



UNIVERSIDADE FEDERAL DO RIO GRANDE
INSTITUTO DE CIÊNCIAS ECONÔMICAS, ADMINISTRATIVAS E CONTÁBEIS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA APLICADA
MESTRADO EM ECONOMIA

**SANEAMENTO BÁSICO E SAÚDE: UMA ANÁLISE ATRAVÉS DE PAINÉIS
ESPACIAIS COM OS MUNICÍPIOS BRASILEIROS**

KELEN ROSSALES DA SILVA

RIO GRANDE

2020

KELEN ROSSALES DA SILVA

**SANEAMENTO BÁSICO E SAÚDE: UMA ANÁLISE ATRAVÉS DE PAINÉIS
ESPACIAIS COM OS MUNICÍPIOS BRASILEIROS**

Dissertação apresentada ao Programa de Pós-Graduação em Economia Aplicada da Universidade Federal do Rio Grande – FURG, em cumprimento às exigências para obtenção do título de Mestre em Economia Aplicada.

Orientador: Prof. Dr. Pedro Henrique Soares Leivas

**RIO GRANDE
2020**

TERMO DE APROVAÇÃO

KELEN ROSSALES DA SILVA

SANEAMENTO BÁSICO E SAÚDE: UMA ANÁLISE ATRAVÉS DE PAINÉIS ESPACIAIS COM OS MUNICÍPIOS BRASILEIROS

Essa dissertação foi apresentada às 19:30 horas do dia 17 de julho de 2020 como requisito parcial para a obtenção do título de Mestre em Economia Aplicada do Programa de Pós-Graduação em Economia Aplicada da Universidade Federal do Rio Grande. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo citados. Após a deliberação, a Banca Examinadora considerou o trabalho aprovado.

Prof. Dr. Pedro Henrique Soares Leivas (Orientador)
PPGE - FURG

Prof. Dr. Vinícius Halmenschlager
PPGE - FURG

Prof. Dr. Anderson Moreira Arístides dos Santos
CMEA - UFAL

RIO GRANDE
2020

AGRADECIMENTOS

Agradeço, primeiramente a Deus, que me deu forças para que eu não desistisse.

Ao Guilherme, meu companheiro de todas as horas e maior incentivador durante essa caminhada. Que sempre esteve ao meu lado, me apoiando, me encorajando a seguir em frente.

Aos meus gatinhos Baphomet, Kali, Pantera Negrinha e Anakin que foram meus únicos companheiros nas longas madrugadas de estudo.

À minha mãe Eva, que mesmo não estando mais ao meu lado, sua lembrança foi meu maior incentivo.

Ao meu pai Edir, pelo apoio que sempre deu a minha educação.

Ao meu orientador Prof. Pedro, pela paciência comigo e por me ajudar a construir esse trabalho.

Aos meus colegas de curso e agora amigos Carolina, Douglas, Claudio, Carlos, Panquene, Semertsides e Iuri, minha trajetória teria sido muito mais árdua sem a presença deles.

Agradeço também, aos professores do PPGE Gibran, Cristiano, Vivian, Cassius, Rodrigo, Tiarajú, Vinicius e Rafael pelos ensinamentos durante todo esse período.

Aos demais amigos e familiares que direta ou indiretamente me auxiliaram nesta jornada.

RESUMO

A relação entre saneamento básico e saúde dos indivíduos está cada vez mais inserida nos estudos das ciências econômicas, tendo em vista a preocupação com o bem-estar social. Nesse sentido, esse estudo analisa as questões espaciais que envolvem os efeitos do acesso a condições de saneamento inadequadas sobre os índices de mortalidade infantil para os estados brasileiros nos anos de 1991, 2000 e 2010. Para isso, foram utilizados modelos de dados em painel espacial para a obtenção dos resultados, sendo utilizadas como variáveis dependentes a taxa de mortalidade infantil (0 a 1 ano) e a taxa de mortalidade na infância (0 a 5 anos). Os resultados indicam que, de maneira geral, o acesso aos serviços de saneamento inadequados impacta positivamente na taxa de mortalidade infantil e na taxa de mortalidade na infância. Da mesma forma, observa-se a presença de efeito transbordamento entre municípios vizinhos sob as taxas de mortalidade infantil no que se refere às condições inadequadas de saneamento.

Palavras-chave: Saneamento Básico; Mortalidade Infantil; Bem-Estar Social; Dados em Painel Espacial.

ABSTRACT

The relationship between basic sanitation and health of individuals is increasingly inserted in studies of the economic sciences, in view of the concern with social well-being. In this sense, this study analyzes the spatial issues that involve the effects of access to inadequate sanitation conditions on infant mortality rates for Brazilian states in the years 1991, 2000 and 2010. For that, spatial panel data models were used to obtain the results, using infant mortality rate (0 to 1 year) and childhood mortality rate (0 to 5 years) as dependent variables. The results indicate that, in general, inadequate access to sanitation services has a positive impact on infant mortality and mortality in childhood. Likewise, there is a spillover effect between neighboring municipalities on infant mortality rates with regard to inadequate sanitation conditions.

Keywords: Sanitation; Infant Mortality; Social Well-Being; Spatial Panel Data.

LISTA DE FIGURAS

| | |
|---|----|
| Figura 1 - Diagrama de Dispersão do <i>I</i> de Moran para a variável de Mortalidade Infantil..... | 22 |
| Figura 2 - Diagrama de Dispersão do <i>I</i> de Moran para a variável de Mortalidade na Infância. | 23 |
| Figura 3 - Diagrama de Dispersão do <i>I</i> de Moran para a variável Saneamento Inadequado... | 24 |
| Figura 4 - Mapas de <i>Clusters</i> para a Mortalidade Infantil..... | 25 |
| Figura 5 - Mapas de <i>Clusters</i> para a Mortalidade na Infância..... | 26 |
| Figura 6 - Mapas de <i>Clusters</i> para o Saneamento Inadequado..... | 27 |

LISTA DE TABELAS

| | |
|--|----|
| Tabela 1 - Descrição das Variáveis..... | 16 |
| Tabela 2 - Estatísticas Descritivas..... | 17 |
| Tabela 3 - Resultados dos modelos de Efeito Fixo e Aleatório, SDM e SAC para a Taxa de Mortalidade Infantil..... | 33 |
| Tabela 4 - Resultados dos modelos de Efeito Fixo e Aleatório, SDM e SAC para a Taxa de Mortalidade na Infância..... | 34 |
| Tabela 5 - Resultados dos dados em painel de Efeitos Fixos do modelo SAC..... | 36 |

LISTA DE SIGLAS

OMS - Organização Mundial da Saúde

SNIS - Sistema Nacional de Informações sobre Saneamento

LDNSB - Lei de Diretrizes Nacionais para o Saneamento Básico

PLANSAB - Plano Nacional de Saneamento

PNAD - Pesquisa Nacional por Amostra de Domicílios

DATASUS - Departamento de Informática do Sistema Único de Saúde

CNEFE - Cadastro Nacional de Endereços para Fins Estatísticos

DNE - Diretório Nacional de Endereços

DRSAI - Doenças Relacionadas ao Saneamento Ambiental Inadequado

PNUD - Programa das Nações Unidas Para o Desenvolvimento

IPEA - Instituto de Pesquisas Econômicas Aplicadas

FJP - Fundação João Pinheiro

IDHM - Índices de Desenvolvimento Humano Municipal

IBGE - Instituto Brasileiro de Geografia e Estatística

AEDE - Análise Explanatória de Dados Espaciais

LISA - Indicador Local de Associação Espacial

SAR - Modelo de Defasagem Espacial

SEM - Modelo de Erro Espacial

SDM - Modelo de Durbin Espacial

SAC - Modelo de Autocorrelação Espacial

AIC - Critério de Informação de Akaike

BIC - Critério Bayesiano de Schwarz

MFV - Máximo da Função de Verossimilhança

POLS - Mínimos Quadrados Empilhados

EF - Efeito Fixo

EA - Efeito Aleatório

SUMÁRIO

| | |
|--|-----------|
| 1 INTRODUÇÃO..... | 9 |
| 2 REVISÃO DE LITERATURA..... | 11 |
| 3 METODOLOGIA..... | 15 |
| 3.1 Dados..... | 15 |
| 3.2 Econometria Espacial..... | 17 |
| 3.2.1 Autocorrelação espacial e heterogeneidade espacial..... | 18 |
| 3.2.2 Matriz de Pesos Espaciais..... | 19 |
| 3.2.3 Análise Explanatória de Dados Espaciais..... | 20 |
| 3.2.3.1 Autocorrelação Global..... | 20 |
| 3.2.3.2 Diagrama de Dispersão de Moran..... | 21 |
| 3.2.3.3 Autocorrelação Espacial Local (Local Indicator of Spatial Association – LISA)..... | 24 |
| 3.3 Estratégia Empírica..... | 27 |
| 3.3.1 Modelo de Defasagem Espacial (Spatial Autoregressive Model - SAR)..... | 28 |
| 3.3.2 Modelo de Erro Espacial (Spatial Error Model - SEM)..... | 29 |
| 3.3.3 Modelo de Durbin Espacial (Spatial Durbin Model – SDM)..... | 29 |
| 3.3.4 Modelo de Autocorrelação Espacial (Spatial Autocorrelation Model – SAC)..... | 30 |
| 4 RESULTADOS..... | 32 |
| 5 CONSIDERAÇÕES FINAIS..... | 38 |
| REFERÊNCIAS..... | 40 |

1. INTRODUÇÃO

Problemas relacionados à saúde da população estão, dentre outros fatores, intrinsecamente ligados às condições de saneamento básico. Isso porque, a ausência de saneamento adequado gera a proliferação de doenças, ocasionando externalidades negativas para a sociedade e afetando o bem-estar social. A má qualidade da água, o destino inadequado do lixo, a deposição incorreta de dejetos e ambientes poluídos são algumas das consequências da falta de saneamento. Desse modo, a ausência ou deficiência nas condições de saneamento básico propiciam a transmissão de bactérias, vírus e parasitas, que estão presente nas fezes, urina ou vômito do doente ou portador, causadores de diversas doenças infectocontagiosas (SCRIPTORE e AZZONI, 2018). Os mais afetados são crianças e jovens que sofrem com o comprometimento do seu desenvolvimento físico e intelectual devido a doenças como diarreia e desnutrição (LEIVAS *et al.*, 2016; MENDONÇA e SEROA DA MOTTA, 2007). Para Scriptore e Azzoni (2018), devido aos seus efeitos sobre a taxa de mortalidade infantil, a diarreia tornou-se a doença mais conhecida causada pelas condições precárias de saneamento. Mas existem outras que também causam danos à população, em específico às crianças, tais como esquistossomose, febre amarela, malária, leptospirose, febre tifoide e cólera.

A plataforma do Instituto Trata Brasil apresenta algumas estatísticas a respeito do acesso à água e ao saneamento adequado para o Brasil e para o restante dos países. Dentre essas estatísticas, estão os dados divulgados em 2017 pela Organização Mundial da Saúde (OMS), no qual mais de 1,5 milhão de crianças com menos cinco anos morrem por ano no mundo por problemas relacionados ao fornecimento inadequado da água, sendo que a diarreia é a segunda causa de morte de crianças até cinco anos. Já no Brasil, entre os dados apresentados estão os divulgados pelo do Sistema Nacional de Informações sobre Saneamento (SNIS) também para o ano de 2017 em que revela que apenas cerca de 53% da população nacional tem acesso a uma rede de coleta de esgoto adequada, e desse total apenas 43% é tratado. No que se refere ao acesso a água, quase 15% das crianças no país não têm acesso, além disso cerca de 13 milhões de crianças e adolescentes não têm acesso a saneamento básico no Brasil. Dessa forma, os serviços de água tratada, coleta e tratamento dos esgotos levam à melhoria da qualidade de vida das pessoas, sobretudo na saúde infantil, com redução da mortalidade infantil, melhor desempenho escolar e acúmulo de capital humano.

No Brasil a redução das taxas de mortalidade infantil, experimentada nas últimas décadas, é o principal componente na melhora dos índices que medem a expectativa de vida do

brasileiro. Contudo, a mortalidade infantil ainda é um desafio para o Brasil, tendo em vista que há fatores socioeconômicos que variam de acordo com as diferentes regiões do país. Apesar da trajetória de queda nos últimos anos, a mortalidade infantil permanece sendo um dos grandes problemas de saúde pública. Nesse sentido, há pontos persistentes a serem superados, tais como desigualdades regionais e interurbanas, além da concentração dos óbitos nas populações mais pobres. Ações efetivas no combate à mortalidade infantil estão ligadas à elaboração de políticas públicas que tratam de saneamento adequado, educação e serviços de saúde. Analisar os fatores socioeconômicos que determinam a redução das taxas de mortalidade infantil torna-se importante na medida em que orienta em quais áreas o governo deve investir, a fim de prevenir óbitos evitáveis no sistema de saúde e garantir um nível maior de bem-estar à população em geral.

Desse modo, o objetivo do presente trabalho é investigar como as condições de saneamento inadequado afetam a saúde, sobretudo de crianças e recém-nascidos. Para tanto utiliza-se como *proxy* de saúde a variável de mortalidade infantil para crianças até um ano de idade e a variável mortalidade na infância para crianças até cinco anos de idade. São analisados dados para todos os municípios do Brasil dos anos de 1991, 2000 e 2010, disponíveis no Atlas Brasil (2013). Para atingir os objetivos foram estimados modelos de dados em painel espacial, a fim de captar o efeito do saneamento inadequado sobre a saúde controlando prováveis efeitos espaciais de transbordamento entre municípios vizinhos. Para Sabater *et al* (2011) os dados espaciais podem ser definidos como sendo a observação de uma variável que se encontra associada a uma localização no espaço geográfico. Nesse sentido, o efeito de transbordamento entre localidades em decorrência da existência de saneamento inadequado pode ocorrer através do deslocamento dos insetos de um distrito pra outro, bem como pela água contaminada dos rios e mares que atravessam várias localidades (SCRIPTORE e AZZONI, 2018).

A contribuição deste trabalho em relação aos anteriores realizados no Brasil é, primeiramente, desenvolver uma análise completa para todos os estados brasileiros; em segundo lugar, além de cobrir um longo período, de 1991 a 2010, é utilizar dados mais atuais e completos disponíveis a partir dos Censos, em terceiro, controlar efeitos espaciais entre os municípios através da dependência espacial e os possíveis transbordamentos entre vizinhos; quarto, e mais relevante, identificar os efeitos diretos e indiretos das variáveis.

Este trabalho está dividido em cinco seções, incluindo esta breve introdução. A próxima seção apresenta a revisão de literatura, que faz um apanhado sobre os trabalhos que tratam da relação entre saneamento e saúde. Na terceira seção descrevemos a metodologia utilizada nas

regressões e a base de dados e seus tratamentos. A seguir são apresentados os resultados e a análise das estimações. Por fim, as considerações finais do estudo.

2. REVISÃO DE LITERATURA

No Brasil o saneamento é um direito previsto na Constituição Federal (1988) e definido pela Lei de Diretrizes Nacionais para o Saneamento Básico (LDNSB) n.º 11.445/2007 como o conjunto de serviços, infraestrutura e instalações operacionais de abastecimento de água, esgotamento sanitário, limpeza e drenagem urbana, manejos de resíduos sólidos e de águas pluviais. Embora esta seja a descrição correta a respeito do saneamento, a definição mais comumente usada é aquela que define saneamento como os serviços de acesso à água potável, bem como à coleta e ao tratamento de esgoto. Nesse sentido, no presente trabalho utilizamos o conceito de saneamento inadequado, no qual representa a proporção de pessoas que vivem em domicílios cujo o abastecimento de água não provém de rede geral e cujo o esgotamento sanitário não é realizado por rede coletora de esgoto ou fossa séptica. No âmbito desse trabalho não serão tratadas questões relacionadas à qualidade dos serviços de saneamento, considerando o acesso à água da rede geral como apta para o consumo e sem diferenciar se o esgoto é tratado ou não. Destaca-se ainda que a LDNSB estabelece que coube a União elaborar e coordenar o atual Plano Nacional de Saneamento (PLANSAB) que visa, dentre outras coisas, estabelecer os procedimentos quanto à execução e às metas de curto, médio e longo prazo para os serviços de saneamento básico para o Brasil entre os anos de 2014 e 2033 (UHR et. al., 2016).

Os avanços nas condições de saneamento básico na história recente, com acesso à água encanada e ao esgotamento sanitário, permitiram ganhos na qualidade de vida das populações. Nesse sentido, a associação entre saneamento básico e saúde é natural, tendo em vista que precárias condições de saneamento acarretam doenças que geram custos ao poder público, afetam a produtividade do país e em última análise podem levar à morte. Dessa forma, conforme apontam Sousa e Leite Filho (2008), a taxa de mortalidade infantil representa um bom indicador da qualidade de vida e do status de saúde da população, por estimar o número de crianças que sobreviverão ao seu primeiro ano de vida. Isso porque, a população infantil é a mais suscetível às doenças e, conseqüentemente, à mortandade em precárias condições de saneamento básico. A taxa de mortalidade infantil é definida como o número de óbitos de menores de um ano de vida a cada mil nascidos vivos em determinada região e período e estima o risco de um recém-nascido morrer durante seu primeiro ano de vida.

Na literatura econômica diversos são os estudos teóricos e empíricos que relacionaram saneamento básico e saúde. Num contexto internacional, Esrey *et al.* (1985) descrevem que o acesso à água encanada, juntamente com o crescimento dos padrões de vida, teve papel importante na redução da ocorrência de diarreia e no controle de doenças como febre tifoide e cólera na Europa e nos Estados Unidos durante 1860 e 1920. Diante disso, os autores analisaram dezenas de estudos que avaliaram como a melhoria na oferta de água e de esgoto impacta na redução da taxa de diarreia em crianças de países em desenvolvimento. O estudo encontrou que a melhoria na disponibilidade da água é o fator de ordem sanitária de maior contribuição para redução da mortalidade infantil, com cerca de 25%, seguido por melhorias no tratamento de esgoto que reduziria, em média, 22% da mortalidade infantil. Outro aspecto apontado pelos autores que reduziria a incidência de diarreia refere-se ao nível de alfabetização da população, utilizada no estudo como *proxy* para o nível de higiene, sobretudo em relação ao hábito de lavar as mãos.

Seguindo no contexto dos Estados Unidos, Cutler e Miller (2005) evidenciaram que a introdução da tecnologia de água encanada, iniciada por volta de 1900, reduziu a taxa de mortalidade infantil em torno de 46%. Já segundo Watson (2006), ao avaliar os efeitos de programas sanitários nos Estados Unidos na década de 1960 em uma comunidade indígena, demonstrou que um incremento de 10 pontos percentuais nos serviços de saneamento acarretou a uma redução da taxa de mortalidade infantil de cerca de 2,5%. Além disso, o estudo identificou que intervenções de saneamento explicaram quase 40% da convergência das taxas de mortalidade infantil em tais reservas desde 1970.

No que tange aos países em desenvolvimento, Jalan e Ravallion (2003) avaliaram o impacto de fatores ambientais, em especial o acesso a água encanada, na morbidade por diarreia em áreas rurais da Índia. Utilizando microdados da população para os anos 1993 e 1994, os autores descobriram que a prevalência e a duração da diarreia entre as crianças menores de cinco anos são significativamente mais baixas, em média, para as famílias com água encanada do que para famílias similares, mas sem água encanada na residência. Já Teixeira *et al.* (2012) buscaram verificar a relação entre a cobertura dos serviços de saneamento básico e indicadores de doenças epidemiológicas para os países da América Latina no período de 2005 a 2010. Utilizando um modelo de regressão linear o estudo verificou que, uma maior cobertura populacional por serviços de saneamento adequado tende a reduzir a mortalidade infantil nos países da América Latina.

No contexto brasileiro, o estudo de Merrick (1983) foi o primeiro a fazer uma análise para o Brasil, utilizando dados do Censo de 1970 e da Pesquisa Nacional por Amostra de Domicílios (PNAD) para o ano de 1976. Utilizando um modelo de equações simultâneas o autor evidenciou a relação positiva entre acesso à água e a redução da taxa de mortalidade infantil.

Alves e Belluzzo (2004), em um estudo mais amplo, analisam as condições de saúde das crianças e as principais causas da mortalidade infantil no Brasil. Os autores utilizaram dados municipais para estimarem um modelo de dados em painel estático e outro dinâmico para o período 1970 a 2000. Constataram que a mortalidade infantil no Brasil diminuiu nesse período, mas que essa taxa continua alta quando comparada a outros países, demonstraram a importância das condições adequadas de saneamento básico, tais como a conexão da casa aos serviços de água encanada tratada e esgotamento sanitário. Dentre os resultados encontrados, os autores apontam que educação, pobreza e saneamento são os determinantes mais importantes da mortalidade infantil.

Sousa e Leite Filho (2008), que analisaram o impacto do saneamento sobre a mortalidade infantil na região Nordeste, concluíram, por meio de metodologia de dados em painel com informações em nível municipal, que a mortalidade de menores de um ano de idade nesta região foi reduzida em 31,8% no período de 1991 a 2000 para os Estados que apresentaram maiores ganhos no acesso à água tratada.

Mendonça e Seroa da Motta (2007) buscaram relacionar, por meio do uso da metodologia de dados em painel com efeito fixo, a incidência da mortalidade infantil, associada a doenças de veiculação hídrica, com indicadores de saneamento e outras variáveis, tais como educação e saúde, para todos os estados do Brasil no período de 1981 a 2001. São utilizadas como variáveis de saneamento os percentuais da população com acesso adequado à água e esgotamento sanitário. Os resultados encontrados apontam que, tanto o acesso aos serviços de saneamento como aos serviços de educação e saúde afetam negativamente a mortalidade infantil.

Barufi *et al.* (2012) analisam o padrão regional da mortalidade infantil no Brasil, avaliando o efeito da infraestrutura de saúde e das variáveis socioeconômicas e demográficas para compreender a distribuição em todo o país. A partir dos dados dos Censos Demográficos para os anos de 1980, 1991 e 2000 é estimado um modelo de dados em painel espacial com efeitos fixos como forma de captar a potencial influência em unidades espaciais vizinhas. Os resultados indicam que a redução na taxa de mortalidade infantil nas duas décadas analisadas está relacionada às características socioeconômicas e a existência de infraestrutura de cuidados

de saúde, água e saneamento. Além disso, o estudo também concluiu que as reduções na taxa de mortalidade infantil ao longo do período analisado tendiam a ser menores em locais que tinham níveis iniciais de infraestrutura e níveis mais baixos de analfabetismo.

O estudo de Leivas *et al.* (2016), busca analisar os impactos do saneamento adequado sobre a saúde infantil no Brasil. Para tanto, os autores coletaram dados das fontes DATASUS, IPEADATA e PNAD para os anos de 2001 a 2010. Como metodologia utilizaram painéis estáticos com efeitos fixos e painel dinâmico para analisar os macrodados, já para a análise de microdados foi estimado um Probit binomial. Dentre os resultados encontrados, os autores verificaram que a presença de saneamento é um importante determinante para a saúde infantil e a redução da mortalidade. Outros resultados obtidos no estudo apontam para a importância da renda e da escolaridade das mães na redução da mortalidade infantil.

O estudo de Uhr *et al.* (2016) avalia como os serviços de saneamento básico no Brasil afetam a saúde da população, tomando como variável dependente as internações hospitalares decorrentes de endemias de veiculação hídrica. A pesquisa traz como principal contribuição uma análise mais ampla das endemias, incluindo dados para internações hospitalares provocadas por ingestão de água contaminada, contato da pele com água contaminada e esgotamento sanitário inadequado. São estimados modelos econométricos utilizando dados em painel para os estados brasileiros, para o período de 2000 a 2011. Os resultados indicam que quanto maior a abrangência dos serviços de saneamento, principalmente da rede coletora de esgotos, maior a redução na morbidade.

Scriptore e Azzoni (2018) fizeram a relação entre saneamento e saúde a partir de uma análise espacial, utilizando dados do Censo Demográfico de 2010, Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE) de 2011, Diretório Nacional de Endereços (DNE) de 2015 e Departamento de Informática do Sistema Único de Saúde (DATASUS) de 2010. A variável de saúde considerada nesse estudo é a taxa de incidência de internações por Doenças Relacionadas ao Saneamento Ambiental Inadequado (DRSAI), os resultados indicaram que o acesso aos serviços de saneamento inadequados causa impacto positivo na taxa de incidência de internações por DRSAI, por mil habitantes. Fazem uma análise dos impactos da falta de saneamento nos indicadores de mortalidade e morbidade para duas faixas de idade, todas as idades e idades de 0-14 anos. Os resultados indicam que os efeitos nocivos da falta de saneamento são mais significativamente sentidos pela população de 0-14 anos. Verificaram também que há um efeito transbordamento na taxa de incidência de internações por doenças relacionadas a saneamento, podendo ser causado tanto pelo deslocamento dos insetos de um

distrito pra outro, quanto pela água contaminada dos rios e mares que atravessam várias localidades.

3. METODOLOGIA

Nesta seção será apresentada a metodologia utilizada no estudo. Primeiramente destaca-se a base de dados e suas fontes, em seguida será abordado os conceitos sobre a econometria espacial e a estratégia empírica utilizada.

3.1 Dados

Para realizar a análise proposta neste estudo, foram utilizados os dados presentes na plataforma Atlas do Desenvolvimento Humano no Brasil (2013), ou simplesmente Atlas Brasil¹. Nele é possível captar informações socioeconômicas de todos os municípios do país e do Distrito Federal. Trata-se de um trabalho elaborado pelo PNUD (Programa das Nações Unidas Para o Desenvolvimento), pelo IPEA (Instituto de Pesquisas Econômicas Aplicadas), e pela Fundação João Pinheiro (FJP). As informações que compõem o Atlas Brasil (2013) são extraídas a partir dos dados dos Censos Demográficos do IBGE de 1991, 2000 e 2010.

A base de dados do Atlas Brasil (2013) foi construída a partir de um processo de compatibilização do Censo 2010 com os censos de 2000 e 1991. Esse cálculo foi realizado pelos idealizadores do Atlas Brasil (2013) para efeito de comparabilidade intertemporal, tendo compatibilizado as malhas municipais existentes em 1991 e 2000 com a existente em 2010. De acordo com os censos do IBGE, em 1991 o país tinha 4.491 municípios, em 2000 tinha 5.507 e em 2010 tinha 5.565 municípios. Durante duas décadas houve divisões político-administrativas dando origem a novos municípios, e, embora esses municípios não existissem do ponto de vista administrativo nos anos anteriores, seu espaço geográfico correspondia a parcelas de um ou mais municípios no passado. Assim sendo, esse espaço geográfico dos municípios de origem foi considerado separadamente com o intuito de estabelecer uma comparação temporal e espacial na construção dos indicadores do Atlas Brasil (2013).

¹O Atlas Brasil (2013) é uma plataforma online utilizada para divulgar os Índices de Desenvolvimento Humano Municipal (IDHM) dos 5.565 municípios brasileiros existentes em 2010, mais o Distrito Federal. Além do IDHM, ainda são apresentados cerca de 200 indicadores socioeconômicos que viabilizam avaliar a evolução do desenvolvimento dos municípios num período de duas décadas. Para tanto, os indicadores correspondem a oito dimensões: saúde, educação, renda, demografia, trabalho, população, habitação e vulnerabilidade social.

Com base na literatura, as variáveis explicativas foram escolhidas por serem consideradas as mais adequadas para servirem como variáveis de controle para fatores observáveis que afetam a saúde infantil. As variáveis buscam contemplar características dos municípios no que tange as condições das crianças, dos pais e do domicílio. Nesse sentido, destacam-se as variáveis de renda e de analfabetismo como características da família. Já para o domicílio utiliza-se variáveis como saneamento inadequado, coleta de lixo, energia elétrica, densidade domiciliar. Além dessas variáveis controles, no presente trabalho utilizamos efeito fixo de tempo. Cabe sempre ressaltar que a variável de maior interesse para esta análise é a variável de saneamento, já que queremos ter como resultado o efeito do saneamento inadequado nos índices de mortalidade infantil e mortalidade na infância. A tabela 1 a seguir apresenta a descrição das variáveis que foram utilizadas para a elaboração deste estudo.

Tabela 1 – Descrição das Variáveis

| NOME DA VARIÁVEL | DESCRIÇÃO |
|---------------------------------|--|
| Variáveis Dependentes | |
| Taxa de Mortalidade Infantil | Número de crianças que não deverão sobreviver ao primeiro ano de vida em cada 1000 crianças nascidas vivas. |
| Mortalidade na Infância | Probabilidade de morrer entre o nascimento e a idade exata de cinco anos, por 1000 crianças nascidas vivas. |
| Variáveis Explicativas | |
| Saneamento inadequado | Proporção de pessoas que vivem em domicílios cujo abastecimento de água não provém de rede geral e cujo o esgotamento sanitário não é realizado por rede coletora de esgoto ou fossa séptica multiplicado por 100. |
| Renda <i>per capita</i> média | Razão entre o somatório da renda de todos os indivíduos do domicílio e o número total dos indivíduos. |
| Taxa de analfabetismo > 15 anos | Proporção da população de 15 anos ou mais de idade que não sabe ler nem escrever multiplicado por 100. |
| Densidade do domicilio > 2 | Percentual da população que vive em domicílios com densidade superior a 2 pessoas por dormitório. |
| Coleta de lixo | Percentual da população que vive em domicílios urbanos com coleta de lixo. |
| Energia elétrica | Percentual da população que vive em domicílios com energia elétrica. |
| População urbana | Razão entre a população que vive em área urbana e o total da população multiplicada por 100. |
| População total | População total do município. |

Fonte: Elaboração própria a partir dos dados do Atlas (2013).

A seguir, na tabela 2 são apresentadas as estatísticas descritivas dos dados utilizados para a obtenção dos resultados a serem analisados na próxima seção. Podemos perceber pelas estatísticas descritivas que houveram algumas alterações nas condições de vida dos brasileiros ao longo do tempo. Nesse sentido, tanto a mortalidade infantil como a mortalidade na infância tiveram uma queda acentuada nos seus índices dentro do período analisado. A taxa de mortalidade infantil caiu de quase 48 crianças a cada mil que não completaram um ano em 1991

para pouco mais de 19 em 2010. Da mesma forma a taxa de mortalidade na infância, que quase atingiu 60 crianças a cada mil que não completaram cinco anos em 1991, foi reduzida para 21 crianças em 2010.

Tabela 2. Estatísticas Descritivas

| | Média | | | Desvio Padrão | | | Mínimo | | | Máximo | | |
|-------------------------------|--------|--------|--------|---------------|-------|-------|--------|-------|-------|---------|---------|---------|
| | 1991 | 2000 | 2010 | 1991 | 2000 | 2010 | 1991 | 2000 | 2010 | 1991 | 2000 | 2010 |
| Mortalidade Infantil | 47,88 | 32,67 | 19,25 | 23,64 | 14,35 | 7,14 | 14,6 | 10,78 | 8,49 | 120,12 | 96,37 | 46,8 |
| Mortalidade na Infância | 59,90 | 39,28 | 21,53 | 31,62 | 18,71 | 7,32 | 17,12 | 12,51 | 9,98 | 151,6 | 106,29 | 50,94 |
| Saneamento inadequado | 17,24 | 13,47 | 9,20 | 24,21 | 15,60 | 12,84 | 0,00 | 9,08 | 14,64 | 100 | 90,93 | 85,36 |
| Renda <i>per capita</i> média | 234,86 | 338,58 | 493,65 | 143,6 | 192,4 | 243,2 | 33,24 | 62,65 | 96,25 | 1185,28 | 1759,76 | 2043,74 |
| Taxa de analfabetismo | 31,08 | 21,75 | 16,16 | 16,92 | 12,47 | 9,84 | 1,81 | 0,91 | 0,95 | 88,39 | 59,95 | 44,4 |
| Densidade domiciliar | 51,18 | 38,13 | 25,12 | 14,93 | 14,92 | 13,00 | 3,26 | 3,21 | 0,65 | 97,21 | 94,47 | 88,64 |
| Coleta de lixo | 50,63 | 79,21 | 94,05 | 34,71 | 25,57 | 11,04 | 0,00 | 0,00 | 0,00 | 100 | 100 | 100 |
| Energia elétrica | 69,58 | 86,61 | 97,19 | 26,06 | 17,03 | 6,02 | 0,00 | 10,3 | 27,41 | 100 | 100 | 100 |
| População urbana | 48,50 | 58,50 | 63,83 | 26,80 | 23,65 | 22,04 | 0,00 | 0,00 | 4,18 | 100 | 100 | 100 |

Fonte: Elaboração própria a partir dos dados do Atlas (2013).

Do mesmo modo, podemos notar também que as condições de saneamento e a renda *per capita* média dos brasileiros também melhorou entre o início e o fim do período analisado. Isso porque a proporção de pessoas que viviam em domicílios com condições de saneamento inadequado diminuiu para menos de 10% em 2010. Já a renda *per capita* média dos domicílios chegou a R\$ 493,65, pouco abaixo dos R\$ 510,00 de salário mínimo vigente em 2010. A taxa de analfabetismo caiu aproximadamente pela metade durante o período, passando de cerca de 31% para pouco mais de 16%. No que tange a coleta de lixo e a energia elétrica, os domicílios brasileiros também apresentaram melhora nesses índices, indicando ampliação ao acesso a serviços básicos. De modo geral, os dados demonstram uma significativa melhora nas condições de vida dos brasileiros durante o período analisado.

3.2 Econometria Espacial

Em pesquisas que tratam sobre economia regional é comum serem observados efeitos de interdependência entre as regiões, isso porque as unidades analisadas passam a ter como uma característica importante a sua localização geográfica. Segundo Vieira (2009), a existência de uma área da econometria denominada de econometria espacial se justifica basicamente por dois aspectos principais: o primeiro é a importância da questão espacial inerente à ciência regional,

em particular, à economia regional. O segundo é que dados distribuídos no espaço podem apresentar dependência ou heterogeneidade em sua estrutura. Os dados espaciais podem ser definidos como sendo a observação de uma variável que se encontra associada a uma localização no espaço geográfico (SABATER *et al*, 2011).

Desse modo, são recorrentes os efeitos encontrados de regiões vizinhas sobre as variáveis de interesse em estudos regionais. Além disso, essa relação tende a se intensificar quanto mais próxima forem as regiões analisadas. Portanto, estudos nos quais utilizem modelos econométricos que não levem em consideração os efeitos da dependência espacial podem vir a produzir estimações viesadas, pois conforme Almeida (2012), a presença de dependência espacial entre as observações, ou heterogeneidade espacial nas relações modeladas ferem os pressupostos básicos de Gauss-Markov.

Nesse sentido, o instrumental da econometria espacial contribui por ser uma alternativa aos modelos tradicionais de estimação, pois não ignora os efeitos da dependência espacial, mais especificamente, a autocorrelação e a heterogeneidade espacial. Para além, essa metodologia permite análises mais robustas sobre questões de desigualdades regionais e possibilita verificar com confiabilidade como um país se desenvolve regionalmente. Esses ganhos advindos com o arcabouço da econometria espacial refletem-se em formulação de políticas públicas mais eficientes, que considerem relevantes os aspectos regionais na sua concepção.

3.2.1 Autocorrelação espacial e heterogeneidade espacial

Como dito anteriormente, no que se refere aos efeitos espaciais, estes podem ser diferenciados em dois tipos: autocorrelação espacial e heterogeneidade espacial. A autocorrelação espacial ocorre quando há uma relação entre o que acontece em determinada unidade no espaço e o que acontece nos demais pontos. Essa relação pode ocorrer quando há erros de delimitação geográfica nas unidades analisadas, problemas em termos de agregação, ou mais frequentemente, quando há externalidades que afetam certas unidades e seus efeitos são extrapolados para unidades geograficamente próximas (VAYÁ *et al.*, 2001). Os efeitos de transbordamento, ou *spillovers*, fazem com que as variáveis dependentes das unidades vizinhas se afetem mutuamente. Ademais, outra forma de autocorrelação que pode ocorrer são os erros estarem espacialmente correlacionados, pois os limites das unidades geográficas podem não ser relevantes para variáveis de interesse. Portanto, a existência de dependência espacial está

relacionada ao fato de variáveis dependentes ou erros em diferentes locais estarem correlacionados entre si (ANSELIN,1992).

O segundo efeito espacial corresponde a heterogeneidade espacial, que se refere a ausência de estabilidade de determinadas variáveis comportamentais. A heterogeneidade ocorre ao se tentar utilizar dados de unidades espaciais muito distintas para explicar um mesmo fenômeno (ANSELIN, 1988). Os efeitos da heterogeneidade podem ser interpretados como erros de mensuração com variáveis relevantes sendo ignoradas, ou, ainda, erros de especificação do modelo.

3.2.2 Matriz de Pesos Espaciais

Ambos os efeitos espaciais apresentados são importantes para o estudo adequado de dados geográficos, para tanto é fundamental determinar o tipo de influência que uma unidade específica exerce nas demais unidades do sistema espacial. Neste caso, é definido o critério de vizinhança e vizinhos próximos, e a partir dessa noção é construída a matriz de pesos espaciais (SABATER *et al.*, 2011). Essa matriz é denominada como matriz W , sendo utilizada para captar padrões de adjacências das unidades geográficas. De acordo com Almeida (2012), a matriz de pesos espaciais é baseada no conceito de contiguidade, podendo ser definida de acordo com a vizinhança, ou seja, pela distância tanto geográfica quanto socioeconômica, ou uma combinação de ambas.

A ideia do efeito espacial de uma localidade sobre as demais é baseada por ponderações, isso significa dizer que a variável observada em cada unidade receberá uma ponderação quando fizer parte da região de vizinhança da localidade analisada. É comumente utilizado como esquema de ponderação a matriz binária de pesos espaciais, sendo baseada na ideia de contiguidade/proximidade, na qual duas regiões serão vizinhas se fizerem fronteira física uma com a outra. Nesse sentido, os elementos recebem valores binários, ou seja, quando existir influência assume-se valor unitário; caso contrário atribui-se valor nulo. Sejam duas unidades vizinhas (i,j) que façam fronteira entre si, sendo consideradas contíguas, o elemento w_{ij} da matriz de ponderação espacial W receberá o valor 1. Por conseguinte, o elemento w_{ii} será nulo, pois nenhuma região pode ser vizinha dela mesma, ou seja, nenhuma região pode exercer influência sobre ela mesma. Portanto, no presente trabalho adotamos como ponderação uma matriz *Queen* binária com vizinhos de primeira ordem. A convenção de contiguidade *Queen*

considera além das fronteiras com extensão diferente de zero, também os vértices (nós), na visualização de um mapa, como contíguos (ALMEIDA, 2012). Formalmente, temos:

$$w_{ij} = \{1 \text{ se } i \text{ e } j \text{ são contíguas; } 0 \text{ se } i \text{ e } j \text{ não são contíguas}\} \quad (1)$$

3.2.3 Análise Explanatória de Dados Espaciais

Para compreender o padrão dos dados espaciais e auxiliar na escolha do modelo econométrico mais adequado para o estudo utiliza-se o método de análise explanatória de dados espaciais (AEDE). De acordo com Perobelli *et al.* (2005), através desse método podemos descrever a distribuição espacial do fenômeno, verificar seus padrões de associação global e local (*clusters*) e sugerir diferentes regimes espaciais. Dessa forma, a AEDE trata-se de uma coleção de técnicas para estudo estatístico das informações geográficas, capaz de auxiliar na elaboração de hipóteses e na compreensão dos padrões espaciais dos dados.

3.2.3.1 Autocorrelação Global

Nesse sentido, quando se realiza a análise explanatória dos dados espaciais primeiramente é testado a aleatoriedade dos dados, ou seja, é verificado se os valores de uma região dependem ou não dos valores das unidades vizinhas. De acordo com Almeida (2012), para verificar a existência de autocorrelação espacial, utiliza-se a estatística I de Moran, que tem por objetivo fornecer uma medida global da autocorrelação espacial, pois mensura o grau de associação linear entre os valores observados no conjunto de dados. Matematicamente, o I de Moran é dado por:

$$I = \frac{z'Wz}{z'z} \quad (2)$$

onde z corresponde aos valores da variável padronizada, Wz denota os valores médios da variável de interesse padronizada nas regiões vizinhas, definidos pela matriz de ponderação espacial W . Nesse sentido, a hipótese nula testada é a da aleatoriedade, ou seja, da independência espacial. Segundo Almeida (2012), o valor esperado do I de Moran é $E(I) = [-1 / (n - 1)]$, no qual n é o número de regiões. Os valores calculados da estatística I de Moran,

maiores ou menores que a esperança, indicam presença de autocorrelação espacial, positiva ou negativa, respectivamente.

A autocorrelação positiva indica a existência de similaridade entre os valores do atributo estudado e sua localização, ou seja, altos valores de uma variável de interesse (y) tendem a estar circundados por altos valores desta mesma variável em regiões vizinhas (Wy), bem como valores baixos tendem a estar rodeados por valores baixos na vizinhança. Para Almeida (2012), esse é o padrão sistemático de distribuição dos valores da variável de interesse quando há o efeito de contágio ou transbordamento de um fenômeno entre regiões. Por outro lado, a autocorrelação negativa revela a presença de dissimilaridade entre os valores do atributo e a sua região espacial, ou seja, altos valores de uma variável de interesse em uma região tende a estar cercado por baixos valores nas regiões vizinhas ou, ainda, um baixo valor da variável de interesse tende a estar rodeado por altos valores nas regiões vizinhas.

De modo geral, a estatística I de Moran fornece algumas informações importantes para a análise dos dados espaciais. Assim, a significância estatística revela a presença de aleatoriedade nos dados ou não. Valores positivos para a estatística I de Moran indicam que os dados estão concentrados através das regiões, por outro lado, o sinal negativo indica a dispersão dos dados. A magnitude do índice mostra o poder da autocorrelação, pois quanto mais próximo de 1 mais forte é a concentração e quanto mais perto de -1 mais dispersos estão os dados (ALMEIDA, 2012).

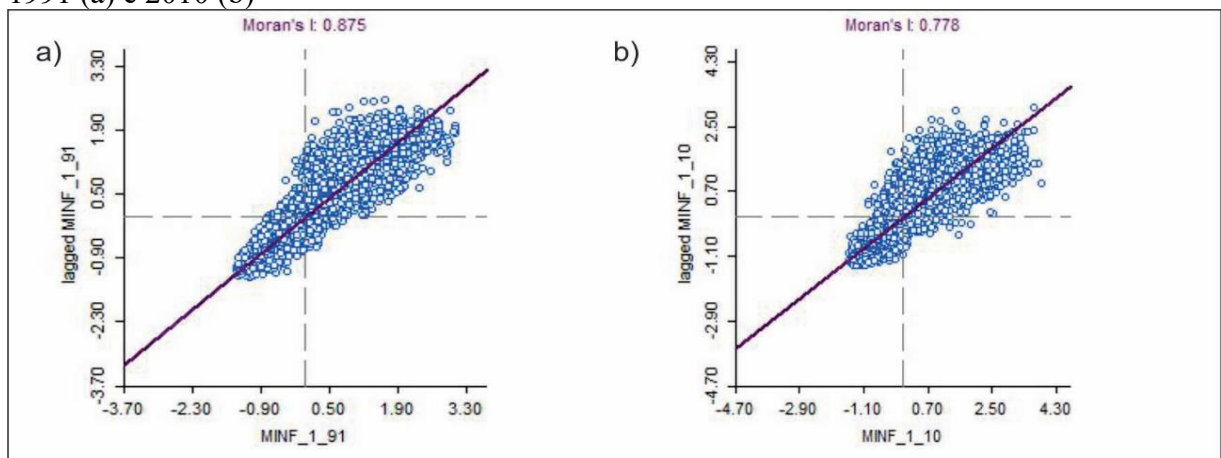
3.2.3.2 Diagrama de Dispersão de Moran

Outra forma de analisar a presença de autocorrelação espacial é através do diagrama de dispersão de Moran, pois ele reflete a estrutura espacial em duas escalas, ou seja, a vizinhança e a tendência. O diagrama de dispersão é uma ferramenta de interpretação gráfica do I de Moran. Dessa forma, a variável de interesse (y) está localizada no eixo horizontal e representa o valor padronizado da variável para cada unidade, já a defasagem espacial da variável de interesse (Wy) está no eixo vertical e representa a média do valor padronizado da mesma variável para os vizinhos.

Conforme Almeida (2004), o diagrama de Moran é dividido em quatro partes, sendo que cada quadrante representa os padrões de associação local espacial entre as regiões e seus vizinhos. O primeiro quadrante refere-se a combinação Alto-Alto (AA) e significa que as unidades espaciais exibem valores altos da variável de interesse cercadas por unidades espaciais

que também possuem valores altos, por outro lado, a combinação Baixo-Baixo (BB) representa os *clusters* das unidades que apresentam baixos valores rodeadas por vizinhos com valores igualmente baixos para a variável de interesse. O agrupamento Alto-Baixo (AB) diz respeito a unidade espacial que possui altos valores da variável, porém rodeadas por outras unidades que apresentam valores baixos para a variável de interesse. Seguindo o mesmo padrão de dissimilaridade, o segundo quadrante demonstra a combinação Baixo-Alto (BA) e corresponde aos *clusters* das unidades com valores baixos para a variável de interesse, mas sendo cercada por outras unidades que apresentam valores altos. Os quadrantes que apresentam as combinações AA e BB possuem autocorrelação espacial positiva, enquanto os quadrantes AB e BA apresentam autocorrelação espacial negativa. Na ausência de correlação espacial entre as unidades, a nuvem de dispersão dos pontos seria distribuída pelos quatro quadrantes. Nesse sentido, apresentamos a seguir o diagrama de dispersão de Moran para as principais variáveis de interesse deste estudo.

Figura 1 - Diagrama de Dispersão do *I* de Moran para a variável de Mortalidade Infantil em 1991 (a) e 2010 (b)

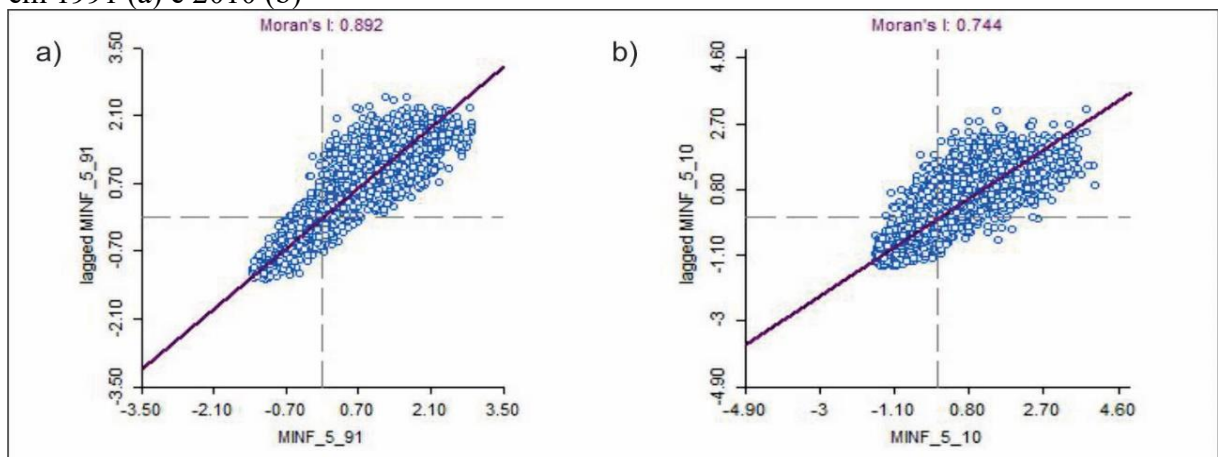


Fonte: Elaboração própria, com base nos dados do Atlas (2013).

A figura 1 apresenta o coeficiente do *I* de Moran bem como o seu diagrama de dispersão para a variável dependente de mortalidade infantil para os anos de 1991 e 2010, de acordo com a matriz de contiguidade do tipo *Queen*. A partir da análise dos coeficientes percebe-se a presença de autocorrelação espacial positiva para ambos os anos entre os municípios brasileiros, sendo de 0,88 e 0,78 para 1991 e 2010, respectivamente. Isso representa a existência de semelhanças entre a variável mortalidade infantil e a localização espacial da variável, ou seja, municípios que possuem altas (baixas) taxas de mortalidade infantil estão rodeados por municípios que também possuem altas (baixas) taxas de mortalidade. Houve ao longo do tempo

uma redução na magnitude do efeito da dependência espacial, porém o valor do coeficiente para o ano de 2010 ainda é considerado alto e representativo para a presença de autocorrelação espacial. Analisando o diagrama de dispersão para cada ano, percebe-se como as observações estão distribuídas segundo seu padrão espacial, assim os pontos que representam os municípios se concentram no primeiro e terceiro quadrantes, reiterando o padrão Alto-Alto e Baixo-Baixo. A inclinação positiva da reta confirma a presença de autocorrelação espacial, tendo em vista que o I de Moran representa o coeficiente angular da mesma.

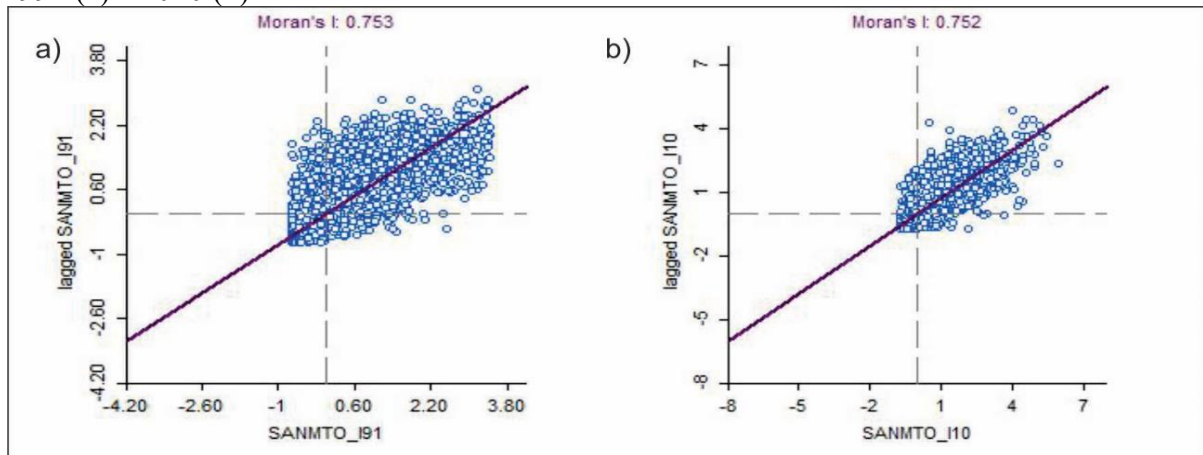
Figura 2 - Diagrama de Dispersão do I de Moran para a variável de Mortalidade na Infância em 1991 (a) e 2010 (b)



Fonte: Elaboração própria, com base nos dados do Atlas (2013).

Da mesma forma, a figura 2 representa o diagrama de dispersão do I de Moran para a variável dependente de mortalidade na infância para os anos de 1991 e 2010. Percebe-se um padrão espacial semelhante ao da variável mortalidade infantil, com os pontos concentrando-se no primeiro e terceiro quadrantes e uma inclinação positiva da reta, para ambos os anos, indicando a presença de autocorrelação espacial positiva. Ademais, o coeficiente do I de Moran para a variável de interesse segue a tendência de queda ao longo dos anos, com 0,89 para 1991 e 0,74 para 2010. Essa tendência de queda dos coeficientes representa uma diminuição no efeito transbordamento entre as regiões para essas variáveis, embora ainda sejam significativamente altos.

Figura 3 - Diagrama de Dispersão do I de Moran para a variável Saneamento Inadequado em 1991 (a) e 2010 (b)



Fonte: Elaboração própria, com base nos dados do Atlas (2013).

Por fim, o coeficiente do I de Moran e o diagrama de dispersão para a variável explicativa de saneamento inadequado estão representados na figura 3, no qual se observa a presença de autocorrelação positiva. O coeficiente do I de Moran manteve-se constante ao longo do período analisado, não havendo alterações entre 1991 e 2010. Analisando o diagrama de dispersão, percebe-se que em ambos os momentos a concentração dos pontos manteve-se no primeiro quadrante, ou seja, na combinação Alto-Alto. Assim sendo, o que podemos inferir diante dessa análise é que o efeito contágio ocorre entre os municípios com piores condições de saneamento adequado.

De modo geral, os índices de autocorrelação global indicam a presença de dependência espacial positiva entre os municípios para as três variáveis testadas. Isso representa a similaridade na distribuição espacial, pois unidades espaciais que apresentam altos (baixos) valores estão cercadas por vizinhos com altos (baixos) valores. Entretanto, podemos concluir que o efeito transbordamento entre os municípios tende a ser maior entre aqueles que possuem piores condições nos indicadores de qualidade de vida. Isso porque a aglomeração entre municípios com elevadas taxas de mortalidade infantil, tanto para até um ano quanto para até cinco anos, é maior do que os municípios com baixas taxas de mortalidade infantil. Do mesmo modo ocorre com a variável de saneamento inadequado, que apresenta um padrão de aglomeração maior entre os municípios com piores índices.

3.2.3.3 Autocorrelação Espacial Local (*Local Indicator of Spatial Association – LISA*)

Os padrões globais de associação linear captam a autocorrelação espacial de todo o espaço analisado, porém podem ocultar padrões locais de associação, como por exemplo a

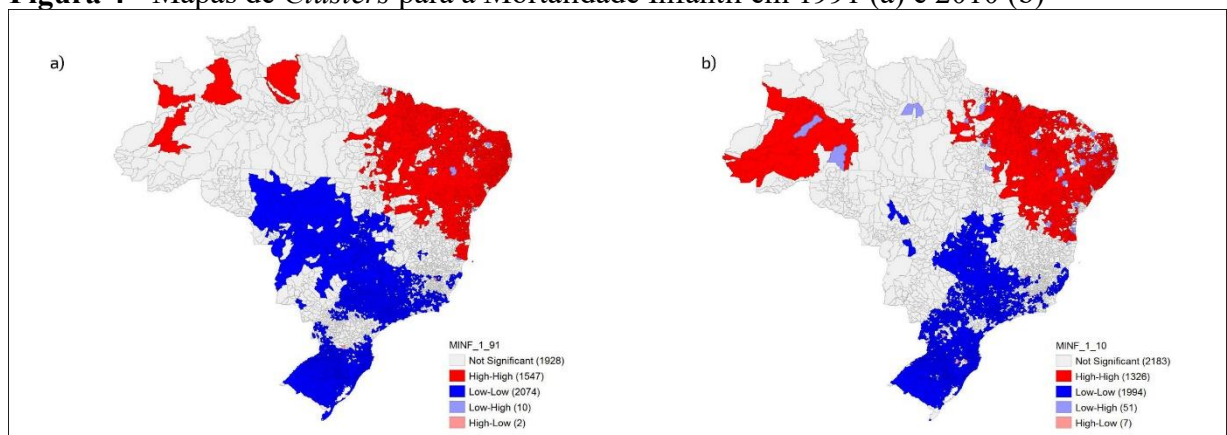
formação de *clusters* ou *outliers* espaciais (Almeida, 2012). Portanto, tem-se a necessidade de se utilizar uma estatística capaz de captar esses efeitos locais. Para tanto, utilizamos o Indicador Local de Associação Espacial (LISA) de forma complementar ao I de Moran global. Assim sendo, o I de Moran local tem a capacidade de obter padrões locais de associação linear estatisticamente significativa, através da decomposição do indicador global de autocorrelação na formação local de cada observação. A estatística LISA permite verificar a existência de *clusters* espaciais locais de valores altos ou baixos, demonstrando quais são as regiões que mais contribuem para a presença de autocorrelação espacial. De acordo com Almeida (2012) o coeficiente I_i de Moran local para uma variável y padronizada Z_i observada para a região i , pode ser expresso pela seguinte equação:

$$I_i = Z_i \sum_{j=1}^j W_{ij} Z_j \quad (3)$$

O cálculo do I_i só considera os vizinhos da observação i , definidos de acordo com uma matriz de pesos espaciais. Dessa forma, para que I_i seja um indicador LISA é necessário indicar *clusters* espaciais significativos estatisticamente, bem como que o somatório dos indicadores locais seja equivalente ao indicador de autocorrelação espacial global correspondente (ALMEIDA, 2012).

A figura 5 a seguir são apresentados os mapas de *clusters* para a mortalidade infantil para os anos de 1991 e 2010, os quais nos permitem visualizar o comportamento dessa variável em cada região do Brasil nesses dois períodos.

Figura 4 - Mapas de *Clusters* para a Mortalidade Infantil em 1991 (a) e 2010 (b)



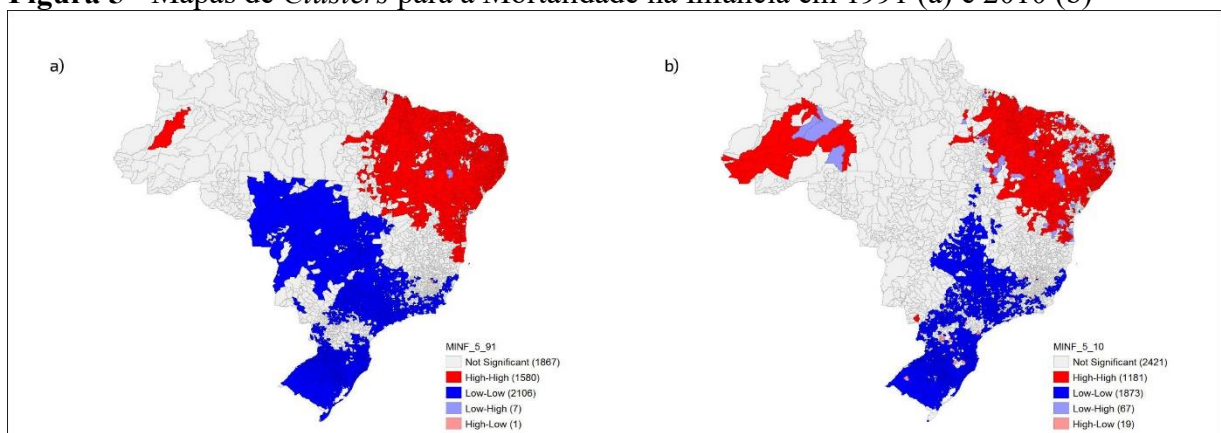
Fonte: Elaboração própria, com base nos dados do Atlas (2013).

Conforme destacado, os mapas de *clusters* permitem visualizar como ocorre a distribuição da variável de mortalidade infantil em cada região do território nacional. Assim, os

mapas da figura acima demonstram a existência de autocorrelação positiva entre as regiões sendo confirmada localmente. Dentre os dados de maior significância destaca-se a predominância da classificação Baixo-Baixo, esse resultado significa que as regiões que apresentam baixas taxas de mortalidade infantil encontram-se próximas de outras regiões também com valores baixos. Nesse sentido, as regiões Sul, Sudeste e Centro-Oeste do Brasil destacam-se por terem baixos índices de mortalidade infantil. Por outro lado, a formação de *clusters* do tipo Alto-Alto concentram-se em parte das regiões Norte e Nordeste, pois trata-se de regiões onde são observadas elevadas taxas de mortalidade cercada por vizinhos que também possuem altas taxas de mortalidade. Vale destacar que para o ano de 2010 parte das regiões Norte e Nordeste apresentaram um aumento na formação de *clusters* do tipo Baixo-Alto, representando uma melhora nos índices de mortalidade infantil para alguns municípios dessas regiões ainda que cercados por outros municípios com elevadas taxas de mortalidade.

Para a variável de mortalidade na infância, representada na figura 5 abaixo, o cenário não se altera muito em relação ao da variável mortalidade infantil. Pois, confirma localmente a existência de autocorrelação espacial positiva, com a predominância de formação de *clusters* do tipo Baixo-Baixo para as regiões Sul, Sudeste e Centro-Oeste e mantendo aglomerados do tipo Alto-Alto em parte das regiões Norte e Nordeste. Da mesma forma, destaca-se o aumento de *clusters* do tipo Baixo-Alto principalmente nas regiões Norte e Nordeste para o ano de 2010.

Figura 5 - Mapas de *Clusters* para a Mortalidade na Infância em 1991 (a) e 2010 (b)

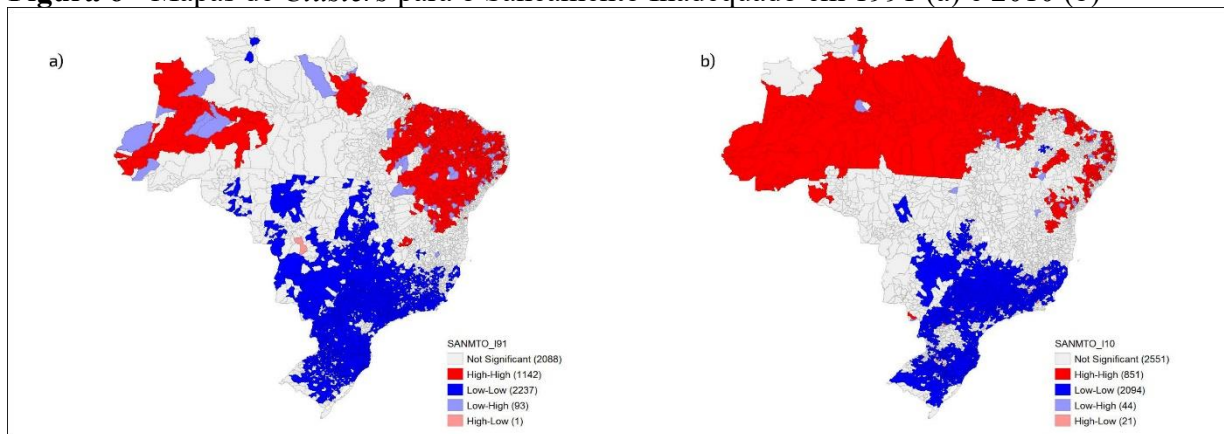


Fonte: Elaboração própria, com base nos dados do Atlas (2013).

A figura 6 apresenta os mapas de *clusters* para a variável de saneamento inadequado para os dados dos anos de 1991 e 2010, tornando-se possível observar a existência de autocorrelação espacial positiva sendo confirmada localmente. A partir da análise dos mapas percebemos a formação de *clusters* tipo Baixo-Baixo para as regiões Sul, Sudeste e parte do

Centro-Oeste, da mesma forma têm-se *clusters* do tipo Alto-Alto concentrados nas regiões Norte e Nordeste. Entretanto, para essas regiões existe também a formação de *clusters* do tipo Baixo-Alto, principalmente para o ano de 1991, indicando que municípios com condições precárias de saneamento estavam cercados por outros municípios com boas condições de saneamento. Todavia, esse cenário se alterou no ano de 2010, onde *clusters* do tipo Baixo-Alto deram lugar a aglomerações do tipo Alto-Alto, principalmente na região Norte do país. Essa mudança nos paradigmas pode indicar a influência do efeito transbordamento entre os municípios.

Figura 6 - Mapas de *Clusters* para o Saneamento Inadequado em 1991 (a) e 2010 (b)



Fonte: Elaboração própria, com base nos dados do Atlas (2013).

3.3 Estratégica empírica

Para atingir os objetivos propostos do presente trabalho foi adotado como modelagem econométrica os dados em painel espacial. De modo geral, os dados em painel são uma combinação de séries de tempo e dados em *cross-section*, no qual séries de tempo analisa a trajetória de um objeto ao longo do tempo e *cross-section* analisa dados para vários objetos em um dado período. Logo, dados em painel espacial é combinação de localização, séries de tempo e os dados em *cross-section*, e permite capturar interações espaciais através da observação e ao longo do tempo. A vantagem neste caso de se empregar uma estrutura de painel espacial é que nos permite analisar um fenômeno que acontece nos municípios brasileiros ao longo do tempo. Os parâmetros dos dados em painel podem ser estimados por duas abordagens, nomeadamente efeito fixo e efeito aleatório. O teste de *Hausman* é usado para selecionar qual abordagem deve ser usada para estimar. Como já mencionado, o *I* de Moran é usado para testar a existência de dependência espacial.

Como já mencionado anteriormente, os modelos de econometria espacial auxiliam na correção das violações dos pressupostos das regressões lineares tradicionais. A existência de efeitos espaciais, como autocorrelação espacial e heterogeneidade espacial, torna a metodologia tradicional inadequada, podendo apresentar parâmetros tendenciosos, ineficientes e inconsistentes. Nesse sentido, ao identificarmos a presença de dependência espacial cabe, então, especificar o modelo econométrico espacial mais adequado. Conforme Anselin (1999), há formas de tratar os efeitos da dependência espacial: primeiramente é introduzida uma variável dependente defasada espacialmente na regressão linear padrão, sendo capaz de captar a existência e o grau de dependência; outra forma é incorporar a dependência espacial no termo de erro, e assim corrigir possíveis vieses causados pela autocorrelação espacial e, por fim, outra maneira de controlar os efeitos espaciais é incluir variáveis exógenas defasadas espacialmente na estimação. A seguir, apresentamos os modelos econométricos que utilizam essas técnicas, ou a combinação delas para controlar os efeitos espaciais.

3.3.1 Modelo de Defasagem Espacial (Spatial Autoregressive Model - SAR)

Esse estudo busca identificar os efeitos espaciais nos dados sobre saneamento inadequado. Nesse sentido, o primeiro efeito espacial ocorre quando choques causam variações em uma observação da variável dependente i e esse efeito transborda para regiões vizinhas, afetando a observação da variável dependente j nessas regiões. Segundo Almeida (2012), esse efeito refere-se ao modelo de defasagem espacial, ou simplesmente o modelo SAR. Esse modelo informa que a variável dependente (y_i) de uma região é influenciada por tal variável dependente observada nas suas regiões vizinhas (Wy_i). O modelo SAR é dado por:

$$y_t = \rho W y_t + X_t \beta + \mu + \varepsilon_t \quad t = 1, \dots, T \quad (4)$$

em que, ρ é o coeficiente de defasagem espacial e mede o efeito médio da variável dependente relativo à vizinhança espacial na região em questão, este assume valores entre -1 e 1 e seu sinal e significância indicam a existência e o sentido das interações espaciais, $W y_t$ é a variável dependente defasada pela matriz de pesos espaciais W , X_t é uma matriz de variáveis explicativas exógenas, β é o vetor de parâmetros a serem estimados, μ em caso de efeito aleatório tem distribuição normal com média zero e variância constante, enquanto que no caso de efeito fixo μ é o vetor a ser estimado e ε_t é um vetor de erro aleatório normalmente distribuído com média zero e variância constante.

3.3.2 Modelo de Erro Espacial (Spatial Error Model - SEM)

Segundo Almeida (2012), o modelo de erro autorregressivo espacial, ou simplesmente o modelo SEM, trata-se do modelo espacial em que os efeitos não modelados não podem estar correlacionados a nenhuma das variáveis explicativas da regressão. Estes efeitos não modelados são considerados os padrões espaciais apresentados no modelo, que são causados por conta da falta adequada de medidas, que por sua vez não estão distribuídas aleatoriamente, porém, estes efeitos não modelados estão espacialmente autocorrelacionados. Conforme Almeida (2004), o modelo de erro espacial centra-se na autocorrelação espacial no termo de erro e é dado em sua forma estrutural por:

$$y_t = X_t\beta + \mu + u_t \quad (5)$$

$$u_t = \lambda W u_t + \varepsilon_t \quad (6)$$

onde, y_t é um vetor de observações da variável dependente, X_t é uma matriz de observações de variáveis explicativas, β é um vetor de coeficientes da regressão e u_t é um vetor nos termos de erro, $W u_t$ é o termo de erro defasado espacialmente. O coeficiente λ é o parâmetro do erro autorregressivo espacial que acompanha o termo de defasagem $W u_t$ e indica a intensidade da autocorrelação espacial entre os resíduos da equação observada; esse parâmetro mensura o efeito médio dos erros dos vizinhos em relação ao resíduo da região em questão e ε_t é um termo de erro não correlacionado com média zero e variância constante, ou seja, neste modelo, os erros associados com qualquer observação são uma média dos erros nas regiões vizinhas somada a um componente de erro aleatório. É claro que este é um caso especial do modelo SAC, mas também pode ser mostrado que é um caso especial da SDM, ambos apresentados a seguir.

3.3.3 Modelo de Durbin Espacial (Spatial Durbin Model – SDM)

Segundo Almeida (2012), o modelo de Durbin espacial, ou modelo SDM, trata de um modelo que além de incorporar a ideia de transbordamento por meio de defasagem das variáveis independentes da regressão (WX) incorpora também a variável endógena defasada espacialmente (Wy) no modelo. Então, o modelo de Durbin espacial consegue nos apresentar

tanto um alcance global como um alcance localizado. Este modelo é uma generalização do modelo SAR que também inclui variáveis independentes espacialmente ponderados como variáveis explanatórias.

O modelo SDM segundo LeSage e Pace (2009), apresenta características que podem gerar uma situação forte em trabalhos empíricos aplicados em econometria, sendo que o modelo trata de omissão de variáveis relevantes que, por sua vez, estão correlacionadas com pelo menos uma variável explicativa incluída na regressão, sendo essa mesma variável relevante omitida autocorrelacionada espacialmente. Nesse sentido, componente espacial a ser apresentado trata da defasagem do vetor de variáveis exógenas, em outras palavras, esse efeito ocorre quando as variáveis explicativas transbordam espacialmente. Com isso, o modelo SDM pode ser escrito por:

$$y_t = \rho W y_t + X_t \beta + W Z_t \theta + \mu + \varepsilon_t \quad (7)$$

onde, θ contém os parâmetros das variáveis explicativas defasadas espacialmente, a variável explicativa Z da equação (8) não é observada, sendo assim omitida da regressão, porém, a mesma variável omitida é autocorrelacionada espacialmente. O modelo pode ser generalizado utilizando de diferentes pesos espaciais para a variável dependente espacialmente defasada ($W y$) e os regressores espacialmente ponderados ($W Z$) ou usando que $Z_t \neq X_t$

3.3.4 Modelo de Autocorrelação Espacial (Spatial Autocorrelation Model – SAC)

Esse modelo combina o modelo SAR, mas com um termo de erro autorregressivo espacial:

$$y_t = \rho W y_t + X_t \beta + \mu + u_t \quad (8)$$

$$u_t = \lambda M u_t + \varepsilon_t \quad (9)$$

onde M é uma matriz de pesos espaciais que pode ou não ser igual a W . A literatura foca na variável de efeito fixo para essa especificação, assim como a variável de efeito aleatório pode ser escrita como um caso especial da especificação do modelo SAR.

De acordo com Belotti *et al.* (2016) uma estratégia comum na literatura que trata de efeitos da dependência espacial é iniciar as estimações de pelo modelo SDM, ou seja, pelo

modelo que inclui em sua estimação a defasagem espacial da variável dependente Wy e as defasagens das variáveis explicativas WX , como uma especificação geral e testar alguns modelos alternativos. Essa estratégia passa a ser importante caso tenha-se interesse em testar se o modelo de SAR ou o SEM é o mais apropriado. Sendo o modelo SDM derivado a partir de um modelo SEM², é facilmente mostrado que se $\theta=0$ e $\rho \neq 0$ o modelo é em SAR, mas se $\theta=-\beta\rho$ o modelo indicado é o SEM. Finalmente, uma vez que o modelo SAC e o modelo SDM são modelos não alinhados, os menores critérios de informação podem ser usados para testar se o modelo mais adequado é o SAC ou SDM, utilizando o Critério de Informação de Akaike (AIC) e o Critério Bayesiano de Schwarz (BIC).

Ao incluir a variável dependente espacialmente defasada como variável explicativa dos modelos SAR, SDM e SAC se faz necessário calcular os efeitos diretos e indiretos de cada uma das variáveis. Os efeitos diretos são uma medida do impacto da alteração de uma variável explicativa sobre a variável dependente em um município, incluindo os efeitos de *feedback*, ou seja, os efeitos que passam pelos municípios vizinhos e retornam para o município que ocorreu a mudança. Por outro lado, os efeitos indiretos medem o impacto de uma mudança em uma variável explicativa em um município sobre a variável dependente em todos os outros municípios. Dessa forma, os efeitos indiretos demonstram a existência ou não do efeito de transbordamento espacial, sem a necessidade de observar a estimativa do coeficiente associado à variável dependente espacialmente defasada, observando apenas o comportamento das variáveis explicativas (ELHORST, 2012).

Para Elhorst (2014), os efeitos tanto diretos quanto indiretos de uma variável explicativa sobre a variável dependente podem ser obtidos por meio da matriz de derivadas parciais do valor esperado de Y com respeito a k -ésima variável explanatória de X na unidade 1 até unidade N , conforme a equação a seguir:

$$\begin{aligned} \left[\frac{\partial E(Y)}{\partial x_{1k}} \right] \cdot \left[\frac{\partial E(Y)}{\partial x_{Nk}} \right] &= \begin{bmatrix} \frac{\partial E(y_1)}{\partial x_{1k}} & \cdots & \frac{\partial E(y_1)}{\partial x_{Nk}} \\ \vdots & \ddots & \vdots \\ \frac{\partial E(y_N)}{\partial x_{1k}} & \cdots & \frac{\partial E(y_N)}{\partial x_{Nk}} \end{bmatrix} \\ &= [I - \delta W]^{-1} \begin{bmatrix} \beta_k & w_{12}\theta_k & \cdots & w_{1N}\theta_k \\ w_{21}\theta_k & \beta_k & \cdots & w_{2N}\theta_k \\ \cdots & \cdots & \ddots & \vdots \\ w_{N1}\theta_k & w_{N2}\theta_k & \cdots & \beta_k \end{bmatrix} \end{aligned} \quad (10)$$

² Para verificar essa demonstração consultar Arbia *et al.* (2010).

sendo que os elementos da diagonal principal da matriz de derivadas parciais representam o efeito direto. Por outro lado, os elementos fora da diagonal principal indicam a presença de efeito indireto. Se uma determinada variável explanatória em uma determinada unidade muda, não só a variável dependente nessa própria unidade muda, mas também as variáveis dependentes de outras unidades. Conforme Elhorst (2014), a primeira mudança é chamada de efeito direto e a segunda de efeito indireto.

Na seção seguinte apresentamos os resultados para o presente estudo. Nesse sentido, considera-se como o melhor modelo de painel espacial, aquele com os menores critérios de avaliação baseados no máximo da função de verossimilhança (MFV): o Critério de Informação de Akaike (AIC) e o Critério Bayesiano de Schwarz (BIC).

4. RESULTADOS

Nesta seção são apresentados os resultados das estimações para o presente estudo. O objetivo do estudo foi encontrar resultados que mostrem a relação entre um saneamento inadequado e as taxas de mortalidade infantil e na infância a partir dos modelos espaciais SAR, SEM, SDM e SAC. Os testes foram importantes para nortear a escolha do melhor modelo. Nesse sentido, as tabelas 3 e 4 apresentam os resultados das estimações para mortalidade infantil e mortalidade na infância, respectivamente. Inicialmente ambas as tabelas mostram os resultados dos estimadores: i) POLS (*Pooling Ordinary Least Squares*), que ignora as características específicas de cada grupo; ii) efeitos fixos, que admite características idiossincráticas de cada município; iii) efeitos aleatórios no qual as características idiossincráticas são aleatórias. Nesse sentido, estimou-se os dados em um modelo de painel, desconsiderando a autocorrelação espacial, para escolher qual o melhor estimador. A partir das estatísticas do teste de Hausman, foi possível rejeitar a hipótese nula, com valor da probabilidade de 0,000, ou seja, o modelo de efeitos aleatórios deve ser descartado pois é inconsistente. Nesse caso, foi indicado como melhor abordagem a utilização dos modelos de efeitos fixos para todas as variáveis dependentes, pois não é possível ignorar a presença de efeitos idiossincráticos.

Tabela 3: Resultados dos modelos de Efeito Fixo e Aleatório, SDM e SAC para a Taxa de Mortalidade Infantil

| VARIÁVEIS | POLS | EF | EA | SDM_EF | SDM_EA | SAC_EF |
|---|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| Saneamento | 0,1507*** (0,0051) | 0,1122*** (0,0053) | 0,1403*** (0,0048) | 0,0282*** (0,0054) | 0,0350*** (0,0049) | 0,0107*** (0,0021) |
| Inadequado | | | | | | |
| Renda domiciliar <i>per capita</i> média | -4,3855*** (0,2690) | -1,0758** (0,4091) | -1,9917*** (0,2755) | -0,6615* (0,3114) | -2,4538*** (0,2506) | -0,4009* (0,1831) |
| Taxa de analfabetismo | 0,7222*** (0,0101) | 1,2207*** (0,0204) | 0,8170*** (0,0109) | 0,2226*** (0,0206) | 0,2793*** (0,0146) | 0,1614*** (0,0099) |
| Densidade domiciliar | 0,0345*** (0,0073) | 0,0677*** (0,0148) | 0,0325*** (0,0082) | 0,0212 (0,0119) | 0,0534*** (0,0093) | 0,0062 (0,0062) |
| Coleta de lixo | -0,0205*** (0,0035) | -0,0039 (0,0041) | -0,0183*** (0,0035) | -0,0113*** (0,0031) | -0,0092** (0,0027) | -0,0051** (0,0019) |
| Energia elétrica | -0,0267*** (0,0057) | -0,0818*** (0,0075) | -0,0687*** (0,0057) | -0,0329*** (0,0071) | -0,0194* (0,0060) | -0,0021 (0,0030) |
| População urbana | 0,0585*** (0,0039) | 0,0796*** (0,0089) | 0,0645*** (0,0044) | 0,0139* (0,0064) | 0,0193*** (0,0038) | 0,0059 (0,0043) |
| Ln população total | 0,7984*** (0,0720) | 0,2283 (0,4179) | 0,6582*** (0,0869) | 0,1098 (0,3469) | -0,5381*** (0,0780) | 0,3370 (0,1764) |
| Constante | 36,9088*** (1,7396) | 10,1330* (4,9045) | 25,4094*** (1,8403) | | 0,0780*** (1,8834) | |
| VARIÁVEIS DEFASADAS ESPACIALMENTE | | | | | | |
| Saneamento | | | | 0,0039 (0,0071) | 0,0128* (0,0063) | |
| Inadequado | | | | | | |
| Renda domiciliar <i>per capita</i> média | | | | -0,9366 (0,5292) | 2,8945*** (0,3485) | |
| Taxa de analfabetismo | | | | 0,3241*** (0,0294) | -0,0012 (0,0182) | |
| Densidade domiciliar | | | | -0,0011 (0,0183) | -0,0627*** (0,0115) | |
| Coleta de lixo | | | | 0,0102 (0,0055) | -0,0048 (0,0045) | |
| Energia elétrica | | | | 0,0302** (0,0102) | -0,0049 (0,0081) | |
| População urbana | | | | -0,0098 (0,0122) | -0,0040 (0,0060) | |
| Ln população total | | | | 0,1221 (0,5544) | 0,9162*** (0,1136) | |
| Mortalidade Infantil | | | | 0,6913*** (0,0076) | 0,7233*** (0,0058) | 0,9106*** (0,0040) |
| e.Mortalidade Infantil | | | | | | -0,8508*** (0,0192) |
| sigma_e cons | | | | 4,6904*** (0,0322) | 4,7182*** (0,0324) | 3,9429*** (0,0308) |
| sigma_u cons | | | | | 3,5762*** (0,0566) | |
| Teste de Hausman | | 2010,10 (0,000) | | 640,42 (0,0000) | | |
| Teste $\theta=0$ | | | | 131,48 (0,0000) | 351,36 (0,0000) | |
| Teste $\theta= -\beta*\rho$ | | | | 761,83 (0,0000) | 876,17 (0,0000) | |
| AIC | | | | 67192,95 | 106764,5 | 66055,12 |
| BIC | | | | 67347,40 | 106934,4 | 66155,50 |
| N | 16683 | 16683 | 16683 | 16683 | 16683 | 16683 |

Obs: Erros padrões robustos entre parênteses. ***Significante a 1%; **Significante a 5%; *Significante a 10%.

Fonte: Elaboração própria a partir dos dados da pesquisa.

Nesse sentido, dando seguimento a estratégia empírica proposta, foi estimado o modelo SDM para efeitos fixos para as variáveis dependentes, de acordo com as tabelas 3 e 4. Esse modelo, conforme mencionado na seção anterior, inclui a defasagem espacial da variável dependente Wy e as defasagens das variáveis explicativas WX . Como forma de indicar qual o melhor modelo capaz de controlar os efeitos espaciais fazemos os testes no parâmetro θ que acompanha as variáveis explicativas defasadas espacialmente. Primeiramente, testamos se $\theta=0$ e $\rho \neq 0$ e, conforme as estatísticas nas estimações realizadas para todas as variáveis dependentes aceitamos a hipótese nula de $\theta \neq 0$, ou seja, o parâmetro que acompanha as variáveis explicativas defasadas espacialmente não é nulo. Em seguida testamos se $\theta = -\beta\rho$ e nesse caso, conforme a seção anterior, o parâmetro sendo igual a $-\beta\rho$ o modelo mais indicado foi o SEM. Dando seguimento a estratégia empírica, como mencionado anteriormente, o modelo SDM pode ser derivado a partir do modelo de Erro Autorregressivo SEM, assim sendo optamos por estimar o modelo SDM por se tratar de um modelo que consegue nos apresentar tanto um alcance global como um alcance localizado.

Tabela 4: Resultados dos modelos de Efeito Fixo e Aleatório, SDM e SAC para a Taxa de Mortalidade na Infância

| VARIÁVEIS | POLS | EF | EA | SDM EF | SDM EA | SAC EF |
|---|------------|------------|------------|------------|------------|-----------|
| Saneamento | 0,2011*** | 0,1529*** | 0,1908*** | 0,0355** | 0,0432*** | 0,0117*** |
| Inadequado | (0,0068) | (0,0072) | (0,0065) | (0,0067) | (0,0059) | (0,0025) |
| Renda domiciliar <i>per capita</i> média | -5,2435*** | -2,5712*** | -2,3273*** | -0,9596** | -2,9502*** | -0,7174** |
| Taxa de analfabetismo | (0,3586) | (0,5504) | (0,3691) | (0,3821) | (0,3056) | (0,2204) |
| Densidade domiciliar | 0,9593*** | 1,7427*** | 1,0737*** | 0,2777*** | 0,3362*** | 0,1975*** |
| Coleta de lixo | (0,0135) | (0,0275) | (0,0145) | (0,0252) | (0,0177) | (0,0119) |
| Energia elétrica | -0,0072 | 0,0996*** | -0,0090 | 0,0291* | 0,0630*** | 0,0076 |
| População urbana | (0,0098) | (0,0200) | (0,0108) | (0,0146) | (0,0114) | (0,0075) |
| Ln população total | -0,0350*** | -0,0103 | -0,0326*** | -0,0149*** | -0,0125*** | -0,0068** |
| Constante | (0,0046) | (0,0055) | (0,0047) | (0,0038) | (0,0033) | (0,0023) |
| | -0,0807*** | -0,1545*** | -0,1396*** | -0,0496*** | -0,0318*** | -0,0077* |
| | (0,0076) | (0,0101) | (0,0077) | (0,0087) | (0,0074) | (0,0037) |
| | 0,0980*** | 0,1345*** | 0,1084*** | 0,0192** | 0,0260*** | 0,0099 |
| | (0,0052) | (0,0119) | (0,0059) | (0,0078) | (0,0046) | (0,0051) |
| | 0,9906*** | -0,1534 | 0,8065*** | 0,0344 | -0,6542*** | 0,3499 |
| | (0,0960) | (0,5624) | (0,1136) | (0,4255) | (0,0943) | (0,2114) |
| | 48,1437*** | 17,7358** | 34,6281*** | | 1,0439 | |
| | (2,3189) | (6,5993) | (2,4531) | | (2,2877) | |
| VARIÁVEIS DEFASADAS ESPACIALMENTE | | | | | | |
| Saneamento | | | | -0,0004 | 0,0110 | |
| Inadequado | | | | (0,0087) | (0,0077) | |
| Renda domiciliar <i>per capita</i> média | | | | -2,0737** | 3,5135*** | |
| Taxa de analfabetismo | | | | (0,6511) | (0,4249) | |
| Densidade domiciliar | | | | 0,4094*** | -0,0239 | |
| Coleta de lixo | | | | (0,0365) | (0,0221) | |
| | | | | -0,0047 | -0,0928*** | |
| | | | | (0,0225) | (0,0140) | |
| | | | | 0,0116 | -0,0076 | |
| | | | | (0,0068) | (0,0054) | |

| | | | | |
|-----------------------------|----------|-----------|-----------|------------|
| Energia elétrica | | 0,0311* | -0,0141 | |
| | | (0,0125) | (0,0099) | |
| População urbana | | 0,0291 | 0,0025 | |
| | | (0,0150) | (0,0072) | |
| Ln população total | | -0,0553 | 1,0451*** | |
| | | (0,6801) | (0,1372) | |
| Mortalidade na Infância | | 0,7291*** | 0,7677*** | 0,9232*** |
| | | (0,0068) | (0,0051) | (0,0034) |
| e.Mortalidade na Infância | | | | -0,8831*** |
| | | | | (0,0185) |
| sigma_e cons | | 5,7535*** | 5,7868*** | 4,7786*** |
| | | (0,0396) | (0,0398) | (0,0369) |
| sigma_u cons | | | 4,2680*** | |
| | | | (0,0691) | |
| Teste de Hausman | 3039,87 | 731,98 | | |
| | (0,0000) | (0,0000) | | |
| Teste $\theta=0$ | | 139,22 | 408,60 | |
| | | (0,0000) | (0,0000) | |
| Teste $\theta= -\beta*\rho$ | | 827,74 | 890,96 | |
| | | (0,0000) | (0,0000) | |
| Aic | | 71913,99 | 113733,6 | 70568,48 |
| Bic | | 72068,44 | 113903,5 | 70668,87 |
| N | 16683 | 16683 | 16683 | 16683 |

Obs: Erros padrões robustos entre parênteses. ***Significante a 1%; **Significante a 5%; *Significante a 10%.

Fonte: Elaboração própria a partir dos dados da pesquisa.

Temos que dado o fato de os modelos SAC e SDM não serem relacionados, uma forma de escolher qual dos modelos apresenta as melhores estimações é através dos menores critérios de informação. Logo, entre o modelo SAC e modelo SDM as estimações que apresentaram melhores Critério de Informação de Akaike (AIC) e o Critério Bayesiano de Schwarz (BIC) foi o modelo SAC. Conforme podemos observar nas tabelas 3 e 4, para as duas estimações das variáveis dependentes os critérios de informação foram menores para o modelo SAC, indicando se tratar do modelo mais adequado para serem feitas as análises.

Definido a partir dos testes qual modelo e abordagem mais indicados para serem feitas as análises, apresentamos na tabela 5 os efeitos marginais para as duas variáveis dependentes através do modelo SAC de Efeitos Fixos. Os resultados começam a ser analisados a partir da variável de interesse no estudo que trata do saneamento inadequado. Para esta variável os resultados referentes a mortalidade infantil e a mortalidade na Infância foram estatisticamente significativos a 1% e o sinal obtido da variável está em conformidade ao esperado e de acordo com a literatura (MERRICK,1983; ALVES E BELLUZZO, 2004; SOUSA e LEITE FILHO, 2008; BARUFI *et al.*, 2012; LEIVAS *et al.*, 2016; SCRIPTORE e AZZONI, 2018). Assim, os resultados sugerem que um sistema de saneamento inadequado acaba por afetar positivamente os índices de mortalidade infantil em geral, corroborando os resultados encontrados na literatura. Os efeitos indiretos dessa variável foram maiores do que o efeitos diretos, esse

resultado demonstra a presença forte do efeito transbordamento entre as regiões. Isso porque, variações nas condições de saneamento inadequado dos municípios vizinhos acabam por afetar as taxas de mortalidade infantil da região de interesse. Isso ocorre porque a difusão de doenças infectocontagiosas transborda entre as regiões e ocorre através de vetores, como por exemplo, o deslocamento de insetos de uma região para a outra ou água contaminada de rios e mares que atravessam várias localidades, (SCRIPTORE e AZZONI, 2018). A magnitude dos efeitos indiretos vai ao encontro dos resultados analisados através da análise explanatória de dados espaciais, que demonstrou uma dependência espacial forte entre as regiões para a variável de interesse.

Tabela 5 – Resultados dos dados em painel de efeitos fixos do modelo SAC

| | SAC Mortalidade Infantil | | | SAC Mortalidade na Infância | | |
|--|--------------------------|-----------------------|-----------------------|-----------------------------|-----------------------|-----------------------|
| | EF. Direto | EF. Indireto | EF. Total | EF. Direto | EF. Indireto | EF. Total |
| Saneamento inadequado | 0,0151*** (0,0029) | 0,1041*** (0,0196) | 0,1193*** (0,0225) | 0,0171*** (0,0036) | 0,1352*** (0,0278) | 0,1523*** (0,0313) |
| Renda domiciliar <i>per capita</i> média | -0,5692* (0,2597) | -3,9140* (1,7848) | -4,4833* (2,0436) | -1,0476** (0,3211) | -8,2985** (2,5393) | -9,3461** (2,8578) |
| Taxa de analfabetismo | 0,2292*** (0,0129) | 1,5757*** (0,0782) | 1,8049*** (0,0876) | 0,2884*** (0,0160) | 2,2847*** (0,1102) | 2,5731*** (0,1219) |
| Densidade domiciliar | 0,0089 (0,0088) | 0,0610 (0,0607) | 0,0699 (0,0695) | 0,0111 (0,0109) | 0,0878 (0,0861) | 0,0988 (0,0970) |
| Coleta de lixo | -0,0073** (0,0027) | -0,0500** (0,0186) | -0,0573** (0,0213) | -0,0100** (0,0033) | -0,0792** (0,0265) | -0,0891** (0,0298) |
| Energia elétrica | -0,0030 (0,0043) | -0,0204 (0,0296) | -0,0234 (0,0339) | -0,0112* (0,0053) | -0,0889* (0,0420) | -0,1001* (0,0473) |
| População urbana | 0,0084 (0,0060) | 0,0578 (0,0415) | 0,0662 (0,0475) | 0,0144 (0,0075) | 0,1144 (0,0591) | 0,1289 (0,0665) |
| Ln população total | 0,4785 (0,2505) | 3,2902 (1,7311) | 3,7687 (1,9809) | 0,5110 (0,3088) | 4,0473 (2,4570) | 4,5583 (2,7653) |

Obs: Erros padrões robustos entre parênteses. ***Significante a 1%; **Significante a 5%; *Significante a 10%.

Fonte: Elaboração própria a partir dos dados da pesquisa.

A variável que corresponde a renda domiciliar *per capita* média apresentou resultados significativos à 10% para mortalidade infantil e 5% para mortalidade na infância e demonstrou efeito negativo para as ambas as variáveis de mortalidade infantil nas estimações. Os resultados obtidos vão ao encontro dos resultados encontrados na literatura (BARUFI *et al.*, 2012; LEIVAS *et al.*, 2016; SCRIPTORE e AZZONI, 2018) e indicam que a renda familiar está relacionada a melhores condições no que tange a saúde. Isso porque, possibilita o pagamento de serviços médicos, aquisição de remédios, uma melhor nutrição às crianças do domicílio, melhores condições de habitação com acesso a saneamento adequado, dentre outros elementos que são afetados pela renda do domicílio e que contribuem para a redução da mortalidade infantil.

Da mesma forma, outro resultado considerável é o efeito da taxa de analfabetismo dos indivíduos maiores de 15 anos sobre a mortalidade infantil. O resultado foi significativo à 1% para as duas variáveis dependentes e apresentou sinal positivo, como esperado, indicando que quanto maior a taxa de analfabetismo dos indivíduos maior será a taxa de mortalidade das crianças. Essa variável buscou captar a escolaridade dos pais, e reflete como indivíduos com baixa escolaridade possuem piores condições de saúde. (BARUFI *et al.*, 2012; LEIVAS *et al.*, 2016; SCRIPTORE e AZZONI, 2018). A escolaridade do pai afeta a mortalidade infantil por ter relação direta com a renda familiar e consequentemente afetar a busca por serviços de saúde. Já a escolaridade da mãe reflete no tempo dedicado ao cuidado com a saúde dos filhos, pela busca por informações, a prevenção de problemas de saúde, bem como a hábitos de higiene.

Demais variáveis que afetam a mortalidade infantil utilizadas no modelo correspondem ao acesso a serviços públicos como coleta de lixo e energia elétrica. Via de regra, essas variáveis apresentaram efeito negativo para a mortalidade infantil, sendo o acesso à energia elétrica estatisticamente significativo apenas para explicar a mortalidade na infância com nível de confiança a 10%. Nesse sentido, a variável com maior poder de explicação foi a que corresponde a coleta de lixo, sendo estatisticamente significativa a 5% para ambas variáveis de mortalidade infantil. Tanto coleta de lixo quanto energia elétrica foram usadas nesse estudo para contemplar características do domicílio que fossem relevantes para diminuir a taxa de mortalidade infantil, porém a magnitude do efeito total dessas variáveis foi menor quando comparada a condição de saneamento inadequada do domicílio. Nesse sentido, para Scriptore e Azzoni (2018) os serviços públicos mais eficientes para a promoção da saúde seriam o esgotamento sanitário e tratamento de água, sobretudo no que se refere a redução de doenças como diarreia, que afeta principalmente crianças.

De modo geral as estimações dos efeitos marginais do modelo SAC demonstraram uma maior magnitude dos efeitos indiretos em detrimento dos efeitos diretos. Esses resultados demonstram o forte componente da dependência especial presente em dados regionais, através do efeito de transbordamento, ou *spillovers*, entre as regiões. Nesse sentido, caso semelhante ocorre no estudo de Scriptore e Azzoni (2018), os autores defendem que caso o efeito espacial não seja considerado nas regressões o efeito positivo do saneamento inadequado sobre as taxas de mortalidade passa ser subestimado. Para os autores, um aumento na estrutura de esgotamento sanitário da vizinhança é capaz de explicar uma queda nos índices de mortalidade infantil de determinada região através das externalidades que são geradas. Seguindo na mesma linha, outro estudo que demonstrou impacto maior nos efeitos indiretos ocorreu no trabalho dos autores

Bastos *et al.* (2018), em que as análises revelaram que a implantação de unidades de Estratégia de Saúde da Família tiveram impacto maior na redução da mortalidade infantil nos municípios vizinhos do que no município que implementou primeiro. Neste caso, os autores defendem que o fato de os efeitos indiretos dominarem os efeitos diretos ressalta a importância do controle espacial nas estimações.

Ademais, uma maior magnitude dos efeitos indiretos em relação aos efeitos diretos presente nesse estudo vai ao encontro dos resultados encontrados na seção 3 através da análise explanatória de dados espaciais (AEDE), bem como com a literatura que indica a econometria espacial como mais adequada para dados regionais (ANSELIN,1992; ELHORST, 2012; ALMEIDA, 2012).

5. CONSIDERAÇÕES FINAIS

A relação entre saúde e as condições inadequadas de saneamento são recorrentes na ciência econômica. Nesse sentido, a mortalidade infantil tornou-se um bom indicador das condições de saúde de uma população (MERRICK, 1985; ALVES E BELLUZZO, 2004; SOUSA e LEITE FILHO, 2008; BARUFI *et al.*, 2012; LEIVAS *et al.*, 2016, SCRIPTORE e AZZONI, 2018). Isso porque, a população infantil é a mais vulnerável em condições precárias de saneamento, tendo em vista que são as crianças que mais sofrem com doenças infectocontagiosas. Dentre essas doenças, a de maior morbidade tem sido a diarreia, que acomete crianças e adultos, mas que os efeitos têm sido mais devastadores entre a população infantil.

Por entender a importância das condições de saneamento sobre a saúde da população, em especial das crianças, que se desenvolveu este estudo com o objetivo de analisar como a falta de acesso ao saneamento básico adequado afeta a saúde dos brasileiros. Para tanto, foram utilizadas variáveis que são consideradas como *proxy* das condições de saúde e bem-estar da população, ou seja, a taxa de mortalidade infantil e a taxa de mortalidade na infância. Por se tratar de dados regionais, optou-se por usar como método mais adequado os dados em painel espacial. A vantagem neste caso de se empregar uma estrutura de painel espacial é que nos permite analisar as interações espaciais de um determinado fenômeno que ocorre entre municípios brasileiros ao longo do tempo. Para chegar aos resultados, foram utilizados a base de dados do Atlas Brasil (2013), que concentra informações dos Censos Demográficos de 1991, 2000, e 2010. Após a realização de testes de especificação e pelos menores critérios de

informação, tais como AIC e BIC, concluímos que o melhor modelo a ser analisado foi o modelo SAC com Efeitos Fixos, que se trata de um modelo que além de incluir a variável dependente espacialmente defasada inclui também um termo de erro autorregressivo espacialmente.

Entre os resultados encontrados cabe destacar o efeito positivo de condições inadequadas de saneamento nas taxas de mortalidade infantil, demonstrando a forte relação entre saneamento e saúde, corroborando com os estudos anteriores (MERRICK, 1983; ALVES E BELLUZZO, 2004; SOUSA e LEITE FILHO, 2008; BARUFI *et al.*, 2012; LEIVAS *et al.*, 2016; SCRIPTORE e AZZONI, 2018). Da mesma forma, a taxa de analfabetismo também obteve resultado positivo nos índices de mortalidade infantil, indicando que a escolaridade dos indivíduos afeta as condições de saúde. Por outro lado, variáveis de renda domiciliar *per capita* média, energia elétrica e coleta de lixo tiveram efeito negativo sobre as taxas de mortalidade infantil, revelando que renda e acesso a serviços básicos interferem nas condições de saúde e bem-estar da população.

A estimação dos efeitos marginais permitiu separar os efeitos diretos, que são aqueles que medem o quanto a variável dependente se modifica em função de uma variação na variável explicativa, dos efeitos indiretos, que são aqueles que medem a mudança na variável dependente decorrente da alteração nas variáveis explicativas dos municípios vizinhos. Nesse sentido, a preponderância da magnitude dos efeitos indiretos, revela a forte dependência espacial entre os municípios. Dependência essa destacada pela análise explanatória de dados espaciais, que revelou a formação de *clusters* em nível local, demonstrando que as regiões Sul, Sudeste e parte do Centro-Oeste mantem melhores índices de mortalidade infantil e saneamento adequado, enquanto em parte das regiões Norte e Nordeste apresentam piores índices para essas variáveis.

Além dos resultados encontrados, uma das contribuições deste estudo foi oferecer um panorama mais atual da relação entre mortalidade infantil e saneamento para todo o território brasileiro. Além disso, a estimação dos efeitos marginais permitiu separar os efeitos diretos e indiretos dos modelos espaciais, demonstrando o forte efeito de transbordamento entre as regiões. Assim, acredita-se que as políticas públicas para melhoria e maior abrangência de saneamento adequado devam considerar as externalidades causadas nos municípios vizinhos. Por exemplo, aqueles municípios que não possuam estrutura adequada de saneamento podem ser beneficiados pelas ações dos vizinhos que possuam. Cabe destacar que os dados utilizados são até o ano de 2010, sendo assim, uma sugestão para trabalhos futuros é a atualização desses

dados com o a realização do Censo de 2021, afim de captar as condições mais atuais de saneamento e saúde após a crise mundial de 2008 e os efeitos da recessão econômica do país.

REFERÊNCIAS

- ALMEIDA, E. S. **Curso de econometria espacial aplicada**. Piracicaba: Esalq-USP, 2004.
- ALMEIDA, E. S. **Econometria espacial aplicada**. Campinas: Editora Alínea, 2012.
- ALVES, D.; BELLUZZO, W. Child health and infant mortality in Brazil. *In: Annual Meeting of Brazillian Econometric Society*, 26. João Pessoa, 2004.
- ANSELIN, L. **Spatial econometrics: methods and models**. Boston: Kluwer Academic, 1988.
- ANSELIN, L. **SpaceStat tutorial: a workbook for using SpaceStat in the analysis of spatial data**. Illinois: Urbana-Champaign, 1992.
- ANSELIN, L. **Spatial econometrics**. Dallas: Bruton Center School of Social Sciences, 1999.
- ARBIA, G.; BATTISTI, M.; DI VAIO, G. Institutions and geography: empirical test of spatial growth models for European regions. **Economic Modelling**, v. 27, n. 1, p. 12-21, 2010.
- BARUFI, A. M.; HADDAD, E; PAEZ, A. Infant mortality in Brazil, 1980–2000: a spatial panel data analysis. **BMC Public Health**, v. 12, n. 1, p. 181, 2012.
- BASTOS, S. Q. A; GOMES, B. S. M; ALMEIDA, E. S. Análise espacial da estratégia saúde da família (1998/2012). *In: ANPEC - Associação Nacional dos Centros de Pós-Graduação em Economia*, 2018.
- BELOTTI, F.; HUGHES, G.; MORTARI, A.P. Spatial panel data models using Stata. **CEIS Research Paper**, v. 14, n 373, 2016.
- BRASIL. Lei nº 11.445, de 5 de janeiro de 2007. Estabelece diretrizes nacionais para o saneamento básico; altera as Leis nºs 6.766, de 19 de dezembro de 1979, 8.036, de 11 de maio de 1990, 8.666, de 21 de junho de 1993, 8.987, de 13 de fevereiro de 1995; revoga a Lei nº 6.528, de 11 de maio de 1978; e dá outras providências. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/lei/111445.htm. Acesso em: 21 de mar. de 2019.
- CUTLER, D.; MILLER, G. The role of public health improvements in health advances: the twentieth-century United States. **Demography**, v. 42, n. 1, p. 1-22, 2005.
- ELHORST, J. P. Dynamic spatial panels: models, methods, and inferences. **Journal of Geographical Systems**, vol. 14, n. 1, p. 5–28, 2012.
- ELHORST, J. P. **Spatial econometrics: from cross-sectional data to spatial panels**. Berlin: Springer, 2014.

ESREY, S. A.; FEACHAM, R. G.; HUGHES, J. M. Interventions for the control of diarrhoeal diseases among young children: improving water supplies and excreta disposal facilities. **Bulletin of the World Health Organization**, v. 63, n. 4, p. 757-772, 1985.

INSTITUTO TRATA BRASIL. **Saneamento**, 2019. Disponível em: <http://www.tratabrasil.org.br/saneamento/principais-estatisticas>. Acesso em: 21 de mar. de 2019.

JALAN, J.; RAVALLION, M. Does piped water reduce diarrhea in rural India? **Journal of Econometrics**, v. 112, n.1, p. 153-173, 2003.

LEIVAS, P. H. S.; GONÇALVES, R. R.; DOS SANTOS, A. M. A.; SOUZA, O. T. Sustentabilidade, saneamento e saúde infantil no Brasil: uma análise a partir de macro e microdados. **Espacios**, v. 37, n. 35, 2016.

LESAGE, J. P.; PACE, R. K. **Introduction to spatial econometrics**. Boca Raton: Taylor & Francis Group, 2009.

MENDONÇA, M. J. C.; SEROA DA MOTTA, R. Saúde e saneamento no Brasil. **Planejamento e Políticas Públicas**, v. 30, p. 15-30, 2007.

MERRICK, T. The effects of pipes water on early childhood mortality in urban Brazil 1970 to 76. **Demography**, v.22, n. 1, p. 1-24, 1985.

PEROBELLI, F. S.; ALMEIDA, E. S.; ALVIM, M. I. S.; FERREIRA, P. G. C. A. Análise espacial da produtividade do setor agrícola brasileiro: 1991-2003. *In: Congresso Brasileiro de Economia e Sociologia Rural*, Ribeirão Preto: 43, 2005.

SABATER, L. A.; TUR, A. A.; AZORÍN, J. M. N. Análise exploratória dos dados espaciais (AEDE). *In: COSTA, J. S.; DENTINHO, T. P.; NIJKAMP, P. Compêndio de Economia Regional: métodos e técnicas de análise regional*. Parede: Principia, 2011.

SCRIPTORE, J. S.; AZZONI, C. R. Impactos do saneamento básico sobre a saúde: uma análise espacial. *In: ANPEC - Associação Nacional dos Centros de Pós-Graduação em Economia*, 2018.

SOUSA, T. R. V.; LEITE FILHO, P. A. M. Análise por dados em painel do status de saúde no Nordeste Brasileiro. **Revista Saúde Pública**, São Paulo, v. 42, n. 5, p.796-804, 2008.

TEIXEIRA, J. C.; GOMES, M. H. R.; SOUZA, J. A. Associação entre cobertura por serviços de saneamento e indicadores epidemiológicos nos países da América Latina: estudo com dados secundários. **Revista Panamericana de Salud Pública**, v. 32, n. 6, p. 419-425, 2012.

UHR, J. G. Z.; SCHMECHEL, M.; UHR, D. A. P. Relação entre saneamento básico no Brasil e saúde da população sob a ótica das internações hospitalares por doenças de veiculação hídrica. **Revista de Administração, Contabilidade e Economia da Fundace**, v. 7, n. 2, p. 01-16, 2016.

VAYÁ, E.; LÓPEZ-BAZO, E.; MORENO, R.; SURIÑACH, J. Growth and externalities across economies: an empirical analysis using spatial econometrics. *In: ANSELIN, L.; FLORAX, R. Advances in spatial econometrics*. Heidelberg: Springer-Verlag, 2001.

VIEIRA, R. S. **Crescimento econômico no estado de São Paulo: uma análise espacial**. São Paulo: Editora UNESP, 2009.

WATSON, T. Public health investments and the infant mortality gap: evidence from federal sanitation interventions on U.S. Indian reservations. **Journal of Public Economics**, v. 90, n. 8/9, p. 1537-1560, 2006.