

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
FACULDADE DE ADMINISTRAÇÃO, CONTABILIDADE E ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA DO DESENVOLVIMENTO

ANDERSON BONETTO CARRARO

**ENSAIOS SOBRE ECONOMIAS DE AGLOMERAÇÃO E MAUP
NO BRASIL: EVIDÊNCIAS A PARTIR DE EQUAÇÕES DE
RENDIMENTO**

Porto Alegre
2015

ANDERSON BONETTO CARRARO

**ENSAIOS SOBRE ECONOMIAS DE AGLOMERAÇÃO E MAUP
NO BRASIL: EVIDÊNCIAS A PARTIR DE EQUAÇÕES DE
RENDIMENTO**

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre pelo Programa de Pós-Graduação da Faculdade de Economia da Pontifícia Universidade Católica do Rio Grande do Sul.

Orientador: Prof. Dr. Paulo de Andrade Jacinto
Coorientador: Prof. Dr. Túlio Antônio Cravo

Porto Alegre
2015

Dados Internacionais de Catalogação na Publicação (CIP)

C313e Carraro, Anderson Bonetto
Ensaio sobre economias de aglomeração e MAUP no Brasil:
evidências a partir de equações de rendimento / Anderson Bonetto
Carraro. – Porto Alegre, 2015.
77 f.

Dissertação (Mestrado em Economia do Desenvolvimento) –
Faculdade de Administração, Contabilidade e Economia, PUCRS.
Orientação: Prof. Dr. Paulo de Andrade Jacinto.

1. Economia – Brasil. 2. Salários – Brasil. 3. Economia de
Aglomeração. 4. Equações Salariais. 5. MAUP. I. Jacinto, Paulo
de Andrade. II. Título.

CDD 331.2981

Ficha Catalográfica elaborada por Ramon Ely – CRB10/2165

Anderson Bonetto Carraro

**Ensaio Sobre Economias de Aglomeração e Maup no
Brasil: Evidências a Partir de Equações de Rendimento.**

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Economia do Desenvolvimento, pelo Programa de Pós-Graduação em Economia, da Faculdade de Administração, Contabilidade e Economia, da Pontifícia Universidade Católica do Rio Grande do Sul.

Aprovado em 30 de março de 2015.

BANCA EXAMINADORA:



Prof. Dr. Paulo de Andrade Jacinto
Presidente da Sessão



Prof. Dr. Hudson da Silva Torrent



Prof. Dr. Gustavo Inacio de Moraes



Prof. Dr. Osmar Tomaz de Souza
Coordenador do Programa de Pós-Graduação em Economia

AGRADECIMENTOS

A Pontifícia Universidade Católica do Rio Grande do Sul- PUCRS, pela excelência de ensino.

Ao corpo docente da Pós-Graduação em Economia do Desenvolvimento da PUCRS, pelo empenho à formação de seus alunos.

Ao Prof. Dr. Túlio Antônio Cravo, pelo incondicional auxílio, disponibilidade e paciência para a construção deste estudo.

Ao Prof. Dr. Paulo de Andrade Jacinto, pelo apoio à construção desta dissertação.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior- Capes, pela bolsa de estudos disponibilizada para a realização deste mestrado.

A minha família.

A Deus por esta oportunidade.

RESUMO

O objetivo dessa dissertação, composta por dois artigos, é analisar a influência que as economias de aglomeração e o Problema da Unidade de Área Modificável (MAUP) têm sobre o nível salarial. Para tanto, dois artigos são necessários para melhor explanação destes dados. O primeiro é dedicado a explorar as economias de aglomerações a partir das equações de rendimentos. Por meio da Análise Exploratória de Dados Espaciais (AEDE) e dos dados do Censo 2010, verifica-se a existência de associações espaciais, nas agregações regionais analisadas, tais como aglomerações espaciais ou regiões homogêneas (clusters), e observações atípicas (outliers). Na segunda parte do artigo, dedicada à estimação do modelo econométrico, utilizam-se equações salariais, com formulação baseada no modelo desenvolvido por Fingleton (2003), cuja principal hipótese aponta para uma relação positiva entre os diferenciais de salários e a densidade de trabalho. No segundo artigo, apresenta-se um caso de MAUP, com dados do Censo 2010 do IBGE, sobre o salário médio dos trabalhadores brasileiros. É discutida e demonstrada a necessidade de se considerarem as diferenças nos resultados obtidos, oriundas da regionalização escolhida pelo pesquisador. A análise se centra em cinco níveis: microdados, municípios, microrregiões, mesorregiões e estados brasileiros, utilizando, para tanto, equações de rendimentos (mincerianas), propostas por Mincer (1974). Ao se analisar equações mincerianas, sob a perspectiva de diferentes escalas regionais, ponto inédito para o Brasil, o papel de suas variáveis torna-se de suma importância, com destaque para a educação. Os resultados mostram, que as conclusões obtidas e as políticas adotadas, a partir de regressões de rendimento, são dependentes da escolha do nível geográfico analisado.

Palavras chave: Economias de aglomeração. Equações salariais. MAUP.

ABSTRACT

The purpose of this dissertation consists of two articles, it is to analyze the influence of agglomeration economies and the Problem of Modifiable Area Unit (MAUP) have on the salary level. For this, two articles are needed to better explanation of these data. The first is dedicated to exploit economies of agglomeration from the earnings equations. Through the Exploratory Spatial Data Analysis (ESDA) and the 2010 Census data, there is the existence of spatial associations in the regional aggregates analyzed, such as spatial agglomerations or homogeneous regions (clusters), and atypical observations (outliers). In the second part of the article dedicated to the estimation of the econometric model, use up wage equations with formulation based on the model developed by Fingleton (2003), whose main hypothesis points to a positive relationship between wage differentials and job density. In the second article, we present a case of MAUP with the IBGE Census 2010 data on the average salary of Brazilian workers. It is discussed and demonstrated the need to consider the differences in results arising from the regionalization chosen by the researcher. The analysis focuses on five levels: micro, municipalities, micro, meso and Brazilian states, using for this purpose, earnings equations (Mincer), proposed by Mincer (1974). When analyzing Mincer equations, from the perspective of different regional scales, a first point to Brazil, the role of its variables becomes very important, especially for education. The results show that the conclusions drawn and the policies adopted, from income regressions, are dependent on the choice of geographically analyzed.

Keywords: Agglomeration economies. Wage equations. MAUP.

FIGURAS

Figura 1 - Diagrama de dispersão de Moran da Taxa Salarial-2010.....	32
Figura 2 - Diagrama de dispersão de Moran da Densidade do Emprego-2010.....	33
Figura 3 - Diagrama de dispersão de Moran da Taxa Salarial x Densidade do Emprego-2010.....	34
Figura 4 - LISA da Taxa Salarial-2010.....	36
Figura 5 - LISA da Densidade do Emprego-2010.....	37
Figura 6 - LISA da Densidade do Emprego x Taxa Salarial – 2010.....	38
Figura 7 - Escalas Espaciais no Brasil.....	61

TABELAS

Tabela 1 - Estatística descritiva dos municípios brasileiros – Censo 2010.....	27
Tabela 2 - Resultados da regressão por MQO.....	40
Tabela 3 - Resultados da regressão por SDM – Queen = 1.....	45
Tabela 4 - Resultados da regressão por SDM – Queen = 10.....	46
Tabela 5 - Resultados da regressão por SDM – Queen = 50.....	47
Tabela 6 - Estatística descritiva dos trabalhadores com carteira de trabalho assinada por agregação geográfica- Censo 2010.....	62
Tabela 7 - Estatística descritiva dos trabalhadores sem carteira de trabalho assinada por agregação geográfica- Censo 2010.....	63
Tabela 8 - Resultados da regressão por MQO do rendimento per capita, por agregação regional – Modelo (1).....	64
Tabela 9 - Resultados da regressão por MQO do rendimento per capita, por agregação regional – Modelo (2).....	65

QUADROS

Quadro 1 - Anos adotados para as variáveis Curso que frequenta e Curso mais elevado que frequentou do Censo 2010.....	29
Quadro 2 - Estudos com abordagem minceriana para o Brasil.....	55

SUMÁRIO

1 INTRODUÇÃO.....	11
2 ECONOMIAS DE AGLOMERAÇÃO NO BRASIL: EVIDÊNCIAS A PARTIR DE EQUAÇÕES DE RENDIMENTOS.....	13
2.1 Introdução.....	13
2.2 Revisão bibliográfica.....	14
2.2.1 Equação de Rendimentos (Minceriana).....	14
2.2.2 Economias de Aglomeração a partir dos Efeitos sobre a Taxa Salarial.....	19
2.3 .Procedimentos Metodológicos.....	21
2.3.1 Modelo Econométrico: Salários e Aglomerações.....	21
2.3.2 Matriz de Peso Espacial.....	26
2.3.3 Base de Dados e Variáveis.....	26
2.4 Análise exploratória de dados espaciais.....	30
2.4.1 Municípios.....	31
2.5 Resultados econométricos.....	39
2.5.1 Modelo Espacial de Durbin (SDM).....	40
2.6 Considerações finais.....	48
3. EQUAÇÕES DE RENDIMENTOS E MAUP NO BRASIL: EVIDÊNCIAS A PARTIR DO CENSO 2010.....	50
3.1. Introdução.....	50
3.2 Revisão da literatura.....	52
3.1.1 Equação Minceriana em estudos para o Brasil.....	54
3.2.2 O Problema da Unidade de Área Modificável – MAUP.....	57
3.3. Metodologia.....	59
3.4 Dados.....	61
3.5 Resultados empíricos.....	63
3.6 Considerações finais.....	66
4. CONCLUSÃO.....	68
REFERÊNCIAS.....	70
APÊNDICE.....	76

1 INTRODUÇÃO

A presente dissertação tem o objetivo de analisar a relação da equação de salários com a dimensão espacial. Existem vários trabalhos na literatura sobre equação de rendimentos no Brasil, contudo a relação com a dimensão espacial é pouco explorada. Dessa forma, o presente estudo busca contribuir para essa discussão incorporando de maneira explícita a dimensão espacial na discussão da equação de salários. O trabalho é composto de dois ensaios que estudam diferentes aspectos dessa relação.

O primeiro artigo dessa dissertação tem como propósito analisar as economias de aglomerações a partir das equações de rendimentos da economia brasileira, com base nos dados do Censo de 2010, disponibilizados pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Durante as duas últimas décadas, uma crescente difusão de técnicas de econometria espacial tem sido observada, entre economistas e pesquisadores, em diversas áreas. Tais técnicas são empregadas para analisar a influência da densidade, em uma equação de salário modificada, visando incorporar aspectos de transbordamento regional. Por meio da Análise Exploratória de Dados Espaciais (AEDE), pretende-se verificar a existência de associações espaciais, sobre as taxas salariais, nas agregações regionais analisadas, tais como aglomerações espaciais ou 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 microeconômico desenvolvido por Fingleton (2003), cuja principal hipótese aponta para uma relação positiva entre os diferenciais de salários e a densidade de trabalho.

Segundo Fingleton (2003), Amarante (2011) obteve 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. A autora utilizou-se de 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.

A limitação, contudo, no estudo de Amarante (2011) 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. Esta diferenciação traz, aos resultados encontrados, uma interpretação mais fidedigna, da realidade salarial brasileira.

No segundo artigo deste trabalho, o objetivo é verificar a presença do Problema de Unidade de Área Modificável (MAUP), no Brasil, para equações de rendimentos (minceriana). A escolha da escala espacial para análise é uma questão problemática na pesquisa aplicada. Neste sentido, o estudo busca investigar em que medida as ambiguidades da escala espacial podem prejudicar, ou informar, o entendimento dos determinantes de nível salarial da população brasileira.

Diferentes desenhos das unidades implicam em diferentes análises, inferências ou mesmo políticas. A situação de um município pobre, em uma área rica, deixa de ser corretamente avaliada quando o nível de agregação espacial é a região, e, assim, a própria delimitação dos recortes regionais, altera os resultados dos indicadores estatísticos.

Sabe-se que, na presença do MAUP, os resultados podem variar de acordo com a escala geográfica, e a sistematização dos diferentes resultados, para as diferentes escalas geográficas, é muito importante para a definição e avaliação de políticas regionais. Segundo Wu e Cutter (2011), haverá perda de informações durante o processo de agregação, o que geralmente impõe viés de agregação e distorcem os resultados. Em geral, o viés de agregação será crescente, na medida em que os agentes, dentro da unidade de agregação, forem mais heterogêneos.

Com exceção de Resende (2011; 2013), os estudos até agora têm apenas investigado os determinantes do crescimento econômico em uma única escala espacial para inferir a consistência de modelos de crescimento espaciais com a realidade. No presente trabalho, analisa-se não apenas uma escala geográfica, mas sim cinco: microdados, municípios, microrregião, mesorregião e estados; além de tratar de um assunto novo para MAUP: a verificação da equação minceriana para o Brasil.

Compreender como uma determinada região comporta-se, e quais suas características, é de suma importância para traçar políticas públicas. Os estudos, para políticas regionais, devem considerar os resultados com base em várias escalas espaciais, e características locais, visando a tomada de decisões políticas, mais eficazes.

A presente dissertação está dividida em quatro seções, incluindo esta introdução. A seguir estão os dois artigos, nas seções dois e três, respectivamente, seguidos da conclusão deste trabalho.

2. ECONOMIAS DE AGLOMERAÇÃO NO BRASIL: EVIDÊNCIAS A PARTIR DE EQUAÇÕES DE RENDIMENTOS

2.1 Introdução

Uma das questões amplamente discutida na teoria econômica diz respeito à concentração de pessoas e atividades econômicas em um determinado espaço geográfico.

Considerando-se as forças aglomerativas, a atratividade dos espaços urbanos está associada tanto às oportunidades oferecidas, tais como variedades de produtos, oferta de trabalho e serviços públicos, quanto à existência de economias de escala urbana ou economias de aglomeração provenientes de externalidades positivas geradas pela concentração de empresas e pessoas em alguns lugares específicos, a despeito das ineficiências tipicamente associadas às grandes cidades.

A literatura sobre a concentração das atividades no espaço remonta desde os escritos de Marshall (1996), o qual identifica nas economias externas um dos fatores determinantes para as aglomerações de firmas e pessoas no espaço. 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, portanto, 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. Um ponto importante neste tipo de estudo refere-se à determinação de qual área se baseará a análise. Arbia (2006) alega que o resultado de qualquer análise de regressão, baseada em dados espaciais, depende, essencialmente, do nível de agregação geográfica escolhido e não pode simplesmente ser transferido de um nível de agregação para outro. O problema de escala espacial, ou seja, a sensibilidade dos resultados ao nível de agregação geográfica dos dados, faz parte do Problema da Unidade de Área

Modificável (MAUP). Além do problema de escala, o MAUP é caracterizado pelo problema de zoneamento, ou seja, os dados apresentam diferentes respostas de acordo com a forma de combiná-los em zonas. Esta característica verificada pelo problema de MAUP, quando agregamos regiões, é o que influencia o presente estudo, a utilizar dados agregados por municípios brasileiros. Pois ao agregar-se por áreas maiores, como microrregiões, mesorregiões ou estados, por exemplo, os resultados encontrados tendem a explicar os problemas da MAUP, não condizendo com a real situação das regiões mais desagregadas espacialmente. Utilizando-se de dados agregados por municípios, pode-se ter uma ideia mais fidedigna sobre o fenômeno estudado e sobre variáveis a serem analisadas de maneira mais clara.

O presente documento é composto por seis seções. Sendo a primeira esta introdução, seguida pela revisão bibliográfica. No terceiro capítulo aprofundam-se os tópicos referentes aos procedimentos metodológicos utilizados; e, em seguida parte-se para a análise espacial das equações de rendimentos, nas agregações regionais propostas. Na quinta seção expõem-se os resultados econométricos verificados. Ao final está a seção de conclusão, onde se apresentam os principais resultados alcançados neste estudo.

2.2 Revisão bibliográfica

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

2.2.1 Equação de Rendimentos (Minceriana)

Segundo Fernandes (2002), a relação entre educação e desigualdade salarial está presente na literatura econômica desde o debate entre Adam Smith e Stuart Mill. Para o primeiro, os salários eram determinados apenas pela oferta e demanda por trabalho, logo, não haveria preocupação com a desigualdade salarial. Para ele, os salários mais elevados eram uma compensação de mercado pelas características não desejáveis dos postos de trabalho ou para o esforço passado que certos indivíduos tiveram de realizar com o objetivo de se habilitar para exercer certas ocupações. Já Stuart Mill enfatizava que se os

salários de mercado fossem ajustados por aspectos não pecuniários, a desigualdade salarial tenderia a se elevar, ao invés de diminuir.

Segundo Pereira e Zavala (2013), as questões envolvidas nas discussões atuais sobre desigualdade salarial são, em essência, as mesmas que estavam presentes no “debate” entre Adam Smith e Stuart Mill. As fontes de desigualdade podem estar associadas às diferenças dos trabalhadores em relação às suas características produtivas e preferências, e às imperfeições do mercado que impedem a mobilidade dos trabalhadores dos postos de trabalho com baixos salários para aqueles com altos salários.

Solow (1956) publicou um artigo analisando os fatores clássicos do crescimento econômico (estoque dos fatores capital e trabalho). Nesse trabalho, o autor demonstra que um modelo de crescimento econômico poderia ser dividido em capital, trabalho e tecnologia, baseando-se na influência da poupança, da depreciação e na variação populacional para explicar a variação do crescimento da economia. Nessa linha, o crescimento econômico é determinado por fatores exógenos, tais como o crescimento populacional e o progresso tecnológico. No entanto, mesmo a variável tecnologia fazendo parte do modelo, tal elemento não é explicado, deixando uma lacuna para inserção de inúmeras pesquisas sobre a origem do progresso tecnológico das nações. Diante disso, inúmeros trabalhos surgiram, com o intuito de justificar a existência do progresso técnico a partir do Capital Humano.

A teoria do Capital Humano, como a conhecemos hoje, surge, em meados da década de 60, da preocupação cada vez maior com problemas de crescimento econômico e melhor distribuição de renda. Partindo de um ambiente neoclássico em que os fatores de produção são remunerados de acordo com a sua produtividade marginal, a teoria do Capital Humano afirma que, à medida que o nível educacional de um indivíduo cresce, cresce também sua renda, pois, a educação afeta direta e positivamente a produtividade destes indivíduos. Daí decorre a tese de que a educação é o principal meio de mobilidade social e é também a principal variável explicativa dos diferenciais de rendimentos.

Uma das instituições que mais contribuiu para o desenvolvimento da Teoria do Capital Humano foi a Universidade de Chicago, a partir dos trabalhos seminais dos economistas Schultz (1963), Becker (1964) e Mincer (1974). Schultz foi um dos primeiros autores a abordar como o fator humano na produção é capaz de criar ganhos de produtividade. Becker expandiu a Teoria do Capital Humano ao considerar que o treinamento/capacitação no trabalho produz ganho salarial ao longo da carreira profissional, posto que a produtividade do trabalhador capacitado seja maior do indivíduo

sem capacitação.

De acordo com Schultz (1964), a qualificação e o aperfeiçoamento da população, advindos do investimento em educação, elevariam a produtividade dos trabalhadores e os lucros dos capitalistas, impactando na economia como um todo. Diante disso, a inclusão do capital humano nos modelos de crescimento econômico é uma questão chave para se compreender a dinâmica da economia no longo prazo, uma vez que, até então, esse fenômeno era explicado somente pelo capital natural e capital manufaturado¹ existente entre regiões e países.

Becker (1993), da mesma forma, alega que o capital humano é um conjunto de capacidades produtivas que uma pessoa pode adquirir, devido à acumulação de conhecimentos gerais ou específicos, que podem ser utilizados na produção de riqueza. Assim, sua principal preocupação é decorrente de que os indivíduos tomam a decisão de investir em educação, levando em conta seus custos e benefícios, atribuindo, entre estes melhores rendimentos, maior nível cultural e outros benefícios não monetários. Desse modo, o nível de capital humano de uma população influencia o sistema econômico de diversas formas, com o aumento da produtividade, dos lucros, do fornecimento de maiores conhecimentos e habilidades, e também por resolver problemas e superar dificuldades regionais, contribuindo com a sociedade de forma individual e coletiva.

Porém, conforme Viana (2010) o precursor da Teoria do Capital Humano foi Mincer, que em 1958 indicou a existência de correlação entre o investimento para a formação das pessoas (trabalhadores) e a distribuição de renda pessoal. Para o autor, era necessário decidir de forma individual e racional entre gastar tempo para obter novos conhecimentos e aplicá-los posteriormente em atividades profissionais ou manter-se no trabalho sem novas formas de treinamento e estudo de novos conhecimentos. Dessa forma, Mincer 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. Neles usaram-se diferentes bases de dados e métodos de estimação, além de períodos distintos. Langoni (1973) foi um dos primeiros

¹ Segundo Denardin e Sulzbach (2005), o capital natural refere-se a toda a espécie de bens e serviços, que a sociedade humana pode converter em produtos úteis. Enquanto que, o capital manufaturado é aquele produzido por meio da atividade econômica e das mudanças tecnológicas (engenhosidade humana), através de interações com o capital natural.

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.

Ao longo do tempo, diferenças de capital humano foram geradas na formação das regiões brasileiras, sendo provavelmente o componente mais importante na determinação das desigualdades regionais (Barros 2011). Na Teoria do Capital Humano, ao adquirir educação e experiência no trabalho, obtém-se investimento em recursos humanos, que impactam positivamente o desenvolvimento econômico. A literatura que aponta a importância da educação na determinação de renda é extensa, começando com Mincer, Becker e Schutz. A partir deles, toda uma linha de pesquisa começou a desenvolver-se, tendo como ponto principal o tratamento econométrico do problema. A principal questão, nesta linha de pesquisa, é qual a função e o impacto da decisão no investimento em capital humano sobre a distribuição e estrutura de ganho individual; outro aspecto diz respeito ao estudo da forma funcional da relação entre renda e escolaridade.

Utiliza-se, no presente artigo, a equação de rendimentos proposta por Jacob Mincer seguida dos modelos que dela derivam, buscando corrigir possíveis problemas de viés de variáveis omitidas, erros de medida e de endogeneidade da variável educação. 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 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 utilizou uma transformação de $exp = idade - S - 6$, 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). Pois a forma funcional tem levantado várias objeções ao utilizar o método de mínimos quadrados ordinários (MMQ) para a estimação. Isso se deve a possibilidade das estimativas dos coeficientes de regressão sofrerem problemas de viés de seleção em virtude de i) erro de medida da variável educação; ii) omissão de variáveis relevantes no modelo, iii) da endogeneidade da escolaridade; e iv) Linearidade.

Segundo Bandeira e Jacinto (2010), o método frequentemente empregado para estimar a equação de rendimentos de Mincer 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 em 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.

Carvalho (2007) afirma que, apesar do fato de que essas estimativas possam ser viesadas e inconsistentes, as estimativas de MQO nos oferecem informações importantes a respeito dos determinantes da renda de um indivíduo. Ou seja, a estimativa de MQO é um bom ponto de partida para depois, em seguida, utilizar-se técnicas de estimação mais sofisticadas, até para se ter uma ideia do tamanho do viés entre as estimativas com diferentes métodos de estimação.

Além dessas observações, de caráter metodológico, uma característica 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. Por exemplo, Fingleton (2003) incluiu a influência das aglomerações na equação de rendimento. Ele testa, a partir de um modelo microeconômico, 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. O objetivo de seu modelo é evidenciar a presença de economias de escala resultantes das externalidades pecuniárias em áreas de elevada densidade econômica. A hipótese é que as áreas mais densas são suscetíveis a ser mais produtivas e a ter salários mais elevados em comparação com áreas de menor concentração da atividade econômica.

Este estudo procura fornecer evidências de tais hipóteses, relacionando as aglomerações geográficas aos salários recebidos por seus habitantes. A seguir apresenta-se melhor esta ideia defendida por Fingleton (2003).

2.2.2 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. Utiliza-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.

Basicamente, esses estudos tentam explicar porque as áreas mais densas são suscetíveis a ter salários mais elevados e a ser mais produtivas do que áreas com menor concentração de atividade econômica. O pressuposto é que, em mercados competitivos, o trabalhador é pago pelo valor de seu produto marginal e que, mesmo em concorrência imperfeita, em locais mais produtivos, os salários serão mais elevados. Considerando-se que as externalidades positivas decorrentes da aglomeração das atividades econômicas elevam a produtividade do trabalho, supõe-se uma relação positiva entre esta e a taxa salarial. Sendo assim, como destacado por Galinari (2006, p.12), tendo em vista que as equações salariais captam os efeitos de atributos urbanos representativos das economias de aglomeração sobre os salários, as mesmas demonstram, indiretamente, os efeitos das economias de aglomeração sobre a própria produtividade.

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) 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. Esta diferenciação traz, aos resultados encontrados, uma interpretação mais fidedigna, da realidade salarial brasileira.

2.3 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 seção está dividida em três sub-seções: na primeira, apresenta-se o modelo econométrico utilizado no trabalho; a seguir, explana-se sobre a matriz de peso espacial necessária para tal estudo; e, por último, analisa-se a base de dados e variáveis, que serão utilizados na pesquisa.

2.3.1 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 (Equação 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 fundamenta-se em densidades, assume-se que $L=I$, 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:

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

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:

$$\begin{aligned} dQ/dL &= f(N)^\alpha L^{1-\alpha} (1-\alpha) / L = (1-\alpha)Q/L \\ r &= (1-\alpha)Q/L \\ rL/Q &= 1-\alpha \end{aligned} \quad (12)$$

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:

$$\ln(A) = b_0 + b_1 H + b_2 T + \rho W \ln(A) + \xi \quad (17)$$

$$\xi = N(0, \sigma^2)$$

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. De acordo com o autor, uma proxy do conhecimento técnico, voltado ao ambiente de trabalho pode ser calculado por meio de um quociente locacional, cuja função é expressar a especialização do trabalho, em atividades ligadas a informática e P&D. Reconhecendo que os trabalhadores se movimentam no espaço e a qualidade de bem público do conhecimento, o autor admite que a comutação dos trabalhadores pode verter informações acerca de tecnologias e conhecimentos próprios de uma área para outra. Portanto, 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:

$$\ln(w) = k_2 + \rho W \ln(w) - \rho W k_1 + (y - 1)[\ln(E) - \rho W \ln(E)] + a_1 H + a_2 T + v \quad (19)$$

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

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 p 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).² 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), no caso, $\ln(w)$ 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.³

² O presente estudo, contudo, reconhece a possível influência da densidade de emprego nesse processo e controla por esses fatores nas regressões estimadas.

³ O modelo a ser estimado nesse trabalho portanto estimará o coeficiente $(y - 1)$ as β_1 , contudo, a interpretação de retornos crescentes da densidade de emprego deve ser feita com cuidado uma vez que o

2.3.2 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. Os elementos não nulos da matriz de pesos espaciais refletem o potencial de interação espacial entre duas observações, enquanto os elementos iguais a zero indicam ausência de correlação espacial, ou seja, as observações não são vizinhas entre si.

Portanto, se,

$$\text{Cov}(y_i, y_j) = E(y_i)E(y_j) \neq 0 \quad \text{para } i \neq j$$

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 grau de conexão expresso nas matrizes de pesos espaciais pode ser classificado de acordo com um critério geográfico ou socioeconômico. 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 ensaio, 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 distância geográfica entre duas regiões, geralmente medida em quilômetros.

2.3.3 Base de Dados e Variáveis

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. Nesse sentido, a

seu resultado não controla explicitamente pelo coeficiente ρ que pode ser endógeno e afetar os coeficientes estimados.

análise compreende um estudo cujo intuito é contribuir para o debate acerca da relação entre salários e aglomerações no Brasil. O método de investigação consiste na verificação empírica da presença de economias de aglomeração baseado em equações salariais espaciais. A análise foi dividida em duas etapas complementares. A primeira parte compreende a Análise Exploratória de Dados Espaciais, empreendida para a taxa salarial, variável dependente do modelo econométrico, e a densidade do emprego, variável representativa das economias de aglomeração. Nesse caso, buscou-se verificar a existência de algum tipo de associação espacial nas variáveis em questão, tais como aglomerações espaciais ou regiões homogêneas (clusters espaciais) e observações atípicas (outliers espaciais). A segunda parte foi dedicada à estimação do modelo econométrico.

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 1 indica tal comparação para os dados dos municípios brasileiros

Tabela 1 – 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 Brancos	0,48	0,46
	Desvio Padrão	0,24	0,25
Sexo	Proporção Homens/ Mulheres	0,72	0,63
	Desvio Padrão	0,10	0,12
Observações	5565	5565	

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

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 verifica-se o número de trabalhadores sem carteira de trabalho assinada, nos municípios brasileiros, chega-se à conclusão de que 71,34% encontram-se nesta situação.

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

1- Taxa salarial ($\ln(w)$): variável dependente do modelo econométrico, calculada como a remuneração média dos trabalhadores ocupados nas atividades econômicas em nível municipal. Optou-se pela utilização da variável *Rendimento domiciliar per capita, em julho de 2010 (em reais)* do Censo 2010. Considerou-se como rendimento nominal, mensal domiciliar, per capita, a divisão do rendimento mensal domiciliar pelo número de moradores do domicílio particular, exclusive aqueles cuja condição no domicílio fosse pensionista, empregado doméstico ou parente do empregado doméstico; expresso em reais.

2- 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). Para a estimativa desta densidade, utilizou-se a variável *Trabalhou ganhando em dinheiro, produtos, mercadorias ou benefícios*, em conjunto com a área municipal, em Km², disponibilizada no site Ipea Data.

3- 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. Para este parâmetro, utilizaram-se as variáveis, *Curso que frequenta* e *Curso mais elevado que frequentou* do Censo 2010.

A seguir apresenta-se o Quadro 1, com os anos adotados, para as diferentes faixas de escolaridade empregadas nas duas variáveis citadas do Censo 2010. Tais modificações fazem-se necessárias, tendo em vista as modelagens econométricas que se deseja fazer.

Quadro 1 - Anos adotados para as variáveis Curso que frequenta e Curso mais elevado que frequentou do Censo 2010⁴

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: Resultados da pesquisa, a partir do Censo 2010.

4 - Conhecimento Técnico (CT): variável de controle, cuja função é parametrizar o potencial local de inovação. Conforme Fingleton (2003), essa variável pode ser calculada por meio do Coeficiente Locacional (CL), o qual expressa a especialização dos trabalhadores locais em atividades de P&D e informática. No presente estudo utilizou-se para estimar esta variável a quantidade de indivíduos com doutorado em relação ao total

⁴ Os anos adotados neste ensaio são uma média dos anos necessários para a conclusão de cada nível escolar especificado.

da população de cada município. A quantidade de doutores é um parâmetro adotado para podermos medir o grau de pesquisa em cada município.

5 - Defasagem espacial da taxa salarial ($Wln(w)$): Essa variável tem a função de mensurar a relação entre a taxa salarial de cada município com aquela apresentada em sua vizinhança, ou seja, o transbordamento dos níveis de eficiência entre as regiões. 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.

A partir destes dados, inicia-se, na próxima seção, com as análises espaciais destes argumentos, verificando a existência de autocorrelação espacial entre o salário e a densidade do trabalho no mercado de trabalho brasileiro.

2.4 Análise exploratória de dados espaciais

Como ressaltado em Galinari (2006), particularmente no estudo das economias de aglomeração, a análise espacial é importante por permitir a identificação de aglomerações econômicas que transcendem os limites geográficos, sugerindo o transbordamento espacial de características econômica locais, tais como os níveis de produtividade. Além disso, 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:

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

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 e 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 e HADDAD, 2003).

Para tanto, foram realizados testes de *I de Moran*, com a utilização da matriz de pesos espaciais. A matriz peso utilizada nesse trabalho refere-se à matriz geográfica de contiguidade por convenção rainha (critério Queen)⁵. 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.

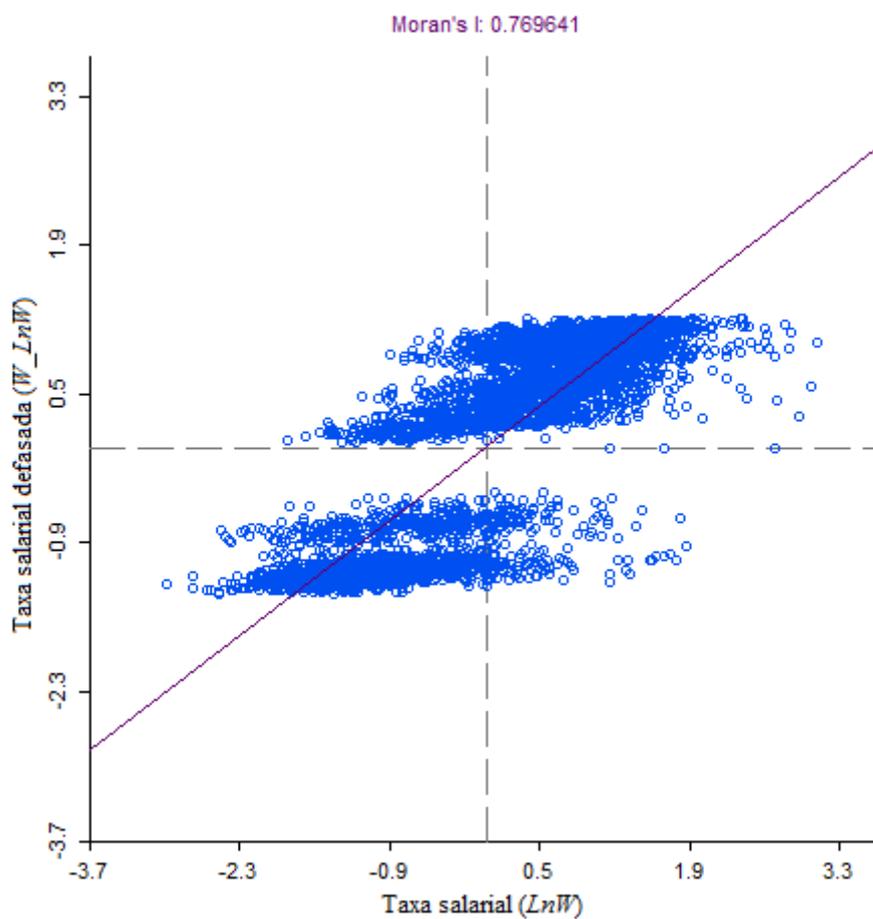
2.4.1 Municípios

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 formal, 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 forte autocorrelação espacial na taxa salarial, chegando-se ao *I de Moran* ser de 0,76.

⁵ O presente artigo 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*.

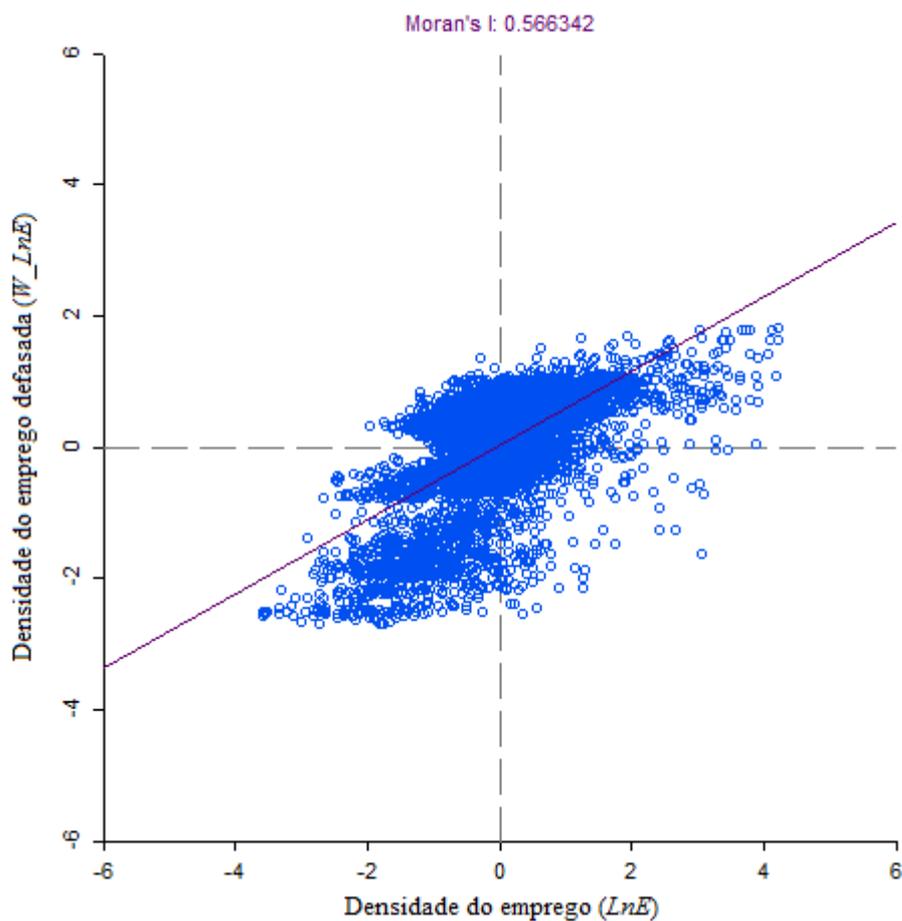
Figura 1 - Diagrama de dispersão de *Moran* da Taxa Salarial-2010



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.

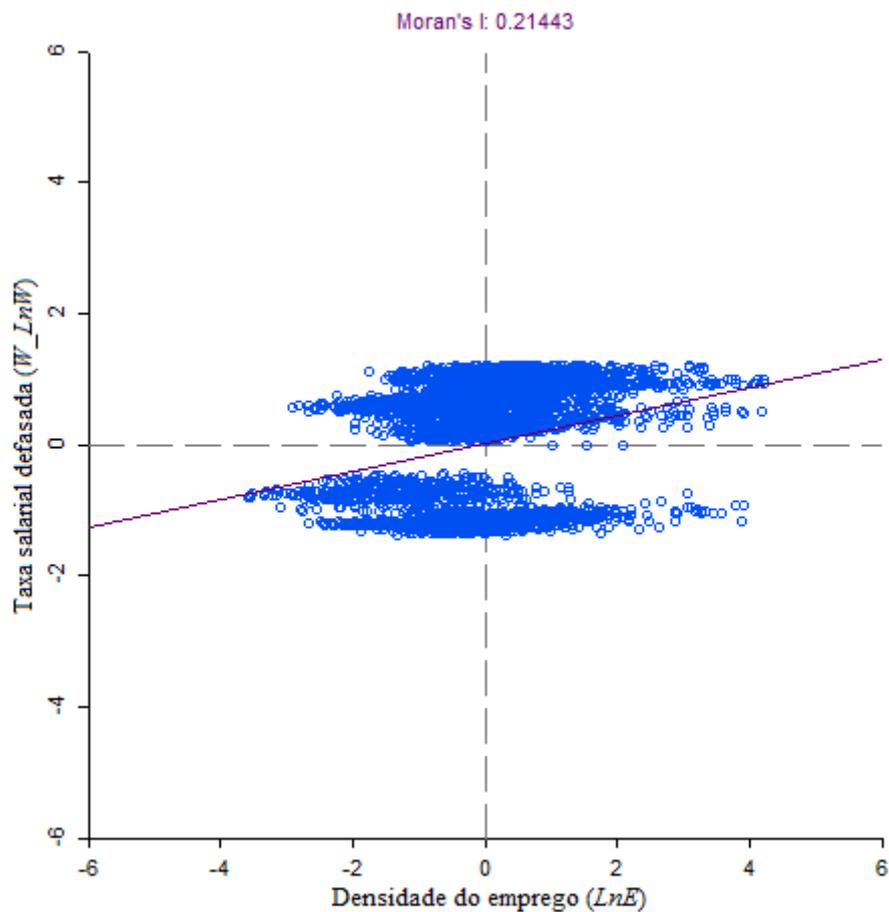
Figura 2 - Diagrama de dispersão de *Moran* da Densidade do Emprego-2010



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, sendo significativo.

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.

Sendo assim, 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 e a densidade do emprego formal, quanto para o teste bivariado, por meio do qual é avaliada a autocorrelação espacial entre a densidade do emprego formal e a média da taxa salarial nas atividades econômicas dos vizinhos de cada município, indicando, assim, a presença de municípios com altos ou baixos valores das variáveis em estudo com vizinhos na mesma condição. Porém, 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 associação espacial, 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.

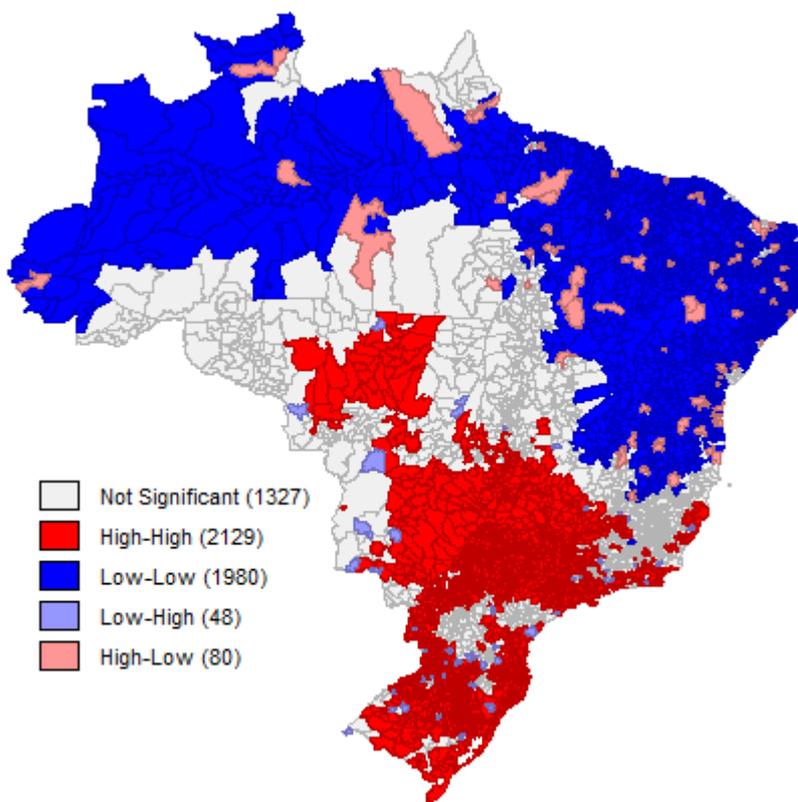
Uma vez que existe uma ligação entre o indicador de associação espacial local e o *I de Moran* global, a identificação de focos de não-estacionariedade espacial, ou seja, a presença de valores discrepantes (outliers) ou regimes espaciais (clusters) será semelhante ao uso do Diagrama de Dispersão de *Moran*. 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, sendo dividido em quatro quadrantes representantes dos diferentes 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 $k = 10^6$. 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.

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.

Figura 4 - *LISA* da Taxa Salarial-2010



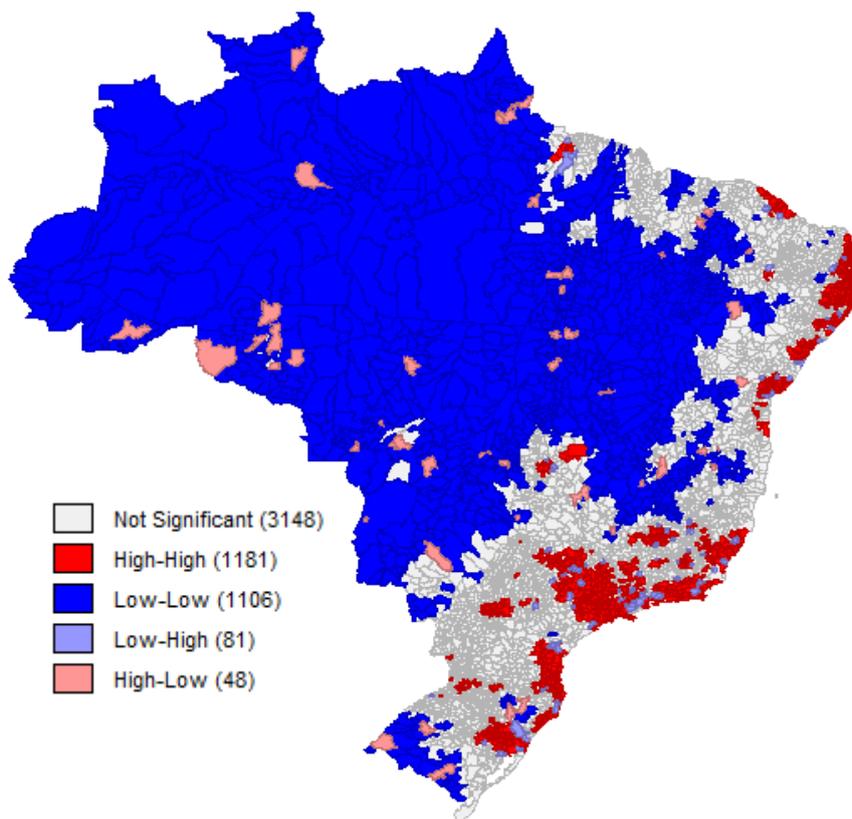
Fonte: Resultados da pesquisa, a partir do Censo 2010, utilizando o software OpenGeoda.

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

⁶ As utilizações de diferentes matrizes de peso espacial fornecem resultados qualitativos similares.

litoral brasileiro, do Sul ao Nordeste. Em contrapartida, as regiões com densidade de emprego baixa-baixa localizaram-se nas regiões Norte, 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.

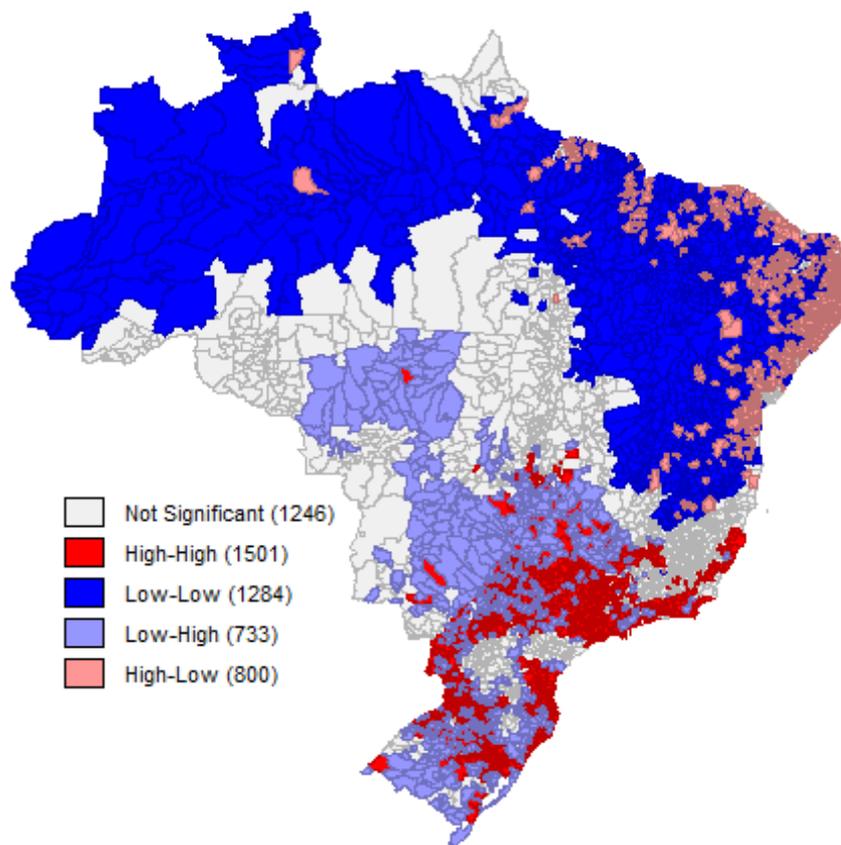
Figura 5 - LISA da Densidade do Emprego-2010



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

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 formal 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 formal, 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 na região Sudeste na região Sul, sobretudo, no estado de São Paulo. 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.

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.

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

2.5 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 2 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 2 - Resultados da regressão por MQO

Variáveis	Matriz de Pesos- <i>Queen</i> = 10
Constante	3,0490*** (0,0250)
Densidade do emprego (<i>LnE</i>)	-0,0214*** (0,0025)
Escolaridade (<i>H</i>)	0,6539*** (0,0058)
Doutores (<i>T</i>)	-21,8501*** (4,8005)
N	5564
Teste F	5961,86
R ²	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.

2.5.1 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 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:

$$y = \rho W y + X\beta - \theta W X\beta + \mu \quad \mu \sim N(0, \sigma^2 I_n) \quad (21)$$

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:

$$\begin{aligned} y &= x\beta + (I - \rho W)^{-1} (x\eta + v), \\ (I - \rho W)y &= (I - \rho W)x\beta + x\eta + v, \\ y &= \rho W y + x(\beta + \eta) + Wx(-\rho\beta) + v, \\ y &= \rho W y + x\beta I + Wx\beta 2 + v \end{aligned} \quad (22)$$

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 $z = \rho W z + u$ deixa os estimadores de mínimos quadrados ordinários (MQO) não viesados e consistentes, porém não eficientes. No entanto, a propriedade de não viés 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 Spatial Durbin Model (SDM), uma especificação que considera tanto lag na variável dependente como nas variáveis independentes: $y = \rho Wy + x\eta + Wx\phi + v$. 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:

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

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

Utilizando-se o modelo SDM para a Equação (23), chega-se aos resultados das Tabelas 3, 4 e 5, 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, $W\ln(w)$, 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 transbordamento 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 escalar, 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), em Galinari (2006) e em 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 formal ao nível de agregação municipal, tais como salário, emprego e escolaridade.

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, contudo com algumas diferenças. 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. Tal diferença entre os resultados encontrados em Amarante (2011) e os aqui especificados, reside no fato de, o primeiro estudo, utilizar como base de dados, as informações acerca do setor formal da RAIS, a qual disponibiliza os dados, para a obtenção de características do mercado de trabalho, ao nível de agregação municipal. 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 utiliza-se dados do Censo 2010, a fim de verificar não

somente o mercado formal, mas também o mercado informal existente no país. Outra diferenciação deste trabalho para o de Amarante (2011), é a metodologia de estimação utilizada.

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. Contudo a densidade de emprego nos municípios vizinhos tem a característica de afetar negativamente o salário. Este fato pode ser explicado pelas externalidades negativas, oriundas da aglomeração. Um município com alta densidade salarial tende a atrair a população dos municípios vizinhos, que veem uma oportunidade de crescimento salarial, migrando seu trabalho, de seu município para o município vizinho; aumentando assim a densidade do trabalho deste e diminuindo a do município de origem. 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 seria mais elevado 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, abaixo, verifica-se que praticamente todos os coeficientes defasados espacialmente são significativos a 5%, a exceção da variável defasada do nível técnico (T) na Tabela 3; de H e T na Tabela 4 e de T na Tabela 5.

Quando utiliza-se a matriz $Queen = 1$, na Tabela 3, o resultado encontrado, de 0,35, para o coeficiente da defasagem espacial da taxa salarial (W_LnW), aponta forte 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 3 - Resultados da regressão por SDM – *Queen* = 1

Variáveis	Matriz de Pesos- <i>Queen</i> = 1
Constante	1,4695*** (0,0385)
Taxa salarial defasada (<i>W_LnW</i>)	0,3522*** (0,0107)
Densidade do emprego (<i>LnE</i>)	0,0021 (0,0022)
Densidade do emprego defasada (<i>W_LnE</i>)	-0,0349*** (0,0037)
Escolaridade (<i>H</i>)	0,3984*** (0,0062)
Escolaridade defasada (<i>W_LnH</i>)	0,1381*** (0,0118)
Doutores (<i>T</i>)	20,5244*** (3,7133)
Doutores defasada (<i>W_LnT</i>)	-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 4, 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 4 - 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 5 - Resultados da regressão por SDM – *Queen* = 50

Variáveis	Matriz de Pesos- <i>Queen</i> = 50
Constante	1,5210*** (0,0553)
Taxa salarial defasada (W_LnW)	0,2879*** (0,0186)
Densidade do emprego (LnE)	0,0082*** (0,0023)
Densidade do emprego defasada (W_LnE)	-0,0419*** (0,0041)
Escolaridade (H)	0,4585*** (0,0062)
Escolaridade defasada (W_LnH)	0,1506*** (0,0184)
Doutores (T)	10,9501*** (3,9915)
Doutores defasada (W_LnT)	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 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.

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.

2.6 Considerações finais

O presente trabalho teve a finalidade de obter evidências empíricas das economias de aglomeração nas atividades econômicas dos municípios brasileiros no ano de 2010, utilizando-se de 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. A análise foi dividida em duas etapas complementares. A primeira parte compreendeu a Análise Exploratória de Dados Espaciais (AEDE), empreendida tanto para a taxa salarial, variável dependente do modelo; quanto para a densidade do emprego, variável explicativa representativa das economias de aglomeração. Nesse caso, buscou-se verificar a existência de algum tipo de associação espacial nas variáveis em questão, tais como aglomerações espaciais ou regiões homogêneas (clusters espaciais) e observações atípicas (outliers espaciais).

Basicamente 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.

Na segunda parte, dedicada à estimação do modelo econométrico, foram utilizadas matrizes de pesos espaciais do tipo *Queen* = 1, 10 e 50. A taxa salarial defasada variou de 0,28 a 0,42 para essas matrizes. 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, corroborando assim com as teorias de Jacobs acerca das economias de urbanização. Quando verificamos a variável defasada da densidade, essa se apresenta com sinal negativo. Uma explicação para isso é a externalidade negativa.

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.

3. EQUAÇÕES DE RENDIMENTOS E MAUP NO BRASIL: EVIDÊNCIAS A PARTIR DO CENSO 2010

3.1. Introdução

No mercado de trabalho atual, o capital humano aparece como um importante diferencial. Esta variável começou a ser estudada e formalizada no final dos anos 50 e começo dos 60. Inicialmente, a introdução do capital humano na análise econômica foi associada à preocupação dos pesquisadores em entender a dinâmica da distribuição de renda entre os indivíduos. Esta foi a principal preocupação de Mincer (1958), em um dos primeiros trabalhos a elaborar o conceito de capital humano na forma em que ele é entendido atualmente. Basicamente, essa teoria afirma que a decisão de um indivíduo de acumular conhecimento por meio da educação vai depender de uma série de variáveis, tais como o aumento de salário esperado, o custo de oportunidade do tempo de ir para a escola, a facilidade de acesso à educação, o custo pecuniário da educação, etc.. Por outro lado, essa decisão de investimento na capacidade humana tem consequências para a economia como um todo, como por exemplo, na desigualdade de renda, na trajetória de crescimento da sociedade ao longo do tempo e nas taxas de fertilidade e de mortalidade de seus habitantes.

Uma importante contribuição para a teoria do capital humano foi feita por Jacob Mincer. Em 1974, ele iniciou os estudos de como medir o impacto do incremento de mais um ano de estudo no salário dos indivíduos, ou seja, a ideia de mensurar o retorno do investimento em educação. Análises das desigualdades de renda per capita (ou de suas taxas de crescimento) entre países revelam a importância da consideração da qualidade da educação na mensuração do capital humano.

Um aspecto chave que explica as variações na taxa de crescimento econômico dos países se refere às diferenças na produtividade dos fatores de produção. A melhoria neste setor decorre, conforme consenso da literatura, da ampliação do estoque de capital físico, juntamente com progresso tecnológico, aliado à ampliação do capital humano. Este último participaria em dois aspectos, favorecendo a invenção tecnológica e tornando a mão de obra mais qualificada (KROTH; DIAS, 2012).

Segundo Júnior e Sampaio (2013), a expansão das habilidades, dos conhecimentos e das capacidades dos indivíduos, ou o crescimento de seu capital humano, influencia diretamente o desenvolvimento do progresso econômico de um país ou de uma região.

Como a educação tem um papel importantíssimo, na expansão do capital humano, governos de todo o mundo têm estipulado frequentes metas para a expansão da educação. Entretanto o alcance de tal resultado demanda um longo processo, permeado de grande esforço de toda a sociedade.

Conforme Cangussu, Salvato e Nakabashi (2010), capital humano é um fator de produção que ganha cada vez mais importância na explicação do diferencial de renda entre os países, seja na literatura teórica ou empírica. Tal capital é importante na determinação da renda por vias diretas e indiretas.

Os efeitos diretos são aqueles que afetam a renda através da melhora na produtividade marginal do trabalho, mantendo todos os outros fatores constantes (capital e tecnologia), isto é, na melhora da habilidade dos trabalhadores para a realização de suas respectivas tarefas. É representado pela introdução do capital humano de forma direta na função de produção. Mas ele também afeta o diferencial de renda através dos seus efeitos indiretos. Tais efeitos são aqueles que afetam a quantidade de tecnologia disponível para ser utilizada no processo de produção. Assim, são, portanto, os fatores que influenciam a criação e propagação de tecnologia. (CANGUSSU, SALVATO E NAKABASHI (2010)).

No Brasil apresenta-se uma alta disparidade socioeconômica entre suas regiões. E muitas destas diferenças são consequências de diferentes níveis de capital humano, encontrados no país. O problema de escala espacial, ou seja, a sensibilidade dos resultados, ao nível de agregação geográfica dos dados, faz parte do Problema da Unidade de Área Modificável (MAUP), que deve ser considerado para entender melhor, a dinâmica dos resultados, da equação minceriana, no contexto brasileiro. Além do problema de escala, o MAUP é caracterizado pelo problema de zoneamento, ou seja, os dados apresentam diferentes respostas de acordo com a forma de combiná-los em zonas. De acordo com Resende, Carvalho e Sakowski, (2013), a escolha da escala espacial para análise é uma questão problemática na pesquisa aplicada.

Resende (2011), por exemplo, investiga os determinantes do crescimento regional da economia brasileira em uma variedade de escalas geográficas, utilizando uma base de dados *cross-section* para a década de 1990. Além disso, Resende (2013) avança na literatura de crescimento estimando modelos de dados em painel em diversas escalas espaciais. Mas é importante ressaltar que o referido estudo apenas utiliza modelos de dados em painel não espacial.

Investigações deste tipo remetem à questão do problema de MAUP e também lança nova luz sobre um problema central na literatura relacionada ao crescimento

econômico regional. Com exceção de Resende (2011; 2013), os estudos até agora têm apenas investigado os determinantes do crescimento econômico em uma única escala espacial para inferir a consistência de modelos de crescimento espaciais com a realidade.

O problema de MAUP, segundo Resende (2013), afeta as regressões de crescimento. Esta abordagem leva à investigação do problema de medida, que pode causar uma variabilidade nas estimativas dos determinantes do crescimento econômico regional, devido à utilização de diferentes escalas espaciais, provavelmente relacionado ao MAUP. No entanto, é importante ter em mente que tais variações podem estar relacionadas a questões teóricas estruturais, subjacentes ao processo de crescimento econômico, em diferentes escalas. Assim como afetam as equações de crescimento, muito provavelmente o problema de MAUP também afete as estimações de salário.

Na busca de ampliar o debate acerca da desigualdade brasileira, o presente estudo tem como objetivo mensurar equações de rendimentos mincerianas para diferentes regiões brasileiras, identificando e verificando o problema de MAUP para tais regiões. Para tal, serão utilizados, como base de dados, os microdados do Censo 2010, disponibilizados pelo Instituto Brasileiro de Geografia e Estatística (IBGE).

O ensaio é organizado em seis seções, incluindo esta introdução. Na seção dois apresenta-se a revisão literária, seguida da terceira seção, onde expõem-se a metodologia utilizada para o estudo. Em seguida apresentam-se os dados, sendo na quinta seção analisados os resultados. Por fim é apresentada a conclusão, a que chega o estudo.

3.2 Revisão da literatura

Conforme Viana (2010), o precursor da Teoria do Capital Humano foi Mincer, que em 1958 indicou a existência de correlação, entre o investimento para a formação das pessoas (trabalhadores) e a distribuição de renda pessoal. Para o autor, era necessário decidir de forma individual e racional, entre gastar tempo para obter novos conhecimentos e aplicá-los posteriormente em atividades profissionais ou manter-se no trabalho, sem novas formas de treinamento e estudo. Dessa forma, Mincer 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.

A teoria do Capital Humano surgiu da preocupação, cada vez maior, com problemas de crescimento econômico e melhor distribuição de renda. Partindo de um ambiente neoclássico em que os fatores de produção são remunerados de acordo com a

sua produtividade marginal, a teoria do Capital Humano afirma que, à medida que o nível educacional de um indivíduo cresce, cresce também sua renda, pois, a educação afeta direta e positivamente a produtividade destes indivíduos. Daí decorre a tese de que a educação é o principal meio de mobilidade social e é também a principal variável explicativa, dos diferenciais de rendimentos.

Uma das instituições que mais contribuiu para o desenvolvimento da Teoria do Capital Humano foi a Universidade de Chicago, a partir dos trabalhos seminais dos economistas Schultz (1963), Becker (1964) e Mincer (1974). Schultz foi um dos primeiros autores a abordar como o fator humano, na produção, é capaz de criar ganhos de produtividade. Becker expandiu a Teoria do Capital Humano, ao considerar que o treinamento/capacitação no trabalho produz ganho salarial ao longo da carreira profissional, posto que a produtividade do trabalhador capacitado seja maior daquele sem esta qualificação.

De acordo com Schultz (1964), a capacitação e o aperfeiçoamento da população, advindos do investimento em educação, elevariam a produtividade dos trabalhadores e os lucros dos capitalistas, impactando na economia como um todo. Diante disso, a inclusão do capital humano, nos modelos de crescimento econômico, é uma questão chave para se compreender a dinâmica da economia no longo prazo, uma vez que, até então, esse fenômeno era explicado somente pelo capital natural e capital manufaturado⁷ existente entre regiões e países.

Becker (1993), da mesma forma, alega que o capital humano é um conjunto de capacidades produtivas que uma pessoa pode adquirir, devido à acumulação de conhecimentos gerais ou específicos, que podem ser utilizados na produção de riqueza. Assim, sua principal preocupação é decorrente de que os indivíduos tomam a decisão de investir em educação, levando em conta seus custos e benefícios, atribuindo, entre estes melhores rendimentos, maior nível cultural e outros benefícios não monetários. Desse modo, o nível de capital humano de uma população influencia o sistema econômico de diversas formas, com o aumento da produtividade, dos lucros, do fornecimento de maiores conhecimentos e habilidades, e também por resolver problemas e superar dificuldades regionais, contribuindo com a sociedade de forma individual e coletiva.

⁷ Segundo Denardin e Sulzbach (2005), o capital natural refere-se a toda a espécie de bens e serviços, que a sociedade humana pode converter em produtos úteis. Enquanto que, o capital manufaturado é aquele produzido por meio da atividade econômica e das mudanças tecnológicas (engenhosidade humana), através de interações com o capital natural.

3.1.1 Equação Minceriana em estudos para o Brasil

Langoni (1973) foi um dos precursores a seguir a abordagem minceriana no Brasil. Basicamente ele estima seu modelo mediante regressões log-lineares tendo como variáveis explicativas a educação, idade, sexo, atividade e região de residência. O citado autor considera que a educação possuía grande importância como variável explicativa da renda. O autor realizou um estudo profundo sobre o agravamento distributivo da renda da década de 1960. De acordo com o mesmo, o aumento da desigualdade de distribuição agregada no período se deve à combinação de ganhos relativamente pequenos (inferiores a 10%) dos grupos de renda média próxima ao salário mínimo, e de ganhos extremamente elevados dos decis superiores (67%). Assim, enquanto a renda relativa no primeiro grupo sofria uma redução da ordem de 22%, a do último grupo aumentava praticamente na mesma proporção. A análise também constatou redução na participação relativa de todos os grupos na renda total, exceto dos 10% mais ricos, que aumentaram sua participação em 20% (LANGONI, 1973, p. 66).

Langoni (1973) destaca a existência de determinadas variáveis (idade, sexo, atividade, região, educação) que contribuem para aumentar ou diminuir o grau de desigualdade de renda dos indivíduos, sendo que, dentre elas, a que mais explica o agravamento distributivo é a variável educação. Entretanto, a principal conclusão de seu estudo foi que o avanço da desigualdade é consequência dos desequilíbrios de mercado peculiares do processo de desenvolvimento. Sendo assim, o autor recomenda que sejam traçadas políticas com objetivo de redistribuir oportunidades, erradicar a pobreza e, ao mesmo tempo, reduzir as possibilidades de ganhos extras decorrentes de desequilíbrios entre oferta e demanda, característicos da aceleração do crescimento.

O Quadro 2 apresenta um resumo de alguns estudos empíricos sobre equações mincerianas realizados no Brasil. Nele, são expostos os períodos, as regiões e as metodologias empregadas, além dos principais resultados de cada pesquisa.

Quadro 2 - Estudos com abordagem minceriana para o Brasil

Fonte	Período	Regiões	Metodologia	Resultado
Langoni (1973)	Década de 1960	Brasil	Regressões log-lineares tendo como variáveis explicativas a educação, idade, sexo, atividade e região de residência.	A variável educação é a que mais explica o agravamento distributivo de renda do Brasil.
Lam e Levison (1990)	1985	Brasil e Estados Unidos.	<i>Cross-section</i> , comparando os perfis de desigualdade de renda por idade e experiência entre homens brasileiros e norte-americanos.	Retornos à educação aumentam com a idade para os dois países, porém com coeficientes mais elevados para o Brasil
Leal e Werlang (1991)	1976-1989	Regiões metropolitanas, exceto Brasília.	MQO	Encontram-se retornos de renda aos investimentos em educação, para o Brasil, de cerca de 16% ao ano.
Chaves (2002)	2000	Região Metropolitana de Porto Alegre	MQO	Ao aumentar a idade do indivíduo, seus rendimentos, uma vez alcançado o máximo, tendem a diminuir.
Sachsida, Loureiro e Mendonça. (2004)	1992-1999	Brasil	i) MQO; ii) Heckman (1979); iii) Garen (1984) e, iv) pseudo painel.	Evidência acerca da hipótese de vantagens comparativas, indícios de endogeneidade na escolha da escolaridade e pouca importância para o viés de variável omitida.
Resende e Wylie (2005)	1996 - 1997	Regiões metropolitanas de Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, outras áreas urbanas da região Sudeste e áreas rurais das regiões Sudeste e Nordeste.	Regressão de dois estágios de Heckman.	Retornos à educação de 12,6% para mulheres e 15,9% para homens, a cada ano de estudo sobre o salário.
Salvato e Silva (2008)	2005	Região Metropolitana de Belo Horizonte	MQO, variáveis instrumentais e procedimento de Heckman.	i) Retornos positivos para educação; ii) retornos diferenciados por atividade na economia; iii) discriminação no mercado de trabalho devido a cor, gênero e filiação ao sindicato; iv) aumentos significativos do retorno da escolaridade após dez anos de educação; v) uso de variáveis instrumentais (MVI)

				aumenta o retorno da escolaridade; vi) procedimento de Heckman mostra que há viés de seleção, na estimação da equação de rendimentos, para região metropolitana de Belo Horizonte.
Cangussu, Salvato e Nakabashi. (2010)	1980-2000	Estados brasileiros	MQO, MQG, MQ2E, MMG, variáveis instrumentais, e MQG2E.	Retorno marginal estimado da educação foi de 15% e os resultados empíricos sustentam a teoria de que o capital humano é um dos principais fatores na determinação do nível de renda.
Rocha e Menezes (2012)	1995, 2002 e 2009.	Brasil	Regressões quantílicas, em diversas categorias profissionais	A educação contribui positivamente para o crescimento do salário em todos os quantis de renda.

Fonte: Resultados da pesquisa.

Segundo Barros (2011), ao longo do tempo, diferenças de capital humano foram geradas na formação das regiões brasileiras, sendo provavelmente o componente mais importante na determinação das desigualdades regionais.

De acordo com Lazzarotto e Lima (2008), o nível de renda de uma dada região é determinado por um amplo conjunto de fatores de naturezas econômica, social e ambiental. Nesse sentido, os principais fatores podem ser enquadrados em alguns pontos principais: a estrutura setorial predominante; o mercado de trabalho; o nível de educação formal; a oferta de infraestrutura básica, de suporte às atividades econômicas; o nível de dependência de programas sociais; e a localização geográfica, juntamente com as condições naturais, características da região.

Segundo Daniel e Oliveira (2014), o papel do poder público em propiciar uma melhor infraestrutura às atividades econômicas em determinada região tem grande relevância para o desenvolvimento e crescimento econômico da mesma. Condições que aumentam o bem-estar da população são fatores que geram externalidades positivas como, por exemplo, um maior poder aquisitivo das pessoas, ou seja, um fator decisivo como determinante da renda para qualquer nação. Portanto, nesse caso uma região melhor equipada poderá acumular mais capital humano e físico, afetando positivamente, assim, a atividade econômica, com reflexos na produção e no emprego devido às vantagens que oferece às empresas e aos indivíduos quando consideradas decisões de localização, criando assim, possibilidades de uma renda justa para a população local (LAZZAROTTO e LIMA, 2008).

Posto esta dicotomia entre regiões, surge o questionamento de qual a melhor forma de análise de equações mincerianas, se por microdados ou pela agregação destes em diferentes regiões. A diferenciação, que muitas vezes surge, ao serem utilizadas estas duas metodologias, parte do Problema da Unidade de Área Modificável (MAUP). Analisar-se-á o MAUP para equações de rendimento, em dados gerados a partir de microdados, e para agregações dos mesmos, em municípios, microrregiões, mesorregiões e estados brasileiros.

3.2.2 O Problema da Unidade de Área Modificável - MAUP

O processo de escolha da base de dados, além de delimitar a abrangência espacial do trabalho, é um importante passo da pesquisa, para lidar com uma questão central que envolve os trabalhos que utilizam dados geográficos agregados: o MAUP. Na análise urbana, o MAUP surge devido ao fato de que um infinito número de sistemas de zoneamento pode ser construído, para subdividir uma cidade em unidades de áreas menores. Dessa forma, os dados reportados para cada unidade de área serão diferentes entre os sistemas de zoneamento (PAÉZ e SCOTT, 2005).

A consequência direta, da existência de diversos sistemas de zoneamento, é que estudos que utilizem uma mesma ferramenta de análise, mas diferentes sistemas de zoneamento de uma mesma região, possivelmente, apontem resultados distintos, tal como afirma Siqueira (2014).

Anselin (1988) trata a questão do MAUP, alegando que os níveis de agregação e a organização espacial, em zonas, afeta a magnitude de várias medidas associadas às unidades, os coeficientes de autocorrelação espacial e os parâmetros do modelo de regressão. O autor alega ser antiga essa questão, afirmando que, na econometria, o problema se caracteriza como sendo o problema de agregação micro-macro.

O MAUP, tal como pontua Siqueira (2014), é composto de dois efeitos, o de escala e o de zoneamento. O efeito de escala é a tendência, dentro de um sistema de unidades de áreas modificáveis, de se obter diferentes resultados estatísticos, para um mesmo conjunto de dados, quando a informação é agrupada em diferentes níveis de resolução espacial. O efeito de zoneamento, por outro lado, é a variabilidade dos resultados estatísticos obtida dentro de um conjunto de unidades de áreas modificáveis em função das várias possibilidades de agrupamentos em uma dada escala, e não em função da variação do

tamanho dessas áreas (por exemplo, a diferença nos resultados, devido à simples alteração das fronteiras ou configurações, dentro de uma mesma escala de análise).

Para Monastério e Ávila (2006), MAUP trata-se, grosso modo, dos problemas decorrentes das distintas formas de delimitar as unidades espaciais. Diferentes desenhos das unidades implicam em diferentes análises, inferências ou mesmo políticas. A situação de um município pobre em uma área rica deixa de ser corretamente avaliada quando o nível de agregação espacial é a região, e a própria delimitação dos recortes regionais altera os resultados dos indicadores estatísticos.

Siqueira (2014) afirma que é necessário se reconhecer, todavia, que qualquer trabalho que lide com dados agregados por áreas estará sujeito ao MAUP, visto que dados individuais, e que seriam capazes de contornar o problema da agregação espacial, são sigilosos e raramente disponibilizados. Por este motivo, tem-se como prática nos trabalhos regionais e urbanos, de cunho espacial, a adoção de técnicas e procedimentos capazes de mitigar o impacto que a agregação espacial causa sob os resultados dos estudos. Dentre estas, destacam-se o uso de dados agregados em reduzidas dimensões espaciais, tal como setores ou zonas censitárias, ou ainda o uso de critérios de agregação e otimização combinatória.

Resende, Carvalho e Sakowski (2013) analisaram o crescimento econômico brasileiro em diferentes escalas regionais, tais como municípios, microrregiões e estados. Os autores afirmam que cada escala gera uma influência no modo da governabilidade de cada região, o que pode influenciar seu desenvolvimento econômico diferentemente, uma da outra.

Resende e Magalhães (2013) afirmam que a heterogeneidade regional brasileira é marcante e persistente ao longo de décadas. Barros (2011) salienta que a própria extensão territorial, em si, seria capaz de gerar heterogeneidades – como se verifica em outros países de grandes dimensões territoriais (por exemplo, Rússia, Canadá, China, Estados Unidos, Austrália e Índia). Entretanto, as diversidades na formação histórica, social e climática contribuíram fortemente para a formação de tais diferenças entre as regiões brasileiras. Shankar e Shah (2003) ao investigarem as disparidades dos PIBs, per capita regionais, para seis países de grandes dimensões territoriais no mundo, em 1997, concluíram que o Brasil apresentava a terceira maior desigualdade regional, ficando atrás de Rússia e China.

Sabe-se que, na presença do MAUP, segundo Resende (2011), os resultados podem variar de acordo com a escala geográfica, e a sistematização dos diferentes

resultados para as diferentes escalas geográficas é muito importante para a definição e avaliação de políticas regionais.

Pouco há de estudos sobre MAUP para o Brasil, e os que existem tratam de apenas algumas regiões específicas do país. Este ensaio visa contribuir para o entendimento do efeito da MAUP, quando se utiliza diferentes escalas geográficas, para analisar o nível salarial brasileiro. Com esse intuito, o presente estudo usará cinco níveis de aglomerações regionais: microdados, municípios, microrregiões, mesorregiões e estados, utilizando, para tanto, equações de rendimentos (mincerianas).

3.3. Metodologia

Na Teoria do Capital Humano, ao adquirir educação e experiência no trabalho, obtém-se investimento em recursos humanos, que impactam positivamente o desenvolvimento econômico. A literatura que aponta a importância da educação na determinação de renda é extensa, começando com Mincer (1974), Becker (1964) e Schutz (1963). A partir deles, toda uma linha de pesquisa começou a desenvolver-se, tendo como ponto principal o tratamento econométrico do problema.

O objetivo, preliminar, do ensaio, será incluir uma base de dados mais recente que contenha informações tanto do mercado de trabalho informal quanto formal. Para tanto, será utilizado o Censo Demográfico de 2010, do IBGE, permitindo, assim, uma identificação mais completa do impacto das economias de aglomeração, nos municípios brasileiros. Além disso, as atividades econômicas podem 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.

A fim de medir o retorno da educação, Mincer (1974) propôs um tipo de equação, que leva em conta a influência da escolaridade no salário dos indivíduos:

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

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:

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

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 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 utilizou uma transformação de $exp = idade - S - 6$, assumindo-se 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 que proposto por Mincer (1974). Pois a forma funcional tem levantado várias objeções ao utilizar o método de mínimos quadrados ordinários (MMQ) para a estimação. Isso se deve a possibilidade dos coeficientes de regressão sofrerem problemas de viés de seleção, em virtude de: i) erro de medida da variável educação; ii) omissão de variáveis relevantes no modelo, iii) endogeneidade da escolaridade; e iv) Linearidade.

Carvalho (2007) afirma que, apesar do fato de que essas estimativas possam ser viesadas e inconsistentes, as estimativas de MQO nos oferecem informações importantes a respeito dos determinantes da renda de um indivíduo. Ou seja, a estimativa de MQO é um bom ponto de partida para depois, em seguida, utilizar-se técnicas de estimação mais sofisticadas, até para se ter uma ideia do tamanho do viés entre as estimativas com diferentes métodos de estimação.

Segundo Bandeira e Jacinto (2010), o método frequentemente empregado para estimar a equação de rendimentos de Mincer 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

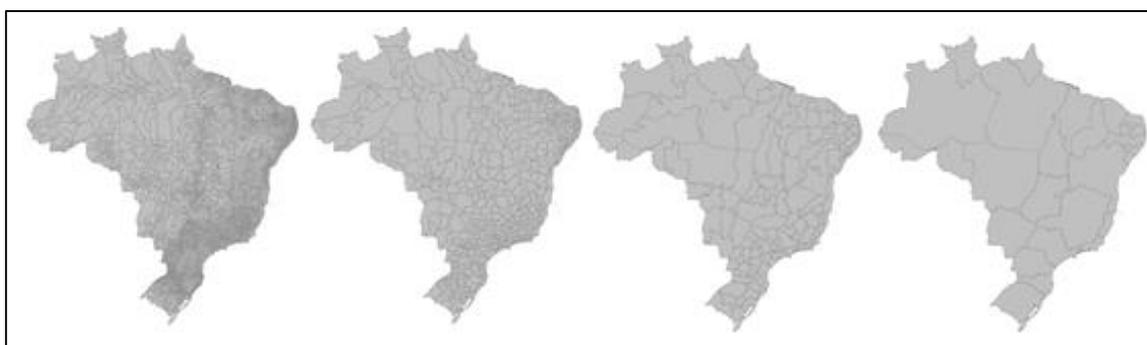
dessa correlação implica em que, os coeficientes estimados serão inconsistentes e não convergirão em probabilidade, para os parâmetros populacionais. Trata-se de um problema de endogeneidade da educação, no caso da equação de rendimentos. Outro problema também comum à estimação dessa equação é o viés gerado pela omissão de variáveis que podem afetar a renda do indivíduo.

Segundo Wu e Cutter (2011) haverá perda de informações durante o processo de agregação, o que geralmente impõe viés de agregação e distorcem os resultados. Em geral, o viés de agregação será crescente, na medida em que os agentes, dentro da unidade de agregação, forem mais heterogêneos. De acordo com autores, uma solução para tal problema é analisar os dados pela menor agregação possível, fazendo assim que os dados sejam mais homogêneos.

3.4 Dados

Para avaliar os resultados, das estimativas de equações de rendimento, em múltiplas escalas espaciais, utilizando modelos de painel espacial, adotou-se procedimento semelhante ao utilizado em Resende (2013). A Figura 7 apresenta quatro das cinco estratificações geográficas, presentes na base de dados, utilizada neste estudo, que são: 27 estados, 137 mesorregiões, 558 microrregiões e 5.565 municípios – a outra escala regional utilizada foi a de microdados. Os dados e aglomerações regionais utilizados referem-se ao Censo 2010, efetuado pelo IBGE.

Figura 7 - Escalas Espaciais no Brasil



Municípios (n=5565)	Microrregião (n=558)	Mesorregião (n=137)	Estados (n=27)
---------------------	----------------------	---------------------	----------------

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

Segundo Resende, Carvalho e Sakowski (2013), o Brasil é dividido em 27 estados, que são as principais unidades político-administrativas do país. Os municípios representam o menor nível administrativo, que lida com a implementação de políticas e gestão local. As microrregiões e mesorregiões são regiões homogêneas, definidas pelo IBGE, como um conjunto de municípios contíguos, dentro de um mesmo estado. As microrregiões foram agrupadas de acordo com características naturais e de produção, e as mesorregiões são áreas maiores do que as microrregiões, definidas de acordo com as seguintes dimensões: aspectos sociais, ambiente natural, e rede de comunicação como um elemento de articulação espacial.

A seguir, apresentam-se os dados relativos à estatística descritiva das variáveis selecionadas, conforme as escalas geográficas em análise, neste estudo. Na Tabela 6 verificam-se os dados dos indivíduos com carteira de trabalho assinada.

Tabela 6 - Estatística descritiva dos trabalhadores com carteira de trabalho assinada por agregação geográfica- Censo 2010

Variáveis	Microdados	Municípios	Microrregião	Mesorregião	UF	
Escarlaridade	Média	7,52	6,92	7,20	7,41	7,67
	Mínimo	0	2,44	4,96	5,92	6,65
	Máximo	21	10,89	9,38	9,38	9,38
	Desvio Padrão	3,89	0,94	0,64	0,57	0,61
Rendimento	Média	1045,40	716,98	811,97	907,97	1019,50
	Mínimo	0	211,01	338,52	468,08	658,74
	Máximo	925000,00	5022,09	2508,01	2508,01	2508,01
	Desvio Padrão	3261,07	283,07	244,93	267,31	346,42
Idade	Média	38,73	38,89	38,77	38,75	38,87
	Mínimo	18	31,45	34,41	35,45	36,06
	Máximo	65	51,07	42,94	41,75	41,14
	Desvio Padrão	10,85	2,15	1,53	1,29	0,98
Raça	Proporção Brancos/ Não Brancos	0,53	0,48	0,46	0,45	0,40
	Desvio Padrão	0,50	0,24	0,21	0,20	0,17
	Proporção Homens/ Mulheres	0,70	0,72	0,72	0,71	0,71
Sexo	Desvio Padrão	0,46	0,10	0,06	0,05	0,04
Observações	1774612	5565	558	137	27	

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

A Tabela 7 expõe a estatística descritiva dos indivíduos que não possuem carteira de trabalho assinada, sendo que dentre estes encontram-se os informais, desempregados,

militares, funcionários públicos e profissionais liberais.

Percebe-se uma distinção de resultados ao comparar-se as variáveis por escala espacial. Estes resultados, díspares, tornam-se mais evidentes quando comparam-se os dois grupos: os trabalhadores com carteira assinada e os sem esta característica.

Tabela 7 - Estatística descritiva dos trabalhadores sem carteira de trabalho assinada por agregação geográfica- Censo 2010

Variáveis	Microdados	Municípios	Microrregião	Mesorregião	UF	
Escolaridade	Média	5,11	4,65	4,83	4,98	5,01
	Mínimo	0	2,12	3,03	3,43	3,87
	Máximo	21	9,25	7,61	7,09	7,09
	Desvio Padrão	3,38	0,88	0,86	0,85	0,76
Rendimento	Média	565,79	477,03	495,58	525,78	526,90
	Mínimo	0,00	94,66	151,25	196,71	253,37
	Máximo	975000,00	2960,85	1335,13	1335,13	1335,13
	Desvio Padrão	1997,14	257,40	242,02	239,94	258,31
Idade	Média	43,42	43,73	43,30	43,16	42,83
	Mínimo	18	31,26	37,81	38,54	39,93
	Máximo	65	51,93	48,42	47,23	46,52
	Desvio Padrão	12,72	2,55	2,13	2,00	1,82
Raça	Proporção Branco/ Não Branco	0,44	0,46	0,44	0,43	0,37
	Desvio Padrão	0,50	0,25	0,22	0,22	0,19
	Proporção Homens/ Mulheres	0,59	0,63	0,62	0,60	0,60
Sexo	Desvio Padrão	0,49	0,12	0,09	0,08	0,06
	Observações	3215298	5565	558	137	27

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

Pela estatística descritiva apresentada, pode-se denotar um claro indício do problema de MAUP, evidente quando comparam-se as diferentes escalas geográficas em estudo. Para análise destes dados, foram selecionados apenas os microdados dos chefes de família, com idade entre 18 e 65 anos.

3.5 Resultados empíricos

A partir de informações extraídas dos microdados do Censo de 2010, apresenta-se, na Tabela 8, inicialmente, a estimação por MQO, da Equação de Rendimento (Minceriana), para as diferentes agregações propostas, com introdução da variável Idade, conforme especificada a seguir:

$$\log_rendimento = \beta_0 + Escolaridade + Idade + Idade^2 + Raça + Sexo + \varepsilon_i \quad (3)$$

Como resultado, observa-se que, conforme se aumenta as agregações, das escalas geográficas utilizadas, mais influência a variável Escolaridade tem sobre o rendimento percebido pela população, daquela região. Verifica-se que, quando se analisa os resultados extraídos dos microdados, o retorno à educação para cada ano adicional de escolaridade é de 13,45% do valor da remuneração percebida. Ao agregarmos tais dados em diferentes escalas geográficas, a influência da educação aumenta, sendo que para os municípios o retorno à educação é de 35,8%; para as microrregiões, 40,34%; mesorregiões, 42,63%; e, chegando, por fim, a 46,40% dos salários da população, aglomerados em Estados brasileiros. Em todos os casos a variável foi significativa.

De acordo com a literatura, espera-se que, quanto maior o estoque de capital humano, maiores sejam as remunerações médias do trabalho, havendo uma significativa associação linear positiva, entre as duas variáveis.

Todas as regressões estimadas controlam por idade, raça e sexo. Contudo, a análise concentra-se no retorno à educação nas diferentes agregações geográficas, que é o objetivo central deste trabalho. As Tabelas 8 e 9, completas, encontram-se no Apêndice desta dissertação.

Tabela 8 - Resultados da regressão por MQO do rendimento per capita, por agregação regional – Modelo (1)

Variáveis	Agregação Regional				
	Microdados	Municípios	Microrregiões	Mesorregiões	Estados
Retorno à Escolaridade	0,1345*** (0,0001)	0,3580*** (0,0038)	0,4034*** (0,0087)	0,4263*** (0,0157)	0,4640*** (0,0311)
R ²	0,2872	0,8272	0,9117	0,9373	0,9672
R ² -ajustado	0,2872	0,8271	0,9109	0,9349	0,9594
N	5867594	5565	558	137	27

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

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.

Para as estimações realizadas na Tabela 9, a seguir ilustrada, optou-se pela introdução da variável Experiência, conforme regressão (4):

$$\log_rendimento = \beta_0 + Escolaridade + Experiência + Experiência^2 + Raça + Sexo + \varepsilon_i \quad (4)$$

Observa-se que a variável Escolaridade, continua representando grande influência sobre o rendimento. Todos os coeficientes são positivos e significativos para essa variável. Em semelhança com os resultados da Tabela 8, o retorno à Escolaridade aumenta seu coeficiente ao aumentar-se as agregações. Ao nível de microdados, seu resultado foi de 14,41%; aos municípios, 37,16%; microrregião, 41,27%; mesorregião, 43,93%; e nos estados, 48,01%. Os resultados da Tabela 9 confirmam os resultados da Tabela 8, e mostram que as regressões indicam maior retorno à educação, advindo de um ano adicional de escolaridade, quanto mais agregado forem os dados. Resultados esses, maiores que os encontrados para o Modelo (1).

Assim como nas regressões da tabela 8, foram utilizadas variáveis padrão da equação minceriana tais como Experiência, Experiência ao quadrado, Idade, e as *dummies* para Raça e para Sexo.

Tabela 9 - Resultados da regressão por MQO do rendimento per capita, por agregação regional – Modelo (2)

Variáveis	Agregação Regional				
	Microdados	Municípios	Microrregiões	Mesorregiões	Estados
Retorno à Escolaridade	0,1441*** (0,0001)	0,3716*** (0,0041)	0,4127*** (0,0098)	0,4393*** (0,0181)	0,4801*** (0,0371)
R ²	0,2827	0,8283	0,9131	0,9385	0,9674
R ² -ajustado	0,2827	0,8281	0,9123	0,9362	0,9596
N	5867594	5565	558	137	27

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

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 resultados, tanto no Modelo (1), quanto no Modelo (2), confirmam os relatos encontrado na literatura, ainda pouco extensa, sobre MAUP. Constata-se a variância dos coeficientes ao analisar-se os mesmos dados, porém em diferentes agregações regionais. Apesar da Equação Minceriana, para os dois modelos, mostrar valores condizentes com os encontrados na literatura, seus coeficientes, em especial para a variável Educação, modificaram-se, consideravelmente, em função do nível geográfico analisado. Esse

resultado pode estar, também relacionado, com o fato de que educação tem características de bem público e gera externalidades positivas para a sociedade. Assim, um ano adicional de educação tem mais benefício para a sociedade, do que para o indivíduo. Os resultados encontrados estão em linha com essa argumentação, uma vez que o retorno à educação é menor no nível do microdado.

3.6 Considerações finais

O presente estudo buscou aplicar a análise econométrica espacial, à questão do rendimento salarial para o território brasileiro, com destaque para o Problema da Unidade de Área Modificável, MAUP.

A Equação de Rendimentos (Minceriana) foi estimada em cinco diferentes aglomerações regionais: microdados, municípios, microrregiões, mesorregiões e estados brasileiros. Ao analisá-la, sob a perspectiva de diferentes escalas regionais, o papel de suas variáveis torna-se de suma importância, com destaque para a educação.

Seguiu-se os prepostos, da equação clássica, de Mincer e formulou-se para este trabalho dois modelos, estimados por MQO; um contendo como variáveis explicativas a Educação, a Idade, a Raça e o Sexo do indivíduo (1); e outro, substituindo a variável Idade pela Experiência (2).

Em ambos os modelos a educação mostrou-se significativa e com sinal positivo, alcançando diferentes patamares, de acordo com a escala espacial analisada. Para o modelo (1), o coeficiente para o retorno à educação variou consideravelmente entre as cinco aglomerações regionais estudadas, passando de 13% para os microdados; 35%, para os municípios; 40%, para as microrregiões; 42%, para as mesorregiões; e 46% para os estados brasileiros. Para o modelo (2), onde a experiência substituiu a idade, os resultados foram parecidos com os anteriores, porém, em um maior grau, com 14% para os microdados; 37% para os municípios; 41%, para as microrregiões; 43%, para as mesorregiões; e 48% para os estados. Esses resultados mostram a discrepância que ocorre quando analisa-se um mesmo dado, porém em diferentes aglomerações regionais.

Os resultados mostram que as conclusões obtidas, a partir de regressões de rendimento são dependentes da escolha do nível geográfico analisado. Os estudos, para políticas regionais, devem considerar os resultados com base em várias escalas espaciais, visando a tomada de decisões políticas mais eficazes. Pesquisas, que usam apenas uma escala geográfica, podem fornecer informações divergentes. A política adotada pode não

estar correta, para uma determinada escala geográfica, podendo até haver resultados opostos aos planejados para aquela região.

O fato de os resultados terem revelado dinâmicas territoriais específicas a cada escala de análise, demonstrou que não existe uma escala de análise capaz de sintetizar todas as respostas procuradas, de forma mais precisa e mais clara que as outras. Ou seja, a melhor escala de análise depende do objetivo da análise e da resposta buscada. Ademais, uma abordagem em diferentes escalas geográficas, demonstra-se útil para um melhor entendimento das disparidades das equações de rendimentos regionais, evitando, assim, conclusões precipitadas, acerca do fenômeno estudado.

4. CONCLUSÃO

Este trabalho teve como objetivo ampliar o conhecimento sobre a influência que as economias de aglomeração e o MAUP têm sobre o salário dos trabalhadores brasileiros.

Em seu primeiro ensaio, essa dissertação apresentou evidências empíricas, das economias de aglomeração, nas atividades econômicas dos municípios brasileiros no ano de 2010, utilizando-se de equações salariais com formulação baseada no modelo microeconômico desenvolvido por Fingleton (2003). 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, por sua vez, valores de 0,76 e 0,56 para a taxa salarial e a densidade do emprego, respectivamente; enquanto que o teste bivariado retornou valor de 0,21, confirmando a autocorrelação existente entre tais variáveis.

Na segunda parte do primeiro artigo, dedicada à estimação do modelo econométrico, foram utilizadas matrizes de pesos espaciais do tipo *Queen* = 1, 10 e 50. A taxa salarial defasada variou de 0,28 a 0,42 para essas matrizes. Os resultados obtidos 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; à presença de spillovers ou transbordamentos espaciais, entre a taxa salarial, dos municípios geograficamente próximos; e, aos níveis de eficiência dos trabalhadores.

O segundo ensaio, deste trabalho, buscou ampliar o debate acerca da desigualdade brasileira. Seu objetivo foi de mensurar equações de rendimentos mincerianas, para diferentes escalas geográficas brasileiras (microdados, municípios, microrregiões, mesorregiões e estados), identificando e verificando o problema de MAUP para tais regiões. Foi utilizado, como base de dados, os microdados do Censo 2010, disponibilizados pelo Instituto Brasileiro de Geografia e Estatística (IBGE).

Para tal, seguiu-se os prepostos, da equação clássica, de Mincer (1974) e formulou-se para o artigo dois modelos, estimados por MQO; um contendo como variáveis explicativas a Educação, a Idade, a Raça e o Sexo do indivíduo; e outro, substituindo a variável Idade pela Experiência.

Em ambos os modelos a educação mostrou-se significativa e com sinal positivo, alcançando diferentes patamares, de acordo com a escala espacial analisada. Para o modelo (1), o coeficiente para o retorno à educação variou consideravelmente entre as cinco aglomerações regionais estudadas, passando de 13% para os microdados; 35%, para os municípios; 40%, para as microrregiões; 42%, para as mesorregiões; e 46% para os estados brasileiros. Para o modelo (2), onde a experiência substituiu a idade, os resultados foram parecidos com os anteriores, porém, em um maior grau, com 14% para os microdados; 37% para os municípios; 41%, para as microrregiões; 43%, para as mesorregiões; e 48% para os estados. Esses resultados mostraram a discrepância que ocorre quando analisa-se um mesmo dado, porém em diferentes aglomerações regionais.

Portanto, este trabalho expôs como as economias de aglomeração têm influência, sobre o nível salarial dos indivíduos, trabalhadores brasileiros. Constatou-se que, em exposição a externalidade positivas, o salário diferencia-se, dependendo de sua região. Outra conclusão, em que esta dissertação chega, é que a forma como são analisados os dados salariais influencia, definitivamente, os resultados finais. O MAUP é responsável pela diferença encontrada, nas equações mincerianas, ao analisá-las em diferentes escalas regionais.

Como sugestão de próximos estudos, podem ser verificadas mais profundamente algumas aglomerações salariais brasileiras. Assim como, utilizados procedimentos econométricos mais avançados, para analisar-se o MAUP, para o salário, em diferentes escalas regionais brasileiras.

REFERÊNCIAS

- ALMEIDA, Eduardo. **Econometria Espacial**. Campinas - SP. Alínea, 2012.
- AMARANTE, Patrícia Araújo. **Economias de Aglomeração nas Atividades Econômicas dos Municípios Brasileiros nos Anos de 2000 e 2009-Evidências a Partir de Equações Salariais**. Dissertação (Mestrado em Economia)- UFPB, João Pessoa-PB, 2011.
- ANSELIN, Luc. **Spatial econometrics: methods and models**. Springer Science & Business Media, 1988.
- ANSELIN, Luc. Local indicators of spatial association - LISA. **Geographical Analysis**, v.27, n.2, p. 91-115, 1995.
- ANSELIN, Luc. **Spatial econometrics**. Dallas: Bruton Center, School of Social Sciences, University of Texas, 1999.
- ARBIA, Giuseppe. **Spatial econometrics: statistical foundations and applications to regional convergence**. Springer Science & Business Media, 2006.
- BADIA, D. B; FIGUEREDO, L. Impacto das externalidades dinâmicas de escala sobre o crescimento do emprego industrial nas cidades brasileiras. In ENCONTRO REGIONAL DE ECONOMIA, 12., 2007, Fortaleza. **Anais...**, 2007.
- BANDEIRA, Luciana; JACINTO, Paulo de Andrade. 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**, 2010, Porto Alegre.
- BARROS, Alexandre Rands. **Desigualdades regionais no Brasil: natureza, causas, origens e soluções**. 2011.
- BECKER, Gary Stanley. **Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education**, by Gary S. Becker. London, 1964.
- CANGUSSU, Ricardo Corrêa; SALVATO, Márcio Antônio; NAKABASHI, Luciano. Uma análise do capital humano sobre o nível de renda dos estados brasileiros: MRW versus Mincer. **Estudos Econômicos (São Paulo)**, v. 40, n. 1, p. 153-183, 2010.
- CARVALHO, Emerson Rildo. Araújo de. Desenvolvimento humano e taxa de retorno da educação. **Pensamento & Realidade. Revista do Programa de Estudos Pós-Graduados em Administração-FEA**. 2007 v.20.
- CATELA, Eva Yamila da Silva; GONÇALVES, Flávio de Oliveira. Economias de localização versus urbanização e os estágios de desenvolvimento dos municípios brasileiros nos anos 1997 e 2007. In: ENCONTRO NACIONAL DE ECONOMIA, 37., 2009, Foz do Iguaçu. **Anais...**, Foz do Iguaçu, 2009.

CHAGAS, André Luis Squarize. **Externalidades da aglomeração: micro fundamentação e evidências empíricas**. 2004. 132f. Dissertação (Mestrado em Economia) – Universidade de São Paulo, São Paulo, 2004.

CHAVES, André Luiz Leite. Determinação dos rendimentos na Região Metropolitana de Porto Alegre: uma verificação empírica da Teoria do Capital Humano. **Ensaio FEE**, v. 23, p. 399-420, 2002.

CRAVO, Túlio Antônio. SMEs and economic growth in the Brazilian micro-regions. **Papers in regional science**, v. 89, p. 711-734, 2010.

CRAVO, Túlio Antônio. RESENDE, Guilherme Mendes. Economic growth in Brazil: a spatial filtering approach. **The annals of regional science**, 2012.

DA SILVA, Magno Vamberto Batista; SILVEIRA NETO, Raul da Mota. Dinâmica da concentração da atividade industrial no Brasil entre 1994 e 2004: uma análise a partir de economias de aglomeração e da nova geografia econômica. **Revista de Economia Aplicada**, v. 13, p. 299-331, 2009.

DA SILVA JÚNIOR, Luiz Honorato; SAMPAIO, Yony. Pobreza e Demanda por Educação no Brasil: Uma Análise à Luz da Teoria do Capital Humano. **Ensaio FEE**, v. 34, n. 1, 2013.

DANIEL, Lingomar Pegorini; OLIVEIRA, Leandro José de. Análise dos principais fatores determinantes do nível de renda no estado de Mato Grosso: um estudo com base nos microdados da PNAD. **Revista de Estudos Sociais**, v. 14, n. 27, 2014.

DE ÁVILA, Rodrigo Peres; MONASTERIO, Leonardo. O Maup e a Análise Espacial: um estudo de caso para o Rio Grande do Sul (1991-2000). **Análise Econômica**, v. 26, n. 49, 2006.

DENARDIN, Valdir F.; SULZBACH, Mayra T. **Capital natural crítico: a operacionalização de um conceito**. 2005. Disponível em: http://www.ecoeco.org.br/conteudo/publicacoes/encontros/vi_en/artigos/ mesa4/capital_natural_critico.pdf. Acesso em: 07/03/2015.

DOMINGUES, Edson Paulo. Aglomerações e Periferias Industriais no Brasil e no Nordeste. **Revista Econômica do Nordeste**, v.36, n.4, p.508-523, 2005.

ELHORST J. P.; PIRAS G.; ARBIA G. Growth and convergence in a multi-regional model with space-time dynamics. **Geographical analysis**, v. 42, n. 3, p. 338-355, 2010. Enfoque da Nova Geografia Econômica. In: ENCONTRO REGIONAL DE ECONOMIA, 9., 2004, Fortaleza. **Anais...** Fortaleza, 2004.

FALCÃO, Natasha de Andrade; NETO, Raul da Mota Silveira. 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.

FERNANDES, Reynaldo. **Desigualdade salarial: aspectos teóricos. Estrutura salarial: aspectos conceituais e novos resultados para o Brasil.** Rio de Janeiro: Ipea, p. 1-50, 2002.

FINGLETON, Bernard. Increasing returns: evidence from local wage rates in Great Britain. **Oxford Economic Papers**, v.55, p.716-739, 2003.

FINGLETON, Bernard. The empirical performance of the NEG with reference to small areas. **Journal of Economic Geography**, v. 11, n. 2, p. 267-279, 2011.

GALINARI, Rangel *et al.* O efeito das economias de aglomeração sobre os salários industriais: uma aplicação ao caso brasileiro. **Revista de Economia Contemporânea**, Rio de Janeiro, v. 11, n. 3, p. 391-420, set./dez. 2007.

GALINARI, Rangel. **Retornos crescentes urbano-industriais e spillovers espaciais: evidências a partir da taxa salarial no estado de São Paulo.** 2006. 162 f. Dissertação (Mestrado em Economia) - UFMG, CEDEPLAR, Minas Gerais, 2006.

GALINARI, Rangel; LEMOS, M. Borges; 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, André Braz. 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: Randon House, 1969.

KROTH, Darlan Christiano; DIAS, Joilson. Os efeitos dos investimentos público e privado em capitais físico e humano sobre o produto per capita dos municípios da região sul: uma análise em painéis de dados dinâmicos. **Nova Economia**, v. 22, n. 3, p. 621-649, 2012.

LAM, David; LEVISON, Deborah. Idade, experiência, escolaridade e diferenciais de renda: Estados Unidos e Brasil. **Pesquisa e Planejamento Econômico**, p. 219-256, 1990.

LANGONI, Carlos Geraldo. Distribuição de Renda e Desenvolvimento Econômico no Brasil. **Rio de Janeiro: Expressão e Cultura**, 1973.

LAZZAROTTO, J. J.; LIMA, J. E. de. Análise espacial dos principais determinantes da renda per capita dos municípios brasileiros. In: XLVI Congresso Brasileiro de Economia e Sociologia Rural, 2008, Rio Branco. **Anais do XLVI Congresso Brasileiro de Economia e Sociologia Rural. Rio Branco: SOBER, 2008.** Disponível em: <http://www.sober.org.br/palestra/9/891.pdf>.

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.

LEAL, Carlos Ivan Simonsen; WERLANG, Sérgio Ribeiro da Costa. Retornos em educação no Brasil: 1976/89. **Pesquisa e planejamento Econômico**, v. 21, n. 3, p. 559-574, 1991.

LEMOS, Mauro Borges *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**. Department of Economics, University of Toledo, Maio, 1999.

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

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

LIMA, R. C. A.; MENEZES, T. A.. **Uma Análise Espacial das Eleições Presidenciais Brasileiras de 2010**. In: 2nd Ibero-American Congress on Regional Development and 12th Congress of the Brazilian Regional Science Association, 2014, Belo Horizonte. ABER, 2014.

MARSHALL, Alfred. **Princípios de Economia: tratado introdutório**. Tradução revista de Rômulo Almeida e Ottolmy Strauch. São Paulo: Nova Cultural, 1996.

MENEZES FILHO, N. A. **A evolução da educação no Brasil e seu impacto no mercado de trabalho**. São Paulo: USP/ Departamento de Economia, março, 2001.

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

MINCER, Jacob A. Schooling and earnings. In: **Schooling, experience, and earnings**. Columbia University Press, 1974. p. 41-63.

MINCER, Jacob A. Schooling and earnings. In: **Schooling, experience, and earnings**. Columbia University Press, 1974. p. 41-63.

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

OLIVEIRA, Cristiano Aguiar. Crescimento Econômico das Cidades Nordestinas: um Enfoque da Nova Geografia Econômica. In: ENCONTRO REGIONAL DE ECONOMIA, 9., 2004, Fortaleza. **Anais...** Fortaleza, 2004.

PÁEZ, Antonio; SCOTT, Darren M. Spatial statistics for urban analysis: a review of techniques with examples. **GeoJournal**, v. 61, n. 1, p. 53-67, 2005.

PEREIRA, Rosângela Saldanha; ZAVALA, Arturo Zavala. Educação e rendimentos do trabalho no Brasil: desafios de um cenário de transformações. **Si Somos Americanos. Revista de Estudios Transfronterizos**, v. 12, n. 2, p. 203-227, 2013.

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

RAMOS, Frederico Roman. Análise espacial de estruturas interurbanas: o caso de São Paulo. **São José dos Campos**, 2002.

RESENDE, Guilherme Mendes. Multiple dimensions of regional economic growth: The Brazilian case, 1991– 2000. **Papers in Regional Science**, v. 90, n. 3, p. 629-662, 2011.

RESENDE, Guilherme Mendes. Spatial dimensions of economic growth in Brazil. **ISRN Economics**, v. 2013, p. 1-19, 2013.

RESENDE, Guilherme Mendes; CARVALHO, Alexandre Xavier Ywata de; SAKOWSKI, Patrícia Alessandra Morita. Avaliando o crescimento econômico no Brasil em múltiplas escalas espaciais com a utilização de modelos de painel espacial (1970-2000). Brasília: Ipea, 2013 (Texto para Discussão, 1830).

RESENDE, Guilherme Mendes; MAGALHÃES, João Carlos Ramos. Disparidades do produto interno bruto (PIB) per capita no Brasil: uma análise de convergência em diferentes escalas regionais (1970-2008). Brasília: Ipea, 2013 (Texto para Discussão, 1833).

RESENDE, Marcelo; WYLLIE, Ricardo. **Retorno para Educação no Brasil: Evidências Empíricas Adicionais**. Texto para discussão n.03, UFRJ, Rio de Janeiro, 2005.

ROCHA, Fabiano Olanda Sales; MENEZES, Francisca Livia Souza. Retornos à Educação e Discriminação no Mercado de Trabalho Brasileiro: Evidências por Regressões Quantílicas em Categorias Ocupacionais. In: **Anais do XL Encontro Nacional de Economia**. ANPEC - Associação Nacional dos Centros de Pós-graduação em Economia, 2014.

RODRIGUES, Adão Aparecido Ferreira. **Impacto da educação no rendimento salarial no Brasil de 2001 a 2008**. 2010.

SACHSIDA, Adolfo; LOUREIRO, Paulo Roberto Amorim; MENDONÇA, Mário Jorge Cardoso de. Um estudo sobre retorno em escolaridade no Brasil. **Revista Brasileira de Economia**, v. 58, n. 2, p. 249-265, 2004.

SHANKAR, R.; SHAH, A. Bridging the economic divide within countries: a scorecard on the performance of regional policies in reducing regional income disparities. **World development**, v. 31, n. 8, p. 1.421-1.441, 2003.

SALVATO, Márcio Antônio; SILVA, Denis Gomes. O impacto da educação nos rendimentos do trabalhador: uma análise para região metropolitana de belo horizonte. **SEMINÁRIO SOBRE A ECONOMIA MINEIRA, XIII, Anais**, 2008.

SCALON, Celi. Desigualdade, pobreza e políticas públicas: notas para um debate. **Contemporânea-Revista de Sociologia da UFSCar**, v. 1, n. 1, p. 49, 2012.

SCHULTZ, Theodore William. **The economic value of education**. New York: Columbia University Press, 1963.

SEN, Amartya Kumar. **Desigualdade reexaminada**. Record, 2001.

SEN, Amartya Kumar; MENDES, Ricardo Doninelli. **Desenvolvimento como liberdade**. São Paulo: Companhia das Letras, 2000.

SILVA, Almir Bittencourt da; MARINHO, Emerson. Capital Humano e Crescimento Econômico: um Reexame da Abordagem Minceriana para Grupos de Países. **Análise Econômica**, v. 26, n. 49, 2008.

SILVEIRA NETO, Raul da Mota. Concentração Industrial Regional, Especialização Geográfica e Geografia Econômica: Evidências para o Brasil no Período 1950-2000. **Revista Econômica do Nordeste**, v.36, n.2, Fortaleza, 2005.

SIQUEIRA, Luiz Paulo Ribeiro. Proposta metodológica para a identificação de subcentros urbanos: estudo de caso na região metropolitana de São Paulo. **Economia Aplicada**, v. 18, n. 1, p. 139-165, 2014.

SOLOW, Robert M. A contribution to the theory of economic growth. **The quarterly journal of economics**, v. 70, n. 1, p. 65-94, 1956.

SULIANO, Daniel Cirilo; SIQUEIRA, Marcelo Lettieri. Retornos da educação no Brasil em âmbito regional considerando um ambiente de menor desigualdade. **Economia Aplicada**, v. 16, n. 1, p. 137-165, 2012.

VIANA, Giomar; DE LIMA, Jandir Ferrera. Capital humano e crescimento econômico. **Revista Internacional de Desenvolvimento Local**, v. 11, n. 2, p. 137, 2010.

VIEIRA, Rodrigo de Souza. **Crescimento Econômico no Estado de São Paulo: uma análise espacial**. 1. ed. Sao Paulo: Fundação Editora Unesp, 2010. 106p

WU, Xiaoyu; CUTTER, Bowman. Who votes for public environmental goods in California?: Evidence from a spatial analysis of voting for environmental ballot measures. **Ecological Economics**, v. 70, n. 3, p. 554-563, 2011.

APÊNDICE

Tabela 8 - Resultados da regressão por MQO do rendimento per capita, por agregação regional – Modelo (1)

Variáveis	Agregação Regional				
	Microdados	Municípios	Microrregiões	Mesorregiões	Estados
Retorno à Escolaridade	0,1345*** (0,0001)	0,3580*** (0,0038)	0,4034*** (0,0087)	0,4263*** (0,0157)	0,4640*** (0,0311)
Idade	- 0,0201*** (0,0001)	-0,0949*** (0,0074)	-0,1362*** (0,0242)	-0,2153*** (0,0552)	-0,4517*** (0,1574)
Idade ²	0,0004*** (9,59e-07)	0,0012*** (0,0001)	0,0016*** (0,0003)	0,0025*** (0,0006)	0,0053** (0,0019)
Dummy para Raça	0,4024*** (0,0008)	0,6130*** (0,0143)	0,4641*** (0,0365)	0,2923*** (0,0683)	0,1399 (0,1678)
Dummy para Sexo	0,0878*** (0,0008)	0,5663*** (0,0361)	1,1351*** (0,1274)	1,4084*** (0,2679)	2,0785*** (0,6650)
Constante	5,0441*** (0,0020)	5,7517*** (0,1487)	6,2971*** (0,4673)	7,7863*** (1,0441)	12,1349*** (2,9487)
R ²	0,2872	0,8272	0,9117	0,9373	0,9672
R ² -ajustado	0,2872	0,8271	0,9109	0,9349	0,9594
N	5867594	5565	558	137	27

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

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.

Tabela 9 - Resultados da regressão por MQO do rendimento per capita, por agregação regional – Modelo (2)

Variáveis	Agregação Regional				
	Microdados	Municípios	Microrregiões	Mesorregiões	Estados
Retorno à Escolaridade	0,1441*** (0,0001)	0,3716*** (0,0041)	0,4127*** (0,0098)	0,4393*** (0,0181)	0,4801*** (0,0371)
Experiência	-0,0068*** (0,0001)	-0,0705*** (0,0051)	-0,1056*** (0,0163)	-0,1602*** (0,0373)	-0,2962*** (0,1019)
Experiência ²	0,0003*** (9.12e-07)	0,0012*** (0,0001)	0,0016*** (0,0003)	0,0025*** (0,0006)	0,0047*** (0,0016)
Dummy para Raça	0,4140*** (0,0008)	0,6103*** (0,0141)	0,4594*** (0,0354)	0,2936*** (0,0652)	0,2055 (0,1487)
Dummy para Sexo	0,0843*** (0,0008)	0,5872*** (0,0362)	1,2108*** (0,1288)	1,5267*** (0,2752)	2,1399*** (0,6764)
Constante	4,7556*** (0,0015)	4,7770*** (0,0715)	4,9337*** (0,2030)	5,5365*** (0,4348)	6,9746*** (1,1617)
R ²	0,2827	0,8283	0,9131	0,9385	0,9674
R ² -ajustado	0,2827	0,8281	0,9123	0,9362	0,9596
N	5867594	5565	558	137	27

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

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.