

Economias de aglomeração no Brasil: evidências a partir de equações de rendimentos

Anderson Bonetto Carraro

Mestre em Economia do Desenvolvimento pelo Programa de Pós-Graduação em Economia do Desenvolvimento (PPGE) da Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS).

E-mail: anderbc@hotmail.com

Paulo de Andrade Jacinto

Professor do Programa de Pós-Graduação em Desenvolvimento Econômico. – Universidade Federal do Paraná - UFPR.

E-mail: paulo.jacinto@ufpr.br

Túlio Antônio Cravo

Economista do Banco Interamericano de Desenvolvimento - BID.

E-mail: tcravo@iadb.org

Recebido em: 29 maio 2018. Aceito: 05 julho 2018.

DOI: <http://dx.doi.org/10.21674/2448-0479.44.558-581>

Resumo

O presente estudo tem como propósito analisar as economias de aglomeração no Brasil, a partir de equações de rendimentos. Por meio da Análise Exploratória de Dados Espaciais (AEDE), pretende-se verificar a existência de associações espaciais salariais, tais como: regiões homogêneas (clusters) e observações atípicas (outliers). Na segunda parte, dedicada à estimação do modelo econométrico, utilizam-se equações salariais, com formulação baseada em um modelo desenvolvido por Fingleton (2003), cuja principal hipótese aponta para uma relação positiva entre os diferenciais de salários e a densidade de emprego.

Palavras-chave: Economias de aglomeração. Equações salariais. Densidade do emprego.

Abstract

Economies of agglomeration in Brazil: evidence from income equations

This study aims to analyze the economies of agglomeration in Brazil, from earnings equations. Using the Exploratory Spatial Data Analysis (ESDA), the paper intends to verify the existence of spatial associations of wages, such as: homogeneous regions (clusters) and atypical observations (outliers). In the second part, dedicated to the estimation of the econometric model, we use wage equations, with formulation based on a model developed by Fingleton (2003), whose main hypothesis points to a positive relationship between wage differentials and employment density.

Keywords: Agglomeration economies. Wage equations. Employment density.

Introdução

Uma das questões amplamente discutida, na teoria econômica, diz respeito à concentração de pessoas e atividades econômicas, em um determinado espaço geográfico. No Brasil, o estudo das economias de aglomeração é particularmente relevante, tendo em vista a constatação do elevado grau de concentração de pessoas e atividades econômicas em poucos espaços geográficos, fato que pode ser explicado pela presença de externalidades positivas. Isso sugere a tendência de um padrão de desenvolvimento econômico, geograficamente diversificado ou especializado entre as diferentes regiões do país, o que também pode gerar desigualdades regionais.

O objetivo deste estudo é verificar se há a existência de algum tipo de associação espacial sobre o rendimento recebido entre os trabalhadores brasileiros, no ano de 2010, nos municípios brasileiros. Para tanto, serão utilizadas equações de rendimentos, semelhantes às propostas por Fingleton (2003), para os municípios brasileiros. Utilizando-se de dados agregados por municípios, pode-se ter uma ideia mais fidedigna, sobre o fenômeno estudado e sobre as variáveis a serem analisadas.

O presente documento é composto por seis seções. Sendo a primeira esta introdução, seguida pela revisão bibliográfica. Na terceira seção aprofundam-se os tópicos referentes aos procedimentos metodológicos utilizados; e, em seguida parte-se para a análise espacial das equações de rendimentos. Na quinta seção, expõem-se os resultados econométricos verificados. Ao final está a seção de conclusão, onde apresentam-se os principais resultados alcançados neste estudo.

Revisão da Literatura

Essa seção visa apresentar uma breve revisão da literatura, acerca das economias de aglomeração, na teoria econômica. Para tanto, o mesmo encontra-se dividido em duas subseções. Na primeira, são destacados os preceitos de Mincer (1958, 1974), assim como apresentada a sua equação de rendimentos. Na segunda subseção, aponta-se o modelo teórico proposto por Fingleton (2003).

Equação de rendimentos (minceriana)

Conforme Viana (2010), o precursor da Teoria do Capital Humano foi Mincer (1958), que indicou a existência de correlação entre o investimento para a formação das pessoas (trabalhadores) e a distribuição de renda pessoal. Mincer (1958) concluiu que, a dispersão entre os rendimentos pessoais estava associada ao volume de investimento efetuado em capital humano, os quais impactariam na produtividade e no crescimento da economia.

No Brasil, inúmeros estudos já estimaram a taxa de retorno da educação através da equação de rendimentos minceriana. Usaram-se, nestes estudos, diferentes bases de dados e métodos de estimação, além de períodos distintos. Langoni (1973) foi um dos primeiros brasileiros a seguir tal abordagem. Basicamente o autor estimou seu modelo mediante regressões, tendo como variáveis explicativas a educação, idade, sexo, atividade e região de residência. Ele considerava que a educação possuía grande importância, como variável explicativa da renda.

A fim de medir o retorno da educação, Mincer (1974) propôs um tipo de equação que leva em conta a influência da educação no salário dos indivíduos, juntamente com o impacto que é causado

nesses ganhos, através do aprendizado que lhe é transmitido pela experiência no trabalho. Com isso o salário de um indivíduo é dado por:

$$\log w = \beta_0 + \beta_1 S_i + \beta_2 \exp_i + \beta_3 \exp_i^2 + \varepsilon_i \quad i=1, \dots, n \quad (1)$$

em que:

$\log w$ é o logaritmo do salário mensal, β_0 é uma constante, S_i e \exp_i são os anos de escolaridade e experiência respectivamente, β_1 é a taxa de retorno da educação, β_2 e β_3 são os efeitos da experiência sobre o salário do trabalhador, e finalmente, ε é um termo aleatório, que capta todas as outras variáveis que influenciam $\log w$ e que não sejam correlacionadas com S e \exp , ou seja:

$$E(\varepsilon_i / S_i, \exp_i, \exp_i^2) = 0 \quad i=1, \dots, n \quad (2)$$

Essa equação (1) permite, de forma simplificada, ter uma ideia dos impactos que a educação e a experiência têm sobre os rendimentos dos indivíduos. Esse impacto é mensurado na forma dos anos adicionais de educação e experiência, expressos, respectivamente, nos coeficientes $\beta_1 > 0$ e $\beta_2 > 0$. A variável experiência ao quadrado indica os aumentos nos rendimentos, ocasionados pelo acúmulo de experiência e na literatura se verifica que estão sujeitos a retornos decrescentes. Por isso o coeficiente tende a apresentar um sinal negativo, isto é $\beta_3 < 0$.

Segundo Rodrigues (2010), uma transformação da idade de cada trabalhador foi utilizada como uma proxy da sua experiência. Mincer (1974) utilizou uma transformação de $\exp = \text{idade} - S - 6$, assumindo que o trabalhador entra na força de trabalho logo após completar a sua educação e que a idade de conclusão da escola é, em média, $S + 6$. Pesquisas sobre o assunto passaram a se basear, sempre e fortemente, nesse método de estimação para obter a magnitude do retorno de educação, porém com algumas alterações do proposto por Mincer (1974).

Segundo Bandeira e Jacinto (2010), o método frequentemente empregado para estimar a equação de rendimentos de Mincer (1974) é o modelo de regressão clássica dos mínimos quadrados ordinários (MQO). Uma importante hipótese desse modelo é a de que os erros não sejam correlacionados com as variáveis explicativas, $E[\varepsilon_i / x_i] = 0$. Por isso o principal problema é a validade dessa hipótese, tendo em vista que a existência dessa correlação implica que os coeficientes estimados serão inconsistentes e não convergirão em probabilidade, para os parâmetros populacionais. Trata-se de um problema de endogeneidade da educação, no caso da equação de rendimentos. Outro problema, também comum à estimação dessa equação, é o viés gerado pela omissão de variáveis que podem afetar a renda do indivíduo.

Outra limitação da equação minceriana é que esta não considera a questão da aglomeração de maneira explícita. As economias de aglomeração consistem em ganhos de produtividade decorrentes da concentração de pessoas e atividades econômicas, em determinados espaços geográficos. Fingleton (2003) incluiu a influência das aglomerações na equação de rendimento. Ele testa, a partir de um modelo econométrico, a tendência de pessoas e empresas de se aglomerar em determinado espaço geográfico, assim como as consequências desse processo para a produtividade e os salários. Este estudo procura fornecer evidências de tais hipóteses, relacionando as aglomerações geográficas aos salários recebidos nos municípios brasileiros em 2010. A seguir apresenta-se melhor esta ideia defendida por Fingleton (2003).

Economias de aglomeração a partir dos efeitos sobre a taxa salarial

Fingleton (2003) desenvolveu um modelo econométrico que visa expressar os efeitos da concentração da atividade econômica, dos *spillovers* (ou transbordamentos) espaciais, do nível de escolaridade dos trabalhadores e do conhecimento técnico aplicado à produção sobre a taxa salarial média de unidades espaciais.

A partir deste modelo, o estudo de Amarante (2011), teve a finalidade de obter evidências empíricas de economias de aglomeração nas atividades econômicas, de forma agregada e na indústria de transformação dos municípios brasileiros nos anos de 2000 e 2009. Utilizam-se equações salariais, com formulação baseada no modelo microeconômico desenvolvido por Fingleton (2003) e aplicado ao Brasil por Galinari (2006) e Galinari *et al.* (2006), cuja principal hipótese indica uma relação positiva entre os diferenciais de salários e a produtividade do trabalho.

No Brasil, os estudos dessa natureza são relevantes, uma vez que se verifica um elevado grau de concentração das atividades econômicas, em poucos espaços geográficos, especialmente nas regiões Sul e Sudeste do País, sugerindo um padrão de desenvolvimento econômico, geograficamente diversificado ou especializado.

Como teorizado por Jacobs (1969), a concentração econômica no espaço amplia a produtividade, beneficiando tanto as empresas quanto os trabalhadores, uma vez que a variedade de bens, serviços, tecnologias e conhecimentos próprios de um centro urbano diversificado estimulam a criatividade e a troca de informações e experiências, impulsionando, assim, a inovação e o crescimento local e regional, dada a possibilidade de propagação dos benefícios pelo espaço.

Amarante (2011) diz que, a partir da constatação do padrão de desenvolvimento regional diferenciado existente no Brasil, fato que se verifica não apenas entre as regiões ou Unidades da Federação, mas também entre os municípios de um mesmo estado, os resultados encontrados podem ser sugestivos para a aplicação de políticas públicas e instrumentos de intervenção. Alguns exemplos disso são a formação de polos regionais locais, aliados ao estímulo à educação; desenvolvimento tecnológico; e, criação e melhoria da infraestrutura urbana e de acessibilidade, que possibilitem uma ampliação dos benefícios gerados, tendo em vista um desenvolvimento regional mais igualitário. A limitação no estudo de Amarante (2011), no entanto, é o uso da Relação Anual de Informações Sociais (RAIS), do Ministério do Trabalho e Emprego, que utiliza dados apenas do mercado de trabalho formal. Visando corrigir esta restrição, o presente artigo utiliza como base de dados, os microdados disponibilizados no Censo 2010, que engloba toda a economia, incluindo o mercado de trabalho informal e os trabalhadores liberais. Esta diferenciação traz, aos resultados encontrados, uma interpretação mais fidedigna da realidade salarial brasileira.

Procedimentos Metodológicos

Nessa seção, apresenta-se os procedimentos metodológicos praticados na elaboração da presente pesquisa, tendo em vista verificar as evidências acerca das economias de aglomeração nos municípios brasileiros, a partir dos dados do Censo 2010. A análise compreende um estudo de caráter bibliográfico e descritivo, cujo intuito é contribuir para o debate acerca da identificação das economias de aglomeração no Brasil.

Modelo econométrico: salários e aglomerações

Galinari (2006) segue as hipóteses formuladas por Fingleton (2003), que deriva a relação entre o produto no setor de bens e serviços finais (Q), em um parâmetro de economias de escala que, por seu turno, é importante para o entendimento da equação salarial. Para isso, assume-se uma função de produção Cobb-Douglas (3) para o segmento de bens e serviços finais em uma área cujos insumos são a terra (L), o nível de eficiência do trabalho, empregado no setor de bens e serviços finais (M), e o nível de produção do segmento de insumos intermediários (I), baseado em uma sub-função de produção CES.

$$Q = (M^\beta I^{1-\beta})^\alpha L^{1-\alpha} \quad (3)$$

Sendo que o modelo se fundamenta em densidades, assume-se que $L=1$, o que leva à equação (4):

$$Q = (M^\beta I^{1-\beta})^\alpha \quad (4)$$

Admitindo-se que, no equilíbrio, cada firma do setor de bens intermediários apresenta um produto igual a $i(t)$ - constante e independente do efetivo total de trabalho (N) da economia de uma área - e que há $x(N)$ firmas, pode-se obter a partir da função de produção CES a seguinte simplificação:

$$I = x^\mu i(t) \quad (5)$$

em que μ é uma medida de retornos de escala, internos à firma produtora de bens intermediários, na posição de equilíbrio. Substituindo a equação (5) em (4), obtém-se:

$$Q = (M^\beta (x^\mu i(t))^{1-\beta})^\alpha \quad (6)$$

e, portanto,

$$Q = M^{\beta\alpha} x^{\alpha(\mu - \mu\beta)} i(t)^{\alpha(1-\beta)} \quad (7)$$

O número de firmas, do setor de bens intermediários x , é igual ao número de trabalhadores empregados neste setor, dividido pelo número de trabalhadores por firma, tal que:

$$X = (1 - \beta)N/ai(t) + s \quad (8)$$

em que $(1 - \beta)$ é a participação relativa dos trabalhadores, do setor produtor de bens e serviços intermediários no mercado de trabalho local, a é o requerimento marginal de trabalho e s o requerimento fixo de trabalho em uma firma representativa.

Substituindo (8) em (7) chega-se à seguinte equação:

$$Q = N^{\alpha(\beta + \mu - \mu\beta)} \beta^{\alpha\beta} (ai(t) + s)^{\alpha\mu(\beta - 1)} i(t)^{\alpha(1-\beta)} (1-\beta)^{-\alpha\mu(\beta - 1)} \quad (9)$$

que, simplificada, torna-se:

$$Q = \varphi N^{\alpha(1+(1-\beta)(\mu-1))} = \varphi N^{\gamma} \quad (10)$$

em que φ é uma constante e $\gamma = \alpha [1 + (1-\beta)(\mu-1)]$ é a elasticidade da produção com relação a N .

Fingleton (2003) explica que, no seu modelo, os retornos crescentes líquidos ($\gamma > 1$) resultam do aumento da variedade de bens intermediários com a densidade da atividade econômica, sujeitos a retornos decrescentes advindos de “efeitos de congestão” ($\alpha < 1$). Além disso, os retornos crescentes dependem da relevância dos insumos à produção de bens finais ($\beta < 1$) e da presença de economias de escala, internas às firmas no setor produtor de insumos intermediários ($\mu > 1$). Com vistas a derivar a equação salarial, o autor supõe que, o produto no setor de bens finais, depende das unidades de eficiência do trabalho (N) e do número de unidades de terra (L):

$$Q = [f(N)]^{\alpha} L^{1-\alpha} \quad (11)$$

Diferenciado (11) com respeito ao fator terra, tem-se:

$$dQ/dL = f(N)^{\alpha} L^{-\alpha} (1-\alpha) / L = (1-\alpha)Q/L$$

$$r = (1-\alpha)Q/L \quad (12)$$

$$rL/Q = 1-\alpha$$

em que, r é a taxa de aluguel por unidade de terra urbana. A parcela do produto final pago ao fator de produção terra ($1-\alpha$) é igual à taxa de aluguel (r), vezes o número de unidades de terra, dividido pelo produto final (Q). Como há apenas dois fatores de produção, a parcela do produto que remunera as unidades de eficiência de trabalho de ambos os tipos (N) é α que, por sua vez, é igual ao produto da taxa salarial por unidade de eficiência de trabalho, pelo total de unidades de eficiência de trabalho, dividido pelo produto final Q :

$$wN/Q = \alpha \quad (13)$$

Ao ser logaritimada, a relação acima passa a ser expressa por:

$$\ln(w) = \ln(Q) + \ln(\alpha) - \ln(N) \quad (14)$$

Sendo E o emprego total, por unidade de área de solo urbano e A o nível de eficiência por área, pode-se substituir (10) e $N=EA$ na equação acima e obter-se:

$$\ln(w) = \ln(\varphi) + \gamma \ln(AE) + \ln(\alpha) - \ln(AE) \quad (15)$$

OU

$$\ln(w) = k_1 + (y-1)\ln(E) + (y-1)\ln(A) \quad (16)$$

em que k_1 é uma constante.

Haja vista que a variável A não pode ser mensurada de maneira direta, o autor delinea hipóteses sobre seus determinantes. Supondo homogeneidade da tecnologia entre as diferentes áreas, a variação do nível de eficiência pode ser atribuída à habilidade dos trabalhadores em seu uso. Assim, uma especificação plausível para o $\ln(A)$ seria:

$$\begin{aligned} \ln(A) &= b_0 + b_1 H + b_2 T + \rho W \ln(A) + \xi \quad (17) \\ \xi &= N(0, \sigma^2) \end{aligned}$$

Conforme a equação acima, Fingleton (2003) admite que o logaritmo do nível de eficiência por unidade de área é influenciado pelo nível educacional da mão-de-obra H ; por um indicador que denomina de “conhecimento técnico” T ; por spillovers dos níveis de eficiência entre áreas vizinhas $W\ln(A)$; e, um termo aleatório ξ , que pode captar os efeitos de variáveis omitidas. Os níveis de eficiência, de uma área, podem depender daquele observado em sua vizinhança. Tal variável é representada pelo produto de $\ln(A)$, pela matriz de pesos espaciais W , que define as relações de vizinhança entre as áreas.

Uma vez que $W\ln(A)$ é desconhecido, o autor determina este termo, em função das variáveis conhecidas. Para tanto, rearranjando a Equação (16) e multiplicando ambos os lados por W , tem-se que:

$$W \ln(A) = W(-k_1/y-1) + (1/y-1)W \ln(w) - W \ln(E) \quad (18)$$

Substituindo (17) e (18) em (16), com a adição de um termo ω , para captar erros de medidas na variável salário, chega-se à equação final do modelo:

$$\begin{aligned} \ln(w) &= k_2 + \rho W \ln(w) - \rho W k_1 + (y-1)[\ln(E) - \rho W \ln(E)] + a_1 H + a_2 T + v \quad (19) \\ v &\sim N(0, \pi^2) \end{aligned}$$

em que k_2 é uma constante e Wk_1 uma variável dependente de valores desconhecidos de φ e α , que pode ser ignorada sem qualquer prejuízo, de acordo com o próprio autor. Fingleton (2003) interpreta o coeficiente $(y-1)$ como sendo o termo que indica se existe ou não retornos crescentes para a densidade de emprego. Esse termo é assim interpretado, pois é resultado da subtração do efeito densidade, menos o efeito densidade induzido pela dependência espacial da taxa de salário. Se o coeficiente $(y-1)$ for positivo, existe indício da existência de retornos crescentes de escala na densidade de emprego. A estimação deste parâmetro é feita em dois estágios pelos autores, mas o procedimento requer a imposição de valores arbitrários de ρ para ser calculado.

Embora a estratégia empírica utilizada esteja baseada nas ideias de Fingleton (2003), opta-se por rearranjar o modelo para ser estimado em um único estágio, e obter coeficientes separados para a influência da densidade, sobre a taxa de salário, e a influência da densidade dos vizinhos, sobre a taxa de salário. Além disso, o objetivo principal desse estudo é analisar a questão da dependência espacial

da taxa de salário, e não de verificar se existem retornos crescentes de densidade de emprego como em Fingleton (2003).¹ Portanto, considerando $(y - 1) = \beta_1$; e $(y - 1)\rho = \beta_2$, o modelo pode ser rearranjado para:

$$\ln(w) = k_2 + \rho W \ln(w) + \beta_1 \ln(E) - \beta_2 W \ln(E) + a_1 H + a_2 T + v \quad (20)$$

$$v \sim N(0, \pi^2)$$

Conforme Vieira (2010), $\ln(w)$ é um vetor n por 1 de observações sobre a variável dependente, W é um vetor n por n de defasagens espaciais para a variável dependente, ρ é o coeficiente autorregressivo espacial, e v é um vetor n por 1 de termos de erro distribuídos aleatoriamente, ou seja, $v \sim N(0, \pi^2 I)$. Esse processo é conhecido como SAR (spatial autoregressive), onde W é a matriz de conectividade que, em geral, contém relações de contiguidade de 1ª ordem ou funções de distância. Em linhas gerais, W é montada de modo a captar a influência dos vizinhos, na variável em consideração. Os elementos de W são usados para especificar a estrutura de dependência espacial, entre as observações.²

Matriz de peso espacial

Em econometria espacial, uma alternativa para verificar a autocorrelação espacial, entre duas variáveis, é a utilização de matrizes de ponderação. Conforme Anselin (1992), a matriz de pesos espaciais é um dos principais elementos que distinguem a análise de dados espaciais da econometria tradicional, determinando quais observações são vizinhas entre si.

Almeida (2012) afirma que uma matriz de ponderação espacial é uma matriz quadrada de dimensão n por n . Os pesos espaciais W_{ij} representam o grau de conexão entre as regiões, segundo algum critério de proximidade, mostrando a influência da região j sobre a região i . Assim, a matriz W é útil por realizar uma espécie de ponderação da influência, que as regiões exercem entre si.

O critério de conexão geográfica, da matriz de pesos espaciais, apoia-se na ideia de proximidade, que por sua vez, pode ser definida de acordo com a contiguidade e/ou com a distância geográfica. Neste artigo, utilizam-se dois tipos de matrizes: a *Queen*, baseada na contiguidade, em que duas regiões são vizinhas, caso elas partilhem de uma fronteira física comum; e, a matriz de k vizinhos mais próximos, onde a matriz binária representa a distância geográfica entre duas regiões, geralmente medida em quilômetros.

Variáveis e estatística descritiva

Nesta subseção são apresentados os procedimentos metodológicos praticados na elaboração da presente pesquisa, tendo em vista verificar as evidências acerca das economias de aglomeração nos municípios brasileiros no ano de 2010.

¹ O presente estudo, contudo, reconhece a possível influência da densidade de emprego nesse processo e controla por esses fatores nas regressões estimadas.

² O modelo a ser estimado nesse trabalho, portanto, estimará o coeficiente $(y - 1)$ como β_1 , contudo, a interpretação de retornos crescentes da densidade de emprego deve ser feita com cuidado, uma vez que o seu resultado não controla explicitamente pelo coeficiente ρ , que pode ser endógeno e afetar os coeficientes estimados.

Para a estimação do modelo de Fingleton (2003), descrito anteriormente, se faz necessário a construção de algumas variáveis, tais como:

- Taxa salarial ($\ln(w)$): variável dependente do modelo econométrico, calculada como a remuneração média dos trabalhadores ocupados nas atividades econômicas em nível municipal.
- Densidade do emprego ($\ln(E)$): variável utilizada para captar o efeito das economias de aglomeração, sobre a taxa salarial, calculada como o logaritmo da razão, entre o total de trabalhadores ocupados, no conjunto das atividades econômicas de cada município, e sua respectiva área urbana (em Km²).
- Escolaridade (H): variável de controle, representativa do grau de habilidades médias dos trabalhadores, obtida a partir de uma média ponderada na qual se considerou o peso de cada grau de escolaridade onde se distribuem os trabalhadores e o número médio de anos de estudo correspondente a cada uma das faixas de escolaridade. O Quadro A1, do Apêndice, expõe estas ponderações.
- Conhecimento Técnico (CT): no presente estudo, utilizou-se para estimar esta variável a quantidade de indivíduos com doutorado, em relação ao total da população de cada município.
- Defasagem espacial da taxa salarial ($W\ln(w)$): Essa variável tem a função de mensurar a relação entre a taxa salarial de cada município com aquela apresentada em sua vizinhança. Para o cálculo dessa variável foram utilizadas matrizes de pesos espaciais, construídas a partir do *software* OpenGeoda: matrizes padronizadas de contiguidade espacial do tipo *Queen*, a qual inclui todos os pontos comuns na sua definição, ou seja, para definir os vizinhos, são levados em conta as fronteiras e os vértices.

Tendo em vista o objetivo deste estudo, em buscar evidências de densidade de emprego em economias de aglomeração, faz-se necessária a observação da análise descritiva de algumas variáveis. A Tabela A1, constante no Apêndice A deste artigo, indica tal comparação para os dados dos municípios brasileiros.

A análise é feita tanto para o grupo formado por trabalhadores com carteira de trabalho assinada, quanto para o contingente de indivíduos sem esta distinção — incluindo-se nestes os informais, desempregados, militares, funcionários públicos e profissionais liberais —, afim de melhor entender as disparidades encontradas entre os dois grupos. Observa-se uma considerável disparidade entre os dois grupos, especialmente entre as variáveis escolaridade e rendimento, sendo que o grupo com carteira de trabalho assinada tem média maior para ambas observações. Quando se verifica o número de trabalhadores sem carteira de trabalho assinada, nos municípios brasileiros, chega-se à conclusão de que 71,34% encontram-se nesta situação.

A partir destes dados, inicia-se, na próxima seção, com as análises espaciais destes argumentos, verificando a existência de autocorrelação espacial entre os salários e as aglomerações de emprego, nos municípios brasileiros.

Análise Exploratória de Dados Espaciais

O modelo econômico de Fingleton (2003), utilizado na estimação das equações salariais, tem como hipótese o transbordamento dos níveis de eficiências entre regiões geograficamente próximas, sugerindo a potencialização das externalidades advindas da concentração das atividades econômicas, por meio do efeito transbordamento ou contágio. Como documentado em Le Galo e Ertur (2000), a Análise Exploratória de Dados Espaciais (AEDE) fornece medidas globais e locais de autocorrelação espacial, as quais podem ser identificadas através de testes formais de associação espacial como as

estatísticas *I* de Moran global e o *Local Indicator of Spatial Association Indicador (LISA)*. Quanto à detecção da autocorrelação global, será utilizado o *I* de Moran global, um dos testes formais mais empregados na literatura empírica. Especificamente, a estatística *I* de Moran identifica a estrutura de correlação espacial, que melhor descreve os dados, fornecendo, para tanto, um único valor como medida de associação espacial, utilizado para caracterizar toda a região de estudo. Formalmente, o *I* de Moran global é expresso da seguinte forma:

$$I_t = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_{i,t} - \bar{y}_t) (y_{j,t} - \bar{y}_t)}{\sum_{i=1}^n (y_{i,t} - \bar{y}_t)^2}$$

em que n é o número de regiões, w_{ij} são os elementos da matriz de pesos espaciais W , $y_{i,t}$ é a observação na região i no período t e \bar{y}_t é a média das observações entre as regiões no período t .

Como o *I* de Moran global, é composto pela média do produto dos desvios das regiões, em relação à média global; e o seu denominador é uma medida de variabilidade dos desvios, o seu valor normalmente varia entre -1 e 1. Dados com baixa associação espacial resultam em valores próximos a zero. Valores positivos (0 a +1) e negativos (-1 a 0) indicam autocorrelação espacial positiva e negativa, respectivamente, oferecendo uma indicação formal do grau de associação linear, entre o vetor de valores observados em um determinado período de tempo t , e o vetor das médias ponderadas dos valores da vizinhança (PEROBELLI; HADDAD, 2003).

Para tanto, foram realizados testes de *I* de Moran, com a utilização da matriz de pesos espaciais. A matriz utilizada nesse trabalho refere-se à matriz geográfica de contiguidade por convenção rainha (critério Queen)³. Essa matriz, além das fronteiras físicas com extensão diferente de zero, considera os vértices (as quinas), na visualização de um mapa, como contíguos (LeSage, 1999). Em suma, são consideradas as interações espaciais existentes apenas entre municípios que apresentam fronteira comum.

Considerando os municípios brasileiros no ano de 2010, os valores obtidos para a estatística *I* de Moran sugerem a existência de dependência espacial tanto para a taxa salarial quanto para a densidade do emprego. O teste bivariado, por meio do qual é avaliada a autocorrelação espacial entre a densidade do emprego e a média da taxa salarial dos vizinhos de cada município, também indica a presença de municípios com altos ou baixos valores das variáveis em estudo, com vizinhos na mesma condição.

Nas figuras a seguir, a primeira e a segunda destacam, respectivamente, o *I* de Moran univariado para a taxa salarial e a densidade do emprego, enquanto na Figura 3 é enfatizado o teste bivariado. Os valores obtidos para essa estatística encontram-se descritos na parte superior de cada um dos diagramas.

Na Figura 1 verifica-se uma alta autocorrelação espacial na taxa salarial, chegando-se ao *I* de Moran de 0,76.

³ O presente teste utiliza-se da matriz de peso espacial *Queen* = 10. Tal escolha foi definida através do maior e mais significativo valor da estatística *I* de Moran.

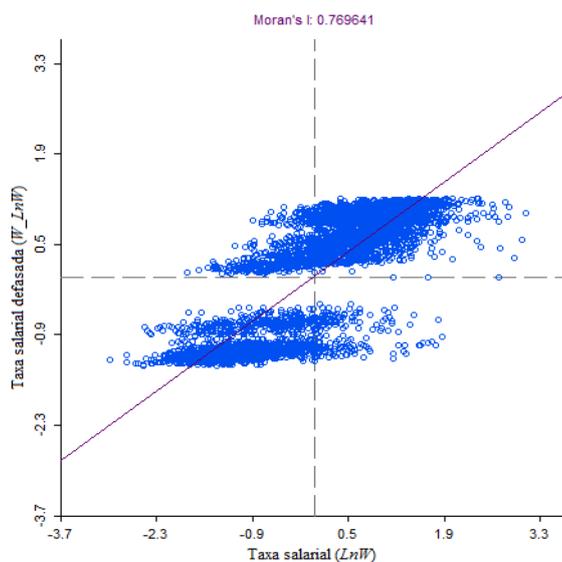


Figura 1 - Diagrama de dispersão de *Moran* da Taxa Salarial-2010

Fonte: Resultados da pesquisa, a partir do Censo 2010, utilizando o *software* OpenGeoda.

Por sua vez, na Figura 2, onde analisa-se a densidade do emprego, o teste tem um valor abaixo do anterior, mas mesmo assim significativo, de 0,56.

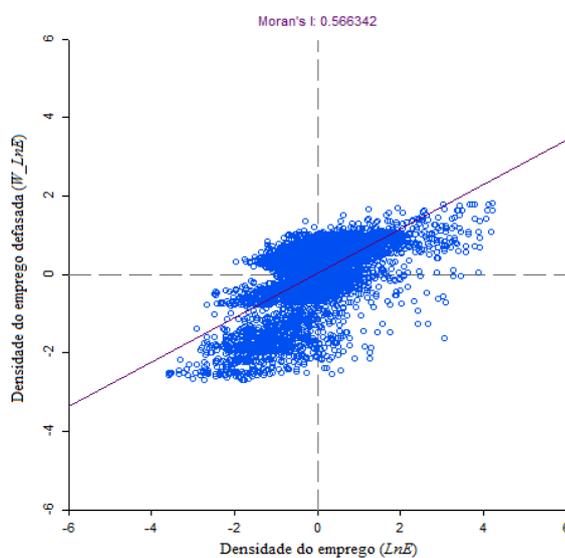


Figura 2 - Diagrama de dispersão de *Moran* da Densidade do Emprego-2010

Fonte: Resultados da pesquisa, a partir do Censo 2010, utilizando o *software* OpenGeoda.

Ao verificarmos o teste bivariado entre taxa salarial e densidade de emprego, na Figura 3, também concluímos haver correlação, porém agora de 0,21.

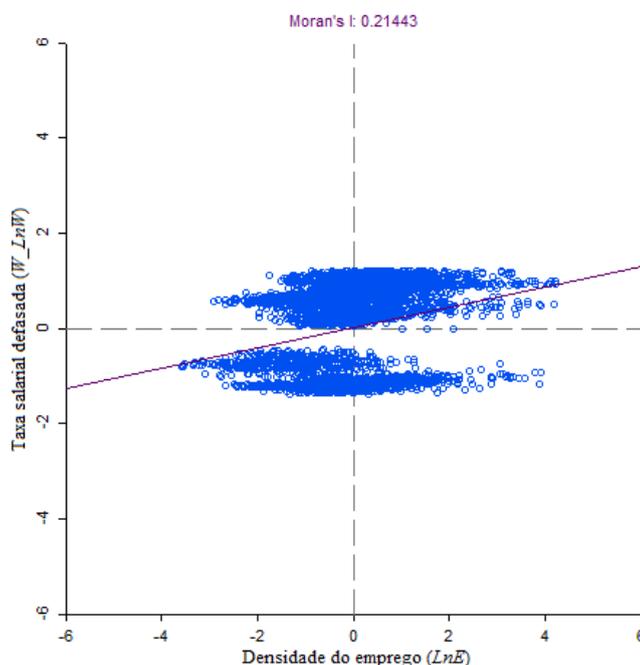


Figura 3- Diagrama de dispersão de *Moran* da Taxa Salarial x Densidade do Emprego-2010

Fonte: Resultados da pesquisa, a partir do Censo 2010, utilizando o *software* OpenGeoda.

Apesar dos diagramas de dispersão de *Moran* indicarem uma tendência à associação espacial positiva, observa-se a existência de pontos nos quadrantes inferiores dos gráficos com direção oposta, ou seja, com associação espacial negativa, os quais sugerem municípios com baixos valores das variáveis em análise, contíguos com municípios com altos valores.

Assim, como ressalta Anselin (1995), o indicador de associação espacial global pode ocultar ou não ser satisfatório na identificação de padrões locais espaciais, como os clusters e outliers. Isso porque, pode haver a necessidade de se examinar padrões espaciais numa escala de desagregação maior, o que, provavelmente, acarretará o surgimento de diferentes regimes de associações espaciais, assim como localidades em que a dependência espacial é mais evidente.

Neste caso, se faz necessária a utilização de indicadores locais de dependência espacial, que possam ser associados a diferentes localizações de uma variável distribuída espacialmente. Sendo assim, a próxima estatística utilizada, o *LISA*, será útil na análise da significância local dos processos espaciais, tanto no que se refere à identificação de clusters espaciais significantes, bem como no diagnóstico de instabilidades locais.

Na análise, da autocorrelação espacial local, será utilizado o indicador local de autocorrelação espacial, *LISA*. Conforme definição de Anselin (1995), a medida *LISA* para cada região *i* e período *t* pode ser expressa da seguinte forma:

$$I_{i,t} = \frac{(y_{i,t} - \bar{y}_t) \sum_{j=1}^n w_{ij} (y_{j,t} - \bar{y}_t)}{m_0}, \text{ em que } m_0 = \frac{\sum_{i=1}^n (y_{i,t} - \bar{y}_t)^2}{n}$$

em que um valor positivo de $I_{i,t}$ indica o agrupamento de valores similares (alto ou baixo), enquanto um valor negativo indica um agrupamento de valores desiguais. Esse diagrama, cuja versão

cartográfica é conhecida como *Moran Map*, representa graficamente a regressão do valor original da variável em análise, sobre o seu valor espacialmente defasado, cujo coeficiente de inclinação é o *I* de *Moran* global. Ele é dividido em quatro tipos de associação espacial: os clusters High-High (HH), região que apresenta alto valor da variável em estudo, circundada por uma vizinhança em que o valor médio da mesma variável também é alto; Low-Low (LL), região de baixo valor na qual a média dos seus vizinhos também é baixa; e os outliers Low-High (LH), região com baixo valor, circunvizinha de uma vizinhança cujo valor médio é alto; e high-low (HL), região com alto valor na qual a média das regiões contíguas é baixa.

Nos testes *LISA*, apresentados nas três figuras a seguir, utiliza-se, ainda, a matriz *W* com $k = 10^4$. Seguindo a mesma sistemática do teste de *Moran*, descreve-se o *LISA* para a taxa salarial na Figura 4, seguida pelo teste da densidade do emprego, na Figura 5, e por último, na Figura 6, pelo teste bivariado *LISA* de ambos.

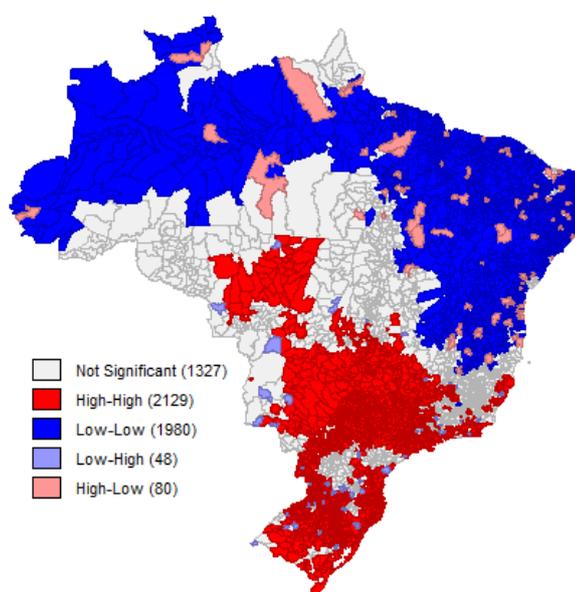


Figura 4- *LISA* da Taxa Salarial-2010

Fonte: Resultados da pesquisa, a partir do Censo 2010, utilizando o *software* OpenGeoda.

No primeiro mapa é possível observar concentrações geográficas de municípios, com associação espacial positiva de altos valores, da variável em análise nas regiões Sul e Sudeste do País e de baixos valores no Norte/Nordeste, indicando uma forte segmentação do território nacional, a semelhança dos resultados encontrados no estudo de Galinari, Lemos e Amaral (2006) e de Amarante (2011) para as taxas salariais observadas.

A Figura 5, que representa o *LISA* da densidade do emprego, mostra que grande parte dos municípios brasileiros não teve valor significativo para este teste em 2010. Porém, os que tiveram autocorrelação alta-alta concentraram-se em regiões perto do litoral brasileiro, do Sul ao Nordeste. Em contrapartida, as regiões com densidade de emprego baixa-baixa localizaram-se nas regiões Norte,

⁴ As utilizações de diferentes matrizes de peso espacial fornecem resultados qualitativos similares.

Centro-Oeste, interior do Nordeste e parte sudoeste do Rio Grande do Sul. Cabe ressaltar que, podem existir distorções nos dados, por conta do tamanho dos municípios do Norte e Nordeste.

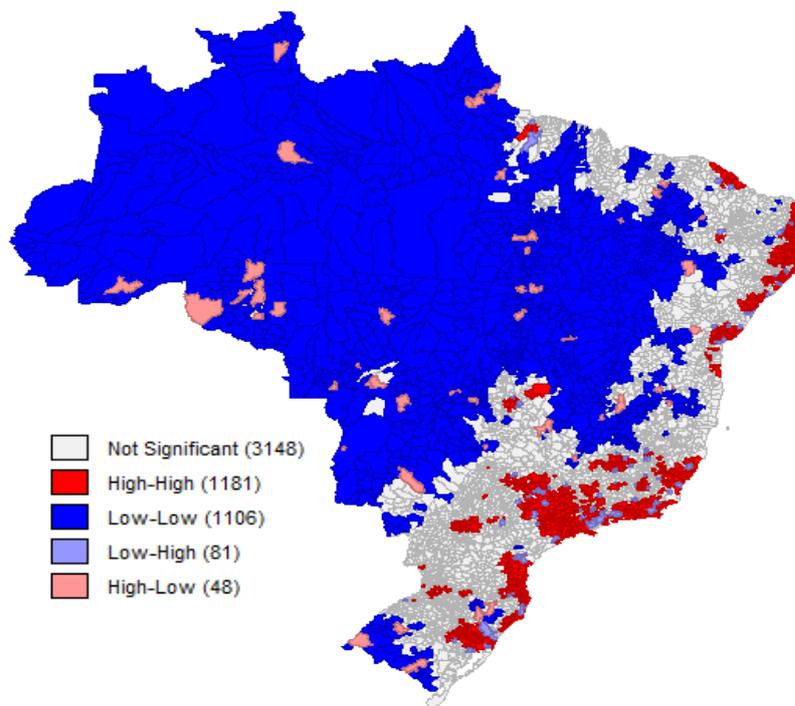


Figura 5- LISA da Densidade do Emprego-2010

Fonte: Resultados da pesquisa, a partir do Censo 2010, utilizando o *software* OpenGeoda.

Como destacado em Lemos et al. (2005), na Análise Exploratória de Dados Espaciais a identificação dos clusters do tipo high-high é bastante relevante, uma vez que expressa a correlação espacial de dois ou mais municípios com elevados valores para uma determinada variável em estudo, sugerindo a existência de transbordamentos e encadeamentos produtivos espaciais, através de complementaridades e integração regional. Já os clusters do tipo low-low são significantes na identificação de regiões que não possuem atividade econômica relevante ou o são excluídos. Por outro lado, os outliers do tipo high-low e low-high revelam a existência de aglomeração localizada em apenas um município, ou uma “ilha” com um entorno de subsistência.

O mapa, destacado na Figura 6, apresenta os resultados para a análise *LISA* bivariada no ano de 2010, por meio da qual se avalia a autocorrelação entre a densidade do emprego e a média da taxa salarial, nos municípios vizinhos. Nesse caso, verifica-se a existência de comunalidades, em que uma das variáveis explicativas para o nível de eficiência produtiva local, a densidade do emprego, se correlaciona positivamente com a produtividade de seus vizinhos.

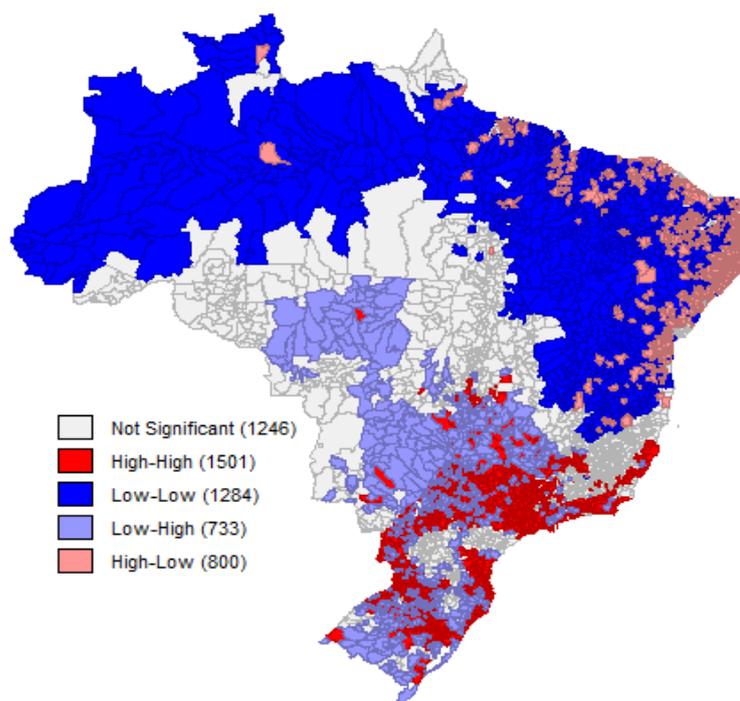


Figura 6- LISA da Densidade do Emprego x Taxa Salarial - 2010

Fonte: Resultados da pesquisa, a partir do Censo 2010, utilizando o *software* OpenGeoda.

Para a aglomeração espacial do tipo alta-alta foram identificadas aglomerações localizadas nas regiões Sudeste, sobretudo no estado de São Paulo, e na região Sul. Os clusters do tipo baixo-baixo foram encontrados principalmente nas regiões Norte/ Nordeste. O terceiro tipo de associação espacial, baixa-alta, foi encontrado distribuída principalmente na região Sul e Centro-Oeste, enquanto o padrão alta-baixa foi verificado, em especial, no litoral do Nordeste brasileiro.

A análise feita sugere a existência de dependência espacial, que será considerada e testada no contexto do modelo que será estimado na próxima seção.

Resultados Econométricos

Na realização das regressões para o modelo proposto por Fingleton (2003), utilizou-se um banco de dados em que, todas as variáveis explicativas e instrumentos defasados espacialmente, foram previamente calculados por meio do *software* espacial OpenGeoda, viabilizando, assim, a estimação dos parâmetros do modelo. Além disso, para cada modelo estimado, foram utilizadas três matrizes de pesos espaciais distintas: *Queen* = 1; *Queen* = 10 e *Queen* = 50.

Inicialmente o modelo foi estimado por Mínimos Quadrados Ordinários (MQO), reportado apenas como base de comparação, assim como para a realização dos testes de autocorrelação espacial nos resíduos, com as matrizes de pesos espaciais de contiguidade. Posteriormente, tendo em vista a presença de variáveis endógenas e/ou omitidas na equação de salários, o modelo foi estimado via Spatial Durbin Model (SDM). Como podemos verificar na Tabela 1, todos os coeficientes são significativos a 5%. No entanto o resultado para o teste de dependência espacial *I de Moran* é altamente significativo para a matriz de peso espacial utilizada, indicando um problema de dependência espacial

nos dados. Nesse caso, as estimativas obtidas por meio de MQO se tornam inconsistentes e viesadas, sugerindo a necessidade de especificações alternativas para os modelos a serem estimados.

Tabela 1- Resultados da regressão por MQO

Variáveis	Matriz de Pesos- Queen = 10
Constante	3,0490*** (0,0250)
Densidade do emprego (LnE)	-0,0214*** (0,0025)
Escolaridade (H)	0,6539*** (0,0058)
Doutores (T)	-21,8501*** (4,8005)
N	5564
Teste F	5961,86
R²	0,7628
I de Moran	z-value = 117,7234 Prob (0,000)
LM (lag)	z-value = 2804,8731 Prob (0,000)
LM Robusto (lag)	z-value = 1178,1648 Prob (0,000)
LM (erro)	z-value = 13457,1714 Prob (0,000)
LM Robusto (erro)	z-value = 11830,4630 Prob (0,000)
LM (sarma)	z-value = 14635,3362 Prob (0,000)

Fonte: Resultados da pesquisa, a partir do Censo 2010, utilizando o *software* OpenGeoda.

Notas: *** estatisticamente significativo ao nível de 1%; ** estatisticamente significativo ao nível de 5%;

* estatisticamente significativo ao nível de 10%. Valores entre parênteses correspondem ao desvio-padrão.

Os testes robustos do Multiplicador de Lagrange (erro e lag) revelam que há tanto auto correlação nos resíduos - um indício de que uma parte da dependência espacial ocorre devido às características não observáveis, que são auto correlacionadas espacialmente - quanto uma dependência espacial na variável dependente. Como ambos os testes são estatisticamente significantes, optamos por seguir o procedimento sugerido por Lesage e Fisher (2008), que consiste em estimar o SDM. Tendo em vista que, o modelo proposto por Fingleton (2003), contém elementos tanto do modelo SAR ($\ln(w) = k_2 + \rho W \ln(w)$), quanto do modelo SDM ($\beta_1 \ln(E) - \beta_2 W \ln(E)$), o presente trabalho modifica o modelo base estimado por Fingleton (2003), acrescentando-lhe as defasagens espaciais aos termos H e T , transformando-o em um modelo totalmente caracterizado na modelagem SDM. A seguir explanam-se os fundamentos do modelo SDM.

Modelo espacial de Durbin (SDM)

Nota-se que, a partir do modelo geral, é possível utilizar a variável WX , isto é, a defasagem espacial das variáveis explicativas. O modelo é conhecido como Modelo Espacial de Durbin (Spatial Durbin Model – SDM) (Anselin, 1999). Este é um modelo de regressão com variáveis espaciais defasadas, dependentes e explicativas, e com um termo de erro não autocorrelacionado.

Segundo Golgher (2012), além dos termos do modelo de lag espacial, o modelo espacial de Durbin (SDM) contém:

$$y = \rho Wy + X\beta - \theta WX\beta + \mu \quad (21)$$

$$\mu \sim N(0, \sigma^2 I_n)$$

Esse modelo inclui a interação endógena, presente no modelo de lag espacial, e a interação exógena, discutida no modelo de erro espacial de Durbin. Manipulando essa equação obtemos o seguinte modelo:

$$y = x\beta + (I - \rho W)^{-1} (x\eta + v),$$

$$(I - \rho W)y = (I - \rho W)x\beta + x\eta + v,$$

$$y = \rho W y + x(\beta + \eta) + Wx(-\rho\beta) + v, \quad (22)$$

$$y = \rho W y + x\beta_1 + Wx\beta_2 + v$$

De modo geral, o modelo de Durbin é um excelente ponto de partida para estimações espaciais, uma vez que os modelos mais simples são apenas casos particulares dele. Outra vantagem desse modelo é que as interações entre vizinhos são levadas em consideração tanto no nível de variáveis endógenas (y defasada) quanto no nível de covariáveis.

O modelo espacial Durbin (SDM-Spatial Durbin Model) é assim chamado em alusão à abordagem de Durbin à regressão com resíduos temporalmente autocorrelacionados. Ele inclui defasamentos espaciais da variável dependente e das variáveis independentes e captura externalidades e spillovers advindos de diferentes fontes. Este modelo incorpora informações da vizinhança da observação i , tanto na variável dependente quanto na variável independente.

Como mostra LeSage e Pace (2009), a desconsideração de uma estrutura do tipo $z = \rho Wz + u$ deixa os estimadores de mínimos quadrados ordinários (MQO) não viesados e consistentes, porém não eficientes. No entanto, a propriedade de não vies e consistência só é válida quando o vetor de erro é não correlacionado com a matriz de variáveis incluídas no modelo, uma hipótese que é demasiadamente restritiva. Por exemplo, no caso da renda, é perfeitamente factível que as variáveis não observáveis (como estado de saúde, os costumes, a cultura, etc.) sejam correlacionadas com as variáveis incluídas no modelo, que são importantes para explicar o nível de renda (como os fatores socioeconômicos e as variáveis de política local), de forma que o termo de erro da equação $y = x\beta + (I_n - \rho W)^{-1}u$ seja dado por $u = x\lambda + v$. Onde, v é um vetor de erros bem-comportados. Combinando as duas últimas equações, acima, obtemos o SDM, uma especificação que considera tanto lag na variável dependente como nas variáveis independentes: $y = \rho W y + x\eta + Wx\phi + v$. Dessa forma, uma estrutura do tipo SDM consegue resolver o problema de variáveis omitidas que são espacialmente correlacionadas, produzindo estimadores consistentes e não viesados. Outra vantagem, desta especificação, é que ela é um caso geral das demais especificações espaciais, SAR e SEM.

O modelo SDM ainda tem a vantagem de flexibilizar a equação original de Fingleton (2003), uma vez que no SDM a dependência espacial da densidade do emprego depende apenas da estrutura de dependência espacial da variável de densidade de emprego.

Portanto, aplicando o SDM à equação (20), o modelo apresenta-se na seguinte especificação:

$$\ln(w) = k_2 + \rho W \ln(w) + \beta_1 \ln(E) - \beta_2 W \ln(E) + a_1 H - a_1 W H + a_2 T - a_2 W T + v \quad (23)$$

$$v \sim N(0, \pi^2)$$

Utilizando-se o modelo SDM para a Equação (23), chega-se aos resultados das Tabelas 2, 3 e 4, descritas a seguir, que usam, respectivamente, matrizes *Queen* com valores 1, 10 e 50.

Observa que, para todos os pesos espaciais utilizados, à exceção do conhecimento técnico, os coeficientes estimados para todas as variáveis explicativas do modelo são significativos e com os sinais esperados. Os resultados encontrados para o coeficiente da defasagem espacial da taxa salarial, $Wln(w)$, positivos e significativos, apontam forte presença dos *spillovers* geográficos, ou seja, a presença de auto correlação espacial da taxa salarial entre municípios vizinhos, ressaltando, assim, a importância dos transbordamentos dos níveis de eficiência entre as regiões.

Como no modelo de Fingleton (2003), o coeficiente auto regressivo, (ρ), está presente tanto na equação representativa dos níveis de eficiência do trabalho local, como na equação dos salários. Tais coeficientes estimados são representativos dos efeitos de contágio entre a produtividade dos municípios brasileiros, geograficamente próximos, indicando que os níveis de eficiência dos trabalhadores, assim como a taxa salarial é positivamente relacionada com aquela observada em sua vizinhança, ou seja, a taxa salarial dos municípios brasileiros é afetada de forma significativa, pela taxa salarial de seus vizinhos.

Como destacado em Galinari (2006), quanto a sua natureza, os *spillovers* mensurados podem ser considerados um mix das externalidades tecnológicas e pecuniárias, uma vez que a proximidade entre firmas pode proporcionar benefícios associados aos custos de transporte e, ainda, em menor escala, pode-se esperar que as firmas possam se beneficiar da oferta e da demanda geradas pelas cidades vizinhas.

Verifica-se que, além de positivos e significativos, os valores obtidos para os coeficientes da defasagem espacial da taxa salarial, os quais variam entre 0,28 e 0,42, são bastante parecidos aos encontrados em Fingleton (2003); Galinari (2006); Galinari, Lemos e Amaral (2006). Evidenciando assim, que os transbordamentos dos níveis de eficiência tendem a ser mais elevados, quando se considera na análise o conjunto de atividades econômicas, bem como as características do mercado de trabalho ao nível de agregação municipal, tais como salário, emprego e escolaridade.

Outra variável de suma importância é a densidade (E), que, como explicada por Galinari (2006) e Galinari *et al.* (2006), tem o objetivo de medir como as taxas de salários modificam-se, quando expostas às variações das aglomerações populacionais. Para Falcão e Neto (2007), espera-se que os trabalhadores menos qualificados se beneficiem duplamente da maior concentração de capital humano. Sua produtividade e, assim, salário seriam mais elevados por conta de potenciais externalidades positivas de capital humano e/ou também devido à substituição imperfeita em relação à mão-de-obra qualificada. Já os trabalhadores mais escolarizados teriam seus salários afetados positivamente pelas externalidades de capital humano, mas negativamente devido à maior oferta de mão-de-obra qualificada no local. O efeito, da maior parcela de qualificados no município, sobre seus salários seria ambíguo e dependeria da magnitude das externalidades.

Nas Tabelas, a seguir, verifica-se que praticamente todos os coeficientes defasados espacialmente são significativos a 5%, a exceção da variável defasada do nível técnico (T) na Tabela 2; de H e T na Tabela 3 e de T na Tabela 4. Quando utiliza-se a matriz *Queen* = 1, na Tabela 2, o resultado encontrado, de 0,35, para o coeficiente da defasagem espacial da taxa salarial (W_LnW), aponta presença dos *spillovers* geográficos, ou seja, a presença de autocorrelação espacial da taxa

salarial entre municípios vizinhos. A densidade defasada (W_LnE) para esta matriz foi significativa, porém negativa, em $-0,03$.

Tabela 2- Resultados da regressão por SDM – *Queen* = 1

Variáveis	Matriz de Pesos- <i>Queen</i> = 1
Constante	1,4695*** (0,0385)
Taxa salarial defasada (W_LnW)	0,3522*** (0,0107)
Densidade do emprego (LnE)	0,0021 (0,0022)
Densidade do emprego defasada (W_LnE)	-0,0349*** (0,0037)
Escolaridade (H)	0,3984*** (0,0062)
Escolaridade defasada (W_LnH)	0,1381*** (0,0118)
Doutores (T)	20,5244*** (3,7133)
Doutores defasada (W_LnT)	-2,4919 (19,6954)
N	5564
R²	0,8633
Teste Breusch-Pagan	z-value = 4728,223 Prob (0,000)
Teste Likelihood Ratio	z-value = 990,072 Prob (0,000)

Fonte: Resultados da pesquisa, a partir do Censo 2010, utilizando o *software* OpenGeoda.

Notas: *** estatisticamente significativo ao nível de 1%; ** estatisticamente significativo ao nível de 5%;

* estatisticamente significativo ao nível de 10%. Valores entre parênteses correspondem ao desvio-padrão.

Já na Tabela 3, que utiliza-se da matriz de peso espacial *Queen* = 10, a Taxa salarial defasada (W_LnW) aumenta para 0,42 e continua significativa. A densidade defasada continua significativa e negativa, praticamente estável.

Tabela 3- Resultados da regressão por SDM – *Queen* = 10

Variáveis	Matriz de Pesos- <i>Queen</i> = 10
Constante	1,3851*** (0,0867)
Taxa salarial defasada (<i>W_LnW</i>)	0,4211*** (0,0355)
Densidade do emprego (<i>LnE</i>)	0,0084*** (0,0023)
Densidade do emprego defasada (<i>W_LnE</i>)	-0,0382*** (0,0042)
Escolaridade (<i>H</i>)	0,4625*** (0,0062)
Escolaridade defasada (<i>W_LnH</i>)	0,0010 (0,0314)
Doutores (<i>T</i>)	9,5514** (4,0021)
Doutores defasada (<i>W_LnT</i>)	32,4664 (21,4614)
N	5564
R²	0,8393
Teste Breusch-Pagan	z-value = 2126,095 Prob (0,000)
Teste Likelihood Ratio	z-value = 208,7771 Prob (0,000)

Fonte: Resultados da pesquisa, a partir do Censo 2010, utilizando o *software* OpenGeoda.

Notas: *** estatisticamente significativo ao nível de 1%; ** estatisticamente significativo ao nível de 5%;

* estatisticamente significativo ao nível de 10%. Valores entre parênteses correspondem ao desvio-padrão.

Ao utilizar-se a matriz *Queen* = 50 para a regressão, o resultado do salário defasado continua significativo, porém diminui significativamente em relação às outras matrizes estudadas, chegando ao valor de 0,28. A densidade defasada, por sua vez, continua a apresentar resultados significativos, porém agora mais negativa: - 0,04.

Tabela 4 - Resultados da regressão por SDM – *Queen* = 50

Variáveis	Matriz de Pesos- <i>Queen</i> = 50
Constante	1,5210*** (0,0553)
Taxa salarial defasada (<i>W_LnW</i>)	0,2879*** (0,0186)
Densidade do emprego (<i>LnE</i>)	0,0082*** (0,0023)
Densidade do emprego defasada (<i>W_LnE</i>)	-0,0419*** (0,0041)
Escolaridade (<i>H</i>)	0,4585*** (0,0062)
Escolaridade defasada (<i>W_LnH</i>)	0,1506*** (0,0184)
Doutores (<i>T</i>)	10,9501*** (3,9915)
Doutores defasada (<i>W_LnT</i>)	31,1244 (21,4435)
N	5564
R²	0,8402
Teste Breusch-Pagan	z-value = 1839,599 Prob (0,000)
Teste Likelihood Ratio	z-value = 2946,437 Prob (0,000)

Fonte: Resultados da pesquisa, a partir do Censo 2010, utilizando o *software* OpenGeoda.

Notas: *** estatisticamente significativo ao nível de 1%; ** estatisticamente significativo ao nível de 5%;

* estatisticamente significativo ao nível de 10%. Valores entre parênteses correspondem ao desvio-padrão.

Em geral, os valores encontrados para H e T são positivos, porém não significativos para T , corroborando, assim com o estudo de Amarante (2011). A variável que mais destoa dos trabalhos de Fingleton (2003); Galinari (2006) e de Amarante (2011) é a que representa a densidade do emprego - E . Seu valor foi positivo, porém menor aos encontrados nos estudos citados. E quando a tomamos pela sua variável defasada, essa se torna com sinal negativo. Esta diferença pode ser explicada pelo fato que este artigo utiliza a área total dos municípios para gerar a variável E , enquanto que os demais estudos mencionados utilizam apenas a área urbana dos municípios brasileiros. Assim sendo, levando em consideração a grande área de terra no Brasil que não é urbanizada, encontrar sinais como, os encontrados neste estudo, para E , evidencia a forte correlação entre a taxa salarial e as aglomerações de empregos existentes no país.

Ao compararmos os resultados encontrados no presente estudo, com o desenvolvido por Amarante (2011), percebemos que seus resultados para a defasagem espacial, para o ano de 2000, variam entre 0,69 e 0,76; e ao se analisar os resultados encontrados para 2009, os valores ficam entre 0,66 e 0,82, uma diferença substancial entre os dois estudos. Ao utilizar-se de dados da RAIS, Amarante (2011) agrega informações somente do mercado formal brasileiro, sendo assim plausível que seus resultados sejam mais elevados dos que encontrados no atual artigo, tendo em vista que neste utilizam-se dados do Censo 2010, a fim de verificar não somente o mercado formal, mas também o mercado informal existente no país, além dos profissionais liberais. Outra diferenciação deste trabalho para o de Amarante (2011), é a metodologia de estimação utilizada.

Os resultados obtidos corroboram as evidências encontradas em estudos que utilizam metodologia semelhante para a identificação das economias de aglomeração, tanto na literatura nacional quanto internacional.

Considerações Finais

Os resultados encontrados sugerem aglomerações geográficas de municípios, com associação espacial positiva de altos valores, das variáveis em análise, na região Centro-Sul do País e de baixos valores no Norte-Nordeste, indicando uma forte segmentação do território nacional. Os testes de *I de Moran* apresentaram valores de 0,76 e 0,56 para a taxa salarial e a densidade do emprego, respectivamente; enquanto que o teste bivariado retornou o valor de 0,21, confirmando a forte autocorrelação existente entre tais variáveis.

Os resultados obtidos, a partir da técnica de variáveis instrumentais, por SDM, apontam que as variações nas taxas salariais, dos municípios brasileiros, são significativamente e positivamente relacionadas com a concentração espacial, da atividade econômica, mensurada a partir da densidade do emprego; e com os transbordamentos dos níveis de eficiência entre áreas geograficamente próximas.

Portanto, as variações na taxa salarial dos municípios brasileiros no ano de 2010 podem estar associadas à existência de economias de aglomeração, a presença de spillovers ou transbordamentos espaciais entre a taxa salarial dos municípios geograficamente próximos e aos níveis de eficiência dos trabalhadores.

Sendo assim, como resultados práticos deste estudo, espera-se ratificar as ideias sobre economia regional e urbana, de que as questões espaciais possuem relevância na determinação do desenvolvimento dos países, auxiliando a proposição de políticas públicas, especialmente no que se refere à redução das desigualdades regionais. Estudos futuros podem utilizar um horizonte temporal

maior, verificando, por exemplo, as mudanças ocorridas ao longo de dois anos distintos. Além disso, as atividades econômicas poderiam ser desagregadas em diferentes setores, no sentido de analisar se as economias de aglomeração são específicas a determinadas atividades produtivas e em quais setores seus efeitos são mais intensos.

Referências

- ALMEIDA, E. **Econometria Espacial**. Campinas: Alínea, 2012.
- AMARANTE, P. A. **Economias de Aglomeração nas Atividades Econômicas dos Municípios Brasileiros nos Anos de 2000 e 2009-Evidências a Partir de Equações Salariais**. Dissertação. (Mestrado) – Programa de Pós-Graduação em Economia, Universidade Federal da Paraíba, João Pessoa, 2011.
- ANSELIN, L. Local indicators of spatial association – LISA. **Geographical Analysis**, v.27, n.2, p. 91-115, 1995.
- ANSELIN, L. **Spatial econometrics**. Dallas: Bruton Center, 1999.
- BANDEIRA, L.; JACINTO, P. de A. Um estudo sobre retorno em escolaridade para a região metropolitana de Porto Alegre. In: XIII ENCONTRO DE ECONOMIA DA REGIÃO SUL, 2010, Porto Alegre. **Anais**: ANPEC/Sul, Porto Alegre, 2010.
- BARROS, A. R. **Desigualdades regionais no Brasil: natureza, causas, origens e soluções**. Rio de Janeiro: Campus, 2011.
- FALCÃO, N. de A.; NETO, R. da M. S. Concentração Espacial De Capital Humano e Externalidades: O Caso Das Cidades Brasileiras. In: **Anais do XXXV ENCONTRO NACIONAL DE ECONOMIA**. ANPEC- Associação Nacional dos Centros de Pós-graduação em Economia, 2007.
- FINGLETON, B. Increasing returns: evidence from local wage rates in Great Britain. **Oxford Economic Papers**, v.55, p.716-739, 2003.
- GALINARI, R. **Retornos crescentes urbano-industriais e spillovers espaciais: evidências a partir da taxa salarial no estado de São Paulo**. 2006. 162 f. Dissertação (Mestrado) - Programa de Pós-Graduação em Economia, Universidade Federal de Minas Gerais, CEDEPLAR, Minas Gerais, 2006.
- GALINARI, R.; LEMOS, M. B.; AMARAL, P. Retornos crescentes urbanos: a influência do espaço na diferenciação da taxa salarial no Brasil. In: DE NEGRI, J. A. *et al.* **Tecnologia, exportação e emprego**. Brasília: IPEA, 2006. Cap.8, p.203-248.
- GOLGHER, A. B. The selectivity of migration and poverty traps in rural Brazil. **Population Review**, v. 51, p. 9-27, 2012.
- JACOBS, J. **The Economy of Cities**. New York-United States: Vintage, 1969.
- LANGONI, C. G. Distribuição da renda e desenvolvimento econômico no Brasil: uma reafirmação. **Ensaio Econômico EPGE**, Rio de Janeiro, n. 7, 1973.
- LE GALLO, J.; ERTUR, C. **Exploratory spatial data analysis of the distribution of regional per capita GPD in Europe, 1980-1995**. Dijon: University of Burgundy, 2000.
- LEMOS, M. B. *et al.* A organização territorial da indústria no Brasil. In: DE NEGRI, J. A., SALERNO, S. S. (Orgs.) **Inovações, padrões tecnológicos e desempenho das firmas industriais brasileiras**. Brasília: IPEA, 2005. v.1, p.325-364.

LESAGE, J. P. **Spatial econometrics**. Toledo: University of Toledo, 1999.

LESAGE, J. P.; FISCHER, M. M. Spatial growth regressions: model specification, estimation and interpretation. **Spatial Economic Analysis**, v. 3, n. 3, p. 275-304, 2008.

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

MINCER, J. A. Investment in human capital and personal income distribution. **The Journal of Political Economy**, Chicago, v. 66, n. 4, p. 281-302, 1958.

MINCER, J. A. Schooling and earnings. In: ____. **Schooling, experience, and earnings**. Nova Iorque: Columbia University Press, 1974. p. 41-63.

MONASTÉRIO, L.; DAMÉ, O.; SALVO, M. Estrutura espacial das aglomerações e determinação dos salários industriais no Rio Grande do Sul. **Ensaio FEE**, v. 28, p. 801-824, 2008.

PEROBELLI, F.; HADDAD, E. Brazilian interregional trade (1985-1996): an exploratory spatial data analysis. In: **Anais...ENCONTRO NACIONAL DE ECONOMIA**, 31, 2003, Porto Seguro: ANPEC, 2003.

RODRIGUES, A. A. F. **Impacto da educação no rendimento salarial no Brasil de 2001 a 2008**. Dissertação (Mestrado) – Programa de Pós-Graduação em Economia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2010.

VIEIRA, R. de S. **Crescimento Econômico no Estado de São Paulo: uma análise espacial**. São Paulo: Fundação Editora Unesp, 2010.

VIANA, G. ; LIMA, J. F. Capital humano e crescimento econômico. **Interações**, v. 11, n. 2, p. 137-148, 2010.

APÊNDICE A

Quadro A1 - Anos adotados para as variáveis Curso que frequenta e Curso mais elevado que frequentou do Censo 2010

Curso mais elevado que frequentou	Anos Adotados
01- Creche; Pré-escolar; Classe de alfabetização- CA	0
02- Alfabetização de Jovens e Adultos	1
03- Antigo Primário (Elementar)	4
04- Antigo Ginásio (Médio- 1º ciclo)	4
05- Regular de Ensino Fundamental ou do 1º Grau	4
06- Supletivo do Ensino Fundamental ou do 1º Grau	4
07- Antigo Científico, Clássico, etc. (Médio- 2º ciclo)	9
08- Regular ou Supletivo do Ensino Médio ou do 2º Grau	9
09- Graduação	12
10- Especialização (mínimo de 360 horas)	16
11- Mestrado	17
12- Doutorado	21
Curso que frequenta	Anos Adotados
01- Creche	0
02- Pré-escolar (maternal e jardim de infância)	0
03- Classe de alfabetização- CA	1
04- Alfabetização de Jovens e Adultos	1
05- Regular de Ensino Fundamental	4
06- Educação de Jovens e Adultos- EJA ou Supletivo do Ensino Fundamental	4
07- Regular do Ensino Médio	9
08- Educação de Jovens e Adultos- EJA ou Supletivo de Ensino Médio	9
09- Graduação	12
10- Especialização (mínimo de 360 horas)	16
11- Mestrado	17
12- Doutorado	21

Fonte: Elaboração dos autores a partir do Censo 2010.

Tabela A1 – Estatística descritiva dos municípios brasileiros – Censo 2010

Variáveis	Com carteira de trabalho assinada		Sem carteira de trabalho assinada	
Escolaridade	Média	6,92		4,65
	Mínimo	2,44		2,12
	Máximo	10,89		9,25
	Desvio Padrão	0,94		0,88
Rendimento	Média	716,98		477,03
	Mínimo	211,01		94,66
	Máximo	5022,09		2960,85
	Desvio Padrão	283,07		257,40
Idade	Média	38,89		43,73
	Mínimo	31,45		31,26
	Máximo	51,07		51,93
	Desvio Padrão	2,15		2,55
Raça	Proporção Brancos/ Não	0,48		0,46
	Brancos			
	Desvio Padrão	0,24		0,25
Sexo	Proporção Homens/	0,72		0,63
	Mulheres			
	Desvio Padrão	0,10		0,12
Observações		5565		5565

Fonte: Resultados da pesquisa, a partir do Censo 2010.