



UNIVERSIDADE DE BRASÍLIA (UnB)
FACULDADE DE ADMINISTRAÇÃO, CONTABILIDADE, ECONOMIA E
GESTÃO DE POLÍTICAS PÚBLICAS
DEPARTAMENTO DE ECONOMIA

VINÍCIUS HIAGO E SILVA GERÔNIMO

DIVERSIDADE INDUSTRIAL E ESTABILIDADE ECONÔMICA
REGIONAL: UMA ANÁLISE ESPACIAL SOBRE UNIDADES
GEOGRÁFICAS BRASILEIRAS

Brasília, DF

2023

VINÍCIUS HIAGO E SILVA GERÔNIMO

**DIVERSIDADE INDUSTRIAL E ESTABILIDADE ECONÔMICA
REGIONAL: UMA ANÁLISE ESPACIAL SOBRE UNIDADES
GEOGRÁFICAS BRASILEIRAS**

Monografia de Conclusão de Curso apresentado à
Universidade de Brasília como exigência para
obtenção do título de bacharel em Ciências
Econômicas.

Orientador: Prof. Dr. Marcelo de Oliveira Torres

Brasília, DF

2023

AGRADECIMENTOS

Agradeço, primeiramente, a todo amor e confiança depositados em mim por minha família. À minha mãe, Michelle, ao meu pai, Higor, às minhas irmãs, Ana Luisa e Júlia, aos meus avós, tios, primos e todos os demais. Eu não seria nada sem a presença da minha família na minha vida e dedico este trabalho e todas as minhas conquistas a vocês. Obrigado por tudo.

Agradeço ao professor Marcelo Torres por todo o ensinamento e ajuda cedidos para a realização deste trabalho. Agradeço a todos os professores da FACE/UnB que tive o prazer de conhecer na graduação pelo imensurável aprendizado ao longo desses anos, especialmente à professora Milene Takasago, por ter me aceito como monitor e orientando no PIBIC.

Agradeço à equipe da Gerência de Economia da CNI, que tive a sorte de conhecer na minha trajetória profissional, pelos ensinamentos passados e pela ideia do tema deste trabalho. Aos professores e toda a equipe educacional do Colégio Madre Carmen Sallés, que formaram o aluno que sou hoje.

Por fim, agradeço a todos os amigos e amigas que fiz na graduação, pelas risadas dadas, pelo companheirismo, pelas manhãs e tardes que estiveram comigo, pelo afeto e pelos desafios compartilhados. Carregarei todos no meu coração e sentirei muita falta dos nossos momentos juntos, do primeiro ao último minuto.

“I wish there was a way to know you're in the good old days before you've actually left them.”

(Andy Bernard, The Office)

RESUMO

A correlação entre diversidade industrial e estabilidade econômica regional é defendida teoricamente como sendo positiva, indicando que uma região que possui uma maior diversificação industrial sofre uma menor interferência de choques exógenos, resultando em uma menor instabilidade econômica. Entretanto, historicamente trabalhos empíricos encontram resultados divergentes, muitas vezes não sendo observado uma significância na correlação. Dentre os possíveis motivos que podem levar a tal divergência, este trabalho foca no Problema de Unidade de Área Modificável (MAUP), enfatizando a relevância que a mudança de escalas geográficas possui sobre os resultados devido, neste caso, à interferência espacial dos dados. Para realizar este estudo, utiliza-se econometria espacial e dados da RAIS e do Censo Demográfico a fim de coletar informações de municípios, microrregiões e mesorregiões brasileiras entre os anos de 2010 e 2019. Os resultados indicam uma influência espacial em todos os níveis geográficos, além de apontarem uma variação na magnitude, significância e no sinal da correlação a depender da escala utilizada, confirmando a hipótese do MAUP no estudo. Além disso, o nível municipal foi o único que apresentou resultados mais coerentes com a teoria no caso brasileiro.

PALAVRAS-CHAVE: diversidade industrial, estabilidade econômica, MAUP, escalas geográficas, econometria espacial.

ABSTRACT

The correlation between industrial diversity and regional economic stability is theoretically advocated as positive, indicating that a region with greater industrial diversification experiences less interference from exogenous shocks, resulting in lower economic instability. However, historically, empirical studies have found divergent results, often observing a lack of significance in the correlation. Among the possible reasons that can lead to such divergence, this study focuses on the Modifiable Areal Unit Problem (MAUP), emphasizing the relevance of geographic scale changes on the results due to spatial interference of data. To conduct this study, spatial econometrics and data from the RAIS and Demographic Census are used to gather information from Brazilian municipalities, micro-regions, and meso-regions between 2010 and 2019. The results indicate spatial influence at all geographic levels, as well as variation in the magnitude, significance, and direction of the correlation depending on the scale used, confirming the hypothesis of MAUP in the study. Furthermore, the municipal level was the only one that presented results more consistent with the theory in the Brazilian case.

KEYWORDS: industrial diversity, economic stability, MAUP, geographic scales, spatial econometrics.

LISTA DE FIGURAS

Figura 1 - Os problemas de formato e de escala do MAUP.....	22
---	----

LISTA DE TABELAS

Tabela 1 - Quadro resumo das pesquisas sobre diversidade econômica.....	20
Tabela 2 - Descrição das variáveis.....	32
Tabela 3 - Estatística descritiva das variáveis HHI e REI.....	33
Tabela 4 - Índices de Moran global (10 vizinhos mais próximos).....	35
Tabela 5 - Resultados da regressão MQO.....	36
Tabela 6 - Comparação do Likelihood Ratio Tests (10 vizinhos mais próximos).....	47
Tabela 7 - Resultados das regressões espaciais (10 vizinhos mais próximos).....	38
Tabela 8 - Estimação dos efeitos diretos, indiretos e totais do HHI (10 vizinhos mais próximos).....	39
Tabela 9 - Índice de Moran global para diferentes matrizes de vizinhança.....	41
Tabela 10 - Estimação dos efeitos diretos, indiretos e totais do HHI para diferentes matrizes de vizinhança.....	42

SUMÁRIO

1. INTRODUÇÃO.....	10
2. REVISÃO TEÓRICA.....	12
2.1 A importância do espaço e da geografia na Ciência Econômica.....	12
2.2 A correlação entre diversidade industrial e estabilidade econômica regional.....	15
2.3 Problema da Unidade de Área Modificável (MAUP).....	20
3. METODOLOGIA.....	23
3.1 Modelos econométricos espaciais.....	23
3.1.1 Das limitações dos modelos econométricos tradicionais.....	23
3.1.2 Da formalização das matrizes de vizinhança.....	25
3.1.3 Da formalização dos modelos econométricos espaciais.....	26
3.2 Índice de Moran Global.....	28
3.3 HHI e REI.....	29
3.4 Dados e modelo empírico.....	30
3.4.1 Estatísticas descritivas.....	32
4. RESULTADOS EMPÍRICOS.....	35
4.1 Análise exploratória espacial.....	35
4.2 Análise das regressões.....	35
4.3 Análise de sensibilidade.....	40
5. CONSIDERAÇÕES FINAIS.....	43
REFERÊNCIAS BIBLIOGRÁFICAS.....	45

1. INTRODUÇÃO

O objetivo deste estudo é analisar a contribuição da diversidade industrial para a estabilidade econômica regional controlando para fatores espaciais e, além disso, examinar se a correlação positiva entre estabilidade econômica e diversidade industrial é observada em diferentes escalas geográficas brasileiras.

Com a modernização dos sistemas de informações geográficas (do inglês, GIS), houve um avanço, nas últimas décadas, na manipulação e análise de dados espaciais, aumentando a demanda por ferramentas metodológicas espaciais em diversas áreas das ciências sociais, como economia, sociologia, geografia, ciência política etc. Na ciência econômica, as contribuições da Nova Geografia Econômica permitiram com que modelos pudessem incorporar efeitos spillovers, competição imperfeita, externalidades espaciais e concentração geográficas (Krugman 1998, 1996).

No ramo da economia regional, é um consenso na literatura que uma performance econômica regional mais estável está correlacionada com uma maior diversidade industrial (e.g. Chinitz, 1961; Conroy, 1974; Kort, 1981). Este fenômeno tende a ocorrer pois na presença de múltiplas especializações, os múltiplos ciclos sazonais das atividades econômicas tendem a balancear as variações da produção e renda das regiões, diferente de áreas com poucas especializações. Além disso, na ocorrência de choques externos, o impacto econômico em regiões mais diversas industrialmente inclina-se a ser menor, pois os choques impactam os setores em momentos distintos, atingindo a produção e o emprego em proporções menores.

Apesar da correlação positiva ser um consenso teórico, muitos trabalhos empíricos divergiram em seus resultados. A principal problemática encontrada nos estudos feitos (à qual esta monografia foca em explorar e corrigir) é a desconsideração do efeito espacial nos dados, não levando em conta a heterogeneidade espacial e o efeito de contágio da mudança nas variáveis. Enquanto pesquisadores como Kort (1981), Malizia e Ke (1993) e Hong (2016) confirmaram a validade de correlação positiva, outros como Attaran (1986), Mizuno (2006) e Deller e Watson (2016) observaram que essa relação não é sempre estatisticamente significativa, sendo que apenas o último trabalho utilizou econometria espacial.

Visando realizar uma análise complementar, além dos municípios, foram incorporados micro e mesorregiões na análise da influência da diversidade industrial na performance econômica de regiões, seguindo o proposto por Anselin (1988) e Chen (2018). A motivação para tal análise se apoia no fato de poder existir diferentes padrões espaciais em cada escala,

podendo alterar a informação presente em cada uma e, assim, distorcer a correlação a depender do tipo de agregação utilizado, fenômeno denominado como Problema de Unidade de Área Modificável (do inglês, MAUP).

Neste trabalho, foram utilizados dados de 5.564 municípios, que posteriormente foram agregados em 558 microrregiões e 137 mesorregiões, visando explorar o MAUP. Para estudar a contribuição da diversidade na estabilidade regional, aplicaram-se métodos econométricos espaciais e foram estimados os modelos SAR, SEM e SDM em cada unidade geográfica, buscando encontrar o melhor modelo para cada unidade geográfica. Os dados referentes à diversidade industrial e às variáveis de controle são do ano de 2010, enquanto a estabilidade econômica foi calculada entre 2010 e 2019.

O segundo capítulo irá apresentar o crescimento do interesse da análise espacial nas ciências sociais, mais especificamente nas Ciências Econômicas. Além disso, serão abordadas questões sobre as diferentes métricas de diversidade industrial regional e o Problema de Unidades de Área Modificável (MAUP). No capítulo 3, é apresentado o método utilizado, enquanto o quarto capítulo examinará os resultados obtidos das regressões e da análise exploratória espacial feita. O quinto e último capítulo irá discutir as conclusões a que o trabalho chega, os cuidados a serem tomados com estas e a contribuição acadêmica da monografia.

2. REVISÃO TEÓRICA

2.1 A importância do espaço e da geografia na Ciência Econômica

Ao longo das últimas décadas, dados espaciais e sistemas de informações geográficas ganharam maior importância nos campos das Ciências Sociais, por permitirem acrescentar aos modelos variáveis locais e interações de espaço e tempo. Com isso, a abordagem espacial ganhou mais atenção nos processos sociais, sendo incorporada na Ciência Econômica, Sociologia, Psicologia, Administração, Política etc (GOODCHILD *et al*, 2000).

Paul Krugman (1991), sendo um dos nomes mais conhecidos na geografia econômica, se voltou a estudar técnicas derivadas da organização industrial incorporando fatores locais em seus modelos formais. Em seu trabalho, o autor derivou modelos de concentração geográfica da indústria de transformação baseados em economias de escala, buscando formalizar matematicamente a teoria do ‘feedback positivo industrial’, que defende que o processo de concentração industrial em poucos locais é, dependendo das hipóteses assumidas, auto estimulante, resultado dos benefícios que os efeitos *spillovers* de indústrias próximas trazem a logística produtiva.

Segundo o modelo de Krugman (1991), assume-se que a economia é dividida em duas regiões, com dois setores em cada região, sendo o número de trabalhadores total distribuído entre elas e permitindo o fluxo de trabalhadores entre regiões a depender da região que possui o maior salário real em comparação a outra. O autor concluiu que o efeito do deslocamento do trabalho para uma determinada região no salário real daquela região vai depender do efeito líquido da influência da maior competição no mercado de trabalho (que tende a diminuir o salário dos trabalhadores) e do maior mercado (que atua elevando o salário do mercado). Caso o efeito sobre o salário real da região seja positivo (negativo), então um processo de divergência (convergência) regional é observado na economia.

Puga *et al* (1996) estudam a importância de se analisar as ligações para trás e para frente dos setores, a variação dos salários e os custos de alocação no processo de industrialização dos países. Em seu modelo, existem N países, que podem produzir bens industriais e agro alimentares, possuindo todos a mesma quantidade de capital e trabalho. A quantidade de firmas operando em cada país é endogenamente determinada. Segundo os autores, existem duas forças que encorajam a dispersão industrial, sendo elas o aumento dos salários (e redução do lucro) quando atuam em grandes mercados e a maior competição entre firmas, enquanto o fato de se aglomerarem em poucas regiões reduz os custos de transporte, os custos com insumos e aumenta as ligações setoriais.

Além do processo de dispersão em si, a velocidade da sua ocorrência também pode variar. Mais especificamente, um número baixo de ligações para frente e para trás¹ diminui as forças de aglomeração, fazendo a indústria se espalhar entre regiões. Além disso, muitas barreiras de entrada e poucas barreiras incentivam também a dispersão geográfica, visto que o primeiro caso tende a dividir produtores e consumidores finais e o segundo caso permite com que indústrias realizem a custos muito baixos. Outro fator que pode levar à dispersão geográfica, segundo Puga *et al* (1996), é o fato de ser mais lucrativo para indústrias que são mais trabalho intensivas se mudarem para regiões com menores mercados, visto que os salários reais tendem a ser menores. Caso essas indústrias tenham pouca dependência da proximidade de seus fornecedores, esse movimento tende a ser reforçado.

Explicitando a importância da localização na competição entre países, Porter (1990) investigou os fatores que levam algumas nações a serem mais bem sucedidas que outras. Dentre os motivos, observou-se que a presença de recursos naturais específicos, restrição de meios de transporte e a distância dos consumidores finais afetam o custo logístico das empresas, gerando (des)vantagens competitivas no processo produtivo. Além disso, o autor mencionou que a distância entre produtores e consumidores influencia o fluxo de informações e a adoção de inovações, interferindo, assim, na prosperidade econômica das nações.

Sob a influência de trabalhos passados e com o maior interesse na geografia econômica, Krugman (1998) formalizou a Nova Geografia Econômica (do inglês, NEG), uma vertente no campo da economia que enfatiza o papel da localização no desenvolvimento das atividades econômicas. Em seu trabalho, o autor explica que o crescente interesse pela geografia é resultado do maior uso de retornos de escala, custos de transporte, concentração industrial e informações assimétricas em modelos econômicos, negando a hipótese de competição perfeita.

Para a vertente, a concentração industrial é um processo que pode ocorrer a depender do resultado líquido da competição entre forças centrípetas e forças centrífugas na região, sendo que a primeira faz referência aos fatores que incentivam a concentração e a segunda a fatores que estimulam a dispersão industrial. Como forças centrípetas, podem ser mencionadas um maior mercado laboral (que é capaz de apoiar trabalhadores com habilidades especializadas), mais ligações setoriais (que facilitam a compra de insumos e reduzem custos

¹ Segundo Miller e Blair (2009), as ligações para trás são usadas para indicar os tipos de interconexões do setor com os setores produtores de seus insumos, enquanto as ligações para frente indicam as conexões do setor como ofertante de recursos aos demais setores da economia, sendo ambas importantes para analisar a interdependência produtiva regional.

de transporte) e efeitos spillovers (beneficiando indústrias pelo transbordamento de informações, por exemplo). As forças centrífugas, por sua vez, tendem a ser fatores imóveis, como recursos naturais, terrenos e local de moradia dos consumidores, que incentivam uma produção mais dispersa.

Sob uma perspectiva microeconômica, Hotelling (1929) dissertou sobre a estabilidade competitiva em um duopólio introduzindo a noção de espacialidade. Em seu trabalho, argumenta-se que o aumento do preço do bem ofertado por uma firma não gera um aumento brusco e total da demanda pelo bem da firma concorrente, como exposto nos trabalhos de Cournot e Edgeworth, em que o “mercado” é representado por um ponto, sem extensão. Segundo o autor, a presença de distâncias entre firmas e seus clientes e custos de transporte com a venda do produto final fazem com que as variações na demanda tenham um aspecto gradual e, por vezes, parcial. Isto pois, introduzidas tais variáveis, as preferências dos consumidores serão alteradas, fazendo com que certos agentes prefiram comprar das empresas mais próximas a si, mesmo estas ofertando o bem a um preço maior que o concorrente. Assim, a noção de equilíbrio se torna menos frágil que em modelos clássicos.

Outro conceito abordado na análise econômica regional e relacionado com a influência espacial é o efeito spillover (ou efeito transbordamento), que pode ser definido como “[...] o impacto de uma transação econômica em um processo que não está diretamente envolvido na transação (CARLSSON, 2010)”. Dito de uma outra maneira, o efeito spillover, nesse contexto, é o impacto indireto de um processo econômico originário em uma certa região/indústria em outras regiões/indústrias. Como abordado por Boschma *et al* (2017), regiões não são entidades independentes, mas sim entes que se relacionam entre si e compartilham externalidades, como a transferência de conhecimentos e inovações tecnológicas entre regiões próximas.

Gleaser *et al* (1992) estudou a interferência de efeitos spillovers no crescimento econômico das cidades nos Estados Unidos, buscando testar a teoria de Marshall-Arrow-Romer (MAR) e Porter e a teoria de Jacobs. Na visão de MAR e Porter, cidades que possuem uma economia mais especializada são mais propensas a exibirem um maior crescimento, pois os setores absorvem com maior facilidade o conhecimento transbordado de indústrias similares vizinhas, ou até mesmo seriam pressionados a buscar inovações tecnológicas frente a maior competição intra-setorial, estimulando o crescimento. Por outro lado, a visão de Jacobs defende que cidades mais diversificadas tendem a apresentar um crescimento maior, devido ao maior fluxo de ideias na economia, propiciando a

internalização de inovações de outros setores. Em linhas gerais, foi concluído que “[...] a especialização fere, a competição ajuda, e a diversidade colabora para o crescimento do emprego nas cidades” (GLEASER *et al*, 1992).

Ainda acerca do efeito transbordamento em economias regionais, De Vreyer e Spielvogel (2005) estudaram a influência dos municípios vizinhos na evolução do PIB per capita nas regiões do Brasil. Os autores constataram que a presença de externalidades espaciais ajuda a explicar o aumento da desigualdade econômica nas regiões Norte e Nordeste e a queda na desigualdade na região Centro-Sul brasileira. O estudo concluiu que municípios que possuem maiores (menores) níveis de PIB per capita tendem a estar localizados próximos a outros municípios na mesma condição (fenômeno explicado tanto por efeito spillover do PIB per capita quanto por políticas públicas implementadas em cada região), além do fato dessa tendência ser reforçada ao longo do tempo, indicando a presença da Armadilha da Pobreza² e de um desenvolvimento polarizado no Brasil.

2.2 A correlação entre diversidade industrial e estabilidade econômica regional

Primeiro introduzido por Chinitz (1961), há um consenso na literatura de que uma maior diversidade industrial está correlacionada com uma maior estabilidade econômica regional, pelo fato de aumentar a capacidade da região de contornar os danos de choques externos e ciclos econômicos. Regiões que possuem múltiplas especializações possuem múltiplos ciclos de negócios e apresentam uma menor dependência entre seus setores, tornando-as mais resilientes em períodos de crises.

O conceito de diversidade industrial possui diferenças a depender do autor abordado. É importante ressaltar que o termo “industrial” se refere a atividades econômicas e não necessariamente à indústria de transformação. Attaran (1986) definiu diversidade como sendo “a presença de um grande número de diferentes tipos de indústrias em uma área”. Por outro lado, desviando da noção quantitativa, Conroy (1975) associa diversidade industrial à distribuição da participação das indústrias na economia da região. Siegel (1995) acrescentou à discussão a ideia da importância de ligações intersetoriais na cadeia produtiva na conceituação de diversidade.

² O conceito de Armadilha da Pobreza é definido por Durlauf (2011) como sendo “a perpetuação da pobreza por certos fatores do ambiente socioeconômico que estão fora do controle do indivíduo, levando indivíduos (ou famílias) com preferências idênticas a acabarem com diferentes níveis de sucesso socioeconômico”. Outra forma de conceituar, segundo o mesmo autor, é “a perpetuação indefinida da pobreza, seja através da vida do indivíduo ou de gerações dentro da família (DURLAUF, 2011)”.

Apesar de um consenso teórico, no que tange a estudos empíricos os resultados sobre a correlação exposta muitas vezes se mostram conflitantes. Testando a hipótese da correlação positiva entre diversidade industrial e estabilidade econômica em 106 MSA's³, Kort (1981) comparou os resultados de regressões simples de Mínimos Quadrados Ordinários (MQO) e de Mínimos Quadrados Ponderados (MQP), a fim de corrigir o problema da heterocedasticidade, utilizando como variável dependente o índice REI (Regional Economic Instability), calculado pela variação média do emprego total de sua tendência, e como variável independente índices de diversidade, sendo eles Entropia, Ogiva, Média Nacional e Percentual de Bens Duráveis, utilizando dados de emprego por setor.

O autor observou que regiões altamente especializadas na produção de setores industriais específicos, como Kenosha e Gary, possuíam maiores índices de instabilidade, enquanto as MSA's de Chicago e Nova York se mostravam bastante diversificadas e estáveis, ou seja, constatou-se que regiões com uma economia mais diversificada tendem a apresentar maiores índices de estabilidade. Além disso, notou-se que o tamanho das zonas metropolitanas influenciava os resultados, aferindo uma correlação positiva estatisticamente significativa entre a área da região e sua estabilidade econômica.

O índice de Média Nacional parte da premissa que uma região economicamente diversa possui a proporção do valor adicionado ou número de ocupados em cada setor idêntica a proporção nacional. A ideia é que quanto mais a proporção dos setores desvia da proporção nacional maior o valor do índice, logo maior a especialização da economia. O índice pode ser calculado como:

$$D = \sum_{j=1}^J \frac{(x_j - x_j^*)^c}{x_j^*} \quad (1)$$

Sendo x_j a proporção do emprego ou valor adicionado do setor j na região, x_j^* a proporção do emprego ou valor adicionado do setor j nacional e c o poder do desvio, que pode variar a depender do autor.

³ MSA refere-se a Metropolitan Statistical Area e pode ser definido, segundo o United States Office of Management and Budget (OMB), como sendo “uma área central contendo um núcleo populacional substancial, juntamente com comunidades adjacentes com alto grau de integração econômica e social com esse núcleo”. A partir dos padrões estabelecidos pela entidade em 2010, uma MSA deve conter ao menos uma área urbanizada de 50.000 habitantes ou mais.

O índice de Ogiva pressupõe que a participação do emprego ou valor adicionado dos setores em uma região seguem uma distribuição equiproporcional, ou seja, são iguais para todos os setores. Quanto maior o desvio em relação ao valor ideal, maior o índice, logo maior a concentração econômica da região. O índice pode ser calculado da seguinte forma, na qual x_j é a proporção do emprego ou valor adicionado do setor j na região e x_{**} é a proporção $1/\text{número de setores}$.

$$Do = \sum_{j=1}^J \frac{(x_j - x_{**})}{x_{**}} \quad (2)$$

O índice de Percentual de Bens Duráveis é a porcentagem simples do emprego, renda ou valor adicionado na produção de bens duráveis, partindo da hipótese que a produção desses bens é geralmente caracterizada por uma maior elasticidade renda da demanda, logo regiões que possuem maiores proporções nesse tipo de bem sofrem de uma maior instabilidade. O índice de HHI (Índice Herfindahl-Hirschman) será melhor abordado futuramente neste trabalho.

De maneira análoga, Conroy (1975) estudou a influência da diversidade industrial na estabilidade do emprego de 52 MSA's, utilizando a abordagem do portfólio industrial para metrificar a variância do emprego nas regiões e, conseqüentemente, obter uma proxy de diversidade. Nesse estudo, o autor mensurou a correlação por meio de regressões simples, sendo observado que as variâncias possuem uma relação fortemente significativa⁴ com os índices de instabilidade (mais especificamente, explicam 42% dos movimentos do emprego), segundo o autor, “[...] aumentando a significância da diversificação da estrutura industrial como política de redução da flutuação econômica [...]”.

Por sua vez, Attaran (1986) estudou a correlação entre diversidade industrial (calculada a partir de uma medida de entropia usando o número de ocupados de oito setores não agrícolas) e crescimento e instabilidade do nível de desemprego e renda per capita dos 50 estados norte-americanos durante os anos de 1972 e 1981. O autor observou que o nível da diversidade dos estados não apresentou uma correlação estatisticamente significativa nem com

⁴ O autor realizou os testes utilizando cinco medidas de diversificação industrial regional: Média Nacional, Ogiva, Percentual de bens duráveis, Requerimentos mínimos e Portfólio industrial. Destas, apenas a última mostrou correlação estatisticamente significativa com instabilidade econômica.

a instabilidade nem com a taxa de crescimento das variáveis a um nível de significância de 5%, porém ambos os coeficientes apresentaram sinais negativos.

Avançando na análise da correlação entre diversidade econômica e estabilidade do nível de desemprego, o trabalho de Deller e Watson (2016) utilizou dados dos condados⁵ de 48 estados norte-americanos utilizando econometria espacial para controlar a dependência espacial dos dados e evitar a distorção dos resultados, representando um avanço metodológico em comparação aos trabalhos de Attaran (1986), Conroy (1975) e Kort (1981). Nesse estudo, os autores utilizaram a técnica de Regressão Geograficamente Ponderada (do inglês, GWR), que permite a variação dos coeficientes para cada indivíduo do modelo, captando a heterogeneidade espacial. As variáveis incluídas no modelo foram: percentual da população acima de 65 anos de idade, renda per capita em comparação a média nacional, percentual do emprego alocado na manufatura, percentual do emprego alocado em serviços, percentual do emprego alocado do setor público, coeficiente de Gini, renda per capita de pagamentos, renda per capita de dividendos e juros, renda per capita total, percentual de população latina, densidade populacional e taxa de desemprego, além das duas variáveis de interesse da correlação.

Foi concluído que a correlação, em grande parte do país, vai de encontro com a teoria econômica, ou seja, apresenta correlação positiva estatisticamente significativa entre diversidade industrial e estabilidade econômica. Entretanto, o autor ressaltou a heterogeneidade espacial da relação, explicitando que “enquanto não encontramos evidência indicando que maiores níveis de diversidade estão associados com aumento da instabilidade, em uma grande parte dos Estados Unidos essa relação não é estatisticamente significativa”.

Trendle e Shorney (2004) também estudaram a correlação estabilidade-diversidade utilizando os dados de Queensland, desagregados em 67 setores. Os autores utilizaram o índice de entropia para quantificar a diversidade industrial e um índice de instabilidade aplicado para desemprego, ocupação e renda per capita. Por meio de um modelo bivariado, observou-se que a diversidade tem um papel positivo e significativo na definição da estabilidade e do crescimento do emprego, mas não na renda per capita, isto é, a correlação entre diversidade e estabilidade mostrou-se positiva. Os autores também calcularam o Índice

⁵ Segundo o United States Census Bureau, condados são, em sua maioria, unidades governamentais funcionais, cujos governos locais possuem responsabilidade administrativa sobre o território. Atualmente, os Estados Unidos possuem 3.143 condados.

de Moran Global em seu trabalho, porém não utilizaram técnicas para controlar a interferência espacial no cálculo da correlação.

Malizia e Ke (1993) usaram dados em painel de áreas metropolitanas americanas durante os anos de 1972 e 1988, usando o índice de entropia para medir a diversidade e calculando a instabilidade do emprego por meio do desvio em relação a sua tendência e dividindo pela tendência do emprego. Os autores adicionaram variáveis de controle referentes ao tamanho da população, a características da população e do mercado de trabalho, educação, fatores geográficos e dummies locais. Utilizando Mínimos Quadrados Ordinários (MQO), os autores também observaram que regiões com maiores índices de diversidade tendem a ser mais estáveis.

A Tabela 1 agrupa as unidades geográficas utilizadas pelos autores, assim como as medidas de diversidade consideradas, o modelo utilizado e se a correlação encontrada foi significativa ou não.

Alguns motivos podem ser elencados como potenciais responsáveis pelas divergências encontradas nos resultados dos estudos. Como apontado por Siegel (1995), Wagner (2000), Kort (2003) e Conroy (1975), a escolha da medida de diversificação industrial utilizada no trabalho possui relevância (visto que cada medida possui pressupostos e metodologias de cálculo diferentes), sendo assim a seleção de métricas distintas é capaz de alterar a significância dos coeficientes, como observado em Conroy (1975).

LeSage (2009) e Deller e Watson (2016) promoveram uma discussão acerca da importância do método econométrico utilizado, criticando a escolha de metodologias convencionais em estudos passados, como o uso de Mínimos Quadrados Ordinários e dados em painel, e defendendo a utilização de métodos de regressão espacial. Segundo os autores, ao desconsiderar a dependência espacial dos dados regionais, ignora-se a violação de independência entre as observações, produzindo estimadores que são viesados e inconsistentes.

Outro possível motivo, abordado por Wang e O'Brien (2005) e Dapena (2018), é a utilização de unidades geográficas altamente desagregadas, que não são recomendadas para análise, devido aos problemas originários da diferença do local de trabalho e local de residência em pesquisas.

TABELA 1 - QUADRO RESUMO DAS PESQUISAS SOBRE DIVERSIDADE ECONÔMICA

Autor	Unidade	Diversidade	Modelo	Significante
Conroy (1975)	MSA	Portfólio	Correlação bivariada	Sim
Kort (1981)	MSA	Entropia, Média Nacional, Ogiva e Porcentagem de Bens Duráveis	Correlação bivariada	Sim
Attaran (1986)	Estado	Entropia	Correlação bivariada	Não
Malizia e Ke (1993)	MSA	Entropia	MQO Empilhado	Sim
Trendle e Shorney (2004)	Estado	Entropia	Correlação bivariada	Sim
Deller e Watson (2016)	Condado	HHI	GWR	Nem sempre

Fonte: elaboração própria.

2.3 Problema da Unidade de Área Modificável (MAUP)

Apesar de uma rica literatura ao redor da relação diversidade-estabilidade, poucos trabalhos atentaram-se a analisar empiricamente o tema alterando a escala geográfica dos dados, limitando a realização dos cálculos realizados a apenas uma unidade espacial.

Ao não incorporar variações na unidade geográfica, os autores acabam por desconsiderar problemas relacionados à escala, como o Problema da Unidade de Área Modificável (do inglês, MAUP) e o problema da Falácia Ecológica. Para este trabalho, o conceito de “escala” está relacionado com a resolução espacial do estudo, ou seja, ao grau de detalhamento do dado espacial (GOODCHILD, 2011).

O Problema da Unidade de Área Modificável pode ser definido como a alteração de resultados estatísticos em consequência da variação da agregação espacial e será o foco deste trabalho. Segundo Openshaw (1984), a origem deste problema está na necessidade de se agrupar dados de indivíduos (pessoas, domicílios, empresas, municípios etc) em zonas ou regiões arbitrárias e modificáveis (como distritos, campos, autoridades locais etc), a fim de viabilizar a implementação de pesquisas estatísticas e não ser necessário o uso de

microdados⁶. Segundo Openshaw (1984), “Se as unidades de área ou zonas são arbitrárias e modificáveis, então o valor de qualquer trabalho baseado nelas deve possuir algum grau de incerteza e pode não possuir validade, independente das unidades que estão sendo estudadas”.

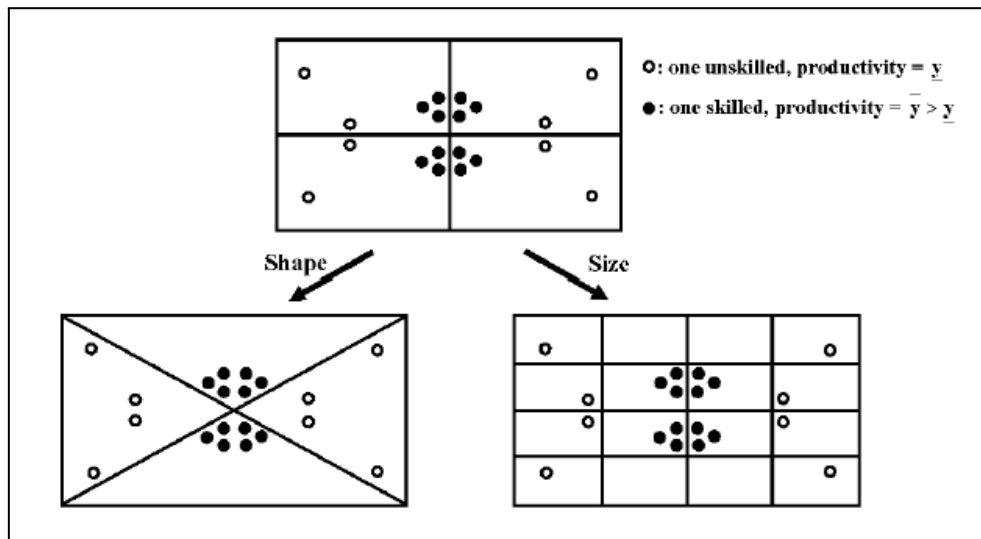
O primeiro problema relacionado ao MAUP é o problema de escala, que é “a variação dos resultados que pode ser obtida quando os dados de uma unidade espacial são progressivamente agregados em maiores unidades para a análise (OPENSHAW, 1984)”. Pode manifestar-se, por exemplo, quando dados de municípios são agregados em Estados ou Regiões.

O segundo problema derivado do MAUP diz respeito ao problema do formato (ou agregação) dos dados, que é evidenciado quando o número de unidades de análise permanece constante, porém o formato dessas unidades se modifica, alterando os resultados anteriores. É razoável admitir que se possui mais liberdade na escolha do formato das unidades do que na escolha da escala delas.

Para ilustrar ambos os problemas presentes no MAUP, a Figura 1 elucida o caso da relação densidade de emprego - produtividade do trabalho, demonstrado por Briant *et al.* (2008). Abaixo, os pontos pretos representam trabalhadores habilitados e os brancos são trabalhadores sem habilidade, permanecendo todos na mesma posição independente da unidade usada. Focando no problema do formato (figura da esquerda), é possível observar que, com essa agregação, surge uma concentração espacial, estando todos os trabalhadores habilitados e sem habilidade em unidades diferentes e, além disso, a correlação entre produtividade e densidade do emprego, que era zero inicialmente, passou a ser positiva (um maior número de empregos está correlacionado com uma maior produtividade do trabalho). Com relação ao problema de escala (figura da direita), a concentração espacial ainda está presente, porém a correlação mostra-se menos evidente, também alterando a informação original.

⁶ Segundo a Cetic, microdados podem ser definidos como a menor fração de um dado coletado em uma pesquisa ou o dado individual de um questionário, que pode ser agrupado em certas unidades gerando dados sobre unidades mais amplas.

FIGURA 1 - OS PROBLEMAS DE FORMATO E DE ESCALA DO MAUP

Briant *et al.* (2008).

Buscando analisar o efeito da correlação espacial do crescimento econômico em municípios, estados, microrregiões e cluster espaciais no território brasileiro, Resende (2011) expôs os efeitos spillovers presentes nas regiões brasileiras e, mais importante, identificou as variações dos coeficientes dos modelos a depender da escala utilizada. O autor concluiu que a significância e a magnitude dos coeficientes se alteram, além de observar diferentes padrões espaciais nas escalas do trabalho.

Também contribuindo para a pesquisa empírica do MAUP, Dapena *et al.* (2018) estudou os efeitos das aglomerações na produtividade regional na Espanha, estimando o efeito da densidade do emprego na geração de externalidades espaciais. O autor decidiu trabalhar com os dados mais desagregados possíveis, buscando agrupá-los em escalas espaciais suficientemente desagregadas a fim de identificar a heterogeneidade presente no país. As conclusões tiradas do estudo foram possíveis apenas pela alta desagregação dos dados, fazendo ser possível agregá-los em áreas que permitissem a adequada análise espacial, abordagem esta que “está mais em consonância com os recentes desenvolvimento do NEG, onde áreas dentro de regiões não administrativas como NUTS-2 ou NUTS-3⁷ possuem um grande papel (DAPENA *et al.*, 2018)”.

⁷ NUTS (Nomenclature of territorial units for statistics) é um sistema de classificação hierárquico que divide os territórios da União Europeia e do Reino Unido. NUTS-2 representam regiões básicas para a aplicação de políticas regionais e NUTS-3, nível ainda mais desagregado, são pequenas regiões para diagnósticos específicos.

3. METODOLOGIA

3.1 Modelos econométricos espaciais

3.1.1 Das limitações dos modelos econométricos tradicionais

Conforme salientado por LeSage (1999), os modelos econométricos tradicionais possuem limitações para tratar dados que contém componentes locais, ignorando o fato de as observações apresentarem dependência e heterogeneidade espacial. Segundo o autor, a dimensão espacial das atividades econômicas, sociais e demográficas é um importante componente para modelar tais problemas, formalizando a noção de transbordamento e interação das observações.

Tendo como base os trabalhos de Deller e Watson (2016) e Chen (2018), parte-se da premissa que a correlação estabilidade econômica/diversidade industrial não é afetada apenas por fatores internos à região, mas depende das características socioeconômicas das unidades próximas e também sofre interferências de fatores distribuídos não aleatoriamente no mapa. Por esses motivos, a utilização de modelos econométricos espaciais é preferível frente aos tradicionais.

Tradicionalmente, as análises empíricas partem de um modelo de regressão linear simples (utilizando o método de Mínimos Quadrados Ordinários) e pode ser representado em seu formato matricial da seguinte maneira:

$$y = X\beta + \varepsilon \quad (3)$$

Onde y é a variável dependente do modelo representada por um vetor linha, com as linhas sendo as unidades geográficas abordadas ($i = 1, \dots$, número de unidades em cada escala). X é o conjunto de variáveis independentes representado por uma matriz de i linhas (sendo i o número de unidades geográficas) e k colunas, sendo k o número de variáveis explicativas. β é um vetor que representa o efeito marginal das variáveis explicativas e ε um vetor de erros do modelo. Os modelos tradicionais quebram dois postulados do Teorema de Gauss-Markov em problemas espaciais:

$$E(\varepsilon\varepsilon') = \sigma^2 I \quad (4)$$

$$E(\varepsilon|X) = 0 \quad (5)$$

Em dados não espaciais, as observações de diferentes indivíduos são consideradas independentes umas das outras. Entretanto, como ressalta Tobler (1970) na sua 1ª Lei da Geografia: “*everything is related to everything else, but near things are more related than distant things*”. Isto é, fenômenos espacialmente distribuídos não são independentes.

Um dos exemplos onde a Equação 2 (referente a exogeneidade estrita) e a Equação 3 (referente a independência dos erros) são quebradas quando assume-se que uma decisão y_i é tomada a partir de fatores X e da decisão de seus vizinhos Wy_i , logo:

$$y_i = \rho W y_i + \beta X + \varepsilon$$

$$(1 - \rho W) y_i = \beta X + \varepsilon$$

$$y_i = (1 - \rho W)^{-1} \beta X + (1 - \rho W)^{-1} \varepsilon$$

$$(1 - \rho W)^{-1} = I + \rho W + \rho W^2 + \dots$$

$$y_i = (I + \rho W + \rho W^2 + \dots) \beta X + (I + \rho W + \rho W^2 + \dots) \varepsilon$$

$$u = (1 - \rho W)^{-1} \varepsilon$$

Logo, temos autocorrelação, $E(u_i u_j) \neq 0$, e quebra-se a independência dos erros, o que causa ineficiência. Além disso, $E(u_i | (I + \rho W + \rho W^2 + \dots) X_i) \neq 0$, ou seja, temos autocorrelação entre as variáveis independentes e o erro, o que quebra a exogeneidade estrita, logo temos viés dos estimadores.

Um segundo caso onde as Equações 2 e 3 são quebradas é o caso de variáveis omitidas com fator espacial. Seja o modelo real:

$$y = \beta X + \theta z$$

Assume-se que z siga um processo autorregressivo espacial.

$$z = \rho W z + r$$

$$z = (I - \rho W)^{-1} r$$

Logo, o modelo real é:

$$y = \beta X + \theta r(I - \rho W)^{-1}$$

$$y = \beta X + (I - \rho W)^{-1}u$$

Com isso, o postulado de exogeneidade estrita é novamente quebrado, causando ineficiência aos estimadores. Assumindo $Cov(z, x) \neq 0$, então $Cov(u, x) \neq 0$, logo:

$$y = \beta X + (I - \rho W)^{-1}(X\gamma + v)$$

$$y = \beta X + (I - \rho W)^{-1}X\gamma + (I - \rho W)^{-1}v$$

Logo, há também a presença de viés nos estimadores.

3.1.2 Da formalização das matrizes de vizinhança

Como visto anteriormente, matrizes de vizinhança estão inseridas em modelos econométricos espaciais, sendo necessárias para o modelo identificar as observações próximas/vizinhas. Uma matriz de vizinhança (ou matriz de pesos) é definida por Anselin (2002) como:

[...] uma matriz positiva $n \times n$ (W) que especifica "conjuntos de vizinhança" para cada observação. Em cada linha i , um elemento não nulo w_{ij} define j como sendo um vizinho de i . Por convenção, uma observação não é vizinha dela mesma, de modo que os elementos diagonais são zero ($w_{ii} = 0$). Note que essa definição é muito mais ampla do que sugere o termo "vizinho". Na maioria das aplicações em econometria aplicada, os vizinhos são unidades espaciais contíguas, [...] mas isso pode ser facilmente generalizado para qualquer estrutura de rede.

A especificação de uma matriz de vizinhança sofre de arbitrariedade e por vezes é a maior fraqueza do estudo em questão (ANSELIN, 2002). Os tipos de matrizes mais utilizados são as com base na contiguidade, ou seja, no compartilhamento de fronteiras, podendo ser divididas em *rook* - considerando as fronteiras em comum - ou *queen* - considerando as fronteiras e os vértices dos polígonos em comum. A principal fraqueza das matrizes de contiguidade é a exclusão do efeito espacial nos casos em que as observações (no caso, unidades geográficas) não fazem fronteira com outras, como é o caso de Ilhabela e Fernando de Noronha.

Um segundo tipo de matriz é a chamada matriz de pesos baseada na distância, que considera como sendo "vizinha" a observação que se encontra a uma distância menor do que um valor crítico especificado. Segundo Anselin (2002), quando uma área geográfica possui um alto nível de heterogeneidade na distribuição espacial dos pontos, como é o caso dos

municípios brasileiros nas regiões Sudeste e Norte do Brasil, pode não haver uma distância única que satisfaça a especificação espacial.

Visando contornar esses problemas, este trabalho utilizará a matriz de pesos baseada nos k próximos, forçando as observações a terem o mesmo número de unidades geográficas que afetam sua instabilidade econômica e evitando as chamadas *ilhas*⁸ (ANSELIN, 2002; CHEN, 2018; RESENDE, 2011; NEVES *et al.*, 2019; SOUZA *et al.*, 2021). O estudo do trabalho foi feito com base nos 10 vizinhos mais próximos, porém ao final será feita uma análise de sensibilidade dos resultados utilizando matrizes com 5 e 15 vizinhos mais próximos.

3.1.3 Da formalização dos modelos econométricos espaciais

Como apontado por Anselin (1988), o nível de agregação dos dados pode afetar o fenômeno espacial e alterar as estimações dos modelos econométricos, sendo assim, é recomendável que diferentes modelos sejam testados em cada unidade geográfica trabalhada, buscando aquele com melhor capacidade explicativa para cada uma. Nesse sentido, foram testados três modelos diferentes: SAR, SEM e SDM.

O modelo espacial auto regressivo⁹ (SAR) é aplicado para incorporar a interação espacial entre a variável dependente de uma região com a variável dependente das regiões próximas, formalizando a noção de transbordamentos (ANSELIN, 2002). O modelo pode ser descrito da seguinte maneira:

$$y = \rho W y + X \beta + \varepsilon \quad (6)$$

Como pode ser observado, adiciona-se o lag espacial da variável dependente, denotado por $\rho W y$, onde ρ é o parâmetro auto regressivo e W a matriz de vizinhança (que será melhor abordada futuramente neste trabalho). Aplicando ao estudo feito por esta monografia, o modelo SAR formaliza a ideia de que o nível da instabilidade econômica presente em uma região é parcialmente explicado pela instabilidade econômica das regiões vizinhas.

⁸ O conceito se refere às unidades geográficas que não possuem vizinhos, ou seja, não sofrem interferência espacial das demais.

⁹ Tradução livre de “spatial autoregressive model”.

O modelo de erros espaciais¹⁰ (SEM) parte da premissa que as observações são interdependentes graças a variáveis não observadas, sendo assim as características não capturadas no modelo capazes de influenciar regiões vizinhas afetam a região em estudo. A descrição formal do modelo é:

$$y = X\beta + u; \quad u = \lambda Wu + \varepsilon \quad (7)$$

O Spatial Durbin Model (SDM) tem a intenção de capturar os transbordamentos tanto das variáveis explicativas quanto da variável dependente. Nele, permite-se incorporar efeitos espaciais diretos e indiretos por meio das defasagens espaciais das variáveis dependentes e independentes. Ao incluir essas defasagens espaciais no modelo, ele captura parte da influência das variáveis omitidas que podem estar correlacionadas tanto com as variáveis dependentes quanto com as variáveis explicativas. É importante salientar que a presença de variáveis explicativas juntamente com seus lags espaciais pode gerar multicolinearidade em caso de elevada correlação, podendo tornar seus estimadores ineficientes (ANSELIN, 2003). Sua representação formal está descrita abaixo.

$$y = \rho Wy + X\beta + WX\theta + \varepsilon \quad (8)$$

No que tange aos métodos de estimação dos modelos, é de conhecimento que a estimação por MQO retorna estimadores viesados e não eficientes na presença de variáveis dependentes defasadas e de autocorrelação serial dos resíduos (ANSELIN, 2002; ELHORST, 2014; LESAGE, 2009). Sendo assim, abordagens alternativas recebem uma maior atenção na econometria espacial, como o caso da estimação por Máxima Verossimilhança (MV) (ANSELIN, 2003; KELEJIAN *et al.*, 1998; LESAGE, 2009).

A estimação por MV busca definir o estimador de máxima verossimilhança de θ , ou seja, o estimador que maximiza a probabilidade de ocorrer a amostra observada. Dentre suas características, pode-se citar a consistência do método para grandes amostras, a menor variância possível dentre as estimativas não-viesadas, a incorporação de heterocedasticidade e a estimação de modelos não lineares. Sob a hipótese de não enviesamento, os estimadores do

¹⁰ Tradução livre de “spatial error model”.

MQO se igualam aos da MV, porém caso tenha uma correlação significativa entre a matriz X e o termo de erro, ambas as estimadores se tornam inconsistentes.

Entretanto, para que a estimação por MV retorne estimadores consistentes e não viesados, é necessário que os parâmetros sejam normalmente distribuídos e que não haja variáveis significantes omitidas relacionadas com as variáveis explicativas, para que não quebre o postulado de exogeneidade estrita. Neste trabalho, assumiu-se que ambas as condições são respeitadas.

Para escolher o melhor modelo para cada unidade geográfica, foi utilizado o teste de Likelihood Ratio, avaliando quando o modelo SDM poderia ser simplificado para os modelos SAR ou SEM.

3.2 Índice de Moran Global

A Análise Exploratória de Dados Espaciais (AEDE) pode ser definida como sendo o conjunto de técnicas que busca descrever e visualizar a distribuição e os padrões espaciais (ANSELIN, 1996). O principal objetivo dessa análise se baseia em estudar a autocorrelação espacial, isto é, o fenômeno onde a similaridade locacional (observações em proximidade espacial) é correspondida pela similaridade de valores (correlação). As principais técnicas para examinar os padrões de estabilidade econômica e diversidade industrial das unidades geográficas são os índices de Moran global e local.

O índice de Moran global I é descrito formalmente como:

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} z_i z_j}{\sum_i z_i^2} \quad (9)$$

Onde n corresponde ao número de observações; w_{ij} são os elementos da matriz de pesos; e z_i e z_j os desvios dos valores dos elementos em relação a média. O índice retorna um valor único para todo o mapa, que varia entre -1 e 1, sendo 1 a característica de uma autocorrelação positiva perfeita, -1 uma autocorrelação negativa perfeita e quanto mais próximo de 0 maior a independência locacional dos atributos. É importante ressaltar que o

índice global não identifica padrões locais presentes no mapa, apenas retorna um único índice indicando a presença de autocorrelação das variáveis.

3.3 HHI e REI

Como explicitado anteriormente, as diferentes conceituações de diversidade propiciam distintas medidas de diversidade industrial (WAGNER, 2000; DISSART, 2003; SIEGEL *et al.*, 1995). Neste trabalho, utilizou-se o Índice Herfindahl-Hirschman (ou HHI) para mensurar a diversidade industrial das unidades geográficas, sendo calculado da seguinte forma:

$$HHI = \sum_{j=1}^N \left(\frac{o_{ij}}{o_i} \right)^2 \quad (10)$$

Onde N representa o número de setores econômicos na região i, o_{ij} representa a quantidade de empregados no setor j da região i e o_i é a quantidade total de empregados da região i (WAGNER, 2000; DELLER e WATSON, 2016; CHEN, 2018).

De acordo com a Eq. (11), quanto mais próximo de 1 for o valor do HHI maior a concentração econômica da região, isto é, maior a quantidade de empregos concentrada em poucos setores relativamente à quantidade existente na unidade geográfica. Consequentemente, quanto menor o valor do índice mais bem distribuídos estão os empregos da região, ou seja, mais diversificada é sua atividade econômica.

A primeira crítica feita a este índice se refere à arbitrariedade de se assumir que regiões com economias diversificadas irão convergir para uma distribuição equiproporcional da atividade econômica (CONROY, 1975). Além disso, Wagner e Deller (1998) ressaltaram a importância de se considerar as ligações interindustriais e o número de setores que pode se transferir para outras regiões, aspectos não incorporados no cálculo de medidas equiproporcionais.

O cálculo da instabilidade econômica regional (do inglês, REI), por sua vez, segue o proposto por Malizia e Ke (1993), sendo calculada como a média do desvio total do emprego em relação a sua tendência linear e dividida pela tendência linear, podendo ser expresso por:

$$REI = \left\{ \sum_{i=1}^N [(O_{it} - O_{it}^T)/O_{it}^T]^2 / T \right\}^{1/2} \quad (11)$$

Onde O_{it} representa a quantidade total de ocupados na região i no período t , O_{it}^T representa a tendência linear do número de ocupados total da região i no período t e T representa a quantidade de períodos analisados, sendo neste caso os anos de 2010 a 2019. Quanto maior for o desvio do número de ocupados em relação a sua tendência, maior o valor do índice REI, isto é, maior será a instabilidade econômica da região. De modo oposto, quanto menor o valor de REI, mais estável economicamente é a região em estudo.

3.4 Dados e modelo empírico

Para a construção das análises, foram utilizadas três escalas geográficas, que incluem 5.563 municípios, 557 microrregiões e 137 mesorregiões¹¹. A construção dos mapas do trabalho foi feita com base nos *shapefiles* dos municípios brasileiros, microrregiões e mesorregiões disponíveis no site IPEADATA. Todos os dados considerados foram obtidos a nível municipal e posteriormente foram agregados para as demais unidades geográficas.

O modelo empírico do trabalho pode ser descrito como:

$$REI_i = f(HHI_i, \text{variáveis de controle}_i) + \varepsilon \quad (12)$$

Onde i representa uma região, REI representa a instabilidade econômica regional, HHI é a variável que mede a diversidade econômica e, por fim, as variáveis de controle do modelo.

As variáveis de controle foram obtidas do “Atlas do Desenvolvimento Humano no Brasil”, projeto existente entre o Programa de Nações Unidas Para o Desenvolvimento (PNUD), o Instituto de Pesquisa Econômica Aplicada (IPEA) e a Fundação João Pinheiro (FJP), que tem o intuito de construir o Índice de Desenvolvimento Humano Municipal brasileiro e consolida dados de saúde, educação, renda, trabalho, habitação, participação

¹¹ A quantidade de unidades geográficas oficiais, segundo o IBGE, em 2010, era de 5.565 municípios, 558 microrregiões e 137 microrregiões, porém algumas unidades foram removidas da análise por falta completa de dados.

política e meio ambiente de 5.570 municípios. Os dados utilizados da plataforma foram retirados do Censo Demográfico de 2010.

O HHI foi calculado em cima dos dados da Relação Anual de Informações Sociais (RAIS) vínculos, do ano de 2010, presentes na plataforma BI do Ministério do Trabalho e Emprego (MTE). Além disso, a quantidade de vínculos foi agregada por CNAE 2.0 divisão, para que o índice pudesse ser construído. A REI também foi calculada com base nos dados da RAIS presentes na plataforma BI do MTE, considerando os períodos de 2010 a 2019, para que a variação do emprego nas regiões pudesse ser calculada.

A RAIS é um relatório feito pelo governo ao final de cada ano com o intuito de coletar dados trabalhistas e acompanhar a situação do mercado de trabalho brasileiro. A base de dados da RAIS contém todos os vínculos ativos e inativos (em 31/12) do mercado de trabalho formal declarados pelos empregadores. A descrição das variáveis utilizadas nas regressões pode ser vista na tabela abaixo.

Destaca-se que a escolha das variáveis de controle foi feita com base em trabalhos anteriores sobre o tema (CHEN, 2018; WATSON e DELLER, 2017; RESENDE, 2011; DELLER e WATSON, 2016), além de considerar também a significância das variáveis levantadas e a disponibilidade dos dados e possibilidade de agregação.

TABELA 2 - DESCRIÇÃO DAS VARIÁVEIS

Variável	Descrição	Fonte
REI	Média do desvio total do emprego em relação a sua tendência linear e dividida pela tendência linear, 2010-2019	RAIS-MTE
HHI	Índice de diversidade industrial utilizando o Índice Herfindahl-Hirschman, 2010	RAIS-MTE
POB	Porcentagem da população com renda domiciliar per capita igual ou inferior a R\$ 140,00 mensais, 2010	Censo-Atlas
SUP25	Porcentagem da população com 25 anos ou mais de idade com ensino superior completo, 2010	Censo-Atlas
POP18	Porcentagem da população total com 18 anos ou mais de idade, 2010	Censo-Atlas
TRABSC	Porcentagem de ocupados de 18 anos ou mais de idade que são empregados sem carteira assinada, 2010	Censo-Atlas
DSPG	Taxa de desocupação - 18 anos ou mais de idade, 2010	Censo-Atlas
NO	Variável dummy para regiões no Norte	Censo-Atlas
NE	Variável dummy para regiões no Nordeste	Censo-Atlas
CO	Variável dummy para regiões no Centro-Oeste	Censo-Atlas
SE	Variável dummy para regiões no Sudeste	Censo-Atlas

Fonte: elaboração própria.

3.4.1 Estatísticas descritivas

A tabela 3 abaixo contém as estatísticas descritivas das variáveis HHI e REI, tanto a nível municipal, quanto das microrregiões e mesorregiões.

TABELA 3 - ESTATÍSTICA DESCRITIVA DAS VARIÁVEIS HHI E REI

Escala geográfica	Município	Microrregião	Mesorregião
Observações	5564	558	137
Média HHI	0.379	0.225	0.176
Mínimo HHI	0.043	0.044	0.043
Máximo HHI	1.000	0.917	0.826
Desvio padrão HHI	0.257	0.184	0.153
Média REI	0.098	0.049	0.039
Mínimo REI	0.010	0.013	0.019
Máximo REI	3.162	0.410	0.249
Desvio padrão REI	0.118	0.037	0.023
Correlação	0.326	0.083	0.110

Notas: Os coeficientes de correlação são significantes a um nível de 1%. 14 municípios apresentaram HHI com valor 1, pelo fato de todos os empregos estarem alocados em “Administração pública, defesa e seguridade social”.

É possível observar com a tabela acima que os valores médios da REI e do HHI aumentam conforme mais desagregada for a escala geográfica, indicando que menores unidades apresentam comparativamente maior instabilidade e menor diversidade industrial que em unidades de maior extensão. De forma similar, o desvio padrão das variáveis também apresenta valores maiores quanto maior a desagregação da escala. Além disso, mesmo exibindo um valor maior no caso municipal, instabilidade econômica e diversidade industrial não possuem uma alta correlação, entretanto é importante notar que todos os coeficientes mostraram valores positivos, indicando que uma maior concentração industrial está relacionada com uma maior instabilidade regional, indo de encontro com o consenso teórico.

Primeiro, analisando a distribuição do HHI nos mapas, é possível notar que esta pode variar a depender do nível de agregação utilizado na escala. Por exemplo, no caso dos municípios, observa-se uma concentração de Baixo-Baixo no Centro-Sul do Brasil, uma faixa de Não Significância no centro do país e uma concentração de Alto-Alto no Norte e no Nordeste, com pontos de Baixo-Alto presentes. Quando analisadas as micro e mesorregiões, o único território que se mantém relativamente inalterado é o Sul, porém torna-se bem mais complexo definir a predominância nas regiões Norte e Nordeste, assim como na região central do Brasil, agora com a presença de Alto-Alto e Baixo-Alto. Outro ponto perceptível é que as regiões sem significância são bem mais presentes do que a nível municipal, dando a indicar

uma menor interferência espacial na determinação da diversidade industrial em micro e mesorregiões.

Com relação a REI, é possível ver, nas três escalas geográficas, mais áreas sem significância comparativamente ao HHI, como indicado com um Índice de Moran global mais próximo de zero. Entretanto, diferente do observado na diversidade industrial, a instabilidade econômica regional aparenta ter uma maior consistência quando analisadas as posições das concentrações ao longo das escalas. Enquanto a região Sul concentra Baixo-Baixo com pontos de Alto-Baixo nas três unidades geográficas, a região Norte concentra as categorias Alto-Alto e Baixo-Alto. Por mais que o Centro-Oeste do país tenha mais pontos de Baixo-Baixo e o Nordeste mais áreas de Alto-Alto e Baixo-Alto a nível municipal, a significância dessas categorias tende a cair conforme aumenta-se a agregação das unidades geográficas.

Interpretando os resultados acima alcançados, visualmente, as áreas mais ao Sul e Sudeste do Brasil são mais diversas industrialmente e encontram-se próximas a áreas que também apresentam maior diversidade industrial. Além disso, essas regiões apresentaram, também, uma menor instabilidade econômica entre os anos de 2010 e 2019, corroborando com a hipótese de correlação diversidade-instabilidade. A região Nordeste e, principalmente, a região Norte possuem uma característica oposta, ou seja, exibem mais áreas com baixa diversidade industrial ao redor de áreas semelhantes e também mais áreas com elevada instabilidade econômica próximas de zonas que também apresentaram uma economia mais instável no período de análise.

Os achados vão de encontro com o abordado por De Vreyer e Spielvogel (2005) acerca da polarização econômica no território brasileiro, mais especificamente sobre a diferença da performance das economias das regiões ao Sul e das regiões ao Norte/Nordeste do Brasil. Essa clusterização observada pode ser resultado do efeito transbordamento de políticas públicas locais, que, por sua vez, têm sua magnitude afetada pela extensão geográfica da jurisdição relevante (RESENDE, 2011). A manifestação da heterogeneidade espacial revela a relevância de reconhecer o MAUP e a ameaça de não considerar diferentes escalas geográficas ao conduzir estudos regionais, além disso, expõe a necessidade de se utilizar modelos econométricos que incorporem os efeitos espaciais no trabalho.

4. RESULTADOS EMPÍRICOS

4.1 Análise exploratória espacial

A tabela 4 mostra os Índices de Moran globais baseados na matriz de vizinhança com os 10 vizinhos mais próximos. Nota-se que ambas as variáveis possuem Índices globais positivos e significantes, porém enquanto o índice de Moran da REI apresenta valores menores e com menor dispersão, o índice do HHI apresenta uma dispersão maior e módulos maiores. Esses fatores indicam que a instabilidade regional possui uma baixa influência espacial nas três unidades geográficas analisadas, enquanto a diversidade industrial exibe uma maior clusterização, principalmente a nível municipal. De qualquer modo, a dependência espacial da diversidade se mostra maior que da estabilidade regional em todas as escalas estudadas.

TABELA 4 - ÍNDICES DE MORAN GLOBAL (10 VIZINHOS MAIS PRÓXIMOS)

Escala geográfica	Município	Microrregião	Mesorregião
HHI	0.540	0.289	0.112
REI	0.176	0.196	0.182

Notas: Todas as estatísticas são significantes a um nível de 1%.

4.2 Análise das regressões

Para fornecer uma ideia inicial da relação entre instabilidade econômica regional e diversidade industrial e as demais variáveis de controle, foi estimado um modelo MQO para as três unidades geográficas. Os resultados da estimação podem ser vistos na Tabela 5.

Nota-se que a significância, magnitude e o sinal dos coeficientes se alteram a depender da escala geográfica utilizada, porém a maior parte dos coeficientes nas micro e mesorregiões perderam sua significância a um nível de 10%, incluindo a variável HHI. Sendo assim, não é possível afirmar que há o MAUP, pois essa mudança pode ser fruto da diminuição do número de indivíduos nas amostras, que eleva os erros padrões e altera a validade dos coeficientes. Enquanto a correlação entre diversidade e estabilidade se mostrou positiva e significativa a um nível de 1% no caso dos municípios, indicando que uma maior concentração industrial contribui para uma maior instabilidade econômica, no caso das micro e mesorregiões esta correlação é negativa e não significativa.

TABELA 5 - RESULTADOS DA REGRESSÃO MQO

Variáveis	Município	Microrregião	Mesorregião
HHI	0.060*** (0.009)	-0.004 (0.008)	-0.011 (0.013)
POB	0.140*** (0.022)	0.079*** (0.029)	0.057 (0.047)
SUP25	-0.152** (0.059)	0.001 (0.067)	0.028 (0.095)
POP18	-0.175*** (0.045)	-0.144* (0.073)	0.064 (0.117)
TRABSC	-0.058*** (0.019)	-0.044 (0.034)	0.053 (0.062)
DSPG	0.085* (0.048)	0.038 (0.069)	0.020 (0.117)
NO	0.017** (0.007)	0.008 (0.007)	0.018* (0.011)
NE	-0.001 (0.006)	-0.001 (0.007)	-0.009 (0.001)
CO	0.0140** (0.006)	0.004 (0.006)	0.003 (0.008)
SE	0.011** (0.004)	0.003 (0.005)	0.000 (0.006)
Constante	0.174*** (0.033)	0.136** (0.055)	-0.031 (0.008)
Observações	5564	558	137
R-quadrado ajust.	0.150	0.227	0.138

Nota: Erros padrões entre parênteses. * significante a 10%; ** significante a 5%; e *** significante a 1%.

Além disso, é possível observar que a maioria das variáveis de controle se mostraram não significantes para micro e mesorregiões, diferentemente do ocorrido em nível municipal. Neste caso, com exceção da variável TRABSC, que representa a porcentagem de ocupados de 18 anos ou mais de idade que são empregados sem carteira assinada, todas as demais variáveis apresentaram sinais condizentes com a teoria (WATSON e DELLER, 2017; CHEN, 2018; HONG, 2016; DELLER e WATSON, 2016).

Como exposto anteriormente, o modelo MQO é considerado inadequado para o estudo, visto que não capta a dependência espacial das observações e sua distribuição heterogênea, como foi possível constatar por meio da análise exploratória espacial. Sendo assim, foram estimados três modelos econométricos espaciais (SAR, SEM e SDM) para cada

unidade geográfica e, por meio do Likelihood Ratio Test, foi escolhido o melhor modelo para cada escala. Mais especificamente, caso ambos os testes rejeitassem a hipótese nula, então o SDM seria escolhido, caso apenas o Likelihood Ratio Lag (LR Lag) rejeitasse a hipótese nula então o modelo SEM seria escolhido e caso apenas o Likelihood Ratio Error (LR Error) rejeitasse a hipótese nula então o modelo SAR seria escolhido (CHEN, 2018).

TABELA 6 - COMPARAÇÃO DOS LIKELIHOOD RATIO TESTS (10 VIZINHOS MAIS PRÓXIMOS)

Testes	Município	Microrregião	Mesorregião
LR Lag	32.41***	19.31**	29.81***
LR Error	34.04***	17.65*	18.77**

Notas: Erros padrões entre parênteses. * significante a 10%; ** significante a 5%; e *** significante a 1%.

Com a Tabela 6 pode-se observar que o modelo SDM foi o escolhido em todas as unidades geográficas. A Tabela 7 exibe os resultados das regressões.

Focando a análise na variável de interesse HHI, é perceptível a influência do nível de agregação na relação instabilidade-diversidade. A nível municipal, novamente o coeficiente da variável HHI se mostrou positivo e significativo, indicando que uma maior concentração industrial corrobora para uma maior instabilidade, ou, em outras palavras, uma maior diversidade industrial corrobora para uma menor instabilidade econômica. A variável $W*HHI$ capta a média dos valores da diversidade industrial dos municípios vizinhos e, como seu coeficiente retornou um valor positivo e significativo na equação, indica que um maior nível de concentração industrial das regiões próximas eleva a instabilidade econômica na região de estudo, incorporando a noção de transbordamento. Além disso, ρ também foi positivo e significativo, insinuando que o nível de instabilidade econômica da região tem uma correlação positiva com o nível da instabilidade nas regiões vizinhas, como sugeriu o Índice de Moran global.

TABELA 7 - RESULTADOS DAS REGRESSÕES ESPACIAIS (10 VIZINHOS MAIS PRÓXIMOS)

Variáveis	Município	Microrregião	Mesorregião
Modelo	SDM	SDM	SDM
HHI	0.046*** (0.009)	-0.002 (0.008)	-0.026*** (0.011)
POB	0.162*** (0.027)	0.086*** (0.031)	0.047 (0.042)
SUP25	-0.216*** (0.062)	0.052 (0.069)	0.018 (0.082)
POP18	-0.088 (0.056)	-0.253*** (0.091)	0.16 (0.109)
TRABSC	-0.002 (0.024)	-0.062 (0.04)	0.072 (0.058)
DSPG	0.101* (0.056)	0.011 (0.08)	-0.063 (0.105)
NO	-0.006 (0.043)	-0.021 (0.025)	0.027** (0.02)
NE	0.022 (0.039)	0.008 (0.024)	-0.007 (0.022)
CO	0.012 (0.035)	-0.006 (0.021)	0.014 (0.018)
SE	0.026 (0.029)	-0.009 (0.019)	0.005 (0.016)
W*HHI	0.074*** (0.021)	-0.022 (0.022)	-0.123*** (0.046)
W*POB	-0.148*** (0.051)	0.048 (0.074)	-0.129 (0.213)
W*SUP25	0.117 (0.143)	-0.061 (0.181)	-0.35 (0.409)
W*POP18	-0.200** (0.099)	0.419** (0.175)	0.277 (0.399)
W*TRABSC	-0.122*** (0.039)	0.062 (0.077)	0.371** (0.171)
W*DSPG	0.046 (0.099)	0.169 (0.163)	-0.51 (0.325)
Constante	0.245*** (0.065)	-0.103 (0.115)	-0.282 (0.309)
ρ	0.236*** (0.027)	0.096 (0.115)	-1.112*** (0.259)
Observações	5564	558	137

Notas: Erros padrões entre parênteses. * significativa a 10%; ** significativa a 5%; e *** significativa a 1%.

No que tange às microrregiões, a conclusão a ser tirada é de que a correlação não é significativa no modelo construído. Para as mesorregiões, o coeficiente da variável de interesse se mostrou negativo e significativo, ou seja, uma maior diversidade industrial eleva a instabilidade econômica da região, resultado oposto ao que é esperado pela teoria. O coeficiente da variável $W*HHI$ também se mostrou negativo e significativo, indicando que a maior diversidade industrial dos vizinhos reforça a instabilidade econômica da região. Outro resultado conflitante acerca das duas escalas geográficas se refere a ρ , que foi negativo e significativo a um nível de 1%, revelando que uma elevação na instabilidade econômica dos vizinhos diminui a instabilidade econômica na região.

Com a Tabela 7, é visível que o tipo de agregação utilizado na análise é capaz de influenciar o sinal e a magnitude dos coeficientes do HHI e do $W*HHI$ (e também de outras variáveis do modelo), dando a indicar que há a ocorrência do MAUP no caso. Além disso, é perceptível que o nível municipal é o que possui os resultados mais alinhados com a teoria econômica.

Como comentado por LeSage e Pace (2009), os coeficientes dos modelos que possuem variáveis defasadas, como é o caso do SDM, não devem ser interpretados como efeitos marginais, sendo assim, buscando avançar na análise, estimou-se os efeitos marginais da diversidade industrial sobre a instabilidade econômica regional. Os resultados podem ser vistos na Tabela 8.

TABELA 8 - ESTIMAÇÃO DOS EFEITOS DIRETOS, INDIRETOS E TOTAIS DO HHI (10 VIZINHOS MAIS PRÓXIMOS)

Estimação	Município	Microrregião	Mesorregião
Modelo	SDM	SDM	SDM
Efeito direto	0.047***	-0.002	-0.02*
Efeito indireto	0.109***	-0.024	-0.05**
Efeito total	0.157***	-0.026	-0.071***

Notas: * significativa a 10%; ** significativa a 5%; e *** significativa a 1%.

No que tange aos municípios, o efeito direto da variável “HHI” foi significativo com um valor de 0.047, enquanto o efeito indireto foi significativo ao valor de 0.109. Isto sinaliza que, enquanto um aumento de uma unidade no HHI na região em estudo aumenta a REI na própria região em, em média, 0.047, o aumento desta unidade no HHI eleva o HHI das demais regiões conjuntamente em 0.109, resultando em um efeito total de 0.157 na REI. Em

outras palavras, o aumento da diversidade industrial em uma região afeta tanto a estabilidade econômica desta região quanto das demais, visto que a diversidade industrial dos vizinhos também influencia a estabilidade destes (como é possível inferir do modelo SDM). Ou seja, a elevação da concentração industrial em uma unidade de medida aumenta a instabilidade econômica das regiões em 0.157.

Com relação às mesorregiões, o efeito direto foi de -0.02 e significativo a 10%, indicando que uma maior concentração industrial reduz a instabilidade econômica, mais especificamente, o aumento de uma unidade no índice HHI reduz o REI em 0.02 unidades. É possível observar, também, que o efeito feedback é positivo no valor de 0.006, indicando que o impacto do aumento da concentração industrial retorna a região de origem com um efeito positivo, ou seja, aumentando minimamente a instabilidade regional. O efeito total chega a -0.071, considerando o impacto em todas as regiões.

É importante mencionar que o efeito indireto já internaliza o efeito transbordamento da variação das variáveis ao longo dos demais vizinhos e o efeito direto já leva em consideração o efeito feedback, isto é, o efeito que a variação do HHI causou nas demais regiões e que teve o impacto retornado a região de origem devido a sua dependência dos valores das regiões vizinhas (LESAGE e PACE, 2009). Menciona-se, também, que a maior magnitude observada nos efeitos indiretos estão de acordo com o trabalho de Chen (2018), que relatou o mesmo fenômeno na maioria de seus testes.

4.3 Análise de sensibilidade

A fim de realizar uma análise de sensibilidade com os resultados encontrados, foram feitos os mesmos procedimentos anteriores com as matrizes de 5 a 15 vizinhos mais próximos e foram reportados os resultados das matrizes com os 5 vizinhos mais próximos e com os 15 vizinhos mais próximos. A partir da Tabela 9, percebe-se que o Índice de Moran global das duas variáveis exibe resultados semelhantes para as três unidades geográficas, ainda indicando um efeito positivo tanto para a REI quanto para HHI, além de apresentar uma interferência espacial mais forte no HHI em comparação à REI em todos os casos. É possível notar que, em geral, quanto maior o número de regiões consideradas vizinhas na matriz, menor o valor do índice.

TABELA 9 - ÍNDICE DE MORAN GLOBAL PARA DIFERENTES MATRIZES DE VIZINHANÇA

Variável	Matriz	Município	Microrregião	Mesorregião
HHI	5 próximos	0.558	0.346	0.229
	10 próximos	0.540	0.289	0.182
	15 próximos	0.530	0.263	0.172
REI	5 próximos	0.169	0.229	0.123
	10 próximos	0.176	0.196	0.112
	15 próximos	0.172	0.179	0.108

Notas: Todas as estatísticas são significantes a um nível de 1%.

Ao analisar os efeitos diretos, indiretos e totais do HHI sobre a REI variando a matriz de vizinhança, nota-se na Tabela 10 que os coeficientes das microrregiões são insignificantes em todas as matrizes, além das matrizes com 6 a 9 e 11 a 14 vizinhos mais próximos. Para o caso dos municípios, o efeito direto se manteve relativamente estável ao variar as matrizes, porém observou-se um aumento do efeito indireto da variável ao considerar mais regiões como sendo vizinhas, intuitivamente.

Com relação às mesorregiões, a Tabela 10 mostra que os efeitos foram mais significantes apenas na matriz com os 10 vizinhos mais próximos, não mostrando significância nas matrizes com 5 e 10 vizinhos. Entretanto, ao analisar os efeitos marginais utilizando as matrizes com 7, 8, 9 e 11 vizinhos mais próximos, também foram obtidos coeficientes significantes, como mostrado no Anexo 1. É interessante notar que em todos os casos os efeitos da variável HHI sobre a REI foram negativos, indicando que uma maior concentração industrial diminui a instabilidade econômica tanto na região de origem quanto em seus vizinhos.

Feita a análise de sensibilidade, foi possível notar a presença do Problema de Unidade de Área Modificável nos dados, na medida em que o sinal, a magnitude e a significância dos coeficientes da variável HHI (e de outras variáveis no modelo) se alteram a depender do tipo de agregação utilizado. Enquanto o trabalho de Chen (2018) apontou que a correlação estudada era sempre positiva e sofria apenas variações na significância e na magnitude dos coeficientes (em todas as escalas geográficas), nesta monografia observou-se que apenas no caso municipal e mesorregional isto ocorreu, visto que em microrregiões nenhum modelo tornou a correlação significativa. A correlação positiva defendida pela teoria foi observada apenas nos municípios, enquanto em mesorregiões a relação foi por vezes negativa e significativa.

TABELA 10 - ESTIMAÇÃO DOS EFEITOS DIRETOS, INDIRETOS E TOTAIS DO HHI PARA DIFERENTES MATRIZES DE VIZINHANÇA

Matriz	Modelo	Município	Microrregião	Mesorregião
5 próximos	Modelo	SDM	SDM	SDM
	Efeito direto	0.049***	-0.001	-0.010
	Efeito indireto	0.086***	-0.015	-0.035
	Efeito total	0.136***	-0.016	-0.046
10 próximos	Modelo	SDM	SDM	SDM
	Efeito direto	0.047***	-0.002	-0.020*
	Efeito indireto	0.109***	-0.024	-0.050*
	Efeito total	0.157***	-0.026	-0.071***
15 próximos	Modelo	SDM	SDM	SDM
	Efeito direto	0.047***	-0.003	-0.007
	Efeito indireto	0.131***	-0.032	0.057*
	Efeito total	0.178***	-0.036	0.049

Notas: * significativa a 10%; ** significativa a 5%; e *** significativa a 1%.

5. CONSIDERAÇÕES FINAIS

Este trabalho se propôs estudar a correlação entre diversidade industrial e estabilidade econômica regional sob a ótica de mais de uma escala geográfica, a fim de buscar captar o Problema de Unidade de Área Modificável. O parâmetro espacial foi positivo e significativo nas três escalas analisadas, além de ter sido observado variações na magnitude, significância e sinal da correlação, indicando que a interferência espacial dos dados interfere nos resultados do estudo, enfatizando a importância de se considerar mais de uma unidade geográfica em estudos regionais.

Verificou-se que a instabilidade regional possui uma baixa influência espacial nas três unidades geográficas analisadas, enquanto a diversidade industrial exibe uma maior clusterização, principalmente a nível municipal. A exploração espacial dos dados evidenciou que as áreas mais ao Sul e Sudeste do Brasil são mais diversas industrialmente e encontram-se próximas a áreas que também apresentam maior diversidade industrial. Além disso, essas regiões apresentaram, também, uma menor instabilidade econômica entre os anos de 2010 e 2019, corroborando com a hipótese de correlação diversidade-instabilidade. A região Nordeste e, principalmente, a região Norte possuem uma característica oposta, ou seja, exibem mais áreas com baixa diversidade industrial ao redor de áreas semelhantes e também mais áreas com elevada instabilidade econômica próximas de zonas que também apresentaram uma economia mais instável no período de análise.

Foi possível constatar que o sinal positivo e significativo esperado da correlação, isto é, uma maior diversidade industrial (mensurada pelo Índice Herfindahl-Hirschman) corrobora para uma menor instabilidade econômica foi observado apenas a nível municipal, enquanto em mesorregiões a correlação obtida foi negativa e significativa e em microrregiões a significância da correlação se mostrou inexistente em todos os modelos. Observou-se nos municípios que a diversidade industrial das regiões vizinhas também apresenta correlação positiva com a estabilidade econômica das regiões estudadas, constatando o efeito transbordamento nesse caso. No caso das mesorregiões, o efeito transbordamento constatado foi contrário ao observado em municípios. Assim sendo, de fato há a presença do MAUP no estudo feito, assim como Resende (2011), ressaltando a importância de se utilizar mais de uma escala geográfica em trabalhos de economia regional.

Por fim, é importante ressaltar que os resultados encontrados no trabalho para o caso brasileiro podem variar a depender do índice utilizado para metrificar a instabilidade e a diversidade industrial, além de depender fortemente também das variáveis incluídas nos

modelos, dos tipos de matrizes de vizinhança e modelos utilizados e do período analisado, podendo afetar as conclusões aqui tiradas em caso de outros estudos da área. Para trabalhos futuros, sugere-se a incorporação de mais de um índice de diversidade e também a introdução de clusters espaciais na análise, conforme sugere Carvalho *et al.* (2007).

REFERÊNCIAS BIBLIOGRÁFICAS

ANSELIN, Luc. **Spatial econometrics: methods and models**. Springer Science & Business Media, 1988.

ANSELIN, Luc. Interactive techniques and exploratory spatial data analysis. 1996.

ANSELIN, Luc. Under the hood issues in the specification and interpretation of spatial regression models. **Agricultural economics**, v. 27, n. 3, p. 247-267, 2002.

ANSELIN, Luc. Spatial externalities, spatial multipliers, and spatial econometrics. **International regional science review**, v. 26, n. 2, p. 153-166, 2003.

ATTARAN, Mohsen. Industrial diversity and economic performance in US areas. **The Annals of Regional Science**, v. 20, n. 2, p. 44-54, 1986.

BRIANT, Anthony; COMBES, P.-P.; LAFOURCADE, Miren. Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations?. **Journal of Urban Economics**, v. 67, n. 3, p. 287-302, 2010.

BOSCHMA, Ron et al. Towards a theory of regional diversification: Combining insights from evolutionary economic geography and transition studies. **Regional studies**, v. 51, n. 1, p. 31-45, 2017.

CARLSSON, B. The Spillover Effects in Brazil Associated With the Gripen NG Project. **Swedish Agency for Growth Policy Analysis**, 2010.

CARVALHO, A.; DA MATA, Daniel; RESENDE, Guilherme Mendes. Clusterização dos municípios brasileiros. **Dinâmica dos municípios**, p. 326-368, 2007.

CHEN, Jing. Economic Diversity and Regional Development: Geographical Scale, Structural Measurement, and Modeling Method. **West Virginia University**, 2018.

CHINITZ, Benjamin. Contrasts in agglomeration: New york and pittsburgh. **The American Economic Review**, v. 51, n. 2, p. 279-289, 1961.

CONROY, Michael E. The concept and measurement of regional industrial diversification. **Southern Economic Journal**, p. 492-505, 1975.

DÍAZ DAPENA, Alberto; FERNÁNDEZ VÁZQUEZ, Esteban; RUBIERA MOROLLÓN, Fernando. Labor density and wages in Spain: Evidence from geographically disaggregated data. **Growth and Change**, v. 49, n. 1, p. 55-70, 2018.

DELLER, Steven; WATSON, Philip. Spatial variations in the relationship between economic diversity and stability. **Applied Economics Letters**, v. 23, n. 7, p. 520-525, 2016.

DE VREYER, Philippe; SPIELVOGEL, Gilles. Spatial externalities between Brazilian municipios and their neighbours. **IAI Discussion Papers**, 2005.

DISSART, Jean Christophe. Regional economic diversity and regional economic stability: research results and agenda. **International Regional Science Review**, v. 26, n. 4, p. 423-446, 2003.

DURLAUF, Steven. Notes on poverty traps and Appalachia. 2010.

ELHORST, J. Paul et al. Spatial econometrics: from cross-sectional data to spatial panels. **Heidelberg: Springer**, 2014.

GLAESER, Edward L. et al. Growth in cities. **Journal of political economy**, v. 100, n. 6, p. 1126-1152, 1992.

GOODCHILD, Michael F. Scale in GIS: An overview. **Geomorphology**, v. 130, n. 1-2, p. 5-9, 2011.

GOODCHILD, Michael F. et al. Toward spatially integrated social science. **International Regional Science Review**, v. 23, n. 2, p. 139-159, 2000.

HONG, Saheum; XIAO, Yu. The influence of multiple specializations on economic performance in US Metropolitan Areas. **Sustainability**, v. 8, n. 9, p. 963, 2016.

HOTELLING, Harold. Stability in competition. **The economic journal**, v. 39, n. 153, p. 41-57, 1929.

KELEJIAN, Harry H.; PRUCHA, Ingmar R. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. **The journal of real estate finance and economics**, v. 17, p. 99-121, 1998.

KORT, John R. Regional economic instability and industrial diversification in the US. **Land Economics**, v. 57, n. 4, p. 596-608, 1981.

KRUGMAN, Paul. Increasing returns and economic geography. **Journal of political economy**, v. 99, n. 3, p. 483-499, 1991.

KRUGMAN, Paul. Urban concentration: the role of increasing returns and transport costs. **International Regional Science Review**, v. 19, n. 1-2, p. 5-30, 1996.

KRUGMAN, Paul. What's new about the new economic geography?. **Oxford review of economic policy**, v. 14, n. 2, p. 7-17, 1998.

LESAGE, James P. The theory and practice of spatial econometrics. **University of Toledo. Toledo, Ohio**, v. 28, n. 11, p. 1-39, 1999.

LESAGE, J.; PACE, R. Introduction to spatial econometrics. crc press, boca raton. **Introduction to spatial econometrics**, v. 1, p. 01, 2009.

MALIZIA, Emil E.; KE, Shanzi. The influence of economic diversity on unemployment and stability. **Journal of Regional Science**, v. 33, n. 2, p. 221-235, 1993.

MILLER, Ronald E.; BLAIR, Peter D. **Input-output analysis: foundations and extensions**. Cambridge university press, 2009.

MIZUNO, Keizo; MIZUTANI, Fumitoshi; NAKAYAMA, Noriyoshi. Industrial diversity and metropolitan unemployment rate. **The Annals of Regional Science**, v. 40, p. 157-172, 2006.

NEVES, Mateus de Carvalho Reis; CASTRO, Lucas Siqueira de; FREITAS, Carlos Otávio de. O impacto das cooperativas na produção agropecuária brasileira: uma análise econométrica espacial. **Revista de Economia e Sociologia Rural**, v. 57, p. 559-576, 2019.

OPENSHAW, Stan. The modifiable areal unit problem. **Concepts and techniques in modern geography**, 1984.

PORTER, Michael E. The competitive advantage of nations The Free Press. **New York**, v. 564, 1990.

PUGA, Diego; VENABLES, Anthony J. The spread of industry: spatial agglomeration in economic development. **Journal of the Japanese and International Economies**, v. 10, n. 4, p. 440-464, 1996.

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.

SIEGEL, Paul B.; JOHNSON, Thomas G.; ALWANG, Jeffrey. Regional economic diversity and diversification. **Growth and change**, v. 26, n. 2, p. 261-284, 1995.

SOUZA, Helson Gomes de et al. Uma análise espacial de como o crescimento e a desigualdade afetam a pobreza no Brasil. **Revista de Administração Pública**, v. 55, p. 459-482, 2021.

TOBLER, Waldo R. A computer movie simulating urban growth in the Detroit region. **Economic geography**, v. 46, n. sup1, p. 234-240, 1970.

TRENDLE, Bernard; SHORNEY, Gianna. The effect of industrial diversification on regional economic performance. **Australasian Journal of Regional Studies**, v. 9, n. 3, p. 355-369, 2003.

WAGNER, John E. Regional economic diversity: action, concept, or state of confusion. **Journal of Regional analysis and policy**, v. 30, n. 1100-2016-90005, 2000.

WAGNER, John E.; DELLER, Steven C. Measuring the effects of economic diversity on growth and stability. **Land Economics**, p. 541-556, 1998.

WATSON, Philip; DELLER, Steven. Economic diversity, unemployment and the Great Recession. **The Quarterly Review of Economics and Finance**, v. 64, p. 1-11, 2017.

WANG, Fahui; O'BRIEN, Van. Constructing geographic areas for analysis of homicide in small populations: Testing herding-culture-of-honor proposition. In: **Geographic information systems and crime analysis**. IGI Global, 2005. p. 84-101.

ANEXO 1

Matriz	Modelo	Mesorregião
6 próximos	Modelo	SDM
	Efeito direto	-0.007
	Efeito indireto	-0.023
	Efeito total	-0.031
7 próximos	Modelo	SDM
	Efeito direto	-0.014
	Efeito indireto	-0.048*
	Efeito total	-0.062**
8 próximos	Modelo	SDM
	Efeito direto	-0.019
	Efeito indireto	0.035
	Efeito total	0.055**
9 próximos	Modelo	SDM
	Efeito direto	-0.018
	Efeito indireto	-0.048
	Efeito total	-0.067**
11 próximos	Modelo	SDM
	Efeito direto	-0.015
	Efeito indireto	-0.043
	Efeito total	-0.059**
12 próximos	Modelo	SDM
	Efeito direto	-0.012
	Efeito indireto	0.039
	Efeito total	-0.027
13 próximos	Modelo	SDM
	Efeito direto	-0.009
	Efeito indireto	0.023
	Efeito total	-0.014
14 próximos	Modelo	SDM
	Efeito direto	-0.008
	Efeito indireto	0.041
	Efeito total	0.033