipótese nula de ausência de dependência espacial, a estatística KR possui distribuição assintótica qui-quadrada, com q graus de liberdade.

10. Nesse caso, as covariâncias não nulas são aquelas para as quais os polígonos i e j são vizinhos, de acordo com a definição de vizinhança utilizada para a análise.

4.3 TESTES ASSINTÓTICOS A PARTIR DE ESPECIFICAÇÕES PARAMÉTRICAS

Nas seções 3.1 e 3.2, foram discutidos dois procedimentos de testes estatísticos para presença de dependência espacial, os quais não dependem de uma especificação paramétrica para a forma de autocorrelação no espaço. Nesta seção, serão revisitados os modelos discutidos na seção 2, para se construírem outros procedimentos de testes, a partir de especificações paramétricas. De forma geral, os procedimentos discutidos são obtidos a partir de três metodologias tradicionais, empregadas para testes de hipóteses em geral. Estas metodologias são:

i) teste de Wald;

ii) teste da razão de verossimilhança (*likelihood ratio – LR*); e

iii) teste dos multiplicadores de Lagrange (*Lagrange multipliers – LM*).

4.3.1 Princípios gerais

Os testes de Wald, LR e LM são baseados nas propriedades dos estimadores de máxima verossimilhança.[11] Mais especificamente, estas propriedades partem do pressuposto de normalidade assintótica dos estimadores. Formalmente, seja θ um vetor de parâmetros e $\hat{\theta}$ suas respectivas estimativas por máxima verossimilhança, satisfazendo a convergência em distribuição

$$\sqrt{n}\left(\hat{\theta} - \theta_0\right) \xrightarrow{L} N(0, [I_1(\theta_0)]^{-1}),$$

onde θ_0 é o valor real do parâmetro no modelo (assumindo um modelo corretamente especificado), o elemento $I_1(\theta_0)$ é a matriz de informação de Fisher para uma observação, e n é o número de observações na amostra. Considere-se então que o conjunto de hipóteses, sobre os parâmetros do modelo a serem testadas, pode ser escrito da forma

$$\begin{cases} H_0: g(\theta_0) = 0 \\ H_a: g(\theta_0) \neq 0 \end{cases}$$

11. O teste de Wald pode ser utilizado em outros contextos que não o de estimação via máxima verossimilhança.

onde g , $g:\Re^k \mapsto \Re^q$, é uma função linear ou não linear do vetor de parâmetros $\theta \in \Theta \subset \Re^k$. Considerem-se, por exemplo, os modelos SAR ou SEM, vistos na seção 2. Como casos especiais de testes de hipóteses para os modelos paramétricos, têm-se os testes individuais dos parâmetros de autocorrelação espacial: $H_0: \rho = 0$ no modelo SAR, ou $H_0: \lambda = 0$ no modelo SEM.

Os testes de Wald, LR e LM são baseados nas distâncias das estimativas para o modelo irrestrito e as estimativas satisfazendo às restrições impostas pela hipótese nula. Por exemplo, se o vetor de parâmetros θ é particionado em dois vetores distintos, da forma $\theta' = [\theta_1', \theta_2']$, e a hipótese nula pode ser escrita da forma $H_0: \theta_1 = 0$, a estimativa $\hat{\theta}_r$ de θ no modelo restrito consistirá das estimativas para θ_2 concatenada com todos os elementos de θ_1 iguais a zero. A estimativa irrestrita $\hat{\theta}_u$ é a estimativa do vetor completo θ . Os testes serão então baseados na medida da diferença entre as estimativas do modelo completo $\hat{\theta}_u$ e o vetor restrito $\hat{\theta}_r$. Intuitivamente, se a distância entre os dois resultados é muito grande, a hipótese nula é rejeitada.

Para a realização dos testes é necessário estimar:

i) Wald: apenas o modelo completo (irrestrito);

ii) RV: o modelo completo (irrestrito) e o modelo restrito (sob a hipótese nula); e

iii) LM: apenas o modelo restrito (sob a hipótese nula).

A seguir se fará uma discussão um pouco mais detalhada dos três tipos de testes. Dadas certas condições de regularidade, e assumindo-se que a hipótese nula é verdadeira, as estatísticas testes comumente empregadas para os três procedimentos possuem distribuição assintótica qui-quadrada χ^2_q , com número de graus de liberdade iguais a q (dimensão da função vetorial $g(\cdot)$).

4.3.2 Teste de Wald

O teste de Wald pode ser expresso na forma geral

$$W = g(\hat{\theta}_u)' [G\hat{\Sigma}G']^{-1} g(\hat{\theta}_u), \tag{29}$$

com $g(\cdot)$ um vetor $q \times 1$ das estimativas obtidas por máxima verossimilhança dos parâmetros irrestritos, G uma matriz de derivadas da função $g(\theta)$ e $\hat{\Sigma}$ uma estimativa consistente da matriz de variâncias e covariâncias do estimador do vetor de parâmetros $\hat{\theta}_u$.

Considere-se, por exemplo, o modelo espacial SARMA, com resíduos homocedásticos, com um parâmetro de autocorrelação igual a ρ , e suponha-se que há interesse em testar se este parâmetro é igual a zero. Para isso, pode-se escrever a hipótese nula como

$$H_0: [1,0'][\rho, \beta', \lambda, \sigma^2]' = \rho = 0 .$$

A derivada $G = \frac{\partial}{\partial \rho}[1,0'][\rho, \beta', \lambda, \sigma^2]' = [1,0']$, e chega-se então a

$$W = \hat{\rho}\{[1,0']\hat{\Sigma}[1,0']'\}^{-1}\hat{\rho} = \frac{\hat{\rho}^2}{\hat{\Sigma}_{11}} \xrightarrow{L} \chi_1^2 ,$$

onde $\hat{\Sigma}_{11}$ é o primeiro elemento da diagonal principal da estimativa $\hat{\Sigma}$.

4.3.3 Teste da razão de verossimilhança

Considere-se o modelo paramétrico indexado pelo parâmetro $\theta \in \Omega$. A partir de uma amostra de tamanho n , constrói-se a função de log-verossimilhança, como função de θ . Seja $l(\hat{\theta}_r)$ o valor da função de log-verossimilhança, computada no ponto $\theta = \hat{\theta}_r$, e seja $l(\hat{\theta}_r)$ o valor da função de log-verossimilhança, computada no ponto $\theta = \hat{\theta}_u$. Conforme discutido anteriormente, $\hat{\theta}_u$ é a estimativa irrestrita do parâmetro θ ,

$$\hat{\theta}_u = \arg\max_{\theta \in \Omega} l(\theta),$$

e $\hat{\theta}_r$ é a estimativa do parâmetro θ , impondo-se a restrição correspondente à hipótese nula, de forma que $g(\theta) = 0$. Ou seja,

$$\hat{\theta}_r = \arg\max_{\theta \in \Omega,\ \text{s.t. } g(\theta)=0} l(\theta).$$

A estatística do teste da razão de verossimilhança é dada por

$$LR = 2[l(\hat{\theta}_u) - l(\hat{\theta}_r)]. \qquad (30)$$

Sob a hipótese nula, e assumindo certas condições de regularidade, tem-se $LR \xrightarrow{L} \chi^2_q$. Considerando-se novamente o modelo SARMA, pretende-se testar a hipótese nula $H_0: \rho = 0$. A função de log-verossimilhança do modelo irrestrito tem expressão

$$l(\rho, \sigma^2, \lambda, \beta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) + log|I - \lambda W_2| + log|I - \rho W_1| = -\frac{1}{2\sigma^2}\{(I - \lambda W_2)[(I - \rho W_1)y - X\beta]\}^T\{(I - \lambda W_2)[(I - \rho W_1)y - X\beta]\},$$

enquanto a função de log-verossimilhança do modelo restrito é dada por

$$l(\sigma^2, \lambda, \beta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) + log|I - \lambda W_2| + log|I| = -\frac{1}{2\sigma^2}\{(I - \lambda W_2)[y - X\beta]\}^T\{(I - \lambda W_2)[y - X\beta]\}.$$

A estatística teste é dada por $LR = l(\rho, \sigma^2, \lambda, \beta) - l(\sigma^2, \lambda, \beta)$, e tem distribuição assintótica χ^2_1.

4.3.4 Teste dos multiplicadores de Lagrange

O teste dos multiplicadores de Lagrange, também conhecido como teste do escore, é baseado na abordagem de otimização, mais precisamente, nas condições de primeira ordem da função lagrangiana da função de log-verossimilhança

$$f(\theta) = l(\theta) + \eta g(\theta),$$

onde η é o vetor dos multiplicadores de Lagrange correspondendo às q restrições em $g(\theta) = 0$. A estatística deste é dada por

$$LM = g(\hat{\theta}_r)'[I(\hat{\theta}_r)]^{-1} g(\hat{\theta}_r)$$

onde $g(\theta_R)$ é o vetor escore do modelo restrito calculado sob a hipótese nula. $\left[I\left(\hat{\theta}_R\right)\right]$ é a matriz de informação de Fisher calculada sob a hipótese nula. A estatística LM terá distribuição χ^2_q.

4.3.5 Teste dos multiplicadores de Lagrange no modelo SEM

No caso do modelo de erros espaciais (SEM), os resíduos são modelados na forma $u = \lambda W u + \epsilon$, e, para se testar a hipótese de ausência de autocorrelação espacial, o interesse reside em se testar a hipótese nula de que $\lambda = 0$. Das três abordagens de testes

(Wald, razão de verossimilhança e multiplicadores de Lagrange), a mais conveniente é a abordagem dos multiplicadores de Lagrange, uma vez que ela requer apenas a estimação do modelo restrito. Neste caso, a partir da estimação dos coeficientes da regressão via mínimos quadrados ordinários, e das estimativas para os erros da regressão, dados por $\hat{u} = y - X(X'X)^{-1}X'y$, pode-se mostrar que a estatística teste tem expressão

$$LM = \frac{[\hat{u}'W\hat{u}]}{T\hat{\sigma}^2}, \tag{31}$$

onde $T =$ traço$[(W' + W)W]$. Caso a matriz W seja simétrica (*i.e.*, $W = W'$), obtém-se $T = n - 1$. Computacionalmente, os testes de Wald e da razão de verossimilhança são mais complexos, uma vez que é necessário o cálculo das estimativas de máxima verossimilhança sem a restrição sobre o parâmetro λ . A estatística teste em (29) converge assintoticamente para uma distribuição qui-quadrada com um grau de liberdade. Note-se que o teste dos multiplicadores de Lagrange constitui-se em um procedimento simples para se testar a hipótese de ausência de dependência espacial nos erros da regressão.

5 ESTIMAÇÃO VIA MÍNIMOS QUADRADOS DE DOIS ESTÁGIOS

Os modelos apresentados na seção 2 tratam de situações nas quais não há variáveis endógenas no lado direito da equação, de forma que a estimação via máxima verossimilhança fornece estimativas consistentes para os parâmetros do modelo. No entanto, em muitas situações, principalmente quando se tem o objetivo de identificar relações de causalidade entre determinadas políticas, o problema de endogeneidade aparece nos modelos espaciais, surgindo a necessidade de se utilizarem abordagens que estendam, por exemplo, os estimadores de variáveis instrumentais para situações com dependência espacial. Kelejian e Prucha, em diversos artigos,[12] exploraram este problema, e propuseram o estimador espacial de mínimos quadrados de dois estágios (S2SLS).

12. Ver Kelejian e Prucha (1997; 1998; 2002; 2007; 2009), e Kelejian, Prucha e Yuzefovich (2004).

Entre as características da abordagem de mínimos quadrados espaciais de dois estágios de Kelejian e Prucha, podem-se citar: *i*) visa à estimação de modelos de regressão linear, com um termo de *lag* espacial da variável resposta do lado direito da equação; *ii*) permite a estimação de modelos com regressores endógenos; *iii*) os coeficientes (inclusive o coeficiente do termo de *lag* espacial da variável resposta) são todos estimados por intermédio do procedimento de mínimos quadrados de dois estágios; *iv*) o coeficiente de *lag* espacial da variável resposta tem como instrumento, para resolver o problema de endogeneidade, os *lags* espaciais dos regressores exógenos; e *v*) o procedimento permite a incorporação de correções para a presença de heteroscedasticidade e autocorrelação espacial residual nos termos de erro da regressão estimada.

Para fazer a exposição de metodologia de mínimos quadrados espacial de dois estágios, considere-se a equação geral a seguir:

$$y = \rho Wy + Yv + X\beta + u, \tag{32}$$

onde y é um vetor coluna contendo as n observações empilhadas para a variável resposta, ρ é o coeficiente do *lag* espacial da variável resposta, W é uma matriz de vizinhança, Y é uma matriz com regressores endógenos, o vetor v é um vetor de coeficientes dos regressores endógenos, X é uma matriz com os regressores exógenos, o vetor β é o vetor com coeficientes dos regressores exógenos, o vetor u é um vetor coluna, de dimensão $n \times 1$ com os resíduos do modelo. Escrevendo-se a equação (32) de forma mais concisa, com $Z = [Wy, Y, W]$, $\gamma = [\rho, v', \beta']'$, tem-se

$$y = Z\gamma + u$$

Seja H uma matriz com observações das variáveis instrumentais para os regressores endógenos em Y. Os instrumentos para a variável endógena Wy são dados pelos *lags* espaciais dos regressores exógenos WX. A matriz com todas as variáveis instrumentais pode ser então representada como:

$$Q = [X, WX, H].$$

O estimador de mínimos quadrados espacial de dois estágios (*spatial two stage least squares* – S2SLS) tem expressão

$$\hat{\gamma}_{S2SLS} = [Z'Q(Q'Q)^{-1}Q'Z]^{-1}Z'Q(Q'Q)^{-1}Qy \ . \quad (33)$$

Na ausência de heteroscedasticidade e autocorrelação espacial dos resíduos, um estimador para a variância assintótica dos estimadores é dada por:

$$\hat{\Sigma}_{\hat{\gamma}_{S2SLS}} = \hat{\sigma}^2[Z'Q(Q'Q)^{-1}Q'Z]^{-1}, \quad 34)$$

com $\hat{\sigma}^2 = (y - Z\hat{\gamma}_{\downarrow}S2SLS)'(y - Z\hat{\gamma}_{\downarrow}S2SLS)/n$.

Na presença de heteroscedasticidade dos resíduos, uma estimativa robusta para a matriz de variância assintótica tem expressão

$$\hat{\Sigma}_{\hat{\gamma}_{S2SLS}} = \left[Z'Q\left(\widehat{Q'\Omega Q}\right)^{-1}Q'Z\right]^{-1}, \quad (35)$$

onde $\widehat{Q'\Omega Q} = Q'SQ$, e S é uma matriz diagonal contendo o quadrado dos resíduos da equacão estimada via S2SLS. Na presença de heteroscedasticidade e autocorrelação espacial, pode-se utilizar um estimador robusto (HAC). Para isso, é preciso estimar $\Psi = Q'\Omega Q$. Uma forma para esta estimativa é dada por

$$\hat{\psi}_{r,s} = \frac{1}{n}\sum_{i}\sum_{j} q_{ir}q_{is}\hat{u}_i\hat{u}_j K\left(\frac{d_{ij}}{d}\right),$$

onde q_{ir} são elementos da matriz Q , e $\hat{u}$ é o vetor de resíduos da equação estimada via S2SLS. O termo $K\left(\frac{d_{ij}}{d}\right)$ é uma função *kernel* (que é uma função de densidade, com integral igual a 1). Algumas alternativas para as funções *kernel* estão apresentadas na tabela 1.

TABELA 1
Alguns tipos de *kernel* a serem utilizados no estimador HAC para a matriz de covariância assintótica do estimador S2SLS

Tipo de *kernel*	Expressão
Kernel triangular ou de Barlett	$K\left(\frac{d_{ij}}{d}\right) = \left[1 - \left(\frac{d_{ij}}{d}\right)\right] \times I_{[d_{ij} \leq d]}$
Kernel de Epanechnikov	$K\left(\frac{d_{ij}}{d}\right) = \left[1 - \left(\frac{d_{ij}}{d}\right)^{2}\right] \times I_{[d_{ij} \leq d]}$
Kernel biquadrado (*bi-squared kernel*)	$K\left(\frac{d_{ij}}{d}\right) = \left[1 - \left(\frac{d_{ij}}{d}\right)^{2}\right]^{2} \times I_{[d_{ij} \leq d]}$

Elaboração dos autores.

Na expressão na segunda coluna da tabela 1, o valor d_{ij} corresponde à distância entre os polígonos (ou demais entidades localizadas em um espaço de coordenadas) i e j. A distância d é uma distância máxima de corte. Pode-se escolher d com um valor fixo para todas as observações, ou d variável, de forma a considerar um número fixo de vizinhos mais próximos de cada observação i (podem-se escolher distâncias variáveis, de forma a incluir os 40 vizinhos mais próximos, por exemplo, de cada observação). A partir da equação anterior para $\Psi = Q'\Sigma Q$, pode-se escrever a variância assintótica, robusta à heteroscedasticidade e à autocorrelação espacial nos resíduos, para os estimadores S2SLS, com a expressão

$$\hat{\Sigma}_{\hat{\gamma}_{S2SLS}} = \left[(Z_q'Z_q)^{-1}Z'Q(Q'Q)^{-1}\hat{\Psi}(Q'Q)^{-1}Q'Z(Z_q'Z_q)^{-1}\right], \quad (36)$$

onde $Z_q'Z_q = Z'Q(Q'Q)^{-1}Q'Z$.

A correção dada pela expressão (36), para contabilizar para desvios em relação à hipótese de homocedasticidade e ausência de correlação entre os resíduos da regressão, baseia-se no trabalho de Conley (1999), que propõe um estimador robusto para correção da *matrix* de variância assintótica no contexto de método de momentos generalizados. Na próxima seção, faz-se uma discussão especificamente sobre a abordagem de Conley, a qual se mostra bastante flexível, permitindo estimar modelos com especificações não lineares. Nesse contexto, será discutido, por exemplo, como a abordagem GMM de Conley pode ser utilizada para estimar modelos *probit*, *logit* etc., quando há correlação espacial entre as observações.

6 MÉTODO DE MOMENTOS GENERALIZADO COM CORREÇÃO PARA DEPENDÊNCIA ESPACIAL

Nesta seção, apresenta-se uma discussão sobre o procedimento de Conley (1999), por meio do qual se permite a estimação de modelos gerais via método de momentos generalizados, na presença de autocorrelação espacial nas observações. Entre as vantagens deste procedimento, podem-se citar: *i*) conta com a flexibilidade da estimação via GMM; *ii*) possibilita a estimação de modelos com especificações não lineares; *iii*) apresenta uma extensão, para o caso espacial, da estimação não paramétrica da matriz de variância, inicialmente proposta, para dados com dependência temporal, por Newey e West (1987); e *iv*) possibilita a estimação de sistemas de equações.

Para simplificar a exposição, serão considerados apenas modelos uniequacionais. Considere-se então a forma geral do modelo de regressão (linear ou não linear)

$$y_i = m(x_i, \beta) + u_i. \tag{37}$$

O termo u_i é um termo de erro que possui média zero. O vetor x_i é um vetor de variáveis explicativas, e β corresponde a um vetor de parâmetros desconhecidos do modelo. Assume-se que pode haver endogeneidade em algumas das variáveis do lado direito da equação. Considere-se então um vetor de instrumentos z_i. No caso de não haver endogeneidade, o vetor de instrumentos é exatamente o vetor de covariáveis; ou seja, $z_i = x_i$.

A partir do vetor de variáveis instrumentais, podem-se então escrever as condições de momento (momentos populacionais)

$$[E[u]_i \times z_i] = E\left[[y_i - m(x_i, \beta)] \times z_i\right] = \mathbf{0}. \tag{38}$$

Para prosseguir a estratégia de estimação, substituem-se os momentos populacionais por seus equivalentes amostrais, obtendo-se

$$\frac{1}{n}\sum_{i=1}^{n}[[y]_i - m(x_i, \beta)] \times z_i = \mathbf{0}. \tag{39}$$

Assumindo-se algumas condições de regularidade, quando o número de coeficientes é exatamente igual ao número de instrumentos, diz-se que o modelo é exatamente identificado e é possível encontrar um vetor $\hat{\beta}$ de coeficientes para o qual a igualdade acima é satisfeita.[13]

No entanto, quando a dimensão de z_i é maior que o número de coeficientes, a probabilidade de se obter uma amostra para a qual a igualdade seja exatamente satisfeita é zero (conjunto de medida nula). Uma alternativa é encontrar o vetor $\hat{\beta}$ que minimiza a forma quadrática

$$J(\beta) = \left[\frac{1}{n}\sum_{i=1}^{n}[[y]_i - m(x_i,\beta)] \times z_i\right]' \Psi \left[\frac{1}{n}\sum_{i=1}^{n}[[y]_i - m(x_i,\beta)] \times z_i\right].$$

A matriz Ψ é uma matriz positiva definida qualquer. O estimador GMM é definido como

$$\hat{\beta}_{GMM} = \underset{\beta \in \Theta}{\arg\min}\ J(\beta)$$

Pode-se mostrar que o estimador GMM é consistente (supondo-se que as devidas condições de regularidade são satisfeitas). Eficiência é obtida utilizando-se a matriz ótima $\Psi = \Omega^{-1}$, onde

$$\Omega = Cov[[[y]_i - m(x_i,\beta)] \times z_i].$$

Na prática, quando não há dependência entre as observações, pode-se estimar Ω por intermédio da expressão

$$\hat{\Omega} = \frac{1}{n}\sum_{i=1}^{n}[[[y]_i - m(x_i,\beta)] \times z_i] \times [[[y]_i - m(x_i,\beta)] \times z_i]'. \quad (40)$$

13. Ver Hamilton (1994) e Matyas (2008).

No entanto, quando há possíveis dependências entre as observações para os vetores correspondentes às condições de momento, o estimador supracitado para Ω não é mais válido. No caso de as observações para y_i, x_i e z_i acontecerem em períodos discretos de tempo igualmente espaçados, Newey e West (1987) propõem uma correção não paramétrica e robusta para o estimador $\hat{\Omega}$. Este estimador foi revisitado em Andrews (1991) e Andrews e Monahan (1992).

Conley (1999) propôs um estimador robusto tanto a heteroscedasticidade quanto autocorrelação espacial, no caso de dados *cross-section*, espacialmente distribuídos, seguindo os mesmos princípios que Newey e West (1987). De maneira geral, o estimador proposto por Conley tem expressão

$$\hat{\Omega} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} K(i,j) \times [[[y]_i - m(x_i,\beta)] \times z_i] \times [[[y]_j - m(x_j,\beta)] \times z_j]', \quad (41)$$

onde

$K(i,j) = \left[1 - \frac{D_H(i,j)}{L_H}\right] \times \left[1 - \frac{D_V(i,j)}{L_V}\right]$, para $D_H(i,j) < L_H$ e $D_V(i,j) < L_V$, e $K(i,j) = 0$, caso contrário. O valor $D_H(i,j)$ corresponde à distância horizontal entre unidades i e j, o valor $D_V(i,j)$ corresponde à distância vertical entre i e j, L_H é a distância de corte horizontal, e L_V é a distância de corte vertical. Em geral, a minimização de $J(\beta)$ não resulta em uma solução explícita, devendo ser feita via algoritmos numéricos. Uma exceção ocorre no caso de modelos lineares; neste caso, o estimador GMM pode ser escrito em forma fechada, sem haver necessidade de minimização numérica.

A flexibilidade da estimação via GMM, na formulação $y_i = m(x_i,\beta) + \epsilon_i$, permite o tratamento de modelos não lineares, com formulações paramétricas comumente encontradas na literatura. A tabela 2 apresenta alguns exemplos de modelos que podem ser incorporados na formulação GMM. Pode-se então proceder com a abordagem de estimação, corrigindo, por exemplo, para problemas de dependência espacial.

TABELA 2
Exemplos de modelos paramétricos enquadrados na formulação GMM, que podem ser estimados corrigindo-se para dependência espacial

Modelos paramétricos	Formulação
Modelos lineares	$m(x_i,\beta) = x_i^T\beta$
Modelos *logit*	$m(x_i,\beta) = \dfrac{e^{x_i^T\beta}}{1+e^{x_i^T\beta}}$
Modelos *probit*	$m(x_i,\beta) = [\Phi(x]_i^T\beta)$
Modelos *complementary log-log*	$m(x_i,\beta) = e^{x_i^T\beta}$
Modelos exponenciais	$m(x_i,\beta) = 1 - \exp(-\exp(x_i^T\beta))$

Elaboração dos autores.

Uma vez estimado o vetor de coeficientes β , pode-se proceder com o processo de inferência a partir da matriz de covariância dos estimadores, estimável a partir da expressão

$$\widehat{Var}(\hat{\beta}) = \left\{ n \times \left[\frac{1}{n}\sum_{i=1}^{n} [\frac{\partial}{\partial\beta([y]_i - m(x_i,\beta)]} \times z_i \right)\right]' \hat{\Omega}^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} [\frac{\partial}{\partial\beta([y]_i - m(x_i,\beta)]} \times z_i \right)\right]\right\}^{-1}$$

Quando o modelo é exatamente identificado, com número de instrumentos igual ao número de parâmetros, a minimização da forma quadrática $J(\beta)$ resulta em $J(\beta) = 0$. Quando o modelo é sobreidentificado, pode ser testada a validade das condições de momento, utilizando-se a estatística de Hansen (1982)

$$J = n \times \left[\frac{1}{n}\sum_{i=1}^{n} [([y]_i - m(x_i,\beta)] \times z_i) \right]^T \hat{\Omega}^{-1} \left[\frac{1}{n}\sum_{i=1}^{n} [([y]_i - m(x_i,\beta)] \times z_i) \right]. \quad (42)$$

Sob a hipótese nula de validade dos instrumentos, pode-se mostrar que a estatística J em (42) tem distribuição assintótica qui-quadrada, com $k-q$ graus de liberdade, sendo q o número de coeficientes e k o número de condições de momento.

7 COMENTÁRIOS FINAIS

Este texto apresenta uma discussão sobre alguns dos modelos econométricos comumente utilizados para modelagem de dados espaciais. Os modelos apresentados estariam mais adequados para estudos empíricos seguindo as abordagens experimentalista e descritiva, nas quais o objetivo é identificar efeitos causais de uma determinada política, ou encontrar relações entre variáveis econômicas. De fato, o estimador de mínimos quadrados de dois estágios, de Kelejian e Prucha, e o estimador de método de momentos generalizado, de Conley (ambos discutidos neste estudo), permitem a estimação de parâmetros na presença de variáveis endógenas do lado direito da equação, contabilizando e/ou corrigindo para a presença de autocorrelação espacial nos resíduos do modelo. Mesmo não tratando diretamente a abordagem estruturalista, as ideias apresentadas neste texto fornecerão ao leitor uma noção dos procedimentos para estimação com dados com presença de dependência especial, o que poderá ser útil para a estimação de parâmetros estruturais em modelos microfundamentados.

Dado o grande avanço recente na literatura de análise de dados espaciais, optou-se por apresentar apenas alguns dos métodos mais comumente utilizados, de forma a transmitir ao leitor uma ideia básica, mas clara, dos fundamentos da estimação de modelos econométricos com dependência espacial. Não foram cobertos modelos para dados de painel,[14] mas apenas para dados *cross-section*. Outro tópico de extrema importância na análise de dados espaciais, que não foi tratado aqui, são os modelos estimados via abordagem bayesiana. O leitor poderá encontrar boas exposições em Banerjee, Carlin e Gelfand (2004), Schabenberger e Gotway (2009), e Tanner (1996), entre outros.

14. Ver, por exemplo, Elhorst (2003), Druska e Horrace (2004), e Egger, Pfaffermayr e Winner (2005).

REFERÊNCIAS

ACKERBERG, D. *et al.* Econometric tools for analyzing market outcomes. *In:* HECKMAN, J. J.; LEAMER, E. E. (Eds.). **Handbook of Econometrics**. Amsterdam: Elsevier, vol. 6A, 2007.

ANDREWS, D. W. K. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. **Econometrica**, vol. 59, p. 817-858, 1991.

ANDREWS, D. W. K.; MONAHAN, J. C. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. **Econometrica**, vol. 60, p. 953-966, 1992.

ANGRIST, J. D.; PISCHKE, J. S. **Mostly harmless econometrics**: an empiricist's companion. New Jersey, Princenton University Press, 2009.

ANSELIN, L. **Spatial econometrics**: methods and models. Kluwer Academic, Dordrecth, 1988.

ANSELIN, L.; FLORAX, R. Advances in spatial econometrics. Heidelberg, **Springer-Verlag**, 2000.

ANSELIN, L., FLORAX, R., REY, S. J. Advances in spatial econometrics – Methodology, Tools and Applications. **Springer: Advances in Spatial Science**, Heidelberg, 2004.

BANERJEE, S.; CARLIN, B. P.; GELFAND, A. E. **Hierarchical modeling and analysis for Spatial Data**. Monographs on Statistics and Applied Probability 101, Chapman & Hall/CRC, Florida, 2004.

BERRY, S.; LEVINSOHN, J.; PAKES, A. Automobile prices in market equilibrium. **Econometrica**, vol. 63, no. 4, pp. 841-890, 1995.

_____. Differentiated products demand systems from a combination of micro and macro data: the new car market. **Journal of Political Economy**, vol. 112, n. 1, 2004.

CAMERON, A. C.; TRIVEDI, P. K. **Microeconometrics:** methods and applications. Cambridge University Press, New York, 2005.

CLIFF, A. D. ; ORD, J. K. **Spatial autocorrelation** Pion, London,1972.

CONLEY, T. GMM estimation with cross-sectional dependence. **Journal of Econometrics**, vol. 92, p. 1-45, 1999.

DAVIS, T. A. **Direct methods for sparse linear systems (Fundamentals of Algorithms)**. Society for Industrial and Applied Mathematics, 2006.

DRUSKA, V.; HORRACE, W. C. Generalized moments estimation for spatial panel data: Indonesian rice farming. **American Journal of Agricultural Economics**, vol. 86, n. 1, p. 185-198, 2004.

ECKSTEIN, Z.; WOLPIN, K. Why youths drop out of High School: the impact of preferences, opportunities, and abilities. **Econometrica**, 67, 1295-1340, 1999.

EGGER, P.; PFAFFERMAYR, M.; WINNER, H. An unbalanced spatial panel data approach to US state tax competition. **Economic Letters**, vol. 88, n. 3, p. 329-335, 2005.

ELHORST, J. P. Specification and estimation of spatial panel data models. **International Reginal Science Review,** vol. 26, n. 3, p. 244-268, 2003.

EPPLE, D.; SIEG, H. Estimating equilibrium models of local jurisdictions. **Journal of Political Economy**, 107, 645-681, 1999.

HAMILTON, J. D. **Time Series Analysis**. Princeton University Press, 1994.

HENDRY, D. F. Dynamic econometrics. **Advanced Texts in Econometrics**, Oxford University Press, Oxford, 1995.

HAHN, J.; TODD, P.; VAN DER KLAAUW, W. Identification and estimation of treatment effects with a regression-discontinuity design. **Econometrica**, 69, 201-209. 2001.

HOLMES, T. J. Structural, experimentalist, and descriptive approaches to empirical work in regional economics. **Journal of Regional Science**, vol. 50, n. 1, p. 5-22, 2010.

KELEJIAN, H. H.; PRUCHA, I. R. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. **The Journal of Real State Finance and Economics**, vol. 17, n. 1, p. 99-121, 1998.

KELEJIAN, H. H.; PRUCHA, I. R. Estimation of spatial regression models with autoregressive errors by two-stage least squares procedures: a serious problem. **International Regional Science Review**, vol. 20, n. 1, p. 103-111, 1997.

KELEJIAN, H. H.; ROBINSON, D. P. Spatial autocorrelation : a new computationally simple test with an application to per capita county police expenditures.**Regional Science and Urban Economics**, vol. 22, issue 3, p. 317-331,1992.

_____. 2SLS and OLS in a spatial autoregressive model with equal spatial weights. **Regional Science and Urban Economics**, vol. 32, n. 6, p. 691-707, 2002.

_____. HAC estimation in a spatial framework. **Journal of Econometrics**, vol. 140, n. 1, p. 131-154, 2007.

_____. Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. Journal of Econometrics. **No prelo**. 2009.

KELEJIAN, H. H.; PRUCHA, I. R.; YUZEFOVICH, Y. Instrumental variable estimation of a spatial autoregressive model with autoregressive disturbances: large and small sample results. *In:* LESAGE, J.; PACE, R. K. **Spatial and Spatiotemporal Econometrics, Advances in Econometrics**, New York: Elsevier, vol. 18, p. 163-198, 2004.

KEANE, M.; WOLPIN, K. I. The career decisions of young men. **Journal of Political Economy**, 105, 473-522, 1997.

LEE, L. GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. **Journal of Econometrics**, vol. 137 (2), p. 489-514, 2007.

LESAGE, J. Bayesian estimation of spatial autoregressive models. **International Regional Science Review**, 20, n. 1 and 2, p. 113-129, 1997.

_____. **The theory and practice of spatial econometrics**. Department of Economics, University of Toledo, 1999.

LESAGE, J., PACE, R. K. Introduction to spatial econometrics. CRC Press, Boca Raton, 2009.

MANSKI, C. Identification of endogenous social effects: the reflection problem. **The Review of Economic Studies**, vol. 60(3), p. 531-542, 1993.

MATYAS, L. **Generalized method of moments estimation - Themes in Modern Econometrics.** Cambridge University Press, 2008.

McMILLEN, D. P. Issues in spatial data analysis. **Journal of Regional Science**, vol. 50, n. 1, p. 119-141, 2010.

NEVO, A. Measuring market power in the ready-to-eat cereal industry. **Econometrica**, vol. 69, n. 2, p. 307-342, 2001.

NEWEY, W. K.; WEST, K. D. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. **Econometrica**, vol. 55, p. 703-708, 1987.

PACE, K.; BARRY, R. Sparse spatial autoregressions. **Statistics and Probability Letters**, 33, 291-7, 1997.

_____. **Simulating mixed regressive spatially autoregressive estimators, computational statistics**. Vol.13, p. 397-418, 1998.

PETRIN, A. Quantifying the benefits of new products: the case of the minivan. **Journal of Political Economy,** vol. 110, n. 4, 2002.

PINKSE, J.; SLADE, M. E. Mergers, brand competition, and the price of a pint. **European Economic Review**, vol. 48, n. 3, p. 617-643, 2004.

_____. The future of spatial econometrics. **Journal of Regional Science**, vol. 50, n. 1, p. 103-117, 2010.

PINKSE, J.; SLADE, M. E.; BRET, C. Spatial price competition: a semiparametric approach. **Econometrica**, vol. 70, n. 3, p. 1111-1153, 2002.

POFAHL, G. **Essays on horizontal merger simulation: the curse of dimensionality, retail price discrimination, and supply channel stage-games**. Tese (Doutorado), Texas A&M, 2007.

SCHABENBERGER, O.; GOTWAY, C. A. Statistical methods for spatial data analysis. **Texts in Statistical Science**, Chapman & Hall/CRC, Florida, 2009.

TANNER, M. Tools for statistical inference, methods for the exploration of posterior distributions and likelihood functions. **Springer Series in Statistics**, 1996.

BIBLIOGRAFIA COMPLEMENTAR

ANSELIN, L.; FLORAX, R. New directions in spatial econometrics. **Springer-Verlag,** Advances in Spatial Science, 1995.

ARBIA, G.; BALTAGI, B. H. Spatial econometrics - Methods and Applications. **Physica-Verlag**, Heidelberg, 2009.

BARRY, R.; PACE, R. **A Monte Carlo estimator of the *log* determinant of large sparse matrices – Linear algebra and its applications**. 289, n. 1-3, p. 41-54, 1999.

CARVALHO, A. X. Y.; ALBUQUERQUE, C. W.; MOTA, J. A.; PIANCASTELLI, M. (Orgs.). **Dinâmica dos municípios**. Brasília: Ipea, 2008.

CHOMITZ, K. M.; DA MATA, D.; CARVALHO, A.; MAGALHAES, J. C. R. **Spatial dynamics of labor markets in Brazil**. World Bank Policy Research Working Paper 3752, 2005.

PINKSE, J.; SLADE, M. E.; SHEN, L. Dynamic spatial discrete choice using one-step GMM: an application to mine operating decisions. **Spatial Economic Analysis**, vol. 1, n. 1, p. 53-99, 2006.

EDITORIAL

Coordenação
Cláudio Passos de Oliveira

Revisão
Luciana Dias Jabbour
Marco Aurélio Dias Pires
Reginaldo da Silva Domingos
Leonardo Moreira de Souza (estagiário)
Maria Angela de Jesus Silva (estagiária)

Editoração
Bernar José Vieira
Cláudia Mattosinhos Cordeiro
Everson da Silva Moura
Luís Cláudio Cardoso da Silva
Renato Rodrigues Bueno
Eudes Nascimento Lins (estagiário)

Capa
Luís Cláudio Cardoso da Silva

Projeto Gráfico
Renato Rodrigues Bueno

Livraria do Ipea

SBS – Quadra 1 - Bloco J - Ed. BNDES, Térreo.
70076-900 – Brasília – DF
Fone: (61) 3315-5336

Correio eletrônico: livraria@ipea.gov.br
Tiragem: 500 exemplares

Ipea – Instituto de Pesquisa Econômica Aplicada

Secretaria de Assuntos Estratégicos da Presidência da República