

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO SOCIOECONÔMICO
DEPARTAMENTO DE ECONOMIA E RELAÇÕES INTERNACIONAIS
CURSO DE GRADUAÇÃO EM CIÊNCIAS ECONÔMICAS

Guilherme Viegas

Capital humano, transbordamento do conhecimento e complexidade econômica:
Uma análise econométrica através de métodos globais e locais de microrregiões do Brasil

Florianópolis
2022

Guilherme Viegas

Capital humano, transbordamento do conhecimento e complexidade econômica:

Uma análise econométrica através de métodos globais e locais de microrregiões do Brasil

Trabalho de Conclusão do Curso de Graduação em Ciências Econômicas do Centro Socioeconômico da Universidade Federal de Santa Catarina como requisito para a obtenção do título de Bacharel em Economia.

Orientador: Professor Dominik Hartmann, Dr.

Florianópolis

2022

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Viegas, Guilherme
Capital humano, transbordamento do conhecimento e
complexidade econômica : uma análise econométrica através
de métodos globais e locais de microrregiões do Brasil /
Guilherme Viegas ; orientador, Dominik Hartmann, 2022.
66 p.

Trabalho de Conclusão de Curso (graduação) -
Universidade Federal de Santa Catarina, Centro Sócio
Econômico, Graduação em Ciências Econômicas, Florianópolis,
2022.

Inclui referências.

1. Ciências Econômicas. 2. Complexidade econômica. 3.
Capital humano. 4. Educação superior. 5. Transbordamento do
conhecimento. I. Hartmann, Dominik. II. Universidade
Federal de Santa Catarina. Graduação em Ciências Econômicas.
III. Título.

Guilherme Viegas

Capital humano, transbordamento do conhecimento e complexidade econômica:

Uma análise econométrica através de métodos globais e locais de microrregiões do Brasil

Este Trabalho Conclusão de Curso foi julgado adequado para obtenção do Título de Bacharel e aprovado em sua forma final pelo Curso de Ciências Econômicas

Florianópolis, 30 de Setembro de 2022

Prof. Dr. Helberte João França Almeida
Coordenador do Curso

Banca Examinadora:

Prof. Dr. Dominik Hartmann
Orientador
UFSC

Prof. Dr. Eva Yamila Amanda da Silva Catela
Avaliadora
UFSC

Prof. Dr. Marcelo Arend
Avaliador
UFSC

Agradecimentos

À minha mãe, Simone, agradeço a presença e suporte, pois sem você nada disso seria possível; à minha madrinha, Solange, agradeço o exemplo de dedicação e esforço, pois todos precisamos de alguém para nos inspirar; à minha avó, Dolores, agradeço todo o apoio neste longo caminho, pois sozinhos não vamos longe; e agradeço à todos os demais familiares e parentes que contribuíram de alguma maneira pra eu chegar até aqui. Obrigado.

Ao meu orientador, Dominik Hartmann, agradeço a excelente orientação, pois por vezes me senti perdido entre tantas possibilidades e você soube me apontar onde eu gostaria de chegar; aos professores Gueibi Peres Souza, Guilherme de Oliveira, Arlei Luiz Fachinello, Andréa Cristina Konrath, Luiz Ricardo Nakamura, Marcelo Arend, Eva Yamila Amanda da Silva Catela, Daniel Vasconcelos, Carmen Rosario Ortiz Gutierrez Gelinski, Fabio Padua dos Santos, Marcos Alves Valente, Helberte João Franca Almeida, Lauro Francisco Mattei, Marialice de Moraes, Helton Ricardo Ouriques, Michele Romanello, Fred Leite Siqueira Campos, Solange Regina Marin, agradeço as excelentes aulas e ideias, pois com suas lições pude enxergar tão longe. Obrigado.

Ao NECODE (Núcleo de Estudos sobre Complexidade, Desenvolvimento e Desigualdade), agradeço o espaço e oportunidade de amadurecer na carreira acadêmica; ao NNQ (Núcleo de Normalização e Qualimetria), agradeço o espaço e contribuição para minha formação acadêmica; à minha querida UFSC, agradeço o espaço, oportunidades, experiências e perspectivas, pois a vida acadêmica me salvou de muita tristeza e me trouxe muitas alegrias. Obrigado.

Aos colegas do curso de economia, Fernanda, Gabriela, Gabriel, Andrei, Vitor, Victor, Ramon, Natália, Luiz, Juan, Wellington, Vinícius, Ben-Hur, Lukas, Lucas, Laís, Morgana, Letícia, Pedro, Isadora, Roberta, Nicoli, entre tantos outros, agradeço as conversas, troca de ideias e a motivação para superar meus próprios limites semana após semana; aos amigos, agradeço os exemplos e experiências; e à minha querida namorada, Oriana, agradeço o carinho e apoio. Obrigado.

Às empresas Aquarela e 4intelligence, agradeço o espaço e oportunidade para aplicar o conhecimento recém adquirido neste estágio inicial de minha carreira profissional, assim como compartilhamento de conhecimento próprio e de clientes, pois na livre iniciativa surgem nossas versões mais produtivas. Obrigado.

Aos autores que fiz referência neste trabalho, assim como aos demais que tem parte na minha formação acadêmica, agradeço a todo o conhecimento compartilhado, pois assim caminho iluminado; e, especialmente, à Alexandra Elbakyan, agradeço as barreiras derrubadas, pois a ciência deve ser livre e para todos. Obrigado.

Resumo

Há clara evidência de que o capital humano pode promover crescimento econômico, contudo pouco se sabe como as áreas do conhecimento da educação superior impactam a distribuição espacial da complexidade econômica em países emergentes, tal como o Brasil. Dada a escassez de fundos, busca-se a melhor alocação de recursos em educação. Naturalmente, não é factível nem efetivo promover cursos custosos, como engenharia ou medicina, em todas as 558 microrregiões do país. Ao invés disso, cursos como administração e direito são mais viáveis de serem oferecidos na maioria das regiões. Ao mesmo tempo, a política de localização dos cursos superiores deve também considerar os efeitos de transbordamento do capital humano sobre regiões vizinhas. Neste trabalho, utilizamos dados do Censo demográfico de 2010 e dados de exportação do Comex Stat para revelar o agrupamento espacial e os efeitos de transbordamento de 8 áreas do conhecimento da educação superior sobre a complexidade econômica nas microrregiões do Brasil. Métodos globais e locais mostram a concentração do capital humano e heterogeneidade no efeito do ensino superior sobre a complexidade econômica. Além disso, para todas as áreas do conhecimento analisadas, o efeito do transbordamento do capital humano sobre a complexidade econômica é maior na região Sudeste, indicando retornos crescentes de escala do capital humano onde habitantes com educação formal se concentram. Isso implica em um problema complexo para tomadores de decisão uma vez que o investimento mais efetivo de recursos em educação aumentaria ainda mais a desigualdade regional.

Palavras-chave: Complexidade Econômica. Capital Humano. Educação. Concentração. Transbordamento do Conhecimento.

Abstract

There is clear evidence that human capital can promote economic growth, however little is known about how higher education knowledge fields impact the spatial distribution of economic complexity in emerging countries, such as Brazil. Due to the scarcity of funds, decision-makers seek the best allocation of resources in education. Naturally, it is neither feasible nor effective to promote expensive courses, such as engineering or medicine, in all 558 microregions of the country. Instead, courses such as business and law are more likely to be offered in most regions. At the same time, the location policy of higher education courses must also consider the spillover effects of human capital over neighboring regions. In this work, we use data from the 2010 Demographic Census and export data from Comex Stat to reveal the spatial clustering and spillover effects of 8 fields of higher education knowledge on economic complexity in Brazil's microregions. Global and local methods show the concentration of human capital and heterogeneity in the effect of higher education on economic complexity. In addition, for all fields of knowledge analyzed, the spillover effect of human capital on economic complexity is greater in the Southeast region, indicating increasing returns to scale of human capital where inhabitants with formal education are concentrated. This implies a complex problem for decision makers since a more effective investment of resources in education would further increase regional inequality.

Keywords: Complexity Economics. Human Capital. Education. Concentration. Knowledge Spillover.

LISTA DE FIGURAS

Figura 1 - Distribuição do ICE e HABES_pc	13
Figura 2 - Quantidade de habitantes com educação superior	36
Figura 3 - Mapa de clusters	37
Figura 4 - Coeficiente de relação causal local de ln_HABES_pc (Total)	41
Figura 5 - Coeficientes de relação causal local de ln_HABES_pc (Áreas do conhecimento)	43
Figura 6 - Matriz de correlação	59
Figura 7 - Mapa de clusters (Áreas do conhecimento)	61
Figura 8 - Coeficiente de determinação local	62
Figura 9 - Predição e Resíduos dos modelos MQO e RGP	63
Figura 10 - Coeficientes causais locais das variáveis de controle	64
Figura 11 - Significância estatística dos coeficientes locais	65

LISTA DE TABELAS

Tabela 1 - Regressões globais	39
Tabela 2 - Sumário causal global e local (Modelo 1 - Total)	44
Tabela 3 - Tabela de grau de ajuste (Modelo 1 - Total)	44
Tabela 4 - Tabela de grau de ajuste (áreas do conhecimento)	45
Tabela 5 - Regressões Globais de Controle	60

LISTA DE ABREVIATURAS E SIGLAS

AC	Estado do Acre
AIC	Cr�terio de Informa�o de Akaike (Akaike Information Criterion)
AM	Estado do Amazonas
API	Interface de programa�o de aplicativo (Application Programming Interface)
BIC	Cr�terio de Informa�o Bayesiano (Bayesian Information Criterion)
BP	Teste estat�stico de Breusch–Pagan
COMEX STAT	Sistema para consultas e extra�o de dados do Com�rcio Exterior brasileiro
dist_costa	Vari�vel de dist�ncia � costa atl�ntica
EMIG	Vari�vel de Emigrantes
ENEM	Exame Nacional do Ensino M�dio
F	Teste estat�stico de Fisher
GWR	Geographically Weighted Regression
HABEF	Habitante com Educa�o Fundamental
HABEM	Habitantes com educa�o M�dia
HABES	Habitantes com educa�o Superior
IBGE	Instituto Brasileiro de Geografia e Estat�stica
ICE	�ndice de Complexidade Econ�mica
ICP	�ndice de Complexidade de Produto
IMIG	Vari�vel de Imigrantes
ln_	Tratamento logar�tmico aplicado na vari�vel
MAPE	Erro Percentual Absoluto M�dio (Mean Absolute Percentage Error)
MI	Coefficiente I de autocorrela�o de Moran (Moran’s I)
MI_res	Autocorrela�o espacial dos res�duos (Moran’s I residual)
MQO	M�nimos Quadrados Ordin�rios
mqo_pred	Predic�o do modelo MQO
mqo_res	Res�duo do modelo MQO
OLS	Ordinary Least Squares
_pc	Tratamento per capita aplicado
PIB	Vari�vel de Produto Interno Bruto
R	Software estat�stico R e RStudio
RG	Vari�vel de efeito fixo por macrorregi�o
RGP	Regress�o Geograficamente Ponderada
rgp_pred	Predic�o do modelo RGP
rgp_res	Res�duo do modelo RGP
RMSE	Raiz do Erro Quadr�tico M�dio (Root Mean Square Error)

RO	Estado de Rondônia
RR	Estado de Roraima
R2	Coefficiente de determinação
R2 adj.	Coefficiente de determinação ajustado
scoef	Coefficiente Causal Local (Spatial Coefficient)
SIDRA	Sistema IBGE de Recuperação Automática
SENAI	Serviço Nacional de Aprendizagem Industrial
SQR	Soma dos Quadrados dos Resíduos
SQT	Soma dos Quadrados Totais
SUS	Sistema Único de Saúde
VCR	Vantagem Comparativa Revelada
W	Matriz de vizinhança (Weighting Matrix)

SUMÁRIO

1	INTRODUÇÃO	13
1.1	OBJETIVOS	15
1.1.1	Objetivo Geral	15
1.1.2	Objetivos Específicos	15
1.2	METODOLOGIA	15
1.3	JUSTIFICATIVA	15
2	REVISÃO DA LITERATURA	16
2.1	CAPITAL HUMANO	17
2.2	EDUCAÇÃO SUPERIOR	19
2.3	DESENVOLVIMENTO ECONÔMICO	21
3	MÉTODOS E DADOS	23
3.1	MÉTODOS GLOBAIS	23
3.1.1	Autocorrelação Espacial Global de Moran	24
3.1.2	Mínimos Quadrados Ordinários (MQO)	24
3.1.3	Coefficiente de Determinação Global	25
3.2	MÉTODOS LOCAIS	26
3.2.1	Autocorrelação Espacial Local de Moran	26
3.2.2	Regressão Geograficamente Ponderada (RGP)	27
3.2.3	Matriz de Ponderação	28
3.2.4	Coefficiente de Determinação Local	30
3.3	DADOS	30
3.3.1	Índice de Complexidade Econômica	31
3.3.2	Habitantes com Educação Superior Completa	32
3.3.3	Variáveis de Controle e Efeito Fixo	32
3.4	ESPECIFICAÇÃO DOS MODELOS	34
4	RESULTADOS E DISCUSSÃO	35
4.1	ANÁLISE DESCRITIVA	35
4.2	ANÁLISE CAUSAL GLOBAL	38
4.3	ANÁLISE CAUSAL LOCAL	41
5	CONSIDERAÇÕES FINAIS	45
5.1	CONCLUSÕES GERAIS	46
5.2	SUGESTÕES PARA TRABALHOS FUTUROS	47
	REFERÊNCIAS	48
	APÊNDICES	59

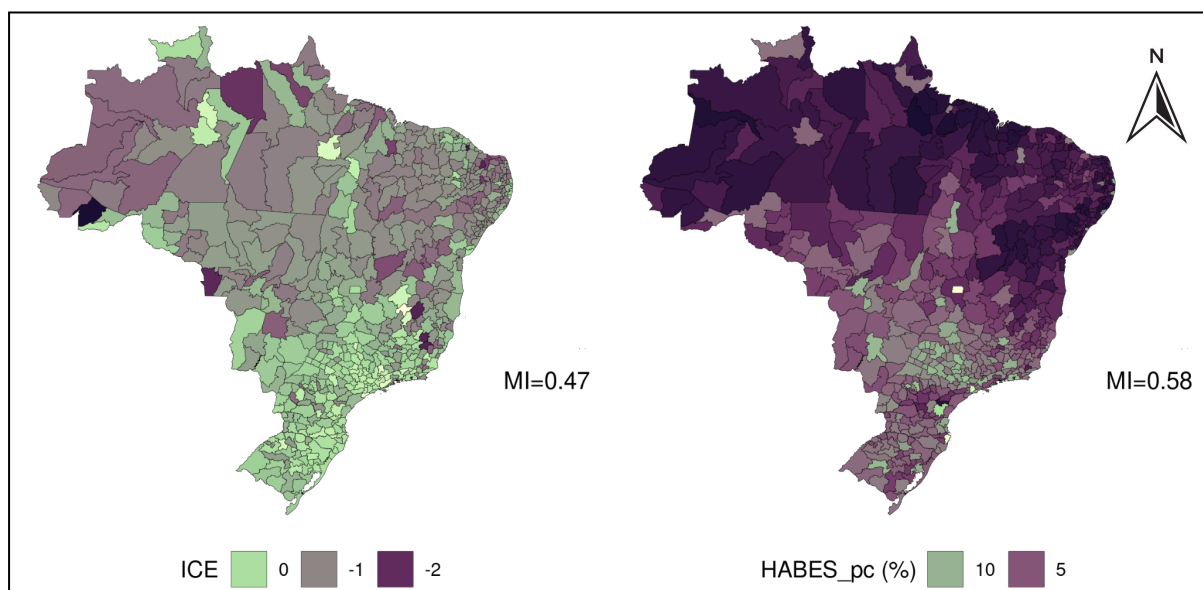
1 INTRODUÇÃO

Há clara evidência de que o capital humano pode promover crescimento econômico ([Romer, 1990](#)). Por exemplo, Schultz ([1971](#)) mostra que o número de estudantes está associado com crescimento significativo do PIB; Audretsch ([1996](#)) e Anselin et. al ([1997](#)) mostram que a pesquisa tem especial importância para inovações tecnológicas; os autores Zhu e Li ([2016](#)) mostram que o capital humano tem efeito positivo de curto e longo prazo sobre o crescimento econômico; Leite ([2019](#)) observa heterogeneidade no efeito do capital humano sobre a complexidade econômica; Laverde-Rojas e Correa ([2019](#)) mostram evidências de que a produtividade científica pode impactar a complexidade econômica; e Komatsu et. al ([2020](#)) mostram heterogeneidade no impacto de diferentes cursos da educação superior.

Contudo, aplicadores de políticas públicas têm à disposição recursos financeiros limitados para o investimento em diferentes tipos de cursos de educação superior, e dado o caráter complexo na alocação do investimento, não se pode esperar que universidades privadas sozinhas preencham essa lacuna. Isto é particularmente verdade em países com fundos escassos e grandes níveis de desigualdade regional (em termos de capital humano e desenvolvimento econômico), como o Brasil.

A figura 1 usa dados do Comex Stat para estimar o Índice de Complexidade Econômica (*ICE*) e do Censo demográfico para ilustrar a proporção de habitantes com educação superior completa (*HABES_pc*), nas microrregiões do Brasil, em 2010. Em ambas as distribuições se observa concentração na macrorregião Sudeste:

Figura 1 - Distribuição do ICE e HABES_pc



Fonte: Comex Stat e IBGE. Elaborado pelo autor.

Assim, aplicadores de políticas públicas e tomadores de decisão precisam priorizar quais disciplinas acadêmicas devem investir e em quais regiões do país. Para este propósito, é necessário saber o que esperar do impacto da educação superior sobre o desenvolvimento econômico regional para fazer o uso mais efetivo dos fundos limitados.

Assim como parte das decisões de investimento no ensino superior, deve-se considerar que, devido aos efeitos do transbordamento espacial, os efeitos do ensino superior podem não se limitar ao local de investimento, mas também às regiões vizinhas. Além disso, a pesquisa mostrou que o capital humano pode ter retornos crescentes, o que implica que não há efeitos marginalmente decrescentes, mas sim crescentes dos investimentos em capital humano.

Além disso, existem limites mínimos para tornar os investimentos no ensino superior eficazes, por exemplo, em termos de custos fixos para a construção de infra-estrutura de ensino. Finalmente, o nível de retornos crescentes, efeitos de transbordamento espacial e investimentos mínimos podem diferir entre as disciplinas que são mais (menos) caras e dependem das capacidades de absorção local da indústria, como tipos específicos de cursos de engenharia aeronáutica ou cursos amplamente aplicáveis, como direito e administração de empresas.

O Brasil é um bom caso para estudar os *spillovers* espaciais e os retornos crescentes do ensino superior. Em primeiro lugar, o Brasil investe cerca de 5% de seu PIB em educação superior. Porém, esse investimento não é distribuído igualmente em seu vasto território, e sim concentrado espacialmente. Além disso, o Brasil produz aviões e vacinas em algumas regiões, sendo economicamente pobre e focando apenas na administração pública, serviços simples e agricultura em outras. Por fim, há a disponibilidade de grandes conjuntos de estatísticas e dados espaciais de baixa granularidade sobre educação superior e complexidade econômica no Brasil.

Para identificar efeitos de *spillovers* espaciais e medir potenciais retornos crescentes, utilizamos regressões geograficamente ponderadas. A principal vantagem dessa metodologia é a possibilidade de identificar se os efeitos de *spillovers* espaciais mútuos da relação entre regiões vizinhas levam ao aumento do retorno esperado.

O restante deste trabalho está estruturado da seguinte forma: o Capítulo 2 revisa a literatura sobre capital humano, educação superior e desenvolvimento econômico; o Capítulo 3 apresenta os métodos globais e locais utilizados; o Capítulo 4 os resultados; e o Capítulo 5 conclui.

1.1 OBJETIVOS

1.1.1 **Objetivo Geral**

Analisar o efeito do transbordamento da educação superior sobre a complexidade econômica no Brasil, para auxiliar tomadores de decisão a fazer uso mais efetivo de recursos em educação e promover desenvolvimento econômico inter-regional.

1.1.2 **Objetivos Específicos**

- I. Analisar a distribuição e concentração espacial de habitantes com educação superior e da complexidade econômica nas microrregiões do Brasil;
- II. Estimar o pressuposto efeito causal do conhecimento da educação superior sobre a complexidade econômica através de métodos globais e locais;
- III. Apresentar meios para que tomadores de decisão possam fazer uso mais efetivo de recursos destinados à educação superior no Brasil.

1.2 METODOLOGIA

O presente trabalho utiliza metodologia descritiva e causal, através de métodos globais e locais, para analisar a distribuição de habitantes com educação superior e estimar a pressuposta causalidade do capital humano sobre a complexidade econômica em escala microrregional. Tomando modelos de crescimento endógeno como referencial teórico para explicar retornos crescentes de escala do capital humano e fazendo uso de métodos que consideram os atributos espaciais dos dados, espera-se apresentar evidências da ocorrência de transbordamento do conhecimento sobre o desenvolvimento econômico no Brasil.

1.3 JUSTIFICATIVA

O tema analisado se mostra relevante para entender o efeito da educação superior sobre a distribuição da complexidade econômica dentro do território nacional. Espera-se que os resultados obtidos possam explicar o desenvolvimento de polos econômicos regionais, assim como auxiliar

agentes públicos e privados na tomada de decisão e na otimização de recursos direcionados à educação superior em diferentes áreas do conhecimento da educação superior no Brasil. O trabalho também pretende promover avanços na literatura de capital humano e complexidade econômica através da aplicação de métodos em econometria espacial.

2 REVISÃO DA LITERATURA

A rica literatura econômica tem muitos trabalhos publicados sobre desenvolvimento econômico. A partir do século XX, a escola neoclássica emergiu com modelos de crescimento econômico que tomaram o chamado *mainstream* econômico. O mais influente deles, o modelo de Solow-Swan, sugere que, no curto prazo, o acréscimo na taxa de crescimento per capita pode ser obtido através de variações nos níveis de capital, trabalho ou tecnologia. Considerando a lei dos retornos decrescentes, todavia, é somente através do progresso tecnológico que seria possível sustentar crescimento per capita no longo prazo. Assim, nos modelos de crescimento exógeno, como o modelo de Solow-Swan, diferenças no produto agregado per capita entre países são entendidas como temporárias. Destarte, países pobres, com pouco estoque de capital, devem ter alto retorno do investimento; enquanto países ricos, que já possuem maior estoque de capital acumulado, devem ter menor retorno do investimento; o que eventualmente levaria a convergência. (SOLOW, 1956, 1957; SWAN, 1956; BRAKMAN, GARRETSSEN, MARREWIJK, 2009)

Seria possível fazer o *catch-up*, então? O que se observa, no entanto, é o distanciamento ainda mais acentuado entre países. Não convergência, como prevê a teoria neoclássica. Esse desencontro com a realidade empírica fez teóricos adicionarem novas variáveis dentro da função de produção aos modelos de crescimento econômico capazes de explicar o crescimento de longo prazo, como o capital humano, o conhecimento e a inovação. Assim, têm-se origem à teoria do crescimento endógeno. (KLENOW et al., 1997)

A teoria do crescimento endógeno presume que o investimento em capital humano, conhecimento e inovação são essenciais para o crescimento econômico de longo prazo. No modelo de Romer (1986), o conhecimento possui produtividade marginal crescente e assume papel fundamental na produção. Assim, o estoque de capital humano determina a taxa de crescimento da economia no longo prazo; Lucas (1988) também faz contribuições à teoria de crescimento endógeno, discorrendo sobre os mecanismos do desenvolvimento econômico, comparando modelos exógenos e endógenos, e destaca a importância do capital humano para o nível de produtividade no uso do capital físico e trabalho.

A partir de uma perspectiva geográfica, contudo, Krugman (1991) faz importantes observações sobre a tendência de concentração da indústria ao se aproximar da demanda e reduzir custos com transporte, promovendo a ocorrência de retornos crescentes de escala, o que reforçaria

padrões centro-periferia. Krugman também observa que, assim como as ações de uma firma podem afetar a demanda de outra em um cenário de competição imperfeita, a pesquisa também pode transbordar de uma firma para fora dela. A partir daí, não demorou para se estabelecer a ideia de que uma fonte de retornos crescentes pode ser derivada da concentração geográfica do capital humano, que favorece a troca de conhecimento, a pesquisa e a inovação.

O caráter geográfico da economia também foi objeto de estudo de pesquisadores posteriores. Moretti (2004) argumenta sobre o retorno social da concentração de capital humano em cidades. Através de modelos empíricos, o autor conclui que o investimento em educação pode superar o retorno individual, podendo surtir efeito também na produtividade, renda e preço da terra. Moretti argumenta que os efeitos do transbordamento do capital humano têm o potencial de aumentar a produtividade agregada além do acréscimo da produtividade individual.

Dada a tendência de concentração do capital humano em grandes centros urbanos e a crescente troca de informação e conhecimento, o termo “transbordamento do conhecimento” ganha importância na literatura econômica. A ideia de que a concentração do capital humano aumenta a troca de conhecimento, criando ambientes favoráveis à inovação e aumento da produtividade, é fundamental para entender o termo. Ainda, é importante observar que o transbordamento pode ocorrer a nível local e, até mesmo, a nível global. (CARLINO, 2014; NORRIS, 2015)

Neste contexto, o capital humano e o transbordamento do conhecimento têm papel fundamental para explicar diferenças no desenvolvimento econômico regional no Brasil. Grandes centros urbanos têm a capacidade de concentrar e multiplicar o conhecimento presente em seus residentes ao ponto de produzir retornos crescentes de escala que potencialmente transbordam para regiões vizinhas. (GALLUP et. al, 1998)

Entretanto, embora seja simples argumentar sobre efeitos de transbordamento do conhecimento em grandes aglomerados urbanos, não é tão simples medir a ocorrência deste de maneira adequada, afinal, não ocorrem uniformemente sobre o espaço. O presente estudo busca contribuir com a literatura econômica ao verificar a ocorrência e mapear o efeito do transbordamento da educação superior sobre o desenvolvimento econômico nas microrregiões do Brasil, em associação com a teoria do crescimento econômico endógeno.

2.1 CAPITAL HUMANO

Escolhas individuais têm o potencial de afetar nosso bem-estar no presente e no futuro. Enquanto algumas escolhas têm caráter mais urgente, outras se assemelham a escolhas de investimento, podendo gerar retornos crescentes no longo prazo. Ainda assim, agora numa perspectiva macro, escolhas de investimento têm o potencial de mudar a realidade de regiões inteiras, dada a

conectividade da vida moderna. Visto a importância da tomada de decisão coletiva, se faz necessário a promoção de escolhas cada vez melhores em políticas públicas.

Schultz (1961) relata que a literatura econômica de sua época já havia repetidamente explicado aumentos no produto agregado através de fatores como terra disponível, horas trabalhadas e capital físico. O autor argumenta que o investimento em capital humano deve ser o fator com maior capacidade explicativa para o desenvolvimento econômico. Schultz contesta a noção clássica de homogeneidade na capacidade laboral de trabalhadores, afirmando que o conhecimento e habilidades estariam associados a maiores níveis de destreza no trabalho, gerando maiores níveis de produto. Naturalmente, nem todo o investimento no ser humano é produtivo, mas Schultz atesta que investimentos em educação, treinamento e migração de jovens têm potencial de retorno de longo prazo, assumindo papel importante na determinação do produto em economias modernas.

Becker (1962) vai além na descrição do que poderia ser considerado investimento no chamado capital humano. O autor indica que investimentos em capital humano podem assumir diferentes formas, como educação formal, treinamento na empresa, cuidados médicos, alimentação nutritiva, ou até mesmo o acesso à informação. Todos esses investimentos diferem-se no efeito sobre a renda e consumo, na quantidade de recursos investidos, no tamanho do retorno e na relação entre investimento e retorno, mas todos têm o potencial de melhorar as habilidades físicas e mentais das pessoas.

Mincer (1984) argumenta que a acumulação de capital humano produz riqueza individual, assim como produto agregado. O autor argumenta que o crescimento do capital humano é ambos condição e consequência do crescimento econômico. Mincer reivindicava que o capital humano deveria ser difundido para além de onde teve origem para gerar maior retorno.

Cabe, também, às instituições promover investimentos em capital humano, assim como reter e atrair capital humano cada vez mais qualificado para manter altos níveis de crescimento e desenvolvimento econômico. Prover acesso à educação e saúde são as maneiras mais fundamentais de se fazer isso, além de atrair imigrantes que possam fazer parte da massa de trabalhadores, assim como reter potenciais emigrantes. Essa tarefa é cada vez mais difícil, no entanto, dada as facilidades de migração no mundo globalizado e contrastes cada vez maiores entre países centrais e periféricos. (DOCQUIER, 2012)

O Brasil absorveu grandes volumes de capital humano durante sua formação. Grande quantidade de africanos foram trazidos como escravos entre os séculos XVII e XIX, especialmente para a região Nordeste e Sudeste do país; Europeus emigraram na formação do país a partir do século XVI, e especialmente nos séculos XIX e XX; japoneses, alemães e italianos fugidos da segunda grande guerra mundial para o Sudeste e Sul do país; mais recentemente, e em menor proporção, latino americanos para as regiões fronteiriças. Dentro do país, especialmente após 1930, habitantes do Nordeste e Norte migraram para regiões onde esperavam por futuros melhores, como Sudeste e Sul. Hoje, porém, o Brasil sofre com altas taxas de brasileiros que desejam emigrar para países mais

desenvolvidos, sobretudo Estados Unidos e Europa. (JÚNIOR, [1997](#); VALE, LIMA, BONFIM, [2004](#); BRAGA, [2006](#))

Para entender como o Brasil chegou até aqui, hoje, talvez ainda mais importante seja observar as mudanças institucionais a partir do período republicano, dado o crescimento vertiginoso registrado desde então. Houve uma verdadeira revolução nas políticas sociais no Brasil após 1889, sobretudo após 1930, provocando o acúmulo de capital humano como jamais antes no país. Diversas políticas sociais poderiam ser citadas, como a Lei Eloy Chaves, de 1923, que é considerada o marco inicial da previdência social no país; o Manifesto dos Pioneiros, de 1932, que propôs uma reforma no sistema educacional brasileiro, definindo a educação como função essencialmente pública, devendo ocorrer de modo laico, obrigatório e gratuito; a instituição da carteira nacional de trabalho, em 1932, com a finalidade de assegurar os direitos trabalhistas; a criação do SENAI (Serviço Nacional de Aprendizagem Industrial), de 1942, com o objetivo de formar capital humano e prestar serviços técnicos à indústria local; o SUS (Sistema Único de Saúde), de 1990, que assegura o direito à saúde como política universal e gratuita; entre outros. (MENICUCCI, GOMES, [2018](#))

A inserção do capital humano nos modelos de crescimento econômico, finalmente, foi essencial para explicar o chamado resíduo de Solow, ou a parte do crescimento de longo prazo não explicada por variações no capital físico e trabalho. Poderia-se argumentar que a palavra resíduo é um eufemismo, neste caso, dada a importância do capital humano para explicar o crescimento econômico. Assim, diversos autores passaram a explicar mudanças tecnológicas de maneira endógena (ROMER, [1990](#); GROSSMAN, HELPMAN, [1991](#); AGHION, HOWITT, [1992](#)).

Diversas características do capital humano foram consideradas em trabalhos empíricos para medir seus impactos sobre o crescimento e desenvolvimento econômico, como: educação formal, seja infantil, fundamental, média, ou superior; a saúde da população, através da saúde infantil, de mulheres e idosos, assim como através da longevidade e mortalidade; a quantidade de emigrantes e imigrantes per capita; dentre outras maneiras de se medir o capital humano disponível. (RAUCH, [1991](#); RUDD, [2000](#); ACEMOGLU et al., [2000](#); MORETTI, [2004](#); PRETTNER, [2012](#); MORETTI et al., [2013](#); BALDANZI et. al, [2017](#); BLOOM et. al, [2018](#))

A partir deste referencial teórico, fica claro que não é apenas o número de trabalhadores que é relevante para o desenvolvimento econômico, sobretudo em economias modernas, mas também o nível de conhecimento disponível. Desse modo, devemos entender que o capital humano pode ser acumulado, é produtivo, deprecia e é rival em seu uso (PRETTNER, [2013](#)).

2.2 EDUCAÇÃO SUPERIOR

Embora a educação superior tenha sido pouco acessível no Brasil até o século XXI, a virada do século trouxe consigo novas possibilidades para a população em geral adquirir formação de nível

superior. Entre 1995 e 2010, houve expressivo avanço da educação superior no Brasil. A expansão no número de graduados ocorreu especialmente devido a 3 fontes: a difusão da educação superior privada; a promoção de universidades federais e estaduais em todas as unidades federativas; e o avanço de cursos à distância. Além disso, os resultados do Censo da Educação Superior registram concentração na quantidade de estudantes matriculados em cursos de educação superior na região Sudeste do país, em menor proporção no Sul e Nordeste, e menor ainda no Centro-Oeste e Norte. (INEP, [2010](#), [2020](#); MANCEBO, [2015](#))

Os tradicionais vestibulares trouxeram simplicidade no ingresso à educação superior. Não somente os vestibulares de cada instituição de educação superior, mas também o ENEM (Exame Nacional do Ensino Médio), que tornou evidente os conhecimentos necessários para avançar para a educação superior. Através das notas médias do ENEM observa-se recorrente média superior dos estudantes das macrorregiões Sudeste e Sul, intermediária do Centro-Oeste, e inferior do Nordeste e Norte. (VIGGIANO, MATTOS, [2013](#); CASTRO, TIEZZI, [2005](#))

Apesar do crescente número de graduados nas últimas décadas, a realidade ainda é bastante marcada por profundas desigualdades no acesso à universidade. Tomando como exemplo experiências de outras nações, o Brasil adotou as chamadas políticas de ação afirmativa, as quais reservam parte das vagas disponíveis para estudantes oriundos de escolas públicas, afrodescendentes e indígenas. (QUEIROZ, [2014](#); ROSA, [2014](#))

A notável distribuição desigual do acesso à educação superior no território brasileiro não surpreende, dada a conjuntura histórica. O Brasil encontra-se historicamente em situação periférica e a reprodução de grandes assimetrias sociais criam bloqueios à inovação e à difusão do progresso técnico que dificultam a endogeneização do movimento de transformação capitalista. O subdesenvolvimento no Brasil é produto de uma situação histórica que divide o mundo em uma estrutura *centro-periferia*, com características políticas que dificultam o desenvolvimento técnico nas regiões periféricas. (FURTADO, [1959](#))

Ainda assim, mesmo em países em desenvolvimento, como o Brasil, há a ocorrência daquilo que Lewis ([1954](#)) ilustrou como “ilhas de modernidade num mar de atraso”. Ou seja, mesmo na condição emergente do caso brasileiro, há o surgimento de polos econômicos de alta produtividade entre regiões de menor produtividade.

Nesse sentido, Fischer e Varga ([2003](#)) argumentam sobre a importância do conhecimento regional de universidades para o transbordamento de conhecimento para o setor produtivo. Para os autores, o transbordamento ocorre pois o conhecimento criado pelas universidades se caracteriza como bem público, e cria valor para firmas e outras organizações. Assim, quanto mais próximo das universidades, maior o transbordamento.

O fato de que as capacidades disponíveis no local determinam o que a região produz no presente e no futuro, o que os economistas chamam de *dependência da trajetória*, já foi repetidamente estudado e verificado (ZHU et al., [2017](#); CONIGLIO et al., [2018](#)). A partir de uma perspectiva

evolucionária, tudo indica que as regiões tendem a diversificar-se na direção de produtos relacionados aos já existentes no seu espaço-produto. (FRENKEN, BOSCHMA, [2007](#); HIDALGO et al., [2007](#); NEFFKE et al., [2011](#)).

Entendendo a educação superior como insumo para o desenvolvimento econômico, dada a necessidade de capital humano especializado para a produção de produtos complexos, a variável explicativa chave neste estudo será a proporção de habitantes com formação de nível superior, para diferentes áreas do conhecimento. Toma-se como hipótese nula a homogeneidade do efeito causal entre a quantidade de habitantes com formação de nível superior per capita sobre a complexidade econômica e procura-se argumentar sobre a ocorrência de heterogeneidade no efeito causal.

Além disso, outros estudos empíricos, como os de Lin ([2004](#)) e Komatsu et. al ([2020](#)), investigaram a relação entre diferentes cursos de educação superior sobre e crescimento econômico, revelando contrastes no impacto de diferentes áreas do conhecimento. Entretanto, o presente estudo parece ser o primeiro a estimar o impacto da quantidade de habitantes de diferentes formações de nível superior per capita sobre a complexidade econômica através de um método que considera o atributo espacial do dado.

Assim, sabendo que a educação superior pode promover desenvolvimento econômico, busca-se apresentar meios para que tomadores de decisão possam fazer escolhas mais bem fundamentadas, sobretudo sobre o fator de localização dos cursos de nível superior no país. O Brasil é um bom estudo de caso para esse problema dada a sua estrutura altamente heterogênea. Ainda que a concentração de capital humano e complexidade econômica sejam claras, os efeitos do transbordamento do conhecimento ainda devem ser melhor entendidos. O presente estudo busca preencher essa lacuna.

2.3 DESENVOLVIMENTO ECONÔMICO

Para representar o desenvolvimento econômico, finalmente, estudos empíricos utilizaram extensivamente variáveis como o Produto Interno Bruto (PIB), a variação do PIB, ou até o PIB per capita. Dado os recentes anseios por sustentabilidade econômica, argumenta-se sobre a inadequação do PIB para representar sucesso econômico regional (BERGH, [2007](#); FLORAMONTI, [2016](#)). No novo século, contudo, dois acadêmicos, César Hidalgo e Ricardo Hausmann, desenvolveram o chamado Índice de Complexidade Econômica (ICE), uma medida que sumariza a sofisticação das atividades produtivas, levando em consideração todo produto exportado com vantagem comparativa revelada, exprimindo o conhecimento produtivo disponível em cada região da amostra (HIDALGO, HAUSMANN, [2009](#)). No presente estudo, temos o ICE como variável dependente nos modelos apresentados, desempenhando o papel de proxy do desenvolvimento econômico regional.

Outros diversos estudos em complexidade econômica tiveram seguidos avanços em maneiras de representar, ou estimar, a capacidade produtiva econômica regional (TACHELLA et al., [2012](#); INOUA, [2016](#); ALBEAIK et al., [2017](#); REYNOLDS et al., [2017](#); GOMEZ-LIEVANO et al., [2021](#)). Todos observam ambiente desigual e hierarquizado, proporcionando a existência de polos econômicos que concentram a complexidade produtiva regional: regiões mais complexas produzem produtos mais complexos (como maquinários, remédios e produtos químicos) e regiões menos complexas produzem produtos menos complexos (como produtos agropecuários, minerais e vestuário).

A literatura de complexidade econômica também fez registros sobre a evolução do desenvolvimento econômico no Brasil. Entre 1930 e 1980, o país cresceu consideravelmente, particularmente devido às políticas econômicas bem-sucedidas, como o PAEG e II PND, havendo crescimento médio substancialmente positivo e diversificação econômica moderada no período. Nos anos 1980, no entanto, o cenário no Brasil mudou radicalmente: foi a chamada década perdida, período quando houve inflação e estagnação econômica recorrentes. A partir dos anos 1990, com a redemocratização e abertura comercial, o Brasil deu início ao processo conhecido como desindustrialização, registrando redução da proporção da indústria no PIB ano após ano. (VASCONCELOS, [2013](#); CARDOSO, [2012](#); SALLES et al., [2018](#); HERRERA et al., [2019](#); TORRES, [2019](#); HARTMANN et al., [2020](#); OREIRO et al., [2020](#))

Mais recentemente, contudo, o Brasil passou por um período de breve sucesso econômico, proporcionado pelo chamado *boom das commodities*, o que gerou taxas de crescimento econômico de 3,7%, em média, entre 2000 e 2011. Houve significativo aumento no padrão de consumo dos brasileiros e redução da desigualdade, embora dependente da exportação de produtos pouco complexos, como soja, açúcar, carne bovina e minério de ferro, em especial. No entanto, com a crise financeira de 2008 e a crise institucional de 2015, o Brasil mergulhou em uma crise profunda e sem precedentes, retomando maiores níveis de desigualdade. Apesar das crises, no entanto, parece evidente a deficiência do modelo macroeconômico brasileiro para sustentar altas taxas de crescimento econômico no longo prazo. (ABDALA, [2015](#); BRESSER-PEREIRA, [2007](#))

Assim, para tornar a economia brasileira mais forte e competitiva, o desafio é superar o estágio intermediário e tornar seu portfólio de produtos mais complexo (HARTMANN et al., [2020](#)). Todavia, atividades econômicas complexas não se distribuem igualmente sobre o espaço. Pelo contrário, se concentram em polos econômicos (FUJITA, THISSE, [2013](#)). Sendo assim, métodos globais se mostram subótimos para análises subnacionais, demandando a superação para métodos de análise local. O questionamento aqui, portanto, é: como a distribuição do capital humano com educação superior e o transbordamento do conhecimento impactam o desenvolvimento econômico dentro do Brasil? Para responder a essa pergunta, precisamos de um instrumental analítico que considere o atributo espacial do dado.

A concentração de atividades economicamente complexas em grandes centros urbanos como consequência da crescente complexidade da economia, assim como aponta Balland et. al ([2018](#)), é

fundamental para entender o presente trabalho. Para que tomadores de decisão façam melhor uso de investimentos em educação superior e permitam o processo de *catch-up* da economia brasileira, o efeito oculto da troca de conhecimento deve ser melhor entendido. O conhecimento desenvolvido nas universidades e transbordado entre os agentes, em diferentes localidades, têm o potencial de transbordar e proporcionar retornos crescentes de escala para o desenvolvimento econômico. O presente trabalho, portanto, busca verificar o impacto da educação superior e a ocorrência do transbordamento do conhecimento para promover desenvolvimento econômico no Brasil.

3 MÉTODOS E DADOS

A metodologia do presente trabalho foi desenvolvida para contrastar métodos globais e locais na estimação do efeito da educação superior sobre a complexidade econômica. Métodos globais estimam parâmetros únicos para toda a amostra; enquanto métodos locais estimam um conjunto de parâmetros, um para cada observação da amostra. Dada a não estacionariedade espacial dos dados do presente estudo, de fato altamente concentrados, mostra-se pertinente o uso de instrumental adequado para se obter os resultados desejados. (LYKOSTRATIS, GIANNOPOULOU, 2018)

Ainda assim, a comparação entre os métodos globais e locais é apresentada para verificação e compreensão dos resultados. Comprovada a relevância de tratamento ponderado entre regiões vizinhas, é então possível argumentar sobre a ocorrência de transbordamento de uma região à outra, e vice-versa. O presente estudo apresenta os métodos locais como uma evolução metodológica dos métodos globais, dada a crescente disponibilidade de dados espaciais em menor granularidade. Entende-se como razoável esperar que regiões próximas umas das outras tenham relações mais intensas do que regiões distantes.

3.1 MÉTODOS GLOBAIS

Para entender as características gerais de variáveis e entre variáveis, inicialmente, é interessante a estimação de métodos de análise globais. Através de medidas únicas para toda a amostra, métodos globais sumarizam características da amostra como um todo. É importante notar a pressuposição de homogeneidade na amostra para aplicar métodos de análise global, o que pode ser argumentado como irrealista para representar a realidade extremamente desigual na extensão do território brasileiro. Ainda assim, útil para entender características da amostra em uma perspectiva ampla.

3.1.1 Autocorrelação Espacial Global de Moran

Diferentemente do coeficiente de correlação de Pearson, que mede a relação entre duas variáveis, o coeficiente I de autocorrelação espacial global de Moran mede a correlação de uma variável com ela mesma sobre o espaço. Ou seja, mede a concentração de uma variável sobre a geografia da amostra. É uma medida normalizada que indica distribuição concentrada quanto mais próximo de +1, distribuição dispersa quanto mais próximo de -1, e distribuição aleatória quanto mais próximo de 0. Através da matriz de ponderação binária W , regiões próximas umas das outras são consideradas no cálculo enquanto regiões distantes são desconsideradas. O método *rainha*, que considera vizinha qualquer região que divida limites consigo, foi selecionado para o cálculo do algoritmo de MI no presente trabalho. (MORAN, [1950](#))

$$MI = \frac{n \sum_r \sum_v w_{rv} (x_r - \bar{x})(x_v - \bar{x})}{(\sum_r \sum_v w_{rv}) \sum_r (x_r - \bar{x})^2}$$

onde MI representa o coeficiente I de autocorrelação espacial de Moran; n representa o tamanho da amostra; W representa a matriz de vizinhança entre os polígonos regionais; w_{rv} representa a ponderação regional da região vizinha v para a região r ; x representa o valor da estatística analisada; r representa a região local; e v representa a região vizinha.

3.1.2 Mínimos Quadrados Ordinários (MQO)

Um dos métodos mais utilizados na literatura econômica para entender a relação entre variáveis é o método de Mínimos Quadrados Ordinários (MQO). Ainda que não seja possível inferir causalidade através de dados em seção cruzada, juntamente com uma boa argumentação teórica é possível tornar mais clara a relação entre duas variáveis.

Na regressão através do método de MQO há uma variável independente (y) e k variáveis independentes (x), assim como expresso na equação abaixo:

$$y_r = \alpha + \sum_k \beta_k x_{rk} + \varepsilon_r$$

onde y_r representa a variável dependente na região r ; α representa o intercepto; β_k representa o coeficiente causal da variável independente k ; x_{rk} representa a variável independente k na região r ; e ε_r representa o termo de erro estocástico.

O método clássico de regressão linear através de MQO é essencialmente um método global, onde estima-se coeficientes únicos para toda a amostra. O modelo de MQO também pode ser expresso em uma equação matricial, assim como expresso abaixo:

$$Y = Xb + \varepsilon$$

onde Y representa um vetor com os valores da variável dependente para cada região; X representa uma matriz de variáveis independentes com n linhas e $k + 1$ colunas; b representa um vetor com o intercepto (α) e coeficientes causais (β_k); ε representa o vetor dos desvios do modelo; e k representa o número de variáveis independentes.

Para se estimar os coeficientes de relação causal através do método de Mínimos Quadrados Ordinários, então, usa-se a seguinte operação de álgebra matricial:

$$\hat{b} = (X^T X)^{-1} X^T Y$$

onde \hat{b} representa um vetor com as estimativas do intercepto ($\hat{\alpha}$) e coeficientes causais ($\hat{\beta}_k$); X representa a matriz de variáveis independentes; e Y representa o vetor da variável dependente. (GUJARATI, [2000](#); WOOLDRIDGE, [2015](#))

3.1.3 Coeficiente de Determinação Global

O coeficiente de determinação global (R^2) é uma medida de ajuste do modelo em relação a sua capacidade de prever corretamente os valores da variável dependente em relação à hipótese nula em que apenas o intercepto tem poder explicativo. O R^2 varia entre 0 e 1, e é comumente apresentado em termos percentuais Assim:

$$R^2 = \frac{(SQT-SQR)}{SQT} = \frac{(y_i - \bar{y})^2 - (y_i - \hat{y}_i)^2}{(y_i - \bar{y})^2}$$

onde R^2 representa o coeficiente de determinação global; SQT representa a Soma do Quadrados Totais; SQR representa a Soma dos Quadrados dos Resíduos; e y representa a variável dependente.

3.2 MÉTODOS LOCAIS

Para entender as características locais de variáveis e entre variáveis, posterior a análise global, podemos estimar medidas locais. É importante notar a pressuposição de heterogeneidade na amostra para aplicar métodos de análise local, no entanto. Dada a realidade socioeconômica extremamente desigual no Brasil, vê-se como necessário o uso de métodos locais para obter resultados satisfatórios para a totalidade da amostra.

3.2.1 Autocorrelação Espacial Local de Moran

O método de autocorrelação espacial local de Moran, proposto por Anselin ([1995](#)), procura identificar clusters locais e regiões discrepantes. O MI local pode ser sumarizado de maneira ponderada para se obter o coeficiente MI global. O cálculo do índice de Moran local pode ser feito da seguinte maneira:

$$MI_r = z \sum_v w_{rv} z_v$$

onde MI_r é o índice I de autocorrelação local de Moran para cada região; z é uma observação da estatística de interesse em desvios padrões da média; e w representa a ponderação atribuída para cada região.

Juntamente com o método de autocorrelação espacial local de Moran, Anselin sugere um método de aleatorização condicional com permutações para produzir os chamados pseudo níveis de significância. Assim, relacionando a região com seus vizinhos, dado certo nível de significância

adotado, é possível verificar clusters e outliers sobre o mapa. No presente trabalho, foram atribuídos limites de significância de 90% (p-valor > 0.1). Então, em uma visualização, podemos identificar clusters de valores altos (Alto-Alto), clusters de valores baixos (Baixo-Baixo), outliers altos cercados de valores baixos (Alto-Baixo), e outliers baixos cercado de valores altos (Baixo-Alto). Assim, portanto, *outlier* neste não se refere ao sentido tradicional da palavra, mas relativo a média de seus vizinhos.

3.2.2 Regressão Geograficamente Ponderada (RGP)

Para superar a limitação metodológica global do método de Mínimos Quadrados Ordinários (MQO), finalmente, se propõe a aplicação da Regressão Geograficamente Ponderada (RGP). O método RGP estima um conjunto de parâmetros locais que variam sobre a extensão da amostra, reduzindo a possibilidade de viés geográfico nas estimativas. O método é geograficamente ponderado pois dá maior peso para observações próximas e menor peso para observações distantes. Assim, trata-se de um método não estacionário nos parâmetros, que mapeia uma distribuição de coeficientes sobre o espaço para melhor representar a heterogeneidade em um pressuposto efeito causal. (BRUNSDON et al., [1996](#); FOTHERINGHAM et al., [2003](#))

O modelo de regressão clássico, ou global, tem como pressuposto a linearidade nos parâmetros. Na abordagem local, através do RGP, no entanto, esse pressuposto é relaxado, permitindo que os parâmetros do modelo possam variar sobre a amostra. Agora, B representa uma matriz com uma sequência de coeficientes causais e interceptos para cada uma das n observações, permitindo que os coeficientes flutuem sobre a amostra. Podemos representar a regressão geograficamente ponderada através da expressão abaixo:

$$y_r = \alpha_r + \sum_k \beta_{rk} x_{rk} + \varepsilon_r$$

onde y_r representa a variável dependente para cada região; α_r representa o intercepto para cada região; β_{rk} representa o coeficiente causal de cada região r e variável independente k ; x_{rk} representa a observação da variável independente k na região r ; e ε_r representa o termo de erro estocástico.

A forma matricial da regressão geograficamente ponderada, no entanto, pode ser expressa como:

$$Y = (B \otimes X)1 + \varepsilon$$

onde Y representa o vetor da variável dependente; B representa uma matriz com n linhas e $k + 1$ colunas; n o tamanho da amostra; k é o número de variáveis independentes; \otimes representa a operação de produto tensorial; X representa uma matriz de variáveis independentes com n linhas e $k + 1$ colunas; 1 representa um vetor com $k + 1$ observações do número 1; e ε representa o vetor de erro estocástico.

A matriz B de coeficientes locais pode ser estimada de maneira semelhante ao método *MQO*. A diferença no método espacial, no entanto, está na utilização da matriz W de ponderação. Assim, tem-se uma sequência de coeficientes para cada variável independente e intercepto, um para cada região da amostra. Para haver melhor compreensão da sequência de coeficientes locais, podemos visualizar estes como uma distribuição de coeficientes sobre o espaço. Cada linha r da matriz B de parâmetros locais pode ser preenchida através da seguinte estimação:

$$\hat{b}_r = \left(X^T W_r X \right)^{-1} X^T W_r Y$$

onde \hat{b}_r representa as estimações dos coeficientes da região r , com uma linha e $k + 1$ colunas, para cada uma das n linhas da matriz B ; X representa a matriz de variáveis independentes; W_r representa o vetor da relação de vizinhança de cada região r com as demais; e Y representa o vetor da variável dependente.

Diferentemente da matriz de vizinhança binária do algoritmo de Moran's I, na RGP tem-se uma matriz que pondera a influência de cada região com as demais através de uma função bi-quadrada que minimiza o erro. Tomando como hipótese a troca de conhecimento inter-regional entre indivíduos e firmas, o modelo geograficamente ponderado tem o potencial de estimar a relação entre as variáveis mais adequadamente em comparação com o modelo global.

3.2.3 Matriz de Ponderação

Podemos representar a relação de ponderação entre as regiões através de uma matriz de ponderação (W) de dimensão $n \times n$. A matriz de ponderação W também é conhecida como matriz de

vizinhança, pois se trata do termo que define a relação entre as regiões vizinhas entre si. (BIDANSET, [2014](#))

Há diversas maneiras como preencher as n^2 observações (w_{rv}) da matriz de ponderação W . No coeficiente I de autocorrelação espacial global de Moran, por exemplo, usa-se uma matriz binária, onde regiões contíguas recebem o valor de 1 e regiões não contíguas recebem o valor de 0.

No método de Mínimos Quadrados Ordinários (MQO), no entanto, podemos entender que todas as observações se relacionam de maneira uniforme entre si. Assim:

$$w_{rv} = 1 \quad \forall \quad rv$$

Na regressão geograficamente ponderada, no entanto, a relação entre as regiões não é uniforme, é justamente geograficamente ponderada. Para atribuir a ponderação entre as regiões neste trabalho, portanto, optou-se por uma função bi-quadrada com kernel adaptativo. Ou seja, a ponderação da região vizinha v para a regressão da região r é obtida através de uma função inversa a distância entre as regiões r e v , assim como definido abaixo:

$$w_{rv} = \begin{cases} \left[1 - \left(\frac{d_{rv}}{bw} \right)^2 \right]^2 & \forall \quad d_{rv} < bw \\ 0 & \forall \quad d_{rv} \geq bw \end{cases}$$

onde w_{rv} representa a ponderação da região v na regressão da região r ; d_{rv} representa a distância entre a região r e a região vizinha v ; e bw representa a banda de relação entre as regiões.

Dada a irregularidade dos polígonos microrregionais do Brasil (Norte e Sudeste possuem polígonos médios bastante distintos, por exemplo), a banda de relação entre as regiões bw é definida de maneira adaptativa. Regiões onde há polígonos regionais maiores bw é maior, enquanto onde há polígonos regionais menores bw é menor. Portanto, usa um kernel adaptativo. A distância ótima de bw para cada região r é calibrada através de um método de validação cruzada com o objetivo de minimizar o erro quadrático médio estimado. Assim, há um processo de ponderação único para cada uma das regiões que considera a relação local entre seus vizinhos. (CLEVELAND, DEVLIN, [1988](#))

3.2.4 Coeficiente de Determinação Local

O coeficiente de determinação local é uma medida do grau de ajuste do modelo local em relação a sua capacidade de prever corretamente os valores da variável dependente, de maneira ponderada sobre espaço. O coeficiente continua variando entre 0 e 1, assim como o global, mas é ponderado espacialmente através da matriz de vizinhança W . Assim como a versão global, é comumente apresentado em termos percentuais. Assim:

$$r_r^2 = \frac{(SQT^w - SQR^w)}{SQT^w} = \frac{[\sum_v w_{rv} (y_v - \bar{y})^2] - [\sum_v w_{rv} (y_v - \hat{y}_v)^2]}{\sum_v w_{rv} (y_v - \bar{y})^2}$$

onde r_r^2 representa o coeficiente de determinação local; SQT^w representa a Soma do Quadrados Totais ponderado; SQR^w representa a Soma dos Quadrados dos Resíduos ponderado; w_{rv} representa a ponderação da região vizinha v para a região r ; e y representa a variável dependente.

3.3 DADOS

A amostra considerada no estudo é toda a extensão do território brasileiro. Para compor a totalidade da amostra, foram utilizadas as 558 microrregiões delimitadas pelo IBGE (IBGE, 1990). No entanto, a microrregião de Fernando de Noronha foi removida da amostra devido à distância da costa, portanto 557 microrregiões foram utilizadas. O conjunto de dados foi organizado em seção cruzada, com 2010 como ano de referência.

Para delimitar os limites das redes urbanas presentes no interior do Brasil, recorreu-se aos polígonos geoespaciais de microrregiões do IBGE. Em sua definição, são considerados vínculos sociais, políticos e econômicos, assim como a circulação de pessoas, de serviços e de informação, para delimitar os limites entre uma microrregião e outra. Assim, compõem um núcleo urbano coeso em si (IBGE, 1990). A divisão microrregional foi selecionada devido a coesão de seus habitantes e devido ao tamanho da amostra. Os polígonos foram obtidos no formato shapefile na seção de Geociências da página web do instituto (IBGE, 2022) e informações geoespaciais foram obtidas através da API de serviços da mesma instituição (IBGE, 2022).

Além disso, devido à falta de dados em algumas microrregiões isoladas, recorreu-se ao preenchimento através das estimções de um modelo de floresta aleatória para preencher estas

observações. Detalhes sobre as transformações e métodos aplicados no estudo podem ser verificadas no repositório do trabalho (VIEGAS, [2022](#)).

3.3.1 Índice de Complexidade Econômica

Como variável dependente, foi estimado o Índice de Complexidade Econômica (ICE), uma medida da estrutura produtiva regional. Para calcular o ICE, recorreu-se ao método subnacional, assim como proposto pelo Observatório da Complexidade Econômica (SIMÕES, HIDALGO, [2011](#); OEC, [2022](#)). Tomando o Índice de Complexidade de Produto (ICP) do Observatório da Complexidade Econômica, podemos chegar a uma fórmula muito mais simples do que o método de reflexões (HIDALGO, HAUSSMANN, [2009](#)). O método do ICE para dados subnacionais, que utiliza o ICP internacional, é, na prática, calculado através da média das complexidades dos produtos que são exportados com Vantagem Comparativa Revelada (VCR) por cada região (BALASSA, [1965](#)). Além disso, utilizando o ICP internacional, é possível fazer uso do índice de complexidade dos produtos em demanda no mundo, não somente na amostra subnacional. O método para o Índice de Complexidade Econômica subnacional pode ser representado através da seguinte fórmula:

$$ICE_r = \frac{\sum_r (VCR_{rp} \times ICP_p)}{\sum VCR_r}$$

onde ICE_r representa o Índice de Complexidade Econômica para cada região; VCR_{rp} representa a ocorrência de vantagem comparativa revelada do produto p na região r ; $\sum VCR_r$ representa a diversidade econômica, ou a quantidade de produtos exportados com vantagem comparativa revelada por cada região; e ICP_p representa a complexidade de cada produto.

Os dados de exportação utilizados no algoritmo foram obtidos através da secretaria de comércio exterior, vinculada ao Ministério da Economia do Governo Federal do Brasil (COMEX STAT, [2022](#)) e estão agrupados de acordo com o Sistema Harmonizado de 1996, de 4 dígitos, método recorrentemente usado para classificação de produtos comercializados mundialmente (UN, [1988](#); WCO, [2018](#)). Produtos exportados com vantagem comparativa revelada por dada região, em 2010, recebem 1, caso contrário recebem 0. A fórmula de vantagem comparativa revelada para dado produto e região pode ser representada pela seguinte expressão matemática:

$$vcr_{rp} = \left(\frac{x_{rp}}{\Sigma x_p} \right) \div \left(\frac{\Sigma x_r}{\Sigma x_{rp}} \right)$$

onde vcr_{rp} representa a vantagem comparativa revelada para dado produto e região; x representa as exportações, em valores do ano de 2010, para toda região r e produto p . A matriz VCR é preenchida com 1 para todo $vcr_{rp} \geq 1$ e 0 para todo $vcr_{rp} < 1$.

3.3.2 Habitantes com Educação Superior Completa

Para representar o estoque de capital humano qualificado fez-se uso da tabela 3543, disponível no Sistema IBGE de Recuperação Automática (SIDRA), de origem do Censo demográfico de 2010, que retrata a quantidade de habitantes com formação superior completa por área do conhecimento (IBGE, 2022). Esse dado nos permite comparar os efeitos de diferentes áreas do conhecimento, aqui divididas nas seguintes áreas: "Total", "Educação", "Humanidades e artes", "Ciências sociais, negócios e direito", "Ciências, matemática e computação", "Engenharia, produção e construção", "Agricultura e veterinária", "Saúde e bem-estar-social", "Serviços".

Além disso, para atender ao pressuposto de linearidade entre as variáveis nos modelos de regressão e possibilitar interpretações de interesse, essas variáveis foram relativizadas pelo total da população da região e aplicado tratamento logarítmico. Neste estudo, portanto, iremos nos referir ao acrônimo ln_HABES_pc como o vetor do logaritmo da quantidade de habitantes com ensino superior completo per capita das microrregiões do país. Essa é a variável independente chave no presente estudo.

3.3.3 Variáveis de Controle e Efeito Fixo

Além da variável dependente (ICE) e variável independente chave (ln_HABES_pc), para evitar viés por variável omitida, também foram estimados modelos com o logaritmo da proporção de habitantes com ensino médio completo (ln_HABEM_pc) e o logaritmo da proporção de habitantes com ensino fundamental completo (ln_HABEF_pc). Assim temos uma análise ainda mais robusta sobre a educação e o capital humano disponível em cada microrregião. Os modelos com estas variáveis foram adicionados ao apêndice deste. (IBGE, 2022)

Para representar o capital físico fez-se o uso da tabela 21 do Sistema IBGE de Recuperação Automática (SIDRA), de origem dos Órgãos Estaduais de Estatística, que retrata o Produto Interno Bruto (PIB), em mil reais, no ano de 2010 (IBGE, [2022](#)). Esse dado nos permite controlar por uma proxy do capital físico disponível na microrregião. Além disso, para atender ao pressuposto de linearidade entre as variáveis nos modelos de regressão e possibilitar interpretações de interesse, as estatísticas foram relativizadas pelo total da população da região e aplicado tratamento logarítmico. Neste estudo, portanto, iremos nos referir ao acrônimo \ln_PIB_pc como o vetor do logaritmo do PIB per capita das microrregiões do país.

Para representar a emigração internacional fez-se o uso da tabela 3173 do Sistema IBGE de Recuperação Automática (SIDRA), de origem do Censo demográfico de 2010, que retrata a quantidade de emigrantes internacionais registrados no país, no ano de 2010 (IBGE, [2022](#)). Esse dado nos permite entender onde há maior perda de capital humano entre as microrregiões do país e potencialmente remover viés por variável omitida. Para atender ao pressuposto de linearidade entre as variáveis nos modelos de regressão e possibilitar interpretações de interesse, as estatísticas foram relativizadas pelo total da população da região e aplicado tratamento logarítmico. Neste estudo, portanto, iremos nos referir ao acrônimo \ln_EMIG_pc como o vetor do logaritmo da quantidade de emigrantes per capita das microrregiões do país.

Para representar a imigração internacional fez-se o uso da tabela 2145 do Sistema IBGE de Recuperação Automática (SIDRA), de origem do Censo demográfico de 2010, que retrata a quantidade de estrangeiros registrados no país, no ano de 2010 (IBGE, [2022](#)). Esse dado nos permite entender onde há maior interesse imigratório e potencialmente remover viés por variável omitida. Para atender ao pressuposto de linearidade entre as variáveis nos modelos de regressão e possibilitar interpretações de interesse, as estatísticas foram relativizadas pelo total da população da região e aplicado tratamento logarítmico. Neste estudo, portanto, iremos nos referir ao acrônimo \ln_IMIG_pc como o vetor do logaritmo da quantidade de imigrantes per capita das microrregiões do país.

$dist_costa$ representa a distância (em mil quilômetros) de cada microrregião em relação à costa atlântica do país, esperando capturar as vantagens para o desenvolvimento econômico das cidades brasileiras que historicamente se concentraram nas proximidades do oceano atlântico. Dada a tendência de maior desenvolvimento das cidades litorâneas do país, essa variável tem o potencial de controlar e remover o efeito por variável omitida e proporcionar modelos mais bem ajustados. Estes dados foram calculados pelo autor através do software R e dos polígonos espaciais disponibilizados pelo IBGE.

RG é uma variável dummy para representar os efeitos fixos da macroregião de cada microrregião da amostra (Centro-Oeste, Nordeste, Norte, Sudeste ou Sul). Ou seja, no intuito de remover possíveis efeitos de endogeneidade das microrregiões analisadas, foram utilizadas variáveis dummy para cada uma das macroregiões do país, potencialmente representando características

endógenas da região. A macrorregião Centro-Oeste foi tomada como base para não cair na chamada *armadilha da dummy*.

3.4 ESPECIFICAÇÃO DOS MODELOS

Para simplificar o modelo proposto e ilustrar de maneira mais clara os fatores considerados como potenciais promotores de complexidade econômica, pode-se recorrer a uma simplificação matemática da equação. No modelo proposto, a variável dependente é o Índice de Complexidade Econômica (*ICE*) e a principal variável independente é o logaritmo da quantidade de habitantes com educação superior per capita (*ln_HABES_pc*). Outras variáveis independentes foram incluídas ao modelo como variáveis de controle, embora seu impacto também seja de interesse.

Abaixo temos a simplificação dos modelos globais e locais propostos no presente trabalho, com a diferença apenas na notação para sinalizar a não estacionariedade nos parâmetros no caso da RGP. Finalmente, chegamos as seguintes fórmulas para estimar a complexidade econômica regional:

Para MQO:

$$ICE_r = \alpha + \beta_1 ln_HABES_pc_r + \beta_2 ln_PIB_pc_r + \beta_3 ln_EMIG_pc_r + \\ + \beta_4 ln_IMIG_pc_r + \beta_5 dist_coast_r + \beta_6 RG_r + \varepsilon_r$$

Para RGP:

$$ICE_r = \alpha_r + \beta_{1r} ln_HABES_pc_r + \beta_{2r} ln_PIB_pc_r + \beta_{3r} ln_EMIG_pc_r + \\ + \beta_{4r} ln_IMIG_pc_r + \beta_{5r} dist_coast_r + \beta_{6r} RG_r + \varepsilon_r$$

onde *ICE* representa o índice de complexidade econômica; α representa o intercepto; β_k representa o coeficiente de relação causal para as *k* variáveis independentes; *ln_HABES_pc* representa o logaritmo da quantidade de habitantes com ensino superior per capita; *ln_PIB_pc* representa o vetor do logaritmo do PIB (em mil reais) per capita; *ln_EMIG_pc* representa a quantidade de emigrantes per capita; *ln_IMIG_pc* representa a quantidade de imigrantes per capita; RG representa a macrorregião (efeito fixo); ε representa o termo de erro estocástico da regressão; e *r* representa a indexação por região (observação) da amostra.

4 RESULTADOS E DISCUSSÃO

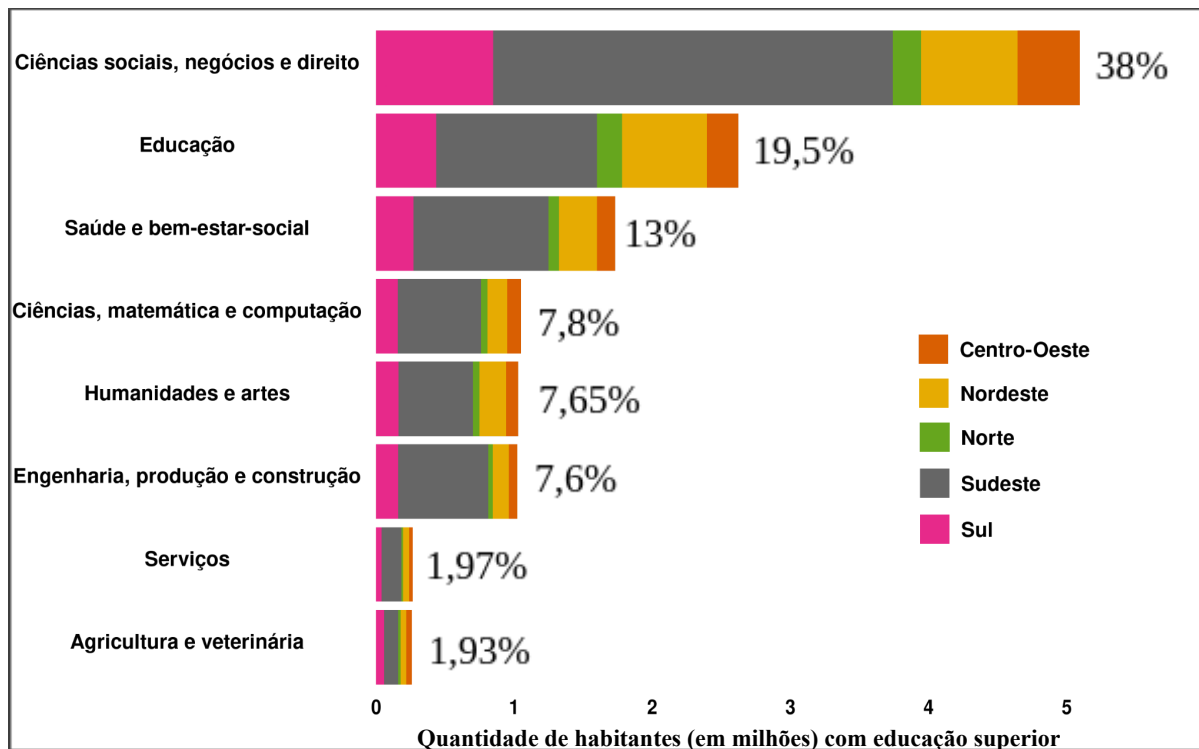
Para descrever os efeitos do capital humano com educação superior sobre a complexidade econômica nas microrregiões do Brasil, assim como argumentar sobre a ocorrência de transbordamento do conhecimento, foram aplicados diversos métodos globais (aespaciais) e locais (espaciais). Após verificadas as características da distribuição do capital humano e desenvolvimento econômico, podemos partir para a análise causal com mais insumos para interpretar as análises.

4.1 ANÁLISE DESCRITIVA

Inicialmente, é importante observar a heterogeneidade na distribuição de estatísticas socioeconômicas no Brasil. Há clara concentração do capital humano e capital físico, seja através de estatísticas absolutas ou relativas, na região Sudeste e, em menor proporção, na região Sul. De maneira mais dispersa, as regiões Centro-Oeste, Nordeste e Norte possuem poucos pontos de desenvolvimento em seus territórios.

O Censo demográfico de 2010 registrou 7,2 milhões habitantes com educação superior na região Sudeste; 2,2 milhões na região Sul; 2,2 milhões na região Nordeste; 1,15 milhão na região Centro-Oeste; e 639 mil na região Norte; totalizando cerca de 13 milhões de habitantes com educação superior, 7,09% da populacional de 190 milhões de habitantes no mesmo período. A macrorregião Sudeste sozinha representa 54% do total de habitantes com formação de nível superior completa no país. Ademais, esse padrão de maior proporção de habitantes com educação superior na macrorregião Sudeste se repete para todas os agrupamentos de áreas do conhecimento acadêmico.

Figura 2 - Quantidade de habitantes com educação superior

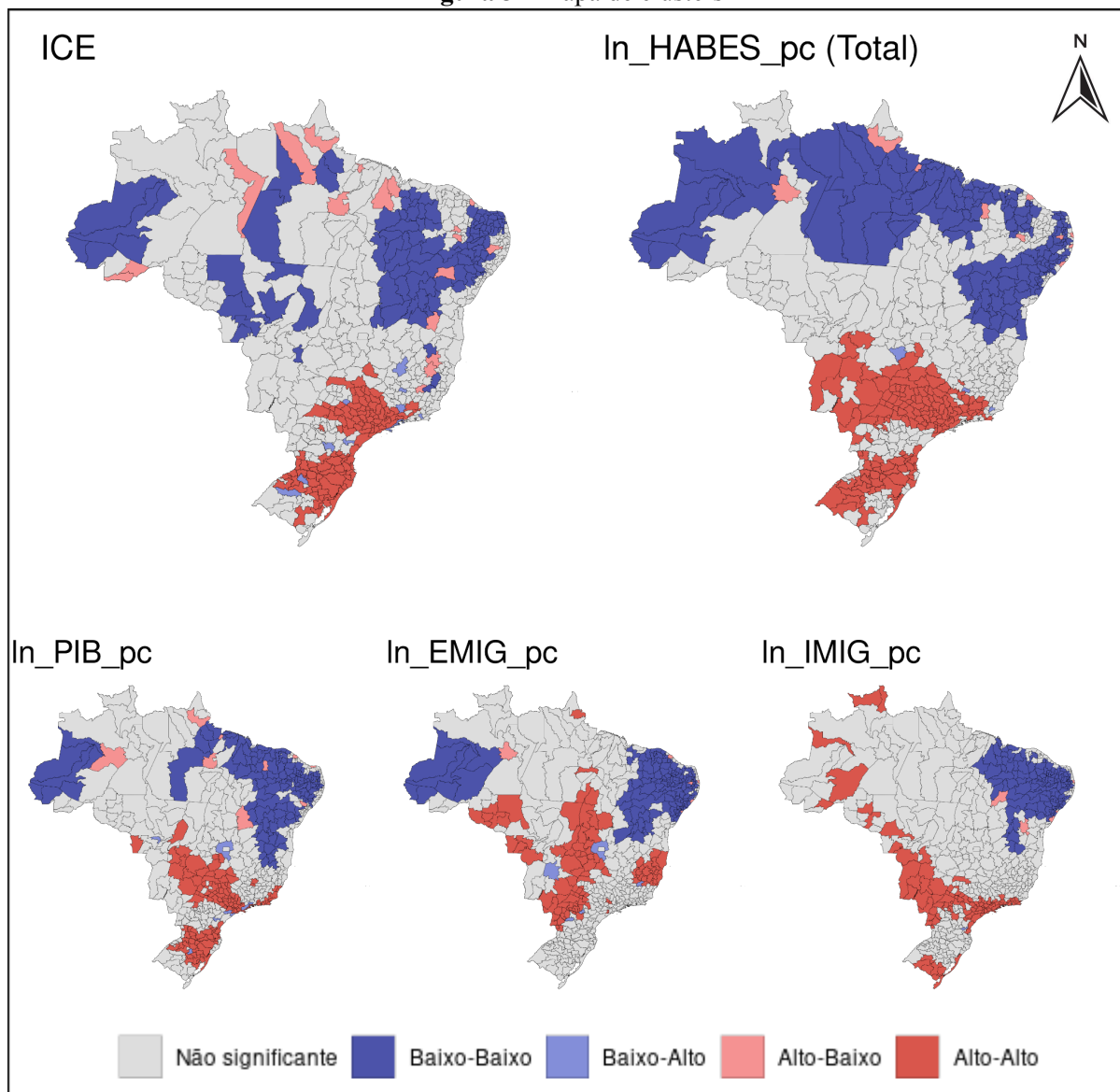


Fonte: IBGE. Elaborado pelo autor.

Além disso, a abrangência dos cursos também varia bastante entre a população. Áreas como “Ciências sociais, negócios e direito” reúnem cerca de 38% do total de habitantes com ensino superior; “Educação” 19,5%; “Saúde e bem-estar-social” 13%; “Ciências, matemática e computação” 7,8%; “Humanidades e artes” 7,65%; “Engenharia, produção e construção” 7,6%; “Serviços” 1,97%; e “Agricultura e veterinária” 1,93%. Isso demonstra clara heterogeneidade na preferência e disponibilidade dos cursos de educação superior no Brasil.

Contudo, ainda que todas as áreas do conhecimento analisadas tenham retratado notável concentração na região Sudeste, habitantes se concentram de maneira distinta dentro delas. Para visualizar clusters de concentração das estatísticas de interesse, recorreu-se ao método de visualização conhecido como mapa de cluster, ou gráfico de autocorrelação espacial local de Moran, onde clusters de observações altas (Alto-Alto) e baixas (Baixo-Baixo), assim como outliers altos em torno de observações baixas (Alto-Baixo) e baixos em torno de observações altas (Baixo-Alto), em relação à média, podem ser identificados. Então, torna-se simples a observação de agrupamentos locais e regiões discrepantes na mesma visualização.

Figura 3 - Mapa de clusters



Fonte: Elaborado pelo autor.

O mapa de autocorrelação local de Moran é um método de associação espacial local que pode ajudar na identificação de agrupamentos de estatísticas e regiões discrepantes. Através do mapa de cluster, é possível observar a concentração de complexidade econômica, capital humano e capital físico nas macrorregiões Sudeste e Sul, divididas pelo planalto paranaense; já as macrorregiões Nordeste e Norte apresentam clusters de baixa concentração, com algumas regiões que se sobressaem no litoral do Nordeste e arredores de Manaus; e o sertão nordestino revelou baixíssimo desenvolvimento regional.

4.2 ANÁLISE CAUSAL GLOBAL

Um dos métodos mais utilizados em econometria para se obter coeficientes causais é o método de mínimos quadrados ordinários. Embora a organização dos dados em seção cruzada não nos permita implicar causalidade, juntamente com uma boa argumentação teórica é possível compreender melhor o pressuposto efeito de uma variável sobre a outra. Abaixo, temos a descrição tabular dos modelos globais estimados. Cada um dos 9 modelos foi estimado com as mesmas variáveis, mudando apenas a área do conhecimento da variável \ln_HABES_pc . A tabela de regressões com as áreas do conhecimento selecionadas permite comparar os resultados com seus semelhantes, servindo como um teste de robustez, e permitindo a comparação do impacto relativo entre áreas do conhecimento.

Tabela 1 - Regressões globais

	ICE								
	Total	Educação	Humanidades e artes	Ciências sociais, negócios e direito	Ciências, matemática e computação	Engenharia, produção e construção	Agricultura e veterinária	Saúde e bem-estar-social	Serviços
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ln_HABES_pc	0.374*** (0.052)	0.196*** (0.061)	0.203*** (0.041)	0.217*** (0.032)	0.191*** (0.034)	0.195*** (0.021)	0.041* (0.021)	0.129*** (0.031)	0.092*** (0.019)
ln_PIB_pc	0.093** (0.042)	0.235*** (0.037)	0.195*** (0.038)	0.074* (0.044)	0.171*** (0.038)	0.015 (0.043)	0.234*** (0.038)	0.179*** (0.040)	0.184*** (0.039)
ln_EMIG_pc	-0.051** (0.023)	-0.021 (0.024)	-0.032 (0.024)	-0.068*** (0.024)	-0.042* (0.023)	-0.048** (0.023)	-0.019 (0.024)	-0.043* (0.025)	-0.028 (0.023)
ln_IMIG_pc	0.046*** (0.014)	0.063*** (0.014)	0.057*** (0.014)	0.046*** (0.014)	0.062*** (0.014)	0.041*** (0.013)	0.065*** (0.014)	0.063*** (0.014)	0.046*** (0.014)
dist_costa	-0.257*** (0.056)	-0.218*** (0.058)	-0.224*** (0.057)	-0.221*** (0.056)	-0.243*** (0.057)	-0.136** (0.054)	-0.211*** (0.059)	-0.244*** (0.058)	-0.204*** (0.057)
rg_Nordeste	-0.004 (0.086)	0.029 (0.089)	0.001 (0.088)	-0.005 (0.087)	0.026 (0.087)	-0.026 (0.084)	0.046 (0.090)	-0.013 (0.089)	0.006 (0.088)
rg_Norte	0.164** (0.075)	0.102 (0.077)	0.115 (0.077)	0.143* (0.076)	0.139* (0.076)	0.076 (0.072)	0.092 (0.079)	0.091 (0.076)	0.063 (0.075)
rg_Sudeste	0.104 (0.075)	0.201*** (0.076)	0.179** (0.075)	0.137* (0.075)	0.139* (0.076)	0.106 (0.072)	0.239*** (0.076)	0.122 (0.079)	0.194*** (0.075)
rg_Sul	0.212*** (0.078)	0.284*** (0.080)	0.280*** (0.079)	0.218*** (0.079)	0.257*** (0.079)	0.250*** (0.075)	0.307*** (0.080)	0.238*** (0.081)	0.296*** (0.078)
Constant	0.466* (0.271)	0.038 (0.343)	0.324 (0.312)	0.193 (0.256)	0.341 (0.295)	0.721*** (0.260)	-0.517** (0.261)	-0.063 (0.278)	-0.148 (0.248)
n	557	557	557	557	557	557	557	557	557
BP	28.135	18.255	15.324	24.214	23.509	26.881	21.196	20.251	17.275
F	59.776	50.554	52.871	57.754	54.855	65.985	49.804	52.328	54.451
R2	0.496	0.454	0.465	0.487	0.474	0.521	0.45	0.463	0.473
R2 adj.	0.445	0.445	0.456	0.479	0.466	0.513	0.441	0.454	0.464
AIC	478.658	518.203	508.85	489.901	498.524	452.904	525.078	514.251	505.974
BIC	526.206	565.751	556.398	537.449	546.072	500.452	572.626	561.799	553.522
RMSE	0.36	0.38	0.37	0.37	0.37	0.36	0.38	0.38	0.37
MAPE	2.32	2.34	2.35	2.27	2.41	2.06	2.34	2.26	2.15
Note:							*p<0.1; **p<0.05; ***p<0.01		

Fonte: Elaborado pelo autor

Os resultados dos modelos globais mostram contrastes no impacto das áreas do conhecimento sobre a complexidade econômica. Naturalmente, o impacto do “Total” de habitantes com educação superior é maior do que o das demais áreas do conhecimento, dado que a análise é feita com valores relativos à quantidade de habitantes na região (per capita). Dessa maneira, a interpretação correta do coeficiente causal é: com a variação de 1% na proporção de habitantes com dada formação superior, espera-se um acréscimo de $\hat{\beta}$ no índice de complexidade econômica (*ICE*). Assim, o coeficiente do “Total” deve sempre ser superior aos coeficientes de áreas específicas.

Destarte, como a proporção de habitantes com formação de nível superior varia entre cursos, o impacto dos coeficientes deve variar também. Assim, áreas do conhecimento menos populares, como “Serviços” e “Agropecuária e veterinária”, devem sinalizar menor impacto com a variação percentual dos mesmos. Por outro lado, áreas do conhecimento mais populares, como “Ciências sociais, negócios e direito” e “Educação”, devem retornar coeficientes maiores devido a sua grande numerosidade pelo país. Desse modo, a comparação do efeito do acréscimo percentual entre diferentes áreas do conhecimento deve ser realizada, quando feita, com cautela, ainda que esperado efeito heterogêneo entre as áreas do conhecimento.

De todo modo, mesmo sem possuir quantidade expressiva (cerca de 1 milhão de pessoas), o efeito da variação percentual da quantidade de habitantes com formação em “Engenharia, produção e construção” per capita se destaca (0.198), sendo, inclusive, um coeficiente que explica melhor a variação da complexidade econômica do que o total de habitantes com educação superior per capita. As métricas de ajuste apresentadas (R^2 , R^2 ajustado, F, AIC, BIC, RMSE e MAPE) podem auxiliar na medição do grau de ajuste dos modelos e na comparação do poder de explicabilidade das áreas do conhecimento dos habitantes.

Para haver maior confiança nas análises, foi necessário a exclusão de outros potenciais fatores que pudessem afetar os resultados da modelagem. Por esse motivo, foram empregadas variáveis de controle e de efeito fixo a fim de se obter resultados mais robustos. Os coeficientes causais globais do logaritmo do PIB per capita (+), do logaritmo da quantidade de emigrantes per capita (-), do logaritmo da quantidade de imigrantes per capita (+), e da distância da costa atlântica (-) foram significativos na maior parte dos 9 modelos globais. As variáveis de efeito fixo também tiveram boa aderência, sobretudo para as macrorregiões Sul e Sudeste. Outras combinações de variáveis são apresentadas na tabela 5 para dar mais confiança às análises.

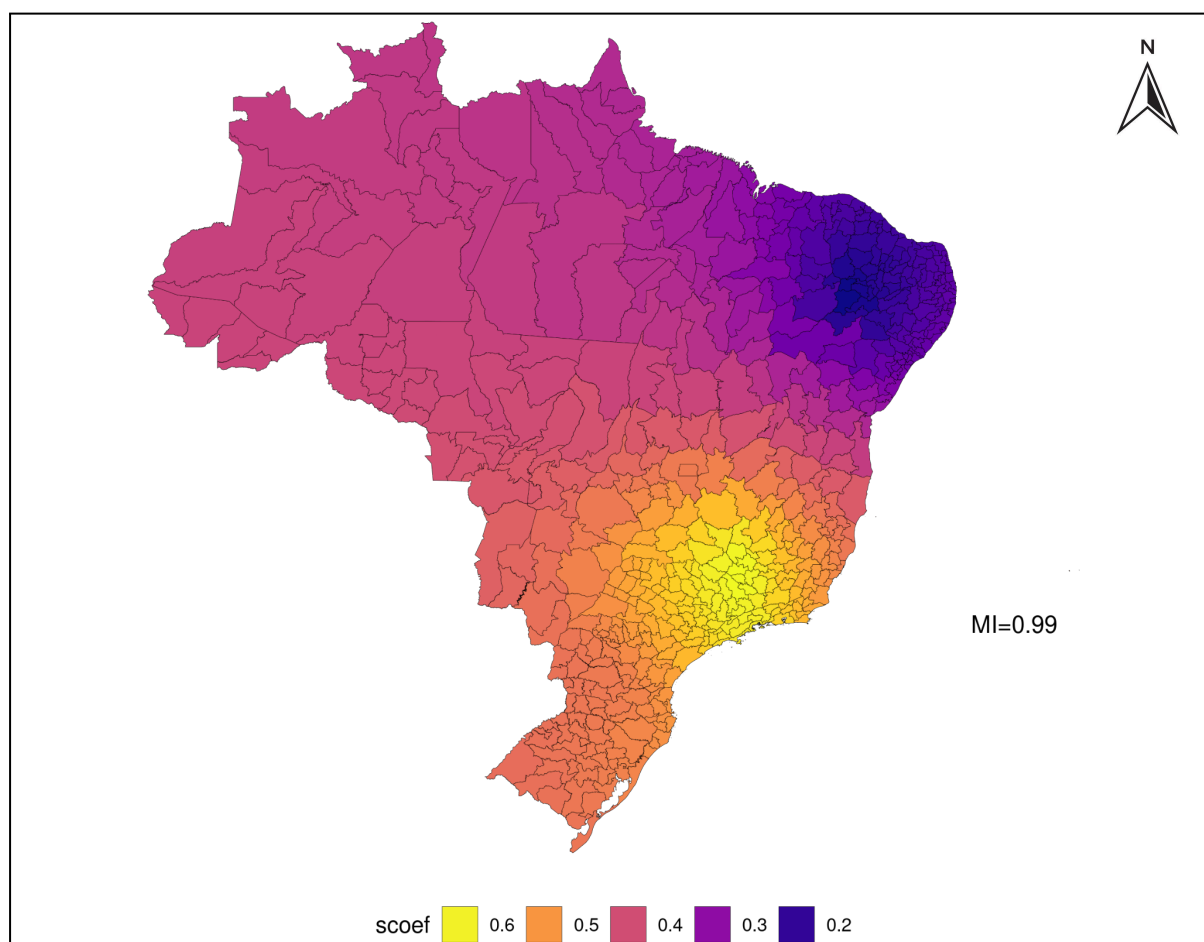
Ademais, através da observação da distribuição dos resíduos do modelo global (Total) sobre o espaço (Figura 9) é possível verificar leve concentração dos resíduos do modelo ($MI_{res} = 0,09$), potencialmente causada por viés de variável omitida. Além disso, algumas regiões se destacam na discrepância de seus desvios da estimativa do modelo global, em especial as microrregiões próximas de Manaus (AM), Rio Branco (AC), Porto Velho (RO) e Boa Vista (RR), onde a proximidade com vizinhos da América Latina contribuem para trocas e desenvolvimento. Por fim, o teste de

Breusch-Pagan (*BP*) foi aplicado e a hipótese nula de homocedasticidade não foi rejeitada, corroborando com a sinalização de modelos bem ajustados.

4.3 ANÁLISE CAUSAL LOCAL

Na abordagem local, no entanto, tem-se uma série de coeficientes locais capazes de retratar a heterogeneidade da relação causal sobre a extensão da amostra. Assim, variações causais locais podem se tornar mais evidentes quando visualizadas no mapa. A interpretação dos coeficientes locais segue a mesma dos globais, com a diferença que variam sobre a amostra. Abaixo temos o coeficiente causal do logaritmo da quantidade de habitantes com educação superior (Total) sobre a complexidade econômica, onde se observa efeito superior na região Sudeste em comparação com regiões periféricas:

Figura 4 - Coeficiente de relação causal local de \ln_HABES_pc (Total)



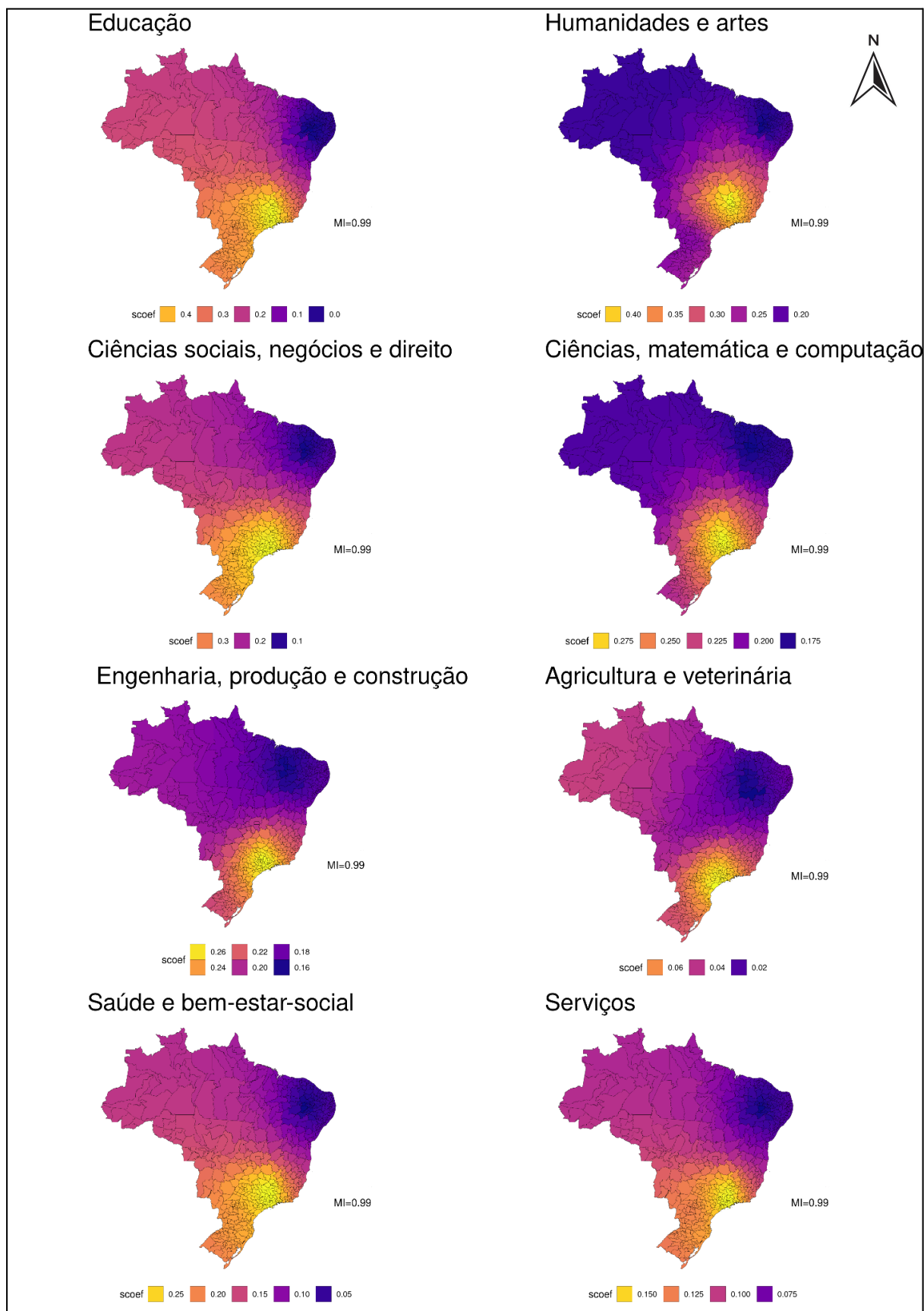
Fonte: Elaborado pelo autor

O coeficiente causal local (*scoef*) do total de habitantes com educação superior sobre a complexidade econômica retrata bem os retornos crescentes de escala do capital humano, supostamente presentes na região, com satisfatório grau de significância estatística (Figura 11). Na região Sudeste, onde habitantes com educação superior se concentram, parece haver retorno superior do que em regiões periféricas; a região Sul aparece em um estágio intermediário; as regiões Centro-Oeste e Norte demonstram menor retorno; e a região Nordeste apresenta um cluster de retorno inferior.

Variáveis de controle foram empregadas para buscar maior robustez nas análises, permitindo também outras interpretações de interesse (Figura 10). O log do PIB per capita, inserido como proxy do capital físico disponível na região, apresentou retorno positivo, significativo e heterogêneo sobre a amostra, explicitando as regiões com maior e menor demanda por capital físico relativo; o log da quantidade de emigrantes per capita apresentou retorno negativo, significativo e heterogêneo sobre a amostra, ilustrando os efeitos da perda de capital humano decorrente da evasão de habitantes; o log da quantidade de imigrantes per capita apresentou retorno positivo, significativo e heterogêneo sobre a amostra, ilustrando os efeitos do acréscimo de capital humano estrangeiro; a distância da costa atlântica apresentou retorno negativo e significativo, ilustrando os efeitos endógenos da histórica formação litorânea da economia brasileira. Ademais, as variáveis de efeito fixo promoveram melhor ajuste dos modelos.

Além disso, deve-se observar a recorrência de efeito superior na região Sudeste para todas as áreas do conhecimento analisadas. Independente da área do conhecimento, as estimativas de relação causal de habitantes com educação superior per capita sobre a complexidade econômica apresentaram impacto superior na macrorregião Sudeste do Brasil, sinalizando a ocorrência de retornos crescentes de escala da educação superior, dada a recorrência verificada. Abaixo é possível visualizar a distribuição dos coeficientes locais das demais áreas do conhecimento.

Figura 5 - Coeficientes de relação causal local de ln_HABES_pc (Áreas do conhecimento)



Fonte: Elaborado pelo autor.

A heterogeneidade nos coeficientes das demais variáveis independentes sobre a complexidade econômica também pode ser verificada (Figura 10), para determinado nível de significância local (Figura 11 no apêndice). Assim, corroborando com as interpretações de interesse sobre os fatores que levam ao desenvolvimento econômico. De maneira tabular, abaixo podemos observar a heterogeneidade dos coeficientes locais e comparar com a estimação do coeficiente global.

Tabela 2 - Sumário causal global e local (Modelo 1 - Total)

	Global	Mínimo	1.Quartil	Mediana	Média	3.Quartil	Máximo
Intercepto	0.466	0.199	0.423	0.621	0.669	0.920	1.305
ln_HABES_pc	0.374	0.172	0.272	0.426	0.404	0.501	0.608
ln_PIB_pc	0.093	0.026	0.058	0.074	0.074	0.093	0.112
ln_EMIG_pc	-0.051	-0.116	-0.085	-0.056	-0.054	-0.023	0.018
ln_IMIG_pc	0.046	0.027	0.036	0.039	0.041	0.048	0.060
dist_costa	-0.257	-0.539	-0.423	-0.301	-0.335	-0.244	-0.213
rg_Nordeste	-0.004	-0.253	-0.161	-0.132	-0.111	-0.055	0.033
rg_Norte	0.164	-0.057	0.032	0.090	0.091	0.141	0.241
rg_Sudeste	0.104	-0.122	-0.060	-0.031	-0.017	0.002	0.131
rg_Sul	0.212	-0.024	0.042	0.095	0.099	0.156	0.235

Fonte: Elaborado pelo autor.

Além disso, comparando medidas de grau de ajuste dos modelos (Total) global e local, é possível verificar o aprimoramento proporcionado pela ponderação geográfica. A ponderação entre as regiões (observações) próximas provocou melhoria no ajuste do modelo especificado (Mais evidências disponíveis nos apêndices). Abaixo é possível verificar a redução do erro percentual absoluto médio (*MAPE*) e da autocorrelação espacial dos resíduos (*MI_res*) com a adoção da regressão geograficamente ponderada.

Tabela 3 - Tabela de grau de ajuste (Modelo 1 - Total)

Modelo	Métrica	MQO	RGP
<i>l</i>	<i>MAPE</i>	2.32	2.24
<i>l</i>	<i>MI_res</i>	0.09	0.07

Fonte: Elaborado pelo autor.

O mesmo pode ser observado para as demais áreas do conhecimento acadêmico, para todas as áreas analisadas houve redução do erro percentual absoluto médio (*MAPE*).

Tabela 4 - Tabela de grau de ajuste (áreas do conhecimento)

<i>Modelo</i>	<i>Métrica</i>	<i>MQO</i>	<i>RGP</i>
2	<i>MAPE</i>	2.34	2.20
3	<i>MAPE</i>	2.35	2.19
4	<i>MAPE</i>	2.27	2.15
5	<i>MAPE</i>	2.41	2.36
6	<i>MAPE</i>	2.06	1.93
7	<i>MAPE</i>	2.34	2.21
8	<i>MAPE</i>	2.26	2.11
9	<i>MAPE</i>	2.15	2.07

Fonte: Elaborado pelo autor.

Aqui foram apresentadas evidências empíricas da importância do atributo espacial dos dados. Ainda que o transbordamento do conhecimento não tenha sido medido, há fortes indícios de que a troca de informação entre os agentes exerce influência sobre o desenvolvimento econômico regional nas microrregiões do Brasil.

5 CONSIDERAÇÕES FINAIS

O presente trabalho de conclusão de curso apresentou análises descritivas e causais, globais e locais, de fatores que levam ao aumento da complexidade econômica regional, dando foco a diferentes áreas do conhecimento acadêmico de habitantes, para explicar a ocorrência de efeitos de transbordamento do conhecimento, em escala microrregional. A partir da teoria do crescimento endógeno e regressões geograficamente ponderadas, argumenta-se sobre a ocorrência de heterogeneidade no pressuposto efeito do capital humano com educação superior sobre a complexidade econômica. Na região Sudeste, onde o capital humano se concentra, estima-se haver

maior retorno da educação superior sobre a complexidade econômica do que em regiões periféricas no Brasil. Caracterizado como um problema complexo, espera-se que os resultados aqui apresentados possam auxiliar tomadores de decisão na melhor alocação de recursos públicos e privados.

5.1 CONCLUSÕES GERAIS

Observada a concentração global do capital humano e complexidade econômica, verificado os clusters locais de habitantes com educação superior completa nas macrorregiões Sudeste e Sul, assim como polarização do pressuposto efeito do transbordamento do conhecimento na macrorregião Sudeste, argumenta-se sobre a existência de retornos crescentes de escala da educação superior sobre o desenvolvimento econômico no Brasil.

O método de regressão geograficamente ponderado foi capaz de entregar resultados melhor ajustados a respeito do efeito do capital humano com educação superior sobre a complexidade econômica do que em comparação com o método tradicional de regressão linear global, sendo possível verificar a heterogeneidade do efeito causal de maneira estatisticamente significativa na maior parte da extensão da amostra. Para todas as áreas do conhecimento analisadas, se verifica um polo de retornos crescentes de escala na região Sudeste do Brasil, onde o capital humano se concentra no país. O transbordamento do conhecimento entre os agentes é apontado como provável causa do maior retorno esperado.

Fica claro, portanto, a caracterização de um problema complexo para aplicadores de políticas públicas. Caso escolham promover igualdade, iriam investir recursos de maneira subótima; caso escolham aplicar de maneira efetiva, iriam promover ainda maiores níveis de desigualdade. Assim, tomadores de decisão devem ter claro o objetivo do investimento para fazer melhor uso dele. Cursos como administração e direito aparentam ser mais factíveis de se oferecer na maioria das regiões do país; por outro lado, cursos custosos, que requerem grandes volumes de investimento para serem oferecidos, devem ser promovidos onde o retorno esperado é maior.

O presente trabalho buscou contribuir com a literatura de capital humano e complexidade econômica superando a limitação de metodologia global recorrentemente verificada em modelos de crescimento econômico. Espera-se que o desenvolvimento deste trabalho possa também transbordar e promover progresso na literatura econômica.

5.2 SUGESTÕES PARA TRABALHOS FUTUROS

Os métodos aqui aplicados, tanto globais como locais, naturalmente, tem suas limitações. Para entender essas limitações, assim como direcionar trabalhos futuros, abaixo estão listadas algumas sugestões para superação em trabalhos futuros.

Primeiramente, no presente trabalho foram utilizados dados em seção cruzada, para o ano de 2010, o que impede a inferência de causalidade. Caso seja feita a mesma análise através de dados em painel poderia-se chegar a resultados mais conclusivos.

Além disso, no presente trabalho foi utilizado o método de estimação subnacional para calcular o índice de complexidade econômica através de dados de comércio exterior. Caso sejam utilizados outros métodos, como ECI-reflexions, ECI-plus ou Fitness-index, utilizando dados de produção, mercado de trabalho ou indústria, poderia-se potencialmente obter resultados mais robustos.

Normalizar a geografia através de hexágonos regulares também poderia ser uma alternativa aos polígonos microrregionais para delimitar as regiões utilizadas na análise (VIEGAS et al., 2020). Caso sejam utilizados polígonos hexagonais regulares, poderia-se usar kernel fixo e potencialmente obter resultados mais robustos e visualmente mais atraentes.

Neste, todas as variáveis foram apresentadas no ano de 2010. Caso seja feita uma especificação dinâmica, de modo a utilizar variáveis defasadas, poderia-se potencialmente obter o efeito de uma variável sobre a outra em períodos distintos no tempo.

No presente trabalho foram usadas variáveis de controle e efeito fixo para remover algum possível viés por variável omitida e endogeneidade. Caso seja utilizada uma variável instrumental, poderia-se obter resultados potencialmente mais robustos.

REFERÊNCIAS

ACEMOGLU, Daron; ANGRIST, Joshua. How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws. **Nber Macroeconomics Annual**, [S.L.], v. 15, p. 9-59, jan. 2000. University of Chicago Press. <http://dx.doi.org/10.1086/654403>.

AGHION, Philippe; HOWITT, Peter. A Model of Growth Through Creative Destruction. **Econometrica**, [S.L.], v. 60, n. 2, p. 323, mar. 1992. JSTOR. <http://dx.doi.org/10.2307/2951599>.

ABDALA, Vitor. Economia brasileira cresceu a uma taxa média anual de 3,7% entre 2000 e 2011. **Agência Brasil**. Mar. 2015. Empresa Brasil de Comunicação. <https://memoria.ebc.com.br/educacao/2015/03/economia-brasileira-cresceu-uma-taxa-media-anual-de-37-entre-2000-e-2011>.

ALBEAIK, Saleh; KALTENBERG, Mary; ALSALEH, Mansour; HIDALGO, Cesar A.. Improving the Economic Complexity Index. **Cornell University**, [S.L.], jul. 2017. <https://arxiv.org/abs/1707.05826>.

ANSELIN, Luc; VARGA, Attila; ACS, Zoltan. Local Geographic Spillovers between University Research and High Technology Innovations. **Journal Of Urban Economics**, [S.L.], v. 42, n. 3, p. 422-448, nov. 1997. Elsevier BV. <http://dx.doi.org/10.1006/juec.1997.2032>.

ANSELIN, Luc. Local Indicators of Spatial Association-LISA. **Geographical Analysis**, [S.L.], v. 27, n. 2, p. 93-115, 3 set. 2010. Wiley. <http://dx.doi.org/10.1111/j.1538-4632.1995.tb00338.x>.

AUDRETSCH, David B.; FELDMAN, Maryann P.. R&D Spillovers and the Geography of Innovation and Production. **The American Economic Review**, [S.L.], v. 86, n. 3, p. 630-640, jun. 1996. American Economic Association. <https://www.jstor.org/stable/2118216>.

BALASSA, Bela. Trade Liberalisation and. **The Manchester School**, [S.L.], v. 33, n. 2, p. 99-123, maio 1965. Wiley. <http://dx.doi.org/10.1111/j.1467-9957.1965.tb00050.x>.

BALLAND, Pierre Alexandre; JARA-FIGUEROA, Cristian; PETRALIA, Sergio; STEIJN, Mathieu; RIGBY, David L.; HIDALGO, Cesar. Complex Economic Activities Concentrate in Large Cities. **Ssrn Electronic Journal**, [S.L.], 2018. Elsevier BV. <http://dx.doi.org/10.2139/ssrn.3219155>.

BALDANZI, Annarita; BUCCI, Alberto; PRETTNER, Klaus. Children's health, human capital accumulation, and R&D-based economic growth. **Hohenheim Discussion Papers in Business, Economics and Social Sciences**, jan. 2017. University of Hohenheim.
<https://ideas.repec.org/p/zbw/hohdps/012017.html>.

BECKER, Gary S.. Investment in Human Capital: A Theoretical Analysis. **Journal of Political Economy**, v. 70, n. 5., out. 1962. The University of Chicago Press.
<https://www.jstor.org/stable/1829103>.

BERGH, Jeroen C. J. M. van Den. Abolishing GDP. **Ssrn Electronic Journal**, [S.L.], 2007. Elsevier BV. <http://dx.doi.org/10.2139/ssrn.962343>.

BIDANSET, Paul E.; LOMBARD, John R.. The Effect of Kernel and Bandwidth Specification in Geographically Weighted Regression Models on the Accuracy and Uniformity of Mass Real Estate Appraisal. **Journal of Property Tax Assessment & Administration**, v. 10, n. 3, 2014.
https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1027&context=publicservice_pubs.

BLOOM, David E.; KUHN, Michael; PRETTNER, Klaus. Health and Economic Growth. **IZA Discussion Paper**, n. 11939, dec. 2018. <https://dx.doi.org/10.2139/ssrn.3301688>.

BRAGA, Fernando. Migração Interna e Urbanização no Brasil Contemporâneo: Um estudo da Rede de Localidades Centrais do Brasil (1980/2000). **XV Encontro Nacional de Estudos Populacionais**, MG, Brasil, 22 set. 2006. ABEP. <https://www.yumpu.com/s/sY1eEgGZ88Vqt0Tv>.

BRAKMAN, Steven; GARRETSEN, Harry; VAN MARREWIJK, Charles. **The New Introduction to Geographical Economics**. ed. 2, The Netherlands, Cambridge University Press, 2009.
<https://doi.org/10.1017/CBO9780511818745>.

BRESSER-PEREIRA, Luiz C.. **Macroeconomia da Estagnação**. São Paulo: Editora 34, 2007.
<https://www.bresserpereira.org.br/index.php/brazilian-economy/main-historical-analysis/8311-2411>.

BRUNSDON, Chris; FOTHERINGHAM, Stewart; CHARLTON, Martin E.. Geographically weighted regression: a method for exploring spatial nonstationarity. **Journal Geographical analysis**, v. 28, n. 4, p. 281-298, out. 1996, Blackwell Publishing Ltd.
<https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1538-4632.1996.tb00936.x>.

CARDOSO, Fernanda G.. A armadilha do subdesenvolvimento: uma discussão do período desenvolvimentista brasileiro sob a ótica da abordagem da complexidade. Tese de Doutorado, **Universidade de São Paulo**. 2012. <http://dx.doi.org/10.11606/t.12.2012.tde-26062012-155604>.

CARLINO A. G.. New ideas in the air: cities and economic growth. **Business Review**, [S.L.], Q4, p.1-7, 2014. <https://ideas.repec.org/a/fip/fedpbr/00014.html>.

CASTRO, Maria H. G.; TIEZZI, Sergio. A reforma do ensino médio e a implantação do Enem no Brasil. **Os Desafios da Educação no Brasil**, p. 115-147, 2005. Nova Fronteira. <http://www.schwartzman.org.br/simon/desafios/4ensinomedio.pdf>.

CLEVELAND, William S.; DEVLIN, Susan J.. Locally Weighted Regression: an approach to regression analysis by local fitting. **Journal Of The American Statistical Association**, [S.L.], v. 83, n. 403, p. 596-610, set. 1988. Informa UK Limited. <http://dx.doi.org/10.1080/01621459.1988.10478639>.

CLEVELARIO JUNIOR, J.. A participação da imigração na formação da população brasileira. **Revista Brasileira De Estudos De População**, v. 14, n. ½, 1997. Rebep. <https://www.rebep.org.br/revista/article/view/421>.

CONIGLIO, Nicola D.; LAGRAVINESE, Raffaele; VURCHIO, Davide; ARMENISE, Massimo. The pattern of structural change: testing the product space framework. **Industrial and Corporate Change**, v. 27(4), p. 763-785, 2018. Oxford University Press. <https://ideas.repec.org/a/oup/indcch/v27y2018i4p763-785..html>.

COMEX STAT. **Estatísticas de Comércio Exterior em Dados Abertos**. Ministério da Economia. Base de dados detalhada por Município da empresa exportadora/importadora e Posição do Sistema Harmonizado (SH4). <https://www.gov.br/produtividade-e-comercio-exterior/pt-br/assuntos/comercio-exterior/estatisticas/ba-se-de-dados-bruta>.

DOCQUIER, Frédéric; RAPOPORT, Hillel. Globalization, Brain Drain, and Development. **Journal Of Economic Literature**, [S.L.], v. 50, n. 3, p. 681-730, 1 set. 2012. American Economic Association. <http://dx.doi.org/10.1257/jel.50.3.681>.

FIORAMONTI, Lorenzo. A Post-GDP World? Rethinking International Politics in the 21st Century.

Global Policy, [S.L.], v. 7, n. 1, p. 15-24, 9 set. 2015. Wiley.

<http://dx.doi.org/10.1111/1758-5899.12269>.

FISCHER, Manfred M.; VARGA, Attila. Spatial knowledge spillovers and university research:

evidence from Austria. **The Annals Of Regional Science**, [S.L.], v. 37, n. 2, p. 303-322, 1 maio 2003.

Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s001680200115>.

FRENKEN, K.; BOSCHMA, R. A.. A theoretical framework for evolutionary economic geography:

industrial dynamics and urban growth as a branching process. **Journal Of Economic Geography**,

[S.L.], v. 7, n. 5, p. 635-649, 13 jun. 2007. Oxford University Press.

<http://dx.doi.org/10.1093/jeg/lbm018>.

FOTHERINGHAM, A. S.; BRUNSDON, M.; CHARLTON, M.. **Geographically Weighted**

Regression: The Analysis of Spatially Varying Relationships, John Wiley & Sons, 2003.

FUJITA, Masahisa; THISSE, Jacques. Economics of Agglomeration. **Cambridge Books**, 2013,

Cambridge University Press. <https://econpapers.repec.org/bookchap/cupcbooks/9780521171960.htm>

FURTADO, Celso. **Formação econômica do Brasil**. Companhia das Letras, ed. 1, fev. 2007.

GALLUP, John L.; SACHS, Jeffrey D.; MELLINGER, Andrew D.. Geography and Economic

Development. **National Bureau of Economic Research**, n. 6849, dez. 1998.

<https://www.nber.org/papers/w6849>.

GOMEZ-LIEVANO, A.; PATTERSON-LOMBA, O.. Estimating the drivers of urban economic

complexity and their connection to economic performance. **Royal Society Open Science**, n. 8, 2021.

<http://www.tinyurl.com/ye7snp6b>

GROSSMAN, Gene M.; HELPMAN, Elhanan. Trade, knowledge spillovers, and growth. **European**

Economic Review, [S.L.], v. 35, n. 2-3, p. 517-526, abr. 1991. Elsevier BV.

[http://dx.doi.org/10.1016/0014-2921\(91\)90153-a](http://dx.doi.org/10.1016/0014-2921(91)90153-a).

GUJARATI, Damodar N. **Basic Econometrics**, Irwin/McGraw-Hill, ed. 5, out. 2008.

HARTMANN, Dominik; ZAGATO, Lgia M. J. C.; GALA, Paulo; PINHEIRO, Flvio L.. Why did some countries catch-up, while others got stuck in the middle? Stages of productive sophistication and smart industrial policies. **FVG**. mai. 2020. <https://hdl.handle.net/10438/29123>

HERRERA, William D. M.; STRAUCH, Julia C.; Bruno, Miguel A. P.. Complexidade Econmica das regies brasileiras no perodo 1998-2017. **Seminrio CEDEPLAR de Diamantina-MG 2019**, ago. 2019. https://diamantina.cedeplar.ufmg.br/portal/download/diamantina-2019/D18_118.pdf

HIDALGO, Csar A.; HAUSMANN, Ricardo. The building blocks of economic complexity. **Proceedings Of The National Academy Of Sciences**, [S.L.], v. 106, n. 26, p. 10570-10575, 30 jun. 2009. Proceedings of the National Academy of Sciences. <http://dx.doi.org/10.1073/pnas.0900943106>.

HIDALGO, Csar A.; HAUSMANN, Ricardo. The building blocks of economic complexity. **Proceedings Of The National Academy Of Sciences**, [S.L.], v. 106, n. 26, p. 10570-10575, 30 jun. 2009. Proceedings of the National Academy of Sciences. <http://dx.doi.org/10.1073/pnas.0900943106>.

HIDALGO, C. A.; KLINGER, B.; BARABSI, A.-L.; HAUSMANN, R.. The Product Space Conditions the Development of Nations. **Science**, [S.L.], v. 317, n. 5837, p. 482-487, 27 jul. 2007. American Association for the Advancement of Science (AAAS). <http://dx.doi.org/10.1126/science.1144581>.

IBGE. **API de Localidades**. Instituto Brasileiro de Geografia e Estatística. <https://servicodados.ibge.gov.br/api/docs/localidades>.

IBGE. **Diviso Regional do Brasil em Mesorregies e Microrregies Geogrficas**. Ministrio da Economia, Fazenda e Planejamento. https://biblioteca.ibge.gov.br/visualizacao/livros/liv2269_1.pdf.

IBGE. **Polgonos Espaciais Microrregionais de 2010**. Instituto Brasileiro de Geografia e Estatística. <https://www.ibge.gov.br/geociencias/downloads-geociencias.html>.

IBGE. **Tabela 21** - Produto interno bruto a preos correntes, impostos, lquidos de subsdios, sobre produtos a preos correntes e valor adicionado bruto a preos correntes total e por atividade econmica, e respectivas participaes - Referncia 2002. Instituto Brasileiro de Geografia e Estatística. <https://sidra.ibge.gov.br/tabela/21>.

IBGE. **Tabela 1554** - Pessoas de 10 anos ou mais de idade, por nvel de instruo - Resultados Gerais da Amostra. Instituto Brasileiro de Geografia e Estatística. <https://sidra.ibge.gov.br/Tabela/1554>.

IBGE. **Tabela 2145** - Naturalizados brasileiros e Estrangeiros, por grupos de anos em que fixaram residência no país, sexo e grupos de idade. Instituto Brasileiro de Geografia e Estatística.
<https://sidra.ibge.gov.br/Tabela/2145>.

IBGE. **Tabela 3173** - Emigrantes internacionais, por sexo, segundo os continentes e países estrangeiros de destino. Instituto Brasileiro de Geografia e Estatística.
<https://sidra.ibge.gov.br/Tabela/3173>.

IBGE. **Tabela 3543** - Pessoas com pelo menos nível superior de graduação concluído, por nível de instrução mais elevado concluído, segundo o sexo e as áreas gerais, específicas e detalhadas de formação do curso de nível mais elevado concluído. Instituto Brasileiro de Geografia e Estatística.
<https://sidra.ibge.gov.br/tabela/3543>.

INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Censo da Educação Superior 2010**. Divulgação dos principais resultados do Censo da Educação Superior 2010. out. 2011. Ministério da Educação.
https://download.inep.gov.br/educacao_superior/centso_superior/documentos/2010/divulgacao_censo_2010.pdf.

INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Censo da Educação Superior 2019**. Divulgação dos resultados. out. 2020. Ministério da Educação.
https://download.inep.gov.br/educacao_superior/centso_superior/documentos/2020/Apresentacao_Censo_da_Educacao_Superior_2019.pdf.

INOUA, Sabiou. Simple Measure of Economic Complexity. **Chapman University**.
<https://doi.org/10.48550/arXiv.1601.05012>.

KLENOW, Peter J.; RODRÍGUEZ-CLARE, Andrés. The Neoclassical Revival in Growth Economics: has it gone too far?. **Nber Macroeconomics Annual**, [S.L.], v. 12, p. 73-103, jan. 1997. University of Chicago Press. <http://dx.doi.org/10.1086/654324>.

KRUGMAN, Paul. Increasing Returns and Economic Geography. **Journal Of Political Economy**, [S.L.], v. 99, n. 3, p. 483-499, jun. 1991. University of Chicago Press.
<http://dx.doi.org/10.1086/261763>.

KOMATSU, Bruno K.; MENEZES-FILHO, Naercio; GANDRA, Pedro. Higher Education and GDP Per Capita in the Brazilian Municipalities. **Inspere**, [S.L.], n. 48, jul. 2020. Policy Paper. <https://www.insper.edu.br/wp-content/uploads/2020/07/Policy-Paper-48.pdf>.

LAVERDE-ROJAS, Henry; CORREA, Juan C.. Can scientific productivity impact the economic complexity of countries? **Scientometrics**, [S.L.], v. 120, n. 1, p. 267-282, 10 maio 2019. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s11192-019-03118-8>.

LEITE, Deyvid W.. Capital humano, investimento em tecnologia e complexidade econômica. Tese de Mestrado, **Universidade Federal de Viçosa**. 2019. <https://www.locus.ufv.br/bitstream/123456789/26588/1/texto%20completo.pdf>

LEWIS, W. Arthur. Economic Development with Unlimited Supplies of Labour. **The Manchester School**, [S.L.], v. 22, n. 2, p. 139-191, maio 1954. Wiley. <http://dx.doi.org/10.1111/j.1467-9957.1954.tb00021.x>.

LIN, Tin-Chun. The role of higher education in economic development: an empirical study of taiwan case. **Journal Of Asian Economics**, [S.L.], v. 15, n. 2, p. 355-371, abr. 2004. Elsevier BV. <http://dx.doi.org/10.1016/j.asieco.2004.02.006>.

LUCAS, Robert E.. On the mechanics of economic development. **Journal Of Monetary Economics**, [S.L.], v. 22, n. 1, p. 3-42, jul. 1988. Elsevier BV. [http://dx.doi.org/10.1016/0304-3932\(88\)90168-7](http://dx.doi.org/10.1016/0304-3932(88)90168-7).

LYKOSTRATIS, Konstantinos. From Global to Local: GWR as an Exploratory Tool for Spatial Phenomena. **Conventional and Fuzzy Regression: Theory and Applications**, New York, Nova Science Publishers, out. 2018. https://www.researchgate.net/publication/339182204_From_Global_to_Local_GWR_as_an_Exploratory_Tool_for_Spatial_Phenomena.

MANCIBO, Deise; VALE, Andréa Araujo do; MARTINS, Tânia Barbosa. POLÍTICAS DE EXPANSÃO DA EDUCAÇÃO SUPERIOR NO BRASIL 1995-2010. **Revista Brasileira de Educação**, [S.L.], v. 20, n. 60, p. 31-50, mar. 2015. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/s1413-24782015206003>.

MENICUCCI, Telma; GOMES, Sandra. **Políticas sociais: conceitos, trajetórias e a experiência brasileira**. Editora Fiocruz, 22. ed. Fundação Oswaldo Cruz, 2018.

MINCER, Jacob. Human capital and economic growth. **Economics Of Education Review**, [S.L.], v. 3, n. 3, p. 195-205, jan. 1984. Elsevier BV. [http://dx.doi.org/10.1016/0272-7757\(84\)90032-3](http://dx.doi.org/10.1016/0272-7757(84)90032-3).

MORAN, P. A. P. Notes on Continuous Stochastic Phenomena. **Biometrika**, v. 37, n. ½, jun. 1950, p. 17-23, Oxford University Press. <https://doi.org/10.2307/2332142>

MORETTI, Enrico. Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. **Journal Of Econometrics**, [S.L.], v. 121, n. 1-2, p. 175-212, jul. 2004. Elsevier BV. <http://dx.doi.org/10.1016/j.jeconom.2003.10.015>.

MORETTI, Enrico. Human capital externalities in cities. **Handbook Of Regional And Urban Economics**, [S.L.], p. 2243-2291, 2004. Elsevier. [http://dx.doi.org/10.1016/s1574-0080\(04\)80008-7](http://dx.doi.org/10.1016/s1574-0080(04)80008-7).

MORETTI, E.; THULIN, P. Local multipliers and human capital in the United States and Sweden. **Industrial And Corporate Change**, [S.L.], v. 22, n. 1, p. 339-362, 24 jan. 2013. Oxford University Press (OUP). <http://dx.doi.org/10.1093/icc/dts051>.

NEFFKE, Frank; HENNING, Martin; BOSCHMA, Ron. How Do Regions Diversify over Time? Industry Relatedness and the Development of New Growth Paths in Regions. **Economic Geography**, [S.L.], v. 87, n. 3, p. 237-265, 2 jun. 2011. Wiley. <http://dx.doi.org/10.1111/j.1944-8287.2011.01121.x>.

REYNOLDS, Christian; AGRAWAL, Manju; LEE, Ivan; ZHAN, Chen; LI, Jiuyong; TAYLOR, Phillip; MARES, Tim; MORISON, Julian; ANGELAKIS, Nicholas; ROOS, Göran. A sub-national economic complexity analysis of Australia's states and territories. **Regional Studies**, [S.L.], v. 52, n. 5, p. 715-726, 16 mar. 2017. Informa UK Limited. <http://dx.doi.org/10.1080/00343404.2017.1283012>.

NORRIS, Ann. How knowledge spillover contributes to economic growth in metro areas. **Monthly Labor Review**, fev. 2015. U.S. Bureau Of Labor Statistics. <https://www.bls.gov/opub/mlr/2015/beyond-bls/pdf/how-knowledge-spillover-contributes-to-economic-growth-in-metro-areas.pdf>

OBSERVATORY OF ECONOMIC COMPLEXITY. **Methods: ECI-Subnational**. Cambridge, 2022. <https://oec.world/en/resources/methods#eci-subnational>.

OREIRO, José Luis; MANARIN, Luciano Luiz; GALA, Paulo. Deindustrialization, economic complexity and exchange rate overvaluation: the case of Brazil (1998-2017). **PSL Quarterly Review**, [S.L.], v. 73, p. 295, 16 jan. 2021. PSL Quarterly Review. http://dx.doi.org/10.13133/2037-3643_73.295_3.

PRETTNER, Klaus; BLOOM, David E.; STRULIK, Holger. Declining fertility and economic well-being: do education and health ride to the rescue?. **Labour Economics**, [S.L.], v. 22, p. 70-79, jun. 2013. Elsevier BV. <http://dx.doi.org/10.1016/j.labeco.2012.07.001>.

PRETTNER, Klaus. Population ageing and endogenous economic growth. **Journal Of Population Economics**, [S.L.], v. 26, n. 2, p. 811-834, 15 set. 2012. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s00148-012-0441-9>.

QUEIROZ, Delcele M.. O Estágio Atual das Políticas Afirmativas nas Universidades Brasileiras. **Revista Tomo**, [S.L.], 1 jun. 2014. <http://dx.doi.org/10.21669/tomo.v0i0.3192>.

RAUCH, James E.. Productivity Gains From Geographic Concentration of Human Capital: Evidence From the Cities. **National Bureau of Economic Research**, n. 3905, nov. 1991. <http://dx.doi.org/10.3386/w3905>.

REYNOLDS, Christian; AGRAWAL, Manju; LEE, Ivan; ZHAN, Chen; LI, Jiuyong; TAYLOR, Phillip; MARES, Tim; MORISON, Julian; ANGELAKIS, Nicholas; ROOS, Göran. A sub-national economic complexity analysis of Australia's states and territories. **Regional Studies**, [S.L.], v. 52, n. 5, p. 715-726, 16 mar. 2017. Informa UK Limited. <http://dx.doi.org/10.1080/00343404.2017.1283012>.

ROMER, Paul M.. Increasing Returns and Long-Run Growth. **Journal of Political Economy**, [S.L.], v. 94, n. 5, p. 1002-1037, oct. 1986. The University of Chicago Press. <https://www.jstor.org/stable/1833190>.

ROMER, Paul M.. Endogenous Technological Change. **Journal Of Political Economy**, [S.L.], v. 98, n. 52, p. 71-102, out. 1990. University of Chicago Press. <http://dx.doi.org/10.1086/261725>.

ROSA, Chaiane M.. Políticas afirmativas para acesso à educação superior brasileira: História e embates. **Plures Humanidades**, v. 15, n. 1, 2014. <http://seer.mouralacerda.edu.br/index.php/plures/article/view/111>.

RUDD, Jeremy B.. Empirical Evidence on Human Capital Spillovers. **Board of Governors of the Federal Reserve System**, Washington, DC, out. 2000. <https://dx.doi.org/10.2139/ssrn.249292>.

SALLES, Fernanda Cimini; ROCHA, Elisa Maria Pinto da; PORTO, Ivana Villefort de Bessa; VASCONCELOS, Felipe Lopes Vieira. A armadilha da baixa complexidade em Minas Gerais: o desafio da sofisticação econômica em um estado exportador de commodities. **Revista Brasileira de Inovação**, [S.L.], v. 17, n. 1, p. 33, 30 out. 2017. Universidade Estadual de Campinas. <http://dx.doi.org/10.20396/rbi.v17i1.8650857>.

SCHULTZ, Theodore W.. Education and Productivity. **National Commission on Productivity**, Washington, DC, ERIC Number ED071152, jun. 1971. U.S. Government Printing Office. <https://eric.ed.gov/?id=ED071152>.

SCHULTZ, Theodore W.. Investment in Human Capital. **The American Economic Review**, [S.L.], v. 51, n. 1, p. 1-17, mar. 1961. American Economic Association. <https://www.jstor.org/stable/1818907>.

SIMÕES, Alexander J. G.; HIDALGO, César A.. The Economic Complexity Observatory: An Analytical Tool for Understanding the Dynamics of Economic Development. **AAAI Publications, Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence**. Ago. 2011. <https://www.aaai.org/ocs/index.php/WS/AAAIW11/paper/viewPaper/3948>.

SOLOW, Robert M.. A Contribution to the Theory of Economic Growth. **The Quarterly Journal Of Economics**, [S.L.], v. 70, n. 1, p. 65, fev. 1956. Oxford University Press (OUP). <http://dx.doi.org/10.2307/1884513>.

SOLOW, Robert M.. Technical Change and the Aggregate Production Function. **The Review Of Economics And Statistics**, [S.L.], v. 39, n. 3, p. 312, ago. 1957. JSTOR. <http://dx.doi.org/10.2307/1926047>.

SWAN, T. W.. ECONOMIC GROWTH and CAPITAL ACCUMULATION. **Economic Record**, [S.L.], v. 32, n. 2, p. 334-361, nov. 1956. Wiley. <http://dx.doi.org/10.1111/j.1475-4932.1956.tb00434.x>.

TACHELLA, Andrea; CRISTELLI, Matthieu; CALDARELLI, Guido; GABRIELLI, Andrea; PIETRONERO, Luciano. A New Metrics for Countries' Fitness and Products' Complexity. **Scientific Reports**, v. 2, n. 723, out. 2012. <https://www.nature.com/articles/srep00723>.

TORRES, Guilherme S.. Complexidade econômica: uma proposta metodológica para identificação de produtos estratégicos. **Dissertação**. Belo Horizonte, MG, UFMG, 2019.

<https://repositorio.ufmg.br/handle/1843/32851>.

UN. **International Convention on the Harmonized Commodity Description and Coding System** (with annex), as amended by the Protocol of Amendment of 24 June 1986. Concluded at Brussels on 14 June 1983. United Nations. 1988.

<https://treaties.un.org/doc/Publication/UNTS/Volume%201503/volume-1503-I-25910-English.pdf>.

VALE, Ana L. F.; LIMA, Luís C.; BONFIM, Maria G.. Século XX: 70 anos de migração interna no Brasil. **Textos e Debates**, [S.L.], n. 7, 2004. <https://doi.org/10.18227/2317-1448ted.v1i7.1027>.

VIEGAS, Guilherme; HOPPEN Joni; NETO, Eurico. **Normalização Geográfica**: O que é e quais suas vantagens?. <https://www.aquare.la/normalizacao-geografica/>.

VIEGAS, Guilherme. Códigos UFSC. **Repositório Institucional**. tcc_gv_dh.

https://codigos.ufsc.br/guilherme.viegas/tcc_gv_dh.

VIGGIANO, Esdras; MATTOS, Cristiano. O desempenho de estudantes no Enem 2010 em diferentes regiões brasileiras. **Revista Brasileira de Estudos Pedagógicos**, Brasília, v. 94, n. 237, p. 417-438, ago. 2013. <https://www.scielo.br/j/rbeped/a/dLdrbB3DSjMvQGm7GnGF7XB/?format=pdf&lang=pt>.

WCO. **The Harmonized System**: A universal language for international trade 1988-2018 30 Years On. World Customs Organization.

<http://www.wcoomd.org/-/media/wco/public/global/pdf/topics/nomenclature/activities-and-programmes/30-years-hs/hs-compendium.pdf>.

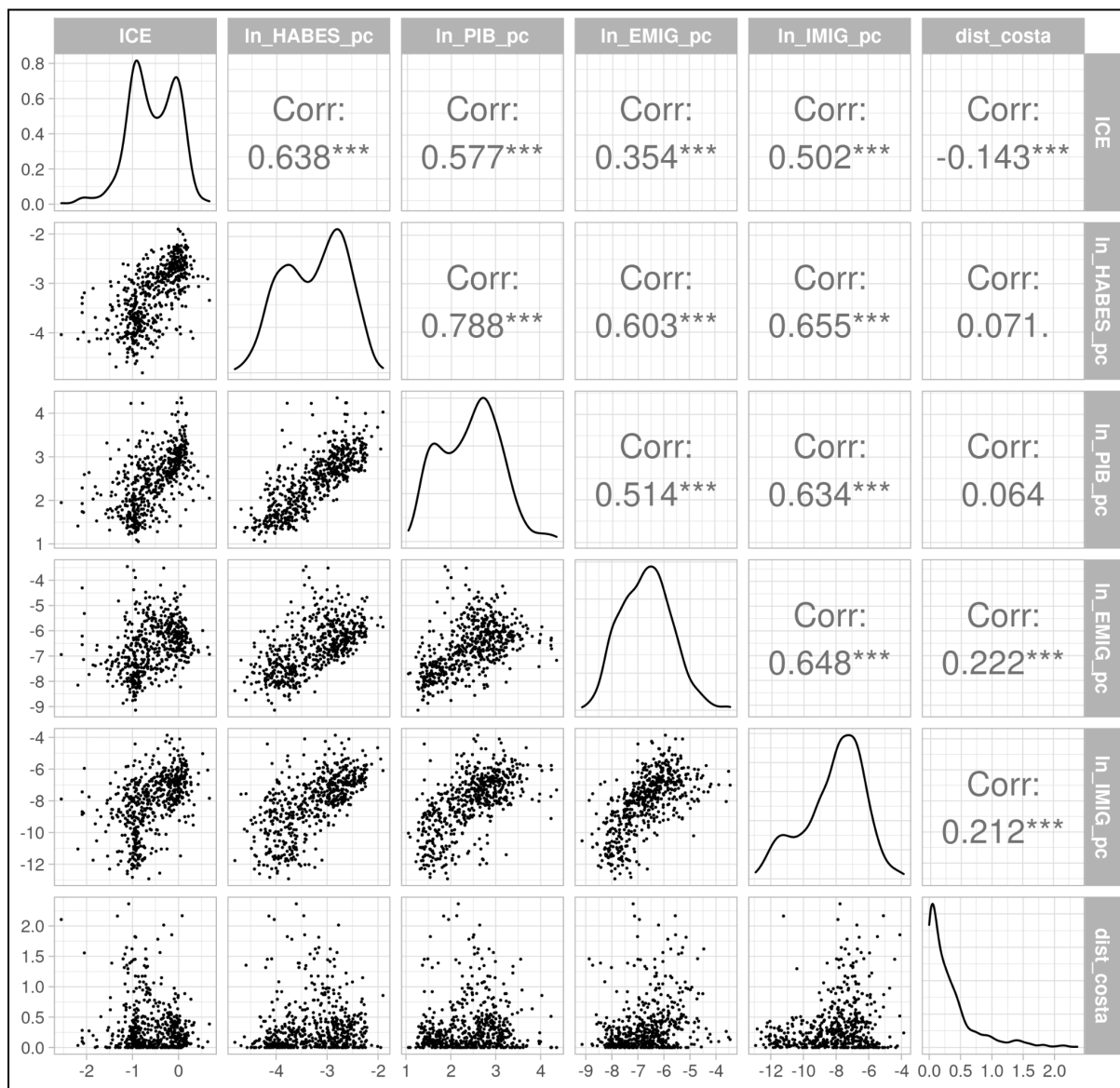
WOOLDRIDGE, Jeffrey M.. **Introductory Econometrics: A Modern Approach**, Cengage Learning, ed. 6, 2015.

ZHU, Shengjun; HE, Canfei; ZHOU, Yi. How to jump further and catch up? Path-breaking in an uneven industry space. **Journal Of Economic Geography**, [S.L.], p. 047, 23 jan. 2017. Oxford University Press. <http://dx.doi.org/10.1093/jeg/lbw047>.

ZHU, Shujin; LI, Renyu. Economic complexity, human capital and economic growth: empirical research based on cross-country panel data. **Applied Economics**, [S.L.], v. 49, n. 38, p. 3815-3828, 30 dez. 2016. Informa UK Limited. <http://dx.doi.org/10.1080/00036846.2016.1270413>.

APÊNDICES

Figura 6 - Matriz de correlação



Fonte: Elaborado pelo autor.

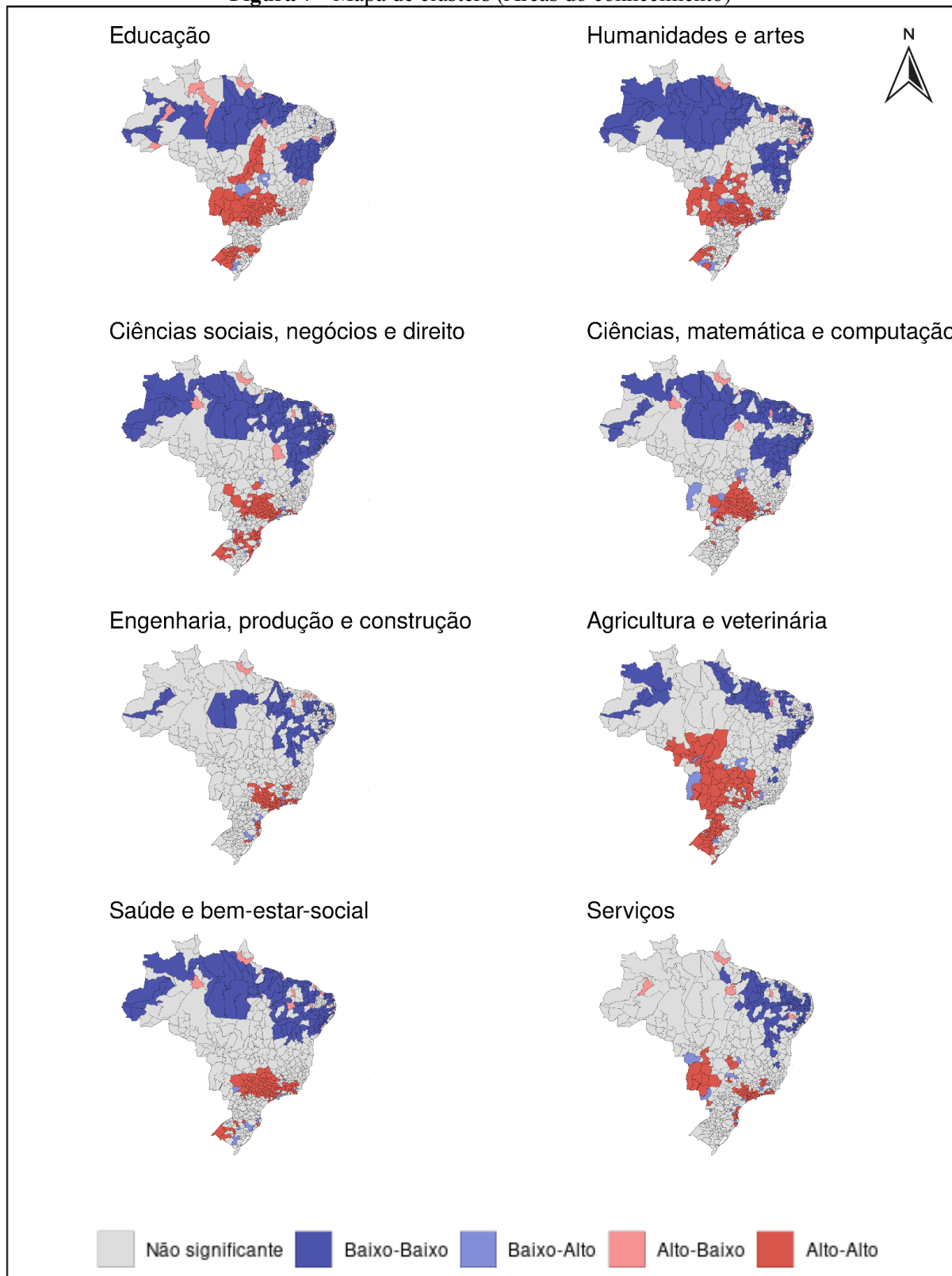
Tabela 5 - Regressões Globais de Controle

	ICE											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ln_HABES_pc	0.276*** (0.069)	0.305*** (0.055)	0.263*** (0.069)	0.376*** (0.052)	0.438*** (0.044)	0.416*** (0.043)	0.475*** (0.039)	0.451*** (0.040)			0.201*** (0.068)	0.217*** (0.067)
ln_HABEM_pc	0.093 (0.131)		0.283** (0.113)						0.710*** (0.060)		0.169 (0.131)	0.246* (0.126)
ln_HABEF_pc	0.500*** (0.179)	0.566*** (0.154)								1.296*** (0.124)	0.518*** (0.182)	0.530*** (0.182)
ln_PIB_pc	0.050 (0.044)	0.059 (0.042)	0.055 (0.044)	0.093** (0.042)							0.086* (0.044)	
ln_EMIG_pc	-0.064*** (0.023)	-0.063*** (0.023)	-0.058** (0.023)	-0.049** (0.023)	-0.054** (0.023)							
ln_IMIG_pc	0.039*** (0.014)	0.041*** (0.014)	0.039*** (0.014)	0.046*** (0.014)	0.051*** (0.014)	0.042*** (0.013)						
dist_costa	-0.227*** (0.056)	-0.231*** (0.056)	-0.239*** (0.057)	-0.263*** (0.056)	-0.283*** (0.056)	-0.282*** (0.056)	-0.287*** (0.056)					
rg_Nordeste	-0.002 (0.090)	0.019 (0.085)	-0.059 (0.088)	-0.005 (0.086)	-0.052 (0.084)	-0.011 (0.082)	-0.083 (0.080)	0.140** (0.068)	-0.012 (0.062)	0.057 (0.066)	0.182** (0.074)	0.139* (0.071)
rg_Norte	0.175** (0.076)	0.186** (0.075)	0.137* (0.075)	0.161** (0.075)	0.142* (0.075)	0.155** (0.075)	0.170** (0.076)	0.200*** (0.077)	0.097 (0.073)	0.115 (0.076)	0.225*** (0.077)	0.207*** (0.077)
rg_Sudeste	0.109 (0.074)	0.109 (0.074)	0.102 (0.075)	0.098 (0.075)	0.071 (0.074)	0.086 (0.074)	0.066 (0.074)	0.291*** (0.061)	0.291*** (0.061)	0.317*** (0.062)	0.298*** (0.060)	0.285*** (0.060)
rg_Sul	0.199** (0.078)	0.192** (0.077)	0.222*** (0.078)	0.207*** (0.078)	0.187** (0.078)	0.211*** (0.077)	0.201** (0.078)	0.423*** (0.066)	0.440*** (0.065)	0.378*** (0.068)	0.400*** (0.065)	0.397*** (0.066)
Constant	1.275*** (0.357)	1.318*** (0.351)	0.647** (0.278)	0.482* (0.272)	0.961*** (0.168)	1.148*** (0.148)	1.035*** (0.145)	0.681*** (0.130)	0.622*** (0.120)	1.851*** (0.246)	1.002*** (0.355)	1.457*** (0.268)
n	557	557	557	557	557	557	557	557	557	557	557	557
BP	34.96	29.19	36.19	28.08	27.44	25.55	24.62	20.55	13.45	19.37	29.67	29.09
F	51.33	56.47	55.01	59.85	66.22	74.34	83.57	90.91	95.08	85.46	64.09	72.34
R2	0.51	0.51	0.5	0.5	0.49	0.49	0.48	0.45	0.46	0.44	0.48	0.48
R2 adj.	0.5	0.5	0.49	0.49	0.48	0.48	0.47	0.45	0.46	0.43	0.48	0.47
AIC	468.71	467.22	474.6	478.92	481.99	485.35	493.82	517.67	506.23	532.98	490.89	492.73
BIC	524.9	519.09	526.47	526.47	525.22	524.26	528.4	547.92	536.49	563.24	534.11	531.63
RMSE	0.36	0.36	0.36	0.36	0.37	0.37	0.37	0.38	0.38	0.39	0.37	0.37
MAPE	2.21	2.2	2.31	2.32	2.47	2.49	2.48	2.57	2.42	2.5	2.21	2.33

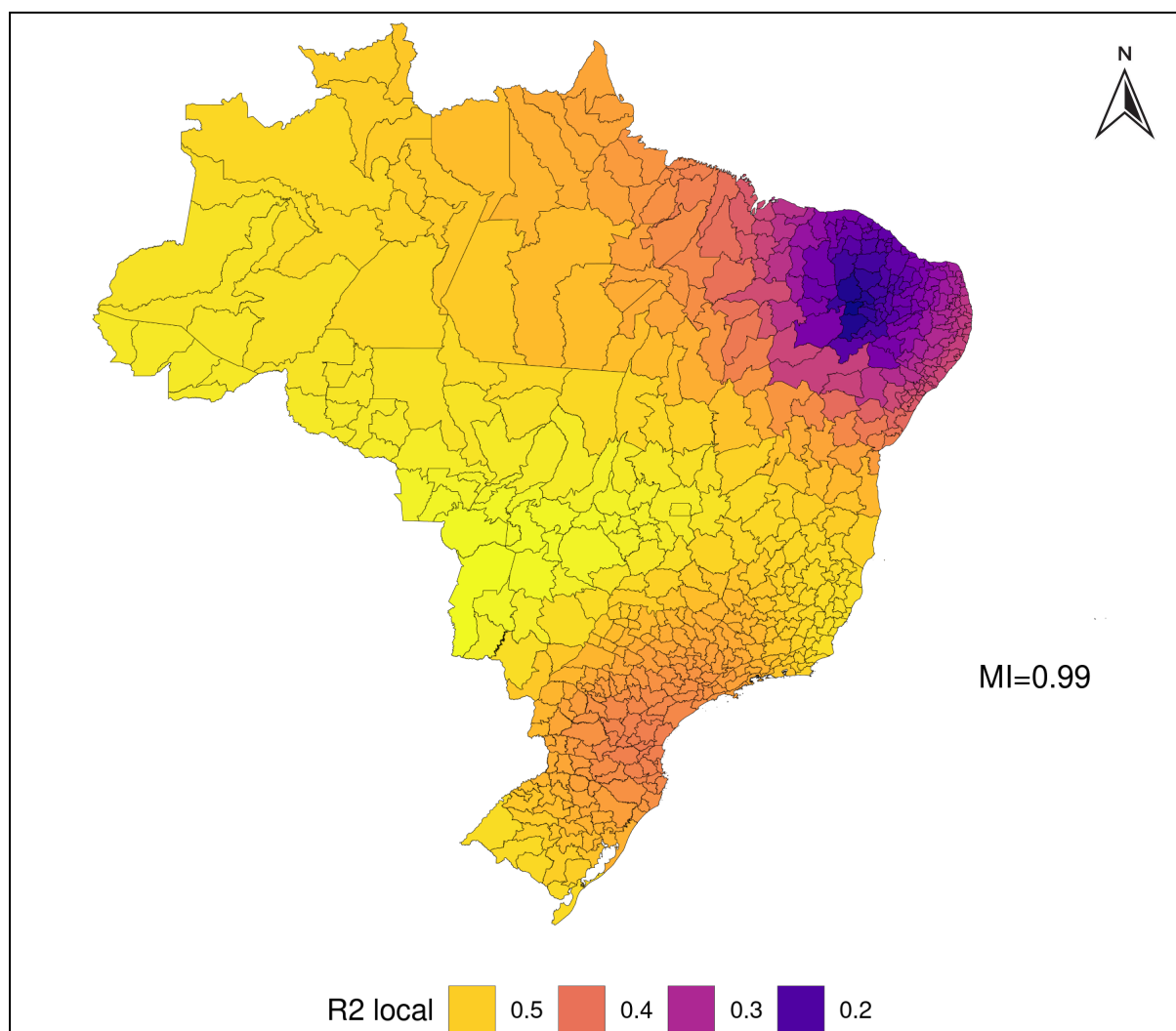
Note: * p<0.1; ** p<0.05; *** p<0.01

Fonte: Elaborado pelo autor.

Figura 7 - Mapa de clusters (Áreas do conhecimento)

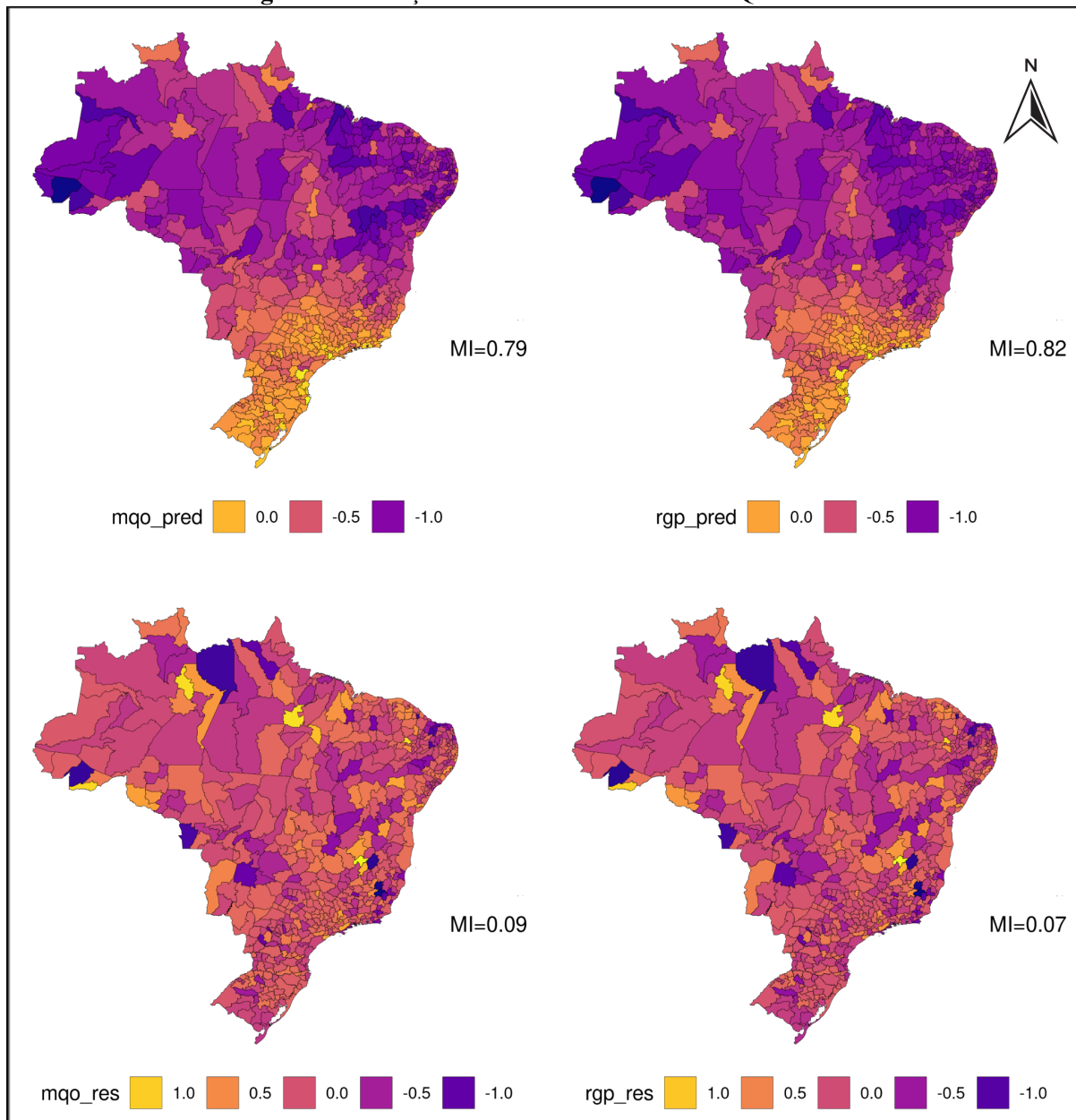


Fonte: Elaborado pelo autor.

Figura 8 - Coeficiente de determinação local

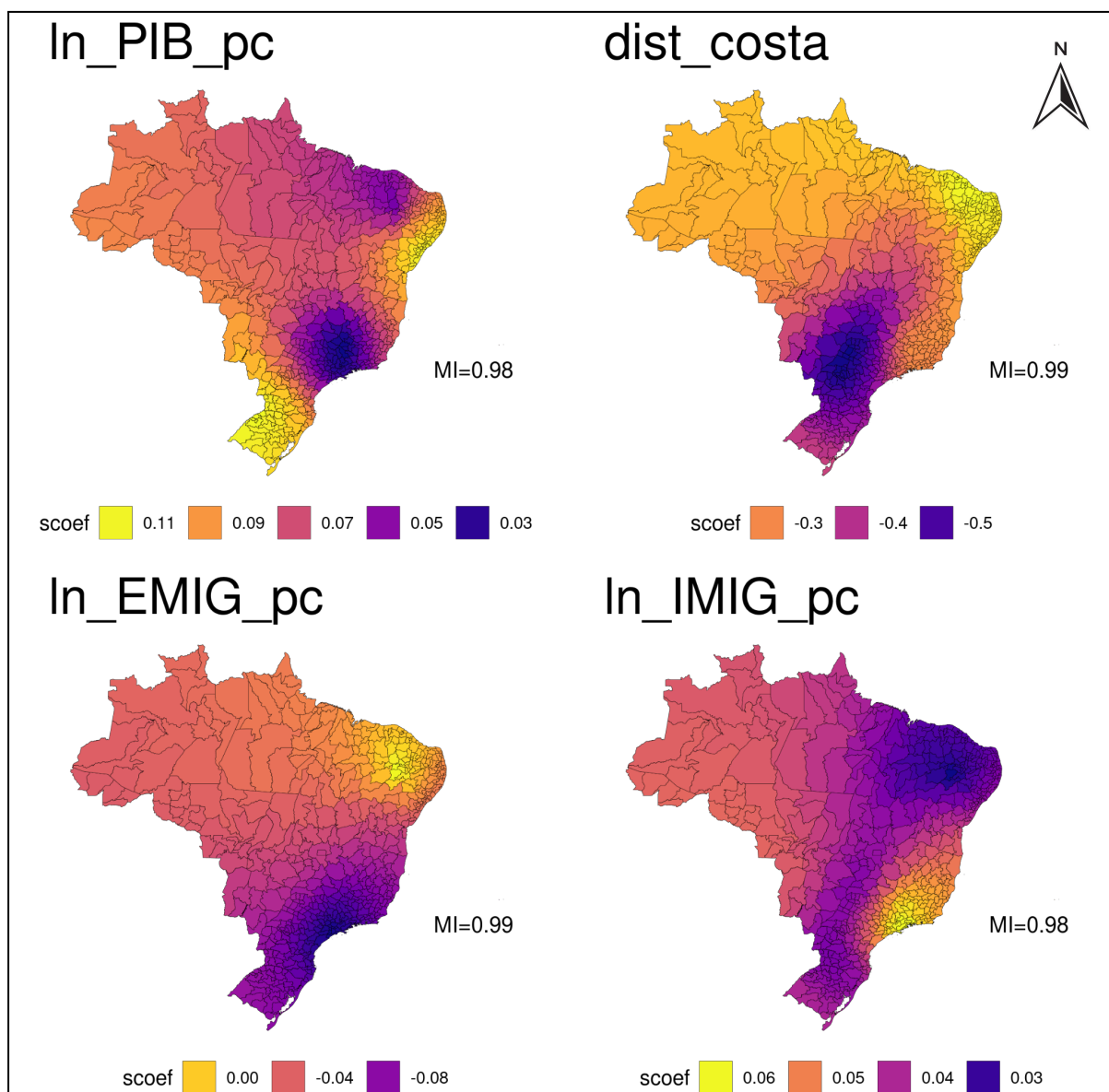
Fonte: Elaborado pelo autor.

Figura 9 - Predição e Resíduos dos modelos MQO e RGP

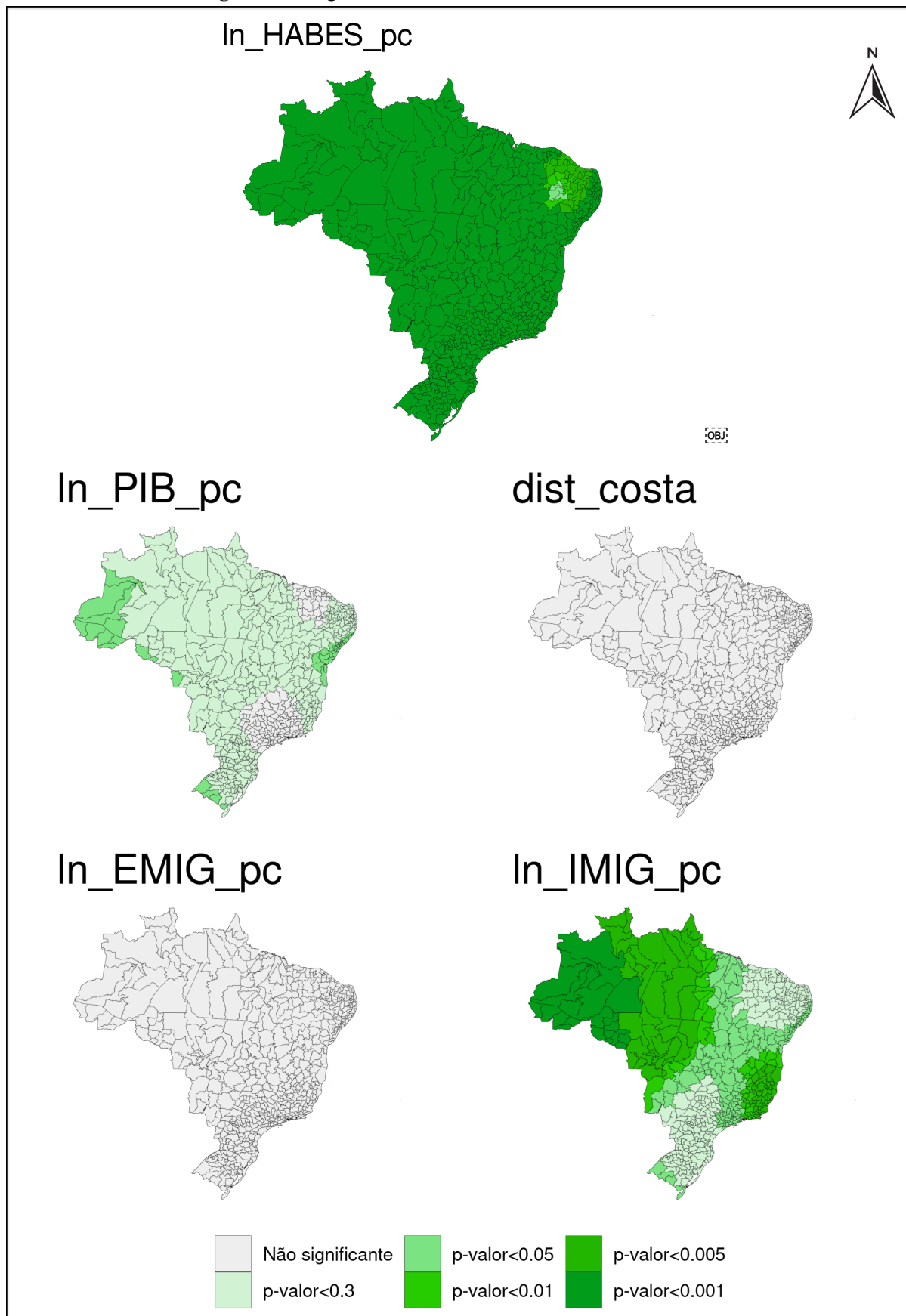


Fonte: Elaborado autor.

Figura 10 - Coeficientes causais locais das variáveis de controle



Fonte: Elaborado pelo autor.

Figura 11 - Significância estatística dos coeficientes locais

Fonte: Elaborado pelo autor.

