

# IMPACTOS DO SANEAMENTO BÁSICO SOBRE A SAÚDE: UMA ANÁLISE ESPACIAL

**Juliana Souza Scriptoro**

Universidade de São Paulo (FEA-USP) e Universidade Federal de Alfenas (UNIFAL-MG)

**Carlos Roberto Azzoni**

Universidade de São Paulo (FEA-USP)

## Resumo

Nesse estudo foi investigado o efeito do saneamento sobre a saúde a partir de uma análise espacial. Essa investigação foi realizada por meio de dados por distrito construídos a partir do Censo Demográfico do Universo por setor censitário (IBGE, 2010), Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE, 2011), Diretório Nacional de Endereços (DNE, 2015) e Departamento de Informática do Sistema Único de Saúde (DATASUS, 2010). Os resultados indicam que, em geral, o acesso aos serviços de saneamento inadequados causa impacto positivo na taxa de incidência de internações por Doenças Relacionadas ao Saneamento Ambiental Inadequado (DRSAI), por mil habitantes, a variável de saúde considerada nesse estudo. Para a população de todas as idades, o impacto total médio do aumento de 1% na parcela de domicílios servidos por água da chuva armazenada em cisterna ou caixa de cimento é 0,217 casos por mil habitantes. Quando se considera a porcentagem dos domicílios que tem banheiro conectado a uma fossa rudimentar, esse impacto é 0,035 casos por 1000 habitantes. Por outro lado, no que diz respeito à variável adequada de saneamento, domicílios conectados à rede de distribuição de água e que, além disso, na quadra onde estão localizados é inexistente condições de esgoto a céu aberto, esse impacto gera uma queda de 0,166 casos, por mil habitantes, nas internações por DRSAI. A variável de maior impacto total médio refere-se à porcentagem de domicílios que não tem banheiro e, além disso, possuem presença de esgoto a céu aberto em suas condições de entorno, pois acarreta aumento de 3,281 na taxa de incidência de internações por DRSAI. Outro resultado desse estudo é que a população mais jovem é a mais atingida pelas condições inadequadas dos serviços de saneamento básico.

**Palavras-chave:** Tobit Durbin Espacial, Econometria Espacial, Saneamento Básico, Saúde.

JEL: O18; I18; C21

## Abstract

This work studies the effect of sanitation on health made use of spatial econometrics techniques. This analysis was conducted applying data at district level, which was constructed from the results from the universe of the Population Census (IBGE) by census tract, the National Register of Addresses for Statistical Purposes (CNEFE, 2011), the National Address Directory (DNE, 2015) and the Department of Informatics of the Unified Health System (DATASUS, 2010). The results indicate that, in general, access to inadequate sanitation services has a positive impact on the incidence of hospitalizations because of Diseases Related to Inadequate Environmental Sanitation (DRIES), per thousand inhabitants, the health variable considered in this study. For the population of all ages, the average total impact of a 1% increase in the share of dwellings served by rain water (stored in cisterns or cement tanks) is 0.217 cases per thousand inhabitants. Considering the percentage of dwellings with toilets connected to rudimentary cesspools, this impact is 0.035 cases per thousand inhabitants. On the other hand, with regard to the appropriate sanitation variable, the average total impact of 1% increase in the share of permanent dwellings connected to the water distribution network and whose block does not present open sewage conditions is -0.166 cases per thousand inhabitants. The variable with highest average total impact refers to the percentage of dwellings that have no toilet and, moreover, have the presence of open sewage in their surrounding conditions. The increase in a percentage unit in this variable generates an increase of 3.281 in the rate of incidence of hospitalizations because of DRIES. Another result of this study is that the younger population is the most affected by the poor conditions to basic sanitation services.

**Key words:** Spatial Durbin Tobit, Spatial Econometrics, Sanitation, Health.

JEL: O18; I18; C21

**Área 10 - Economia Regional e Urbana**

## 1. Introdução

A importância do saneamento básico e da sua associação com a saúde remonta a tempos mais antigos. É reconhecida e consolidada a literatura que trata dos efeitos positivos do acesso aos serviços de abastecimento de água, tratamento e coleta de esgoto tanto em relação aos impactos sobre mortalidade infantil, evidenciados por Cutler e Miler (2005), Watson (2006), Geruso e Spears (2015), quanto no que diz respeito aos indicadores de morbidez retratados por Bleakley (2007), Duflo et al. (2015), Fewtrell et al. (2005).

As precárias condições de saneamento básico propiciam a transmissão de bactérias, vírus e parasitas, que estão presentes nas fezes, urina ou vômito do doente ou portador, causadores de diversas doenças infectocontagiosas. A diarreia é a mais conhecida dentre elas, pelos seus efeitos devastadores sobre a taxa de mortalidade infantil em diversos países do mundo. No entanto, existem diversas outras doenças que também merecem preocupação quanto aos danos causados à população, em especial às crianças, tais como a esquistossomose, febre amarela e malária. A grande importância em diminuir as taxas de incidência de tais doenças levou a Fundação Nacional da Saúde (Funasa, 2004) a classificá-las como Doenças Relacionadas ao Saneamento Ambiental Inadequado (DRSAI) baseada na proposta de classificação das doenças de Cairncross e Feachem (1990; 1993) e Mara e Feachem (1999).

A partir do reconhecimento milenar desta associação entre saneamento e saúde, o presente artigo investiga, a partir de uma análise espacial que controla a censura dos dados por meio de um *Spatial Durbin Model Tobit*, as consequências das DRSAI sobre a saúde de toda população e das crianças de 0 a 14 anos. Os dados utilizados, não disponíveis em domínio público, foram construídos em nível distrital para o ano de 2010 com o objetivo de comprovar de forma mais realista as hipóteses dos efeitos espaciais estabelecidas neste trabalho. O artigo está dividido em seis seções incluindo essa introdução. A próxima seção trata da revisão bibliográfica, a terceira da metodologia e dos dados utilizados, a quarta apresenta os resultados e a última tece as conclusões.

## 2. Revisão bibliográfica

O objetivo desta seção é apresentar a literatura acerca da relação entre o acesso aos serviços de saneamento básico e as DRSAI que, não levam o indivíduo a óbito, mas que podem gerar um estado de saúde debilitado e gerar preocupantes impactos socioeconômicos no longo prazo.

Fewtrell et al. (2005) concluíram que abastecimento de água, coleta de esgoto, práticas higiênicas e também a combinação destas ações foram efetivas na redução de doenças diarreicas, com destaque para as intervenções associadas à qualidade da água, que se mostraram ainda mais eficientes. Porém, estudo de Esrey (1991) destaca superioridade da intervenção de esgotamento sanitário comparativamente à oferta de água na queda de incidência de diversas doenças relacionadas ao saneamento. Duflo et al (2015) estimaram o impacto de um programa integrado de oferta dos serviços de água encanada e fossas higiênicas, com custo razoável, para cada residência de um conjunto de vilas na área rural da Índia. A complementariedade de tais serviços, juntamente com as externalidades que geram, mostrou-se efetiva para a redução de doenças transmitidas pela água. Os resultados encontrados pelos autores indicaram redução em torno de 30 a 50% dos episódios de diarreia aguda e, ainda que em menor magnitude e precisão, também houve diminuição da incidência de malária.

Em relação à malária, dentre as principais causas dessa doença estão a má qualidade da água e as poças de água parada que afetam principalmente as crianças, que se tornam adultos menos produtivos (OMS, 2004). Gallup e Sachs (2001) mostraram que países como Costa do Marfim ou Zâmbia, em que 50% da população está exposta a essa doença, tem rendimentos per capita que são um terço daqueles países em que as pessoas estão livres da contaminação pela mesma. Bleakley (2010) encontrou evidências, em vários países da América Latina, de que uma criança que cresce livre da malária ganha 50% mais por ano durante toda sua vida adulta, em comparação com uma criança que tenha contraído a doença.

Bleakley (2007) encontrou que crianças infectadas com o parasita da ancilostomose (“amarelão”) no Sul dos Estados Unidos, que se desenvolveu em condições sanitárias inadequadas, tinham uma probabilidade 20 por cento menor de matricular-se na escola. Além disso, os efeitos de longo prazo

mostraram que a população infectada durante a infância teve uma redução de 40% em seus salários após se tornarem adultos. Miguel e Kremer (2004) encontraram um efeito positivo da medicação contra helmintos (“vermes”) intestinais sobre a frequência escolar de crianças quenianas. Os autores ressaltam a importância da adequada provisão de bens públicos de saneamento, uma vez que a transmissão de tais agentes patogênicos ocorreu pela matéria fecal não tratada ou não disposta corretamente pelos mesmos.

Para o caso brasileiro, Heller (1997) mostrou que um número importante de estudos epidemiológicos aponta para a associação entre condições de saneamento e saúde. Andreazzi et al. (2007) encontraram em 15 estudos associação positiva entre a variável de acesso a saneamento e a melhora da variável de saúde investigada. Kassouf (1994) destacou o papel da infraestrutura residencial sobre a saúde das crianças em idade escolar. A água encanada foi importante para a melhoria da saúde das crianças na região Nordeste e no setor urbano. Na região Sul e no setor rural, o sistema de esgoto afetou positivamente a saúde das crianças.

Barreto et al. (2010) apontaram queda de 21% na prevalência de diarreia nas crianças com menos de 3 anos de idade entre o período pré e pós introdução de um programa de saneamento em Salvador no ano de 1997. O estudo ainda ressaltou que essa queda ocorreu via aumento da cobertura na vizinhança, já que o programa conseguiu diminuir a poluição fecal dos vizinhos como um todo. A partir de três regiões pré-estabelecidas em Salvador, os resultados de Moraes (2003) *et al.* apontaram que a incidência de diarreia nas crianças de regiões que contavam apenas com os serviços de drenagem de águas pluviais foi de dois terços da incidência de diarreia nas crianças de regiões que não tinham esse serviço. Esse valor para as regiões que tinham drenagem e esgotamento sanitário foi menos que um terço em comparação a regiões que não tinham nenhum desses dois serviços. Rego et al. (2004) mostrou que crianças expostas ao lixo no ambiente em que vivem tiveram probabilidade 3,98 vezes maior de ter diarreia, quando comparadas a crianças que não são expostas. De forma geral, os estudos apontam o efeito positivo de ações de saneamento sobre a queda da incidência de DRSAI, com destaque para a diarreia.

### **3. Metodologia**

#### **3.1 Dados**

Os dados deste trabalho são provenientes do Censo Demográfico do Universo por setor censitário (IBGE, 2010), Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE, 2011), Diretório Nacional de Endereços (DNE, 2015) e Departamento de Informática do Sistema Único de Saúde (DATASUS, 2010).

Os dados por setor censitário compreendem características dos domicílios particulares e das pessoas que foram investigadas para a totalidade da população e são denominados, por convenção, resultados do universo (IBGE, 2011). O setor censitário é a menor unidade territorial, formada por área contínua, integralmente contida em área urbana ou rural, com dimensão adequada à realização da coleta de dados por um pesquisador que vai a campo por ocasião do censo. O setor constitui um conjunto de quadras, no caso de área urbana, ou uma área do município, no caso de uma área não urbanizada.

Nesse estudo utilizaram-se dados por setor censitário, agregados por distritos<sup>1</sup>, com o objetivo de se alcançar um nível de desagregação menor que o municipal.

O menor nível de desagregação é relevante para o problema em análise, pois o deslocamento dos agentes patogênicos pode ocorrer mais facilmente entre distritos, mas menos provavelmente entre municípios. Além disso, a justificativa dessa desagregação, além do menor raio de deslocamento dos agentes patogênicos, encontra-se na percepção de que as condições de saneamento básico de um município tendem a ser muito heterogêneas entre diferentes distritos, em especial nas grandes cidades<sup>2</sup>. Logo, nesse estudo foram utilizados dados agregados em nível distrital que exigiu a construção de uma nova variável dependente, que não está disponível nas bases de dados comumente

---

<sup>1</sup>Os distritos são unidades administrativas dos municípios e criados por legislação municipal, devendo observar a continuidade territorial e os requisitos previstos em lei complementar estadual (IBGE, 2011).

<sup>2</sup> Por exemplo, São Paulo, o maior município brasileiro em população, possui 96 distritos, que apresentam distintas taxas de acesso a saneamento básico.

utilizadas pelos pesquisadores. Esta variável é proveniente dos registros de Autorização de Internação Hospitalar (AIH) extraídos do Sistema de Informações Hospitalares do Sistema Único de Saúde (SIHSUS-DATASUS, 2010). Esses dados são informados mensalmente por todos os estabelecimentos públicos, conveniados e contratados, que realizam internações financiadas pelo SUS. Entre os diversos tipos de internações disponíveis, foram selecionadas as internações cujo diagnóstico principal fosse típico das DRSAIs, de acordo com classificação da FUNASA (2004). Tais doenças foram identificadas pelo capítulo 10 da Classificação Internacional de Doenças, CID-10 (OMS, 1997). Cada registro de internação acompanha o Código de Endereçamento Postal (CEP) do local da residência de cada paciente. Para a construção da nova variável dependente foi necessário estabelecer uma lista de CEPs para cada distrito com o objetivo de realizar um “merge” e identificar em quais distritos os casos de DRSAI ocorreram. As faixas de CEP de cada distrito foram construídas utilizando-se duas bases de dados: o CNEFE e o DNE. O Cadastro Nacional de Endereços para fins Estatísticos (CNEFE) é uma base de dados disponibilizada também pelo IBGE, que contém aproximadamente 80 milhões de endereços urbanos e rurais distribuídos pelos 310 mil setores censitários. No entanto, os CEPs do CNEFE não cobrem todo o território brasileiro<sup>3</sup>. De forma a complementar e identificar os CEPs das regiões faltantes, utilizou-se a base de dados dos Correios, o Diretório Nacional de Endereços (DNE). O DNE é um banco de dados que contém mais de 900 mil CEPs de todo o Brasil que contém os Códigos de Endereçamento Postal (CEP). A tabela A.1 do Anexo apresenta o número de distritos por cada Unidade da Federação bem como o número de agregações realizadas utilizando essas duas bases de dados. Por fim, a tabela 3.1 contém a descrição de todas as variáveis utilizadas nesse estudo e a Tabela A.2 do Anexo as suas estatísticas descritivas.

**Tabela 3.1 – Descrição variáveis dependentes e das variáveis independentes (Modelo 1 e 2).**

Siglas	Descrição das variáveis	Unidade	Fonte
Inc_todas_idades	Taxa de incidência de internações por DRSAI da população de todas as idades	casos/mil hab	DataSUS
Inc_0a14	Taxa de incidência de internações por DRSAI da população de 0 a 14 anos	casos/mil hab	DataSUS
dcist	Domicílios particulares permanentes com abastecimento de água da chuva armazenada em cisterna ou caixa de cimento	%	Censo (IBGE)
sbceuab	Domicílios particulares permanentes que não tinham banheiro ou sanitário e que suas condições de entorno apresentavam esgoto a céu aberto	%	Censo (IBGE)
dfrud	Domicílios particulares permanentes com banheiro de uso exclusivo dos moradores ou sanitário ligado a uma fossa rudimentar, ou seja, fossa negra, poço ou buraco.	%	Censo (IBGE)
esg_ao_ceuab	Domicílios particulares permanentes que tinham banheiro ou sanitário e esgotamento sanitário via rede geral de esgoto, rede pluvial ou fossa séptica e onde não existia em sua face/quadra esgoto a céu aberto.	%	Censo (IBGE)
ag_ao_ceuab	Domicílios particulares permanentes com rede geral de distribuição de água e onde não existia em sua face/quadra esgoto a céu aberto.	%	Censo (IBGE)
lixo_mor_inad	Domicílios particulares permanentes com moradia inadequada e onde em sua face/quadra existia lixo acumulado nos logradouros	%	Censo (IBGE)
d_nbueiro	Domicílios particulares permanentes cedidos, alugados e próprios e onde em sua face/quadra não existia bueiro/boca-de-lobo.	%	Censo (IBGE)
rendapc	Rendimento nominal mensal dos domicílios particulares permanentes pela população residente	R\$	Censo (IBGE)
gini	Índice de Gini	0-1	Censo (IBGE)
mu_resp_alfab	Número de pessoas do sexo feminino, responsáveis pelo domicílio, alfabetizadas	%	Censo (IBGE)
tpopurb	número de pessoas residentes em setores censitários urbanos	%	Censo (IBGE)
dom_kmq	Número total de domicílios particulares permanentes pela área em quilômetros quadrados.	Dom/km <sub>q</sub>	Censo (IBGE)
est_saude	Número de estabelecimentos de saúde.	Est. Saúde	CNEFE (IBGE)
dummy_capital	Variável binária que indica se o distrito é capital	0;1	Censo (IBGE)
dummy_rm	Variável binária que indica se o distrito pertence a uma região metropolitana oficial	0;1	Censo (IBGE)

Fonte: Elaboração própria

### 3.2 O modelo teórico

A grande questão que permeou os estudos da economia da saúde desde a segunda metade do século passado refere-se a como os cuidados com a saúde diferem entre os indivíduos. Uma das

<sup>3</sup> Há uma grande quantidade de CEPs iguais para distritos diferentes e outros com códigos errados e inexistentes. Isso ocorre, pois, os mesmos foram obtidos a partir de declaração dos próprios moradores. A verificação da validade de tais informações por parte do IBGE foi realizada apenas para os municípios pequenos de CEP único.

características dos cuidados de saúde, de acordo com Grossman (1972), é que a procura por cuidados médicos é derivada da demanda por saúde, que é fundamental para os indivíduos participarem do mercado de trabalho e auferirem rendimentos no futuro.

O estudo de Sousa e Leite Filho (2008) introduziram variáveis relativas aos serviços de saneamento para explicar o *status* de saúde de uma população. No presente trabalho, a variável explicada de saúde será medida pela taxa de incidência de internações por Doenças Relacionadas ao Saneamento Ambiental Inadequado (DRSAI) e, em consonância com os autores, serão introduzidas variáveis de saneamento na função de produção de dias saudáveis também denominada de função de produção de saúde ou investimento bruto em saúde ( $I_i$ ). Em sua forma original, Grosman (1972) considera quatro variáveis na função de produção de saúde: a primeira delas diz respeito ao consumo de bens e serviços médicos apropriados, denominados cuidados médicos ( $M_i$ ), a segunda é referente ao tempo e ao esforço que cada pessoa empreende para melhorar a saúde ( $TH_i$ ), a terceira ( $A_i$ ) demonstra que a idade do indivíduo pode afetar a capacidade de produção de saúde e, por fim, a última variável ( $E_i$ ) evidencia que pessoas educadas produzem investimento em saúde de forma mais eficiente e, portanto, cuidam-se mais. No entanto, outras variáveis exógenas também podem ser consideradas nessa função de produção. Berndt *et al.* (2000) ressaltam que a função de produção de saúde apresenta uma série de argumentos que podem ser interdependentes entre si. Um dos elementos que os autores relatam afetar a saúde refere-se ao ambiente. As mudanças ambientais (poluição do ar, água, clima, entre outros) podem melhorar ou retardar a saúde e, portanto, devem ser contempladas nesta função.

O acesso aos diversos tipos de serviços de saneamento (água, esgoto, lixo e drenagem de águas pluviais) será considerado variável exógena importante na função de produção de saúde. Dessa forma, a variável que será adicionada na mesma é denominada “tipo do serviço de saneamento básico ( $ST_i$ )” no período de tempo  $i$ . A outra variável adicionada diz respeito ao vetor de características socioeconômicas ( $SOC_i$ ) dos distritos. A partir dessas variáveis, a função de produção de saúde assume a especificação expressa na equação abaixo:

$$(3.1) \quad I_i = I_i(M_i, TH_i; E_i, A_i, ST_i, SOC_i)$$

A quantidade de equilíbrio de  $H_i$  é encontrada pela maximização da utilidade familiar<sup>4</sup> sujeita a um conjunto de restrições. A saúde é considerada por Grosman (1972) como um bem de consumo, pois entra diretamente na função utilidade e proporciona, por si só, bem-estar aos indivíduos. Além disso, pode ser vista como um bem de investimento, porque as pessoas demandam saúde com o intuito de ganhar renda no futuro ao participarem do mercado de trabalho. Por último, a saúde é também bem de capital, pelo fato da mesma durar mais que um período de tempo e constituir-se em estoque acumulado, que se deprecia com o aumento da idade.

Quando a aquisição de saúde é considerada como um bem puro de investimento, os custos de capital de saúde são os custos de se abdicar de tempo e dinheiro para se obter saúde. Desde que o estoque de saúde inicial herdado pelo indivíduo e as taxas de depreciação são dadas, a quantidade ótima de investimento bruto em saúde determina a quantidade ótima de capital de saúde ( $H^*$ ).

A relação desse custo com o estoque de saúde do indivíduo descreve o padrão das taxas de retorno do investimento em saúde, ou seja, a eficiência marginal do investimento em saúde (EMH), que exhibe retornos marginais decrescentes para a função de produção de dias saudáveis. A curva que mostra a relação entre o estoque de saúde e o custo de capital é infinitamente elástica. O ponto ótimo de demanda de saúde é obtido pela interseção da curva EMH com a curva acima.

À medida em que a educação, o acesso aos serviços adequados de saneamento básico e as variáveis socioeconômicas (tal como a renda) aumentam o capital de saúde do indivíduo é afetado. Isso ocorre pelo fato das pessoas utilizarem os insumos de produção de forma mais eficiente (educação), devido à presença de externalidades ambientais que reforçam os retornos obtidos dos dias saudáveis (saneamento) ou em decorrência do investimento em novos insumos de saúde (variáveis socioeconômicas). No caso da renda, duas situações podem ocorrer: o maior valor da hora do

---

<sup>4</sup> Nesse caso, na função utilidade intertemporal de um consumidor típico,  $U = U(\phi_0 H_0, \dots, \phi_n H_n; Z_0, \dots, Z_n)$ ,  $H_0$  é o estoque de saúde inicial herdado pelo indivíduo,  $H_i$  é o estoque de saúde em qualquer período de tempo,  $\phi_i$  é o fluxo de serviços de saúde por unidade de estoque de saúde, em qualquer período de tempo,  $\phi_i H_i$  é o consumo total de serviços de saúde e  $Z_i$  é consumo total de outras mercadorias não relacionadas com a saúde.

indivíduo no mercado de trabalho o leva a desejar não perder tempo com a falta de saúde e, portanto, investe na mesma. De forma alternativa, por trocar um pouco do seu rendimento por lazer terá mais tempo para investir em saúde que, conseqüentemente, reforça o primeiro efeito e ainda lhe causa bem-estar.

### 3.3 Estratégia empírica

#### 3.3.1 Os modelos e testes espaciais

O modelo geral para a relação entre a taxa de incidência de DRSAI (*IncDr Sai*) e as variáveis de saneamento (*San*) é expresso por:

$$(3.2) \text{IncDr Sai} = \text{San}\beta + \varepsilon$$

Em que *IncDr Sai* é um vetor  $n \times 1$ , *San* é uma matriz  $n \times n$  de observações das variáveis de interesse de saneamento e outras variáveis independentes de controle,  $\beta$  é o vetor  $k \times 1$  de parâmetros e  $\varepsilon$  é um vetor de erros aleatórios independente e identicamente distribuídos.

Esse estudo tratará dos efeitos espaciais presentes nos dados sobre saneamento básico. Há três tipos deles. O primeiro efeito espacial ocorre quando alterações em uma observação da variável incidência de DRSAI (*IncDr Sai<sub>i</sub>*), causadas por choques endógenos e/ou exógenos ao modelo, extravasam para os vizinhos, causando alterações nas observações dos mesmos (*IncDr Sai<sub>j</sub>*). Este modelo pode ser representado em termos matriciais, por (Anselin, 1988):

$$(3.3) \text{IncDr Sai} = \rho W \text{IncDr Sai} + \text{San}\beta + \varepsilon; \quad |\rho| < 1$$

$$(3.4) \varepsilon/X \approx i. i. d. N(0, \sigma_{un}^2 I_n)$$

Em que,  $\rho$  é o parâmetro autoregressivo espacial, que mede a influência média dos vizinhos sobre observações no vetor *IncDr Sai*, *W* é a matriz  $n \times n$  de pesos espaciais exógenos e *WIncDr Sai* é o termo de defasagem espacial, que representa, para cada distrito, a taxa de incidência de internações por DRSAI dos distritos vizinhos.

O segundo efeito espacial é a heterogeneidade espacial, que significa que podem existir respostas distintas decorrentes de um estímulo do fenômeno em estudo, dependendo da região onde ocorrem. Neste caso, ainda de acordo com Anselin e Bera (1998), a especificação mais comum desse efeito assume a seguinte forma:

$$(3.5) \text{IncDr Sai} = \text{San}\beta + \varepsilon; \quad \varepsilon = \lambda W \varepsilon + \xi$$

Em que,  $\lambda$  é o coeficiente autoregressivo espacial para a defasagem do termo de erro e  $\xi$  é o termo de erro não correlacionado e homocedástico.

Por fim, o último componente espacial a ser apresentado trata da defasagem do vetor de variáveis exógenas (*San*) que contém as variáveis de saneamento e de outras variáveis de controle. Em outras palavras, esse efeito ocorre quando as variáveis contidas na matriz *San* transbordam espacialmente.

$$(3.6) \text{IncDr Sai} = W \text{San}\theta + \text{San}\beta + \varepsilon$$

Em que,  $\theta$  contém os parâmetros das variáveis explicativas defasadas espacialmente.

Os três tipos de componentes espaciais mencionados acima dão origem, respectivamente, ao Modelo de Defasagem Espacial (SAR), ao Modelo de Erro Espacial (SEM) e ao Modelo Regressivo Cruzado Espacial (SLX). A combinação do modelo SAR ( $\rho \neq 0$ ) com o modelo SLX ( $\theta \neq 0$ ) resulta no modelo Durbin Espacial (SDM), que inclui dois tipos de defasagem espacial: da variável dependente e das variáveis independentes. O modelo Durbin Espacial de Erro (SDEM), por sua vez, é a combinação do modelo SEM ( $\lambda \neq 0$ ) com o modelo SLX ( $\theta \neq 0$ ).

Segundo Anselin (1998), a análise exploratória dos dados espaciais (ESDA) descreve distribuições espaciais, identifica localidades atípicas, agrupamentos de observações semelhantes (*clusters*) e aponta formas de heterogeneidade espacial. Logo, antes de se estimar um modelo espacial deve-se fazer tal análise. A primeira estatística de teste global elaborada para a análise em questão é o *I* de Moran, coeficiente de autocorrelação espacial cuja hipótese nula é de aleatoriedade espacial<sup>5</sup>. O valor positivo e estatisticamente significativo da estatística *I* de Moran é global, pois leva em conta todas

<sup>5</sup> Moran (1948) propôs esta estatística de teste utilizando uma medida de autocovariância na forma de produto cruzado

da seguinte forma: 
$$I = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

as regiões. No entanto, é possível encontrar alguns agrupamentos espaciais locais que apresentam autocorrelação espacial negativa. Segundo Anselin (1995), um *Local Indicator of Spatial Association* (LISA) tem o objetivo de detectar padrões locais de autocorrelação espacial nas variáveis. Estas estatísticas identificam, para cada observação, os clusters espaciais significativos estatisticamente, em função da média dos vizinhos. Neste estudo, calculou-se o indicador LISA para a taxa de incidência de DRSAI por mil habitantes para o ano de 2010<sup>6</sup>. A Figura A.1 do Anexo identifica tais *clusters*.

A decisão de escolha entre os modelos para especificação final envolve a análise dos testes de Multiplicador de Lagrange (LM), em suas versões tradicionais e robustas.

O teste LM-erro assume que  $\rho = 0$  e testa a hipótese nula  $H_0: \lambda = 0$ . Se esta hipótese nula for rejeitada, a escolha recai sobre o modelo SEM. Analogamente, o teste LM-lag assume que  $\lambda = 0$  e a hipótese nula a ser testada é  $H_0: \rho = 0$ . Neste caso, se a hipótese nula for rejeitada, existem evidências de que se trata de um modelo de defasagem espacial (FLORAX et al., 2003). As versões robustas de ambos os testes estabelecem, no primeiro caso, a mesma hipótese nula, porém não assume  $\rho = 0$ . Ou seja, permite que esse parâmetro seja de qualquer valor e, portanto, admite a presença da defasagem espacial, ainda que o valor desta seja desconhecido. No segundo caso, não assume  $\lambda = 0$ . A interpretação dos coeficientes em um modelo espacial é diferente da forma usual, ou seja, não se interpreta mais como o impacto marginal da variável independente sobre a variável dependente. Quando os efeitos da interação entre os distritos são levados em consideração, o efeito marginal passa a contar com a matriz da equação 3.7, tomando como exemplo o modelo SDM:

$$(3.7) \quad \frac{\partial IncDr sai}{\partial San} = (I - \rho W)^{-1} [W\theta + \gamma I]$$

O cálculo das derivadas parciais desse modelo são feitos por uma matriz não diagonal, que é produto de duas outras, também não diagonais:  $(I - \rho W)^{-1}$  e  $[W\theta + \gamma I]$ . Quando há *spillovers* no modelo ( $\rho \neq 0$ ), os termos fora da diagonal principal são diferentes de zero. Logo, quando a taxa de internações por DRSAI de um distrito varia, todas as observações do sistema são influenciadas. O diferencial do modelo SDM é que além dos efeitos associados à correlação espacial endógena há também os efeitos da correlação espacial exógena. Esses dois efeitos implicam em *spillovers* espaciais para os demais distritos, tanto via variável dependente diretamente como via variável independente.

LeSage e Pace (2009) definem o efeito direto (ED) como a média dos elementos da diagonal principal da matriz de derivadas parciais (efeitos da variação de uma variável explicativa na variável dependente da própria observação). Logo, é o impacto médio na taxa de incidência de internações por DRSAI ( $IncDr sai_i$ ) quando ocorre variação de 1% na variável de saneamento ( $San_i$ ). O efeito indireto (EI) é a média das somas dos elementos não diagonais em cada uma das linhas (ou colunas) da matriz (efeitos da variação de uma variável explicativa na variável dependente de outras observações). Dessa forma, tem-se o impacto médio na taxa de incidência de internações por DRSAI ( $IncDr sai_i$ ) quando ocorre variação de 1% na variável de saneamento ( $San_j$ ). O efeito total (ET) consiste na soma dos efeitos diretos e indiretos.

### 3.3.2 O modelo Spatial Durbin Tobit

Em muitas análises empíricas, as informações da variável dependente não são observadas por causa de um determinado tipo de censura. Em relação ao fenômeno de estudo aqui apresentado, a censura dos dados deve-se à não ocorrência de casos de internações por DRSAI nos distritos. Nesse contexto, são distritos saudáveis os que não registraram nenhuma internação por DRSAI para qualquer uma das

---

<sup>6</sup> Os distritos que possuem altas taxas de incidência de DRSAI e que são cercados por distritos cuja média das taxas de incidência de DRSAI também é alta fazem parte da aglomeração ou cluster denominado Alto-Alto. O cluster Baixo-Baixo indica que os distritos pertencentes a estas aglomerações têm baixas taxas de incidência de DRSAI e possuem vizinhos cuja média também é baixa. Logo, nesses dois casos, há autocorrelação espacial positiva. Ao contrário, a autocorrelação espacial negativa ocorre quando um determinado distrito possui alta (baixa) taxa de incidência e é cercado por vizinhos que, conjuntamente, apresentam um valor baixo (alto) para a média das taxas de incidência de DRSAI. Ou seja, trata-se do cluster Alto-Baixo (Baixo-Alto).

duas populações analisadas. Conforme descrito anteriormente, a base de dados utilizada nesse estudo apresenta 3.703 observações (distritos), de um total de 10.225 observações, que apresentam valor igual a zero para a taxa de internações por DRSAI no ano de 2010 em relação à população de todas as idades. Quando se trata da população de 0 a 14 anos esse valor é de 4.174. Para tais situações, LeSage e Pace (2009) sugerem o uso de um modelo Tobit Espacial.

De acordo com Tobin (1958), a variável explicada neste caso é contínua ao longo dos valores positivos, mas parte da amostra é composta por observações que têm valores nulos para essa variável. Se a variável latente ( $y^*$ ) supera um nível crítico ( $c$ ) os valores da variável latente são observados; caso contrário, observam-se valores nulos.

$$(3.8) y = \begin{cases} y^* & \text{se } y^* > c \\ 0 & \text{se } y^* \leq c \end{cases}$$

A estimação consistente dos parâmetros do modelo é realizada por meio do método de máxima verossimilhança (MV). Almeida (2012) ressalta que existem duas complicações para um modelo Tobit com dados espaciais: a primeira delas refere-se à dificuldade computacional de se utilizar o estimador de máxima verossimilhança num contexto de dependência espacial. A segunda complicação refere-se ao fato do modelo Tobit ser muito sensível a hipóteses sobre a distribuição dos termos de erro. Se houver heterocedasticidade ou não normalidade dos erros, a estimação torna-se inconsistente.

Se existe incerteza acerca de um vetor de parâmetros de, por exemplo, um modelo Durbin Espacial (SDM),  $\phi = (\rho, \beta, \theta, \sigma^2)$ , é possível representá-lo por uma distribuição de probabilidade, denominada **distribuição de probabilidade a priori**,  $f(\phi)$ . O principal objetivo da abordagem Bayesiana é atualizar essa incerteza inicial por meio dos dados empíricos ( $D$ ) a fim de produzir uma **distribuição de probabilidade a posteriori**,  $f(\phi/D)$ , com menor incerteza (LYNCH, 2006). A partir do teorema de Bayes, a seguinte relação pode ser obtida:

$$(3.9) f(\phi/D) = \frac{f(D/\phi)f(\phi)}{f(D)}$$

Em que,  $f(D/\phi)$  é a probabilidade de ocorrência dos dados condicionada nos parâmetros e  $f(D)$  é a probabilidade marginal (ou probabilidade simples) dos dados empíricos.

De acordo com Golgher (2015), se for possível e, teoricamente plausível, deve-se escolher o formato da distribuição a priori de forma que a distribuição a posteriori resultante seja de um formato conhecido. No entanto, quando a distribuição a posteriori não pode ser obtida analiticamente ou se a utilização desta é complicada, em alguns casos utiliza-se o algoritmo de Gibbs<sup>7</sup>, o mais básico dentre os métodos bayesianos com Monte Carlo via cadeias de Markov (MCMC), para gerar as estimativas de forma indireta. A partir de uma distribuição conjunta, o algoritmo encontra as distribuições condicionais em formato fechado. A implementação desse algoritmo parte de um valor inicial de uma das variáveis e, por meio de um valor conhecido para  $\rho$  obtém-se um conjunto de valores de  $x$  e outro de  $y$ . Dado que ambos os valores de cada grupo dependem do valor inicial, deve-se descartar os primeiros valores de cada grupo.

A estimativas bayesianas do modelo heterocedástico Tobit Durbin Espacial (SDMT) foram obtidas a partir das funções disponibilizadas no site do LeSage e Pace (2009). É importante destacar que a distribuição  $f(\phi)$  para o parâmetro de lag espacial é uniforme, definida no intervalo  $\rho \in (-1; 1)$  e, portanto, todos os valores nesse intervalo têm a mesma probabilidade inicial,  $f(\rho) = 1/2$ . O procedimento foi realizado 2.500 vezes, mas, destas, as 500 primeiras foram descartadas. Logo, o procedimento de obter um número aleatório de uma distribuição com média e variância conhecidas foi realizado 2000 vezes.

### 3.3.3 Hipóteses

Nesse artigo, são estimados dois tipos de especificação: Modelo 1 e Modelo 2. Para cada um deles, haverá duas variáveis dependentes representativas de duas faixas etárias da população: a de todas as idades e a de 0 a 14 anos. Com exceção das variáveis de saneamento do vetor  $San$ , as outras variáveis de controle ( $X$ ) são as mesmas para os dois modelos. No primeiro deles, o vetor relativo às variáveis

<sup>7</sup> Para maiores detalhes ver Lacombe (2008).



que expressam o tipo de acesso aos serviços de saneamento básico contém três formas inadequadas de acesso à água e esgoto (“d\_cist”, “d\_frud” e “sbceuab”) sendo que uma delas refere-se às piores condições em termos de serviços de esgotamento sanitário (“sbceuab”). O **Modelo 1** é expresso conforme a equação 3.10:

$$(3.10) \text{ IncDr}_{sai_i} = D\_cist_i \gamma_1 + Sbceuab_i \gamma_2 + D\_frud_i \gamma_3 + X_i \beta + \varepsilon_i$$

O segundo modelo contém duas variáveis representativas das melhores condições de acesso aos serviços de água e esgoto (“ag\_ nao\_ ceuab” e “esg\_ nao\_ ceuab”) e duas outras representativas das condições inadequadas de acesso à coleta de lixo e moradia inadequada e drenagem de águas pluviais (“lixo\_ mor\_ inad” e “d\_ nao\_ bueiro”). A equação (3.11) expressa a especificação do **Modelo 2**:

$$(3.11) \text{ IncDr}_{sai_i} = Ag\_nao\_ceuab_i \alpha_1 + Esg\_nao\_ceuab_i \alpha_2 + Lixo\_mor\_inad_i \alpha_3 + D\_nao\_bueiro_i \alpha_3 + X_i \beta + \varepsilon_i$$

No primeiro modelo, a primeira e a terceira variáveis expressam formas alternativas de inadequada provisão dos serviços de abastecimento de água e coleta de esgoto, respectivamente. A segunda variável, por sua vez, reflete ausência total do serviço de esgotamento sanitário: além dos domicílios não terem banheiro, estão inseridos em condições adversas de esgoto a céu aberto<sup>8</sup>.

O modelo teórico desenvolvido na seção 3.1 também corrobora essa associação. Em relação à variável de saneamento a implicação do modelo é que, por ser uma variável exógena, pode melhorar ou retardar a saúde (BERNDT, 2000). Nessa situação, por se tratar de variáveis “adversas” de infraestrutura, a variável de saúde é negativamente impactada via curva de eficiência marginal do investimento em saúde (EMH). Logo, de acordo com os estudos mencionados e com o modelo teórico, é esperado que piores condições de acesso a tais serviços correspondam a uma maior propagação de DRSAIs e, conseqüentemente, maiores taxas de internação, o que justifica o esperado sinal positivo para os coeficientes estimados para essas variáveis.

No segundo modelo, há dois efeitos a serem analisados: o efeito das variáveis que refletem as melhores condições de saneamento (“ag\_ nao\_ ceuab” e “esg\_ nao\_ ceuab”) e o efeito daquelas que refletem ausência de políticas de manejo de resíduos sólidos e drenagem de águas pluviais. Em relação às duas primeiras variáveis, a análise é complementar à realizada para o primeiro modelo, porém com sinal contrário. Espera-se que os domicílios que tenham acesso a água e esgoto e não estejam circundados por esgoto a céu aberto apresentem menor taxa de internações por DRSAI.

As duas últimas variáveis, o acesso ao lixo e à drenagem de águas pluviais, representam dois problemas pouco estudados na literatura que tratam da associação de tais variáveis com melhores indicadores de saúde<sup>9</sup>. Em relação aos efeitos espaciais, as hipóteses estabelecidas nesse estudo são:

$$(3.12) H_0: \rho \neq 0 \text{ e } \lambda = 0$$

$$(3.13) H_0: \theta \neq 0$$

$$(3.14) H_0: \lambda \neq 0 \text{ e } \rho = 0$$

O primeiro deles, expresso na equação 3.12, refere-se ao efeito transbordamento, associado à taxa de incidência de internações por DRSAI de um distrito. A própria definição de tais doenças, infectocontagiosas, pressupõe a possibilidade de existir interações endógenas entre diferentes distritos. Uma dessas interações pode ocorrer por meio do deslocamento do inseto de um distrito a outro. Os insetos são considerados como vetores mecânicos que carregam uma diversidade de microrganismos transmissores e podem invadir áreas habitadas próximas e transmitir doenças. A outra forma de interação entre os distritos é via água contaminada dos rios e mares que atravessam diversas localidades. Em geral, a transmissão dos agentes patogênicos que causam as DRSAI ocorre pela ingestão de água contaminada por fezes, urina ou vômitos de doente ou portador.

<sup>8</sup> Em relação aos serviços de abastecimento de água (via quantidade ou qualidade), resultados nessa direção também foram encontrados para outros países por Snow (1990) e Fewtrell et al. (2005) e, para o caso brasileiro, por Kassouf (1994) e Andreazzi et al. (2007). Além disso, Duflo et al. (2015), Fewtrell et al. (2005), Phillippi e Malheiros (2005) e Heller (1997), este último para o Brasil, apontam que as duas formas principais de intervenção em saneamento (água e esgoto) devem ser combinadas em ação conjunta e efetiva para a redução de incidência de DRSAIs.

<sup>9</sup> Segundo Rego et al. (2004) e Moraes et al. (2003), acredita-se que a falta ou inadequado fornecimento desses dois serviços gere maiores taxas de internações por DRSAI, dado que também são serviços inseridos na concepção abrangente de saneamento básico.

Outro feito espacial que será considerado é o representado na equação 3.13. A contribuição para a incidência das internações por DRSAI de um determinado distrito pode estar associada à infraestrutura de saneamento básico do distrito vizinho. Natal *et al.* (2005) defendem que, do ponto de vista ecológico, a cidade não está isolada e tanto pode importar materiais e alimentos como pode devolver seus resíduos para as áreas vizinhas.

O terceiro efeito espacial possível, que também será testado neste estudo, é a presença da heterogeneidade espacial, equação 3.14. De acordo com Golgher (2015), regiões próximas tendem a ter características similares, o que pode ser causado por externalidades ou podem representar um equilíbrio de longo prazo devido a fatores históricos em comum ocorridos no desenvolvimento das localidades. Ou seja, a heterogeneidade espacial é relacionada com a estabilidade observada no espaço com relação a variáveis que apresentam estrutura espacial desigual, fato observado em diversos dos indicadores socioeconômicos.

A situação acima descrita pode se aplicar ao setor de saneamento, pois as empresas (públicas ou privadas) que oferecem os serviços à população atuam em diversos municípios, com diferentes níveis de eficiência e abrangências geográficas, o que pode fazer com que desigualdades estruturais se propaguem e afetem tanto os indicadores de acesso e, conseqüentemente, as taxas de internações por DRSAI.

#### 4. Resultados

Os testes LM foram importantes para nortear a escolha do melhor modelo e para verificar se os resultados iriam variar ao se utilizar diferentes matrizes de ponderação. O modelo SDM com a matriz de contigüidade se mostrou o mais adequado, na maior parte dos casos e foi escolhido para servir de comparação ao SDM Tobit. Além disso, as expectativas teóricas também justificam a escolha desse modelo.

**Tabela 4.1: Testes LM para o Modelo 1 e 2 – pop. de todas as idades e de 0 a 14 anos.**

	INC DRSAI (pop. todas as idades)				INC DRSAI (pop. de 0 a 14 anos)			
	Wcontigüidade		Wdistancia_inversa		MQO - Wcont		MQO - Wdist_inv	
	Estatística	p-valor	Estatística	p-valor	Estatística	p-valor	Estatística	p-valor
<b>Modelo 1</b>								
<b>LM-erro</b>	243,741	0,00	1417,56	0,00	85,693	0,00	447,71	0,00
<b>LM-lag</b>	243,979	0,00	1314,92	0,00	79,708	0,00	415,27	0,00
<b>RLM-erro</b>	1,698	0,19	133,53	0,00	8,882	0,00	34,70	0,00
<b>RLM-lag</b>	1,937	0,16	30,89	0,00	2,897	0,09	2,27	0,13
<b>Modelo 2</b>	Estatística	p-valor	Estatística	p-valor	Estatística	p-valor	Estatística	p-valor
<b>LM-erro</b>	196,328	0,0000	239,27	0,000	215,55	0,000	449,985	0,000
<b>LM-lag</b>	187,225	0,0000	241,51	0,000	224,385	0,000	422,706	0,000
<b>RLM-erro</b>	12,563	0,0004	0,83	0,362	16,951	0,000	30,242	0,000
<b>RLM-lag</b>	3,459	0,0629	3,08	0,079	25,786	0,000	2,964	0,085

Fonte: Elaboração própria a partir de dados da pesquisa.

A estimação do SDM Tobit foi realizada para os Modelos 1 e 2 e, cada um deles, considera as duas populações de interesse. Nas próximas tabelas, foram realizadas comparações entre o modelo MQO (sem qualquer dependência espacial), o modelo SDM (dependência espacial na forma de defasagem da variável dependente e das exógenas) e o modelo SDMT (SDM com controle dos dados censurados estimado por um Tobit Bayesiano). A Tabela 4.2 apresenta e compara tais modelos para a primeira especificação (Modelo 1), somente com variáveis inadequadas de saneamento, para a população de todas as idades.

**Tabela 4.2: MQO, SDM e SDMT - Pop. de todas as idades e de 0 a 14 anos do Modelo 1**

MODELO 1	Pop de todas idades			Pop. de 0 a 14 anos		
	MQO	SDM	SDMT	MQO	SDM	SDMT
<b>dcist</b>	<b>0,0823***</b>	<b>0,0577**</b>	<b>0,1429***</b>	<b>0,1372***</b>	<b>0,1055*</b>	<b>0,243***</b>
<b>sbceuab</b>	<b>0,7554***</b>	0,2074	-0,0015	<b>0,879***</b>	0,1315	-0,444
<b>dfrud</b>	<b>0,023***</b>	<b>0,0133***</b>	-0,0024	<b>0,0403***</b>	<b>0,0405***</b>	0,004
rendapc	-0,0012**	0,0002	-0,0003	-0,0019**	-0,0001	-0,001
gini	12,3859***	4,3069**	15,9892***	18,5096***	6,0992	43,309***
mu_resp_alfab	-0,0307***	0,0079	0,2269***	-0,0264	0,0647**	0,334***
tpopurb	0,0536***	0,0607***	0,1821***	0,0996***	0,1082***	0,43***
dom_kmq	-0,0003	-0,0003	-0,0001	-0,0005	-0,0015***	-0,001**
est_saude	0,0003	-0,0016	-0,0027	0,0025	0,0012	0,005
dummy_capital	-1,6921**	-1,7087	-3,658**	-2,5055*	-7,6224***	-13,102***
dummy_rm	-2,7477***	-0,9583	-3,7953***	-4,8263***	-3,366	-9,753***
<b>Wdcist</b>		0,0096	0,0216		-0,0051	0,128
<b>Wsbceuab</b>		<b>1,3223***</b>	<b>2,481***</b>		<b>1,7543***</b>	<b>4,827***</b>
<b>Wdfrud</b>		0,0058	<b>0,0283***</b>		-0,0127	<b>0,04**</b>
<b>Wrendapc</b>		-0,0009	-0,0006		-0,0017	-0,002
<b>Wgini</b>		5,695**	6,1999*		10,622*	7,813
<b>Wmu_resp_alfab</b>		-0,0225	-0,2444***		-0,0847**	-0,37***
<b>Wtpopurb</b>		-0,0391***	-0,0908***		-0,0614***	-0,203***
<b>Wdom_kmq</b>		0,0003	0,0002*		0,0023***	0,001***
<b>West_saude</b>		-0,0028	0,0042		-0,0134*	0,001
<b>Wdummy_capital</b>		0,7381	1,7968		6,8834**	11,027***
<b>Wdummy_rm</b>		-1,0273	0,3859		-0,8319	2,051
Constante		-4,2876***	-19,2627***		-6,7549*	-47,407***
Rho		0,2251***	0,2446***		0,1447***	0,2***

\*\*\*Significante a 1%; \*\*Significante a 5%; \*Significante a 10%.

Fonte: Elaboração própria a partir de dados da pesquisa.

A análise da comparação entre modelos SDM e SDMT, para a população de todas as idades, mostra que o sinal e a significância da maior parte dos coeficientes se mantiveram, com exceção da variável *dfrud*, que não foi significativa no coeficiente representativo do efeito direto, mas o foi em relação à variável defasada. O contrário se evidencia no modelo SDM para essa variável: o coeficiente da variável não defasada foi significativo, mas o da defasada não. A magnitude dos coeficientes se alteraram em todos os casos. Ou seja, as primeiras estimações, sem o controle dos zeros, estão subestimando o efeito do saneamento sobre a saúde. Segundo LeSage e Pace (2009), as estimativas de um modelo com dependência espacial na forma de defasagem tendem a ter os coeficientes subestimados pela existência de um viés para baixo quando não se faz o controle da censura dos dados.

A mesma conclusão obtida anteriormente pode ser aplicada à análise dos dados para a população de 0 a 14 anos. Os sinais e significância da maior parte das estimativas são mantidos, mas a magnitude das mesmas é maior para o modelo SDMT. Em geral, a população de 0 a 14 anos é mais afetada pelos efeitos deletérios da inadequada provisão dos serviços de saneamento básico.

O efeito de transbordamento da taxa de incidência de internações por DRSAI ( $\rho$ ) foi positivo e significativo com maior magnitude para o modelo SDMT para ambas populações. Esse resultado era esperado pela própria definição de doenças infectocontagiosas e, portanto, a hipótese estabelecida na equação 3.12 foi confirmada. A transmissão dos agentes patogênicos causadores das DRSAI que pode ocorrer tanto pelo deslocamento dos insetos de um distrito para o outro quanto pela água contaminada dos rios e mares que atravessam diversas localidades, gera o efeito de *spillover* espacial de tal forma

que a incidência de DRSAI em um determinado distrito seja influenciada pela média da taxa de incidência de DRSAI dos distritos vizinhos. Além disso, a hipótese estabelecida nessa mesma seção, e expressa na equação 3.13, de que a defasagem do vetor de variáveis explicativas pode importar para explicar a taxa de incidência de internações por DRSAI foi confirmada. As condições de infraestrutura de saneamento básico da vizinhança de um determinado distrito são capazes de afetar essa variável, conforme o modelo Durbin prevê. De acordo com Natal *et al.* (2005), há diversos exemplos de distritos ou municípios que lançam águas servidas sem tratamento, dejetos e outros resíduos coletados em suas regiões para fora de seus limites territoriais e que, conseqüentemente, atingem seus vizinhos mais próximos.

Os coeficientes de um modelo espacial não são informativos em termos de análise do impacto marginal. Nesse sentido, os efeitos marginais também são calculados para o SDM e SDMT para os dois modelos e duas populações. A Tabela 4.3 apresenta os resultados dos efeitos marginais da população de todas as idades do Modelo 1.

**Tabela 4.3: Comparação entre efeitos marginais (SDM e SDMT)**

Inc DRSAI (todas idades)	ED		EI		ET	
	SDMT	SDM	SDMT	SDM	SDMT	SDM
dcist	0,145***	0,0586***	0,072	0,0282	0,217***	0,08684
sbceuab	0,113	0,2653**	3,168***	1,7087***	3,281***	1,97399***
dfrud	-0,001	0,0137***	0,036***	0,011	0,035***	0,02467*
rendapc	-0,0003	0,0002	-0,0009	-0,0011	-0,001	-0,00089
gini	16,41***	4,589**	12,864***	8,3185	29,274***	12,90743
mu_resp_alfab	0,218***	0,007	-0,242***	-0,0258	-0,024	-0,01879
tpopurb	0,18***	0,0597***	-0,059***	-0,0318*	0,121***	0,02788
dom_kmq	-0,0001	-0,0002	0,0002	0,0003	0,0001	0,00003
est_saude	-0,0026	-0,0017	0,005	-0,0039	0,002	-0,00567
dummy_capital	-3,632***	-1,6937**	1,152	0,4412	-2,48	-1,25248
dummy_rm	-3,807***	-1,0109	-0,693	-1,5515	-4,499***	-2,56243*

\*\*\*Significante a 1%; \*\*Significante a 5%; \*Significante a 10%.

Fonte: Elaboração própria a partir de dados da pesquisa.

Os efeitos marginais para a população de todas as idades do modelo SDMT também apresentaram magnitudes superiores em relação ao modelo SDM. Porém, na coluna dos efeitos diretos, comparando esses dois modelos, duas das variáveis de interesse perderam significância: *sbceuab* e *dfrud*. Apesar disso, a variável *dfrud* foi significativa para os efeitos indiretos e, para esses mesmos efeitos, a variável *sbceuab* manteve-se significativa e com maior magnitude. A variável *dcist* tornou-se significativa nos efeitos totais, dado que no SDM não era.

Os três efeitos (direto, indireto e total) para a variável *sbceuab* vão ao encontro ao resultado de Barreto *et al* (2007). Os autores apontam que ter banheiro no interior do domicílio (80% dos domicílios da cidade de Salvador possuem banheiro) não explica a redução da diarreia. O aumento da cobertura de esgotamento sanitário de toda a vizinhança foi capaz de reduzir a poluição fecal do domínio público e, portanto, explicar a queda nesse indicador de morbidade. Ou seja, segundo os autores, a prevenção da transmissão da diarreia se deu principalmente na vizinhança e não no ambiente doméstico dos domicílios cujos fatores de risco apresentaram pequena importância. As contribuições do estudo de tais autores e dos resultados apresentados acima encontram-se na amplitude espacial das externalidades geradas aos vizinhos pela ausência de esgotamento sanitário. Caso esse efeito espacial não seja considerado, o impacto positivo do saneamento inadequado sobre o indicador de morbidade será subestimado.

Além destes, outros autores tais como Bleakley (2007), Miguel e Kremer (2004), Esrey (1991), Kassouf (1994), Moraes *et al.* (2003), defendem a relação positiva entre variáveis de saúde (na maior

parte medida pela incidência de doenças diarreicas) e o acesso aos serviços de esgotamento sanitário<sup>10</sup>.

O modelo teórico da seção também corrobora essa associação. Uma das variáveis que pode determinar a produção de saúde dos indivíduos é aquela relacionada ao ambiente em que o mesmo reside, ou seja, as condições de infraestrutura de saneamento básico do domicílio (variável “*ST<sub>i</sub>*”). Em relação a essa variável, a implicação do modelo é que, por ser uma variável exógena, pode melhorar ou retardar a saúde (BERNDT, 2000). No contexto dos resultados apresentados, pelo fato das variáveis de condições de saneamento desse estudo se tratarem de variáveis “adversas” para a saúde humana, o impacto de mudanças na eficiência marginal do investimento em saúde (EMH), provoca queda da quantidade ótima de capital de saúde ( $H_i^*$ ).

Em relação ao efeito total, algumas variáveis de controle tornaram-se importantes para explicar a taxa de incidência de internações por DRSAI: *gini*, *tpopurb* e *dummy\_rm*. As duas primeiras apresentaram coeficientes positivos e significativos e, a última, coeficiente negativo e significativo. Quanto maior o grau de urbanização do distrito em questão e dos distritos vizinhos, maior a incidência de internações por DRSAI desse distrito. A maior concentração da renda, medida pelo coeficiente Gini, leva a maior incidência de DRSAI em valores significativos e de grande magnitude. Por fim, o efeito da *dummy* de região metropolitana se mostrou significativo e negativo, como era esperado, dado que em tais regiões a provisão dos serviços de saneamento e saúde tende a ser melhor e mais homogênea. Na Tabela 4.4, a mesma comparação foi repetida para a população de 0 a 14 anos de idade para o mesmo Modelo 1.

**Tabela 4.4: Comparação entre efeitos marginais (SDM e SDMT)**

Inc DRSAI (de 0 a 14 anos)	ED		EI		ET	
	SDMT	SDM	SDMT	SDM	SDMT	SDM
dcist	0,25***	0,1049*	0,215	0,0137	0,465***	0,1185*
sbceuab	-0,267	0,1827	5,733***	2,0347***	5,466***	2,2174***
dfrud	0,006	0,0381***	0,049**	-0,0055	0,055***	0,0325***
rendapc	-0,002	-0,0001	-0,003	-0,0019	-0,004	-0,002
gini	43,908***	6,1355	20,266	13,0663**	64,174***	19,2018***
mu_resp_alfab	0,324***	0,0559**	-0,368***	-0,0826**	-0,044	-0,0267
tpopurb	0,426***	0,1075***	-0,142***	-0,0516***	0,284***	0,0559***
dom_kmq	-0,001	-0,0015***	0,001***	0,0024***	0,001***	0,0009
est_saude	0,005	0,0009	0,002	-0,0153*	0,007	-0,0144
dummy_capital	-12,8	-7,4961***	10,169**	6,6577**	-2,631	-0,8384
dummy_rm	-9,815***	-3,4182	0,138	-1,4363	-9,677***	-4,8545***

\*\*\*Significante a 1%; \*\*Significante a 5%; \*Significante a 10%.

Fonte: Elaboração própria a partir de dados da pesquisa.

A comparação dos efeitos marginais dos dois modelos (SDM e SDMT) seguiu a mesma tendência da análise dos coeficientes realizada na tabela anterior com a população de todas as idades no tocante à magnitude e significância dos coeficientes. Logo, a hipótese de que a população de zero a catorze anos é mais atingida por DRSAI por estar mais sujeita às doenças infecto-parasitárias transmitidas pela água, esgoto não tratado ou pelos insetos vetores foi verificada pela magnitude dos coeficientes apresentados.

Esse resultado está em consonância com o estudo de Geruso e Spears (2015) que, apesar de ser aplicado para outro país (Índia) e a variável de análise se tratar da mortalidade infantil, investigam as externalidades do esgoto a céu aberto e mostram que é mais provável que crianças muçulmanas sobrevivam até seu primeiro aniversário em comparação às crianças hindus, pois a vizinhança das

<sup>10</sup> Alguns desses estudos, tais como Bleakley (2007), Miguel e Kremer (2004) e Kassouf (1994), indicaram o impacto de saneamento em educação, mas por meio do estado de saúde do indivíduo que deteriora o perfil de morbidade do mesmo.

crianças muçulmanas é formada em grande parte por vizinhos muçulmanos, que são menos propensos a defecar a céu aberto<sup>11</sup>. A comparação dos coeficientes para o Modelo 2 foi realizada somente para a população de 0 a 14 anos, de acordo com a Tabela 4.5<sup>12</sup>.

**Tabela 4.5: MQO, SDM e SDMT - Pop. de 0 a 14 anos do Modelo 2**

Variáveis	MQO	SDM	SDMT
<b>esg_nao_ceuab</b>	<b>-0,0356***</b>	-0,0056	0,0044
<b>ag_nao_ceuab</b>	<b>-0,0486***</b>	<b>-0,0588***</b>	<b>-0,0583***</b>
<b>lixo_mor_inad</b>	0,0517	-0,1026	-0,5544
<b>d_nbueiro</b>	0,0154	-0,0182	0,0107
rendapc	-0,0009	0,0002	-0,0009
gini	16,2374***	1,6238	12,8452***
mu_resp_alfab	-0,038**	0,0919***	-0,0089
tpopurb	0,1385***	0,1557***	0,1417***
dom_kmq	-0,0004	-0,0015***	0,0003
est_saude	0,0027	0,0001	0,0065*
dummy_capital	-3,1044**	-7,683***	-0,8451
dummy_rm	-5,4576***	-3,5105	-2,3734**
<b>Wesg_nao_ceuab</b>	-	<b>-0,0347*</b>	<b>-0,0495**</b>
<b>Wag_nao_ceuab</b>	-	0,0304	-0,0172
<b>Wlixo_mor_inad</b>	-	0,9136	0,1172
<b>Wd_nbueiro</b>	-	<b>0,0523**</b>	-0,0002
<b>Wrendapc</b>	-	-0,0007	-0,0028
<b>Wgini</b>	-	10,443	57,3316***
<b>Wmu_resp_alfab</b>	-	-0,1361***	-0,2743***
<b>Wtpopurb</b>	-	-0,0888***	0,3249***
<b>Wdom_kmq</b>	-	0,0026***	-0,0014**
<b>West_saude</b>	-	-0,0129*	-0,0055
<b>Wdummy_capital</b>	-	6,5008	-4,5744*
<b>Wdummy_rm</b>	-	-1,0463	-10,8401***
Constante	-	-0,8823	-42,1422***
Rho	-	0,14212***	0,6237***

\*\*\*Significante a 1%; \*\*Significante a 5%; \*Significante a 10%.

Fonte: Elaboração própria a partir de dados da pesquisa.

Nesse caso, não há diferenças significativas, com exceção do coeficiente de “*Wd\_nbueiro*” que perdeu significância no modelo SDMT. Os efeitos marginais para essa população de 0 a 14 anos do modelo 2 são apresentados na Tabela 4.6.

<sup>11</sup> Os autores solucionam um "puzzle" dos dados indianos: é mais provável que crianças muçulmanas sobrevivam até seu primeiro aniversário do que crianças hindus, mesmo que as primeiras tenham menor riqueza, consumo, frequência escolar e pior acesso aos serviços públicos em comparação à maioria hindu. Isso ocorre por dois motivos. Um deles deve-se ao fato da população hindu considerar "impuro" o uso de banheiros e fossas, defendendo que defecar a céu aberto faz parte de um estilo de vida: mais da metade da população indiana, 600 milhões de pessoas, tem essa prática. O segundo motivo deve-se ao efeito vizinhança já relatado.

<sup>12</sup> Na análise da escolha entre modelos, o SDM para a população de todas as idades foi estimado com a matriz inversa da distância. Dessa forma, não foi possível compará-lo com o SDMT estimado com a matriz de contiguidade.

**Tabela 4.6: Comparação entre efeitos marginais (SDM e SDMT) – Modelo 2**

INC DRSAI (de 0 a 14 anos)	ED		EI		ET	
	SDMT	SDM	SDMT	SDM	SDMT	SDM
esg_ao_ceuab	-0,0032	-0,002	-0,1174	-0,0443**	-0,1206	-0,0463***
ag_ao_ceuab	-0,0674***	-0,0663***	-0,1339	0,0429	-0,2013**	-0,0234
lixo_mor_inad	-0,5859	-0,0547	-0,5616	0,9661	-1,1476	0,9114
d_nbueiro	0,0117	-0,0211	0,0162	0,0683***	0,0279	0,0472**
rendapc	-0,0015	0,0002	-0,0085	-0,0006	-0,01	-0,0004
gini	23,4857***	4,1053	163,1399***	10,7831	186,6257***	14,8883**
mu_resp_alfab	-0,0533**	0,0576**	-0,6952***	-0,1076***	-0,7485***	-0,05*
tpopurb	0,2089***	0,1644***	1,0317***	-0,0991***	1,2406***	0,0653**
dom_kmq	0,0002	-0,0015***	-0,003	0,0026***	-0,0029	0,0012*
est_saude	0,0064	0,0001	-0,0028	-0,0131	0,0036	-0,013
dummy_capital	-1,6916	-8,2409***	-12,7709	6,9385**	-14,4625	-1,3024
dummy_rm	-4,3958***	-3,7947*	-30,8404***	-1,3476	-35,2362***	-5,1423***

\*\*\*Significante a 1%; \*\*Significante a 5%; \*Significante a 10%.

Fonte: Elaboração própria a partir de dados da pesquisa.

Os efeitos das variáveis relativas à lixo e drenagem de águas pluviais passaram a ser não significantes, na estimação do SDMT. Os resultados do modelo SDMT apontam que localidades que têm acesso à água por rede geral e não apresentam em suas condições de entorno esgoto a céu aberto têm menores taxas de incidência de internações, considerando o efeito direto e total.

Esse resultado pode estar relacionado ao fenômeno apontado por Phillippi Jr. e Mallheiros (2005). Segundo os autores, o abastecimento de água potável, quando não acompanhado da implantação conjunta e concomitante do esgotamento sanitário, pode vir a ter um saldo negativo para os vizinhos, em face do lançamento de efluentes sem tratamento ou com tratamento inadequado que poder ser devolvido ao meio ambiente em condições nocivas que ainda apresentem agentes patogênicos. A variável “ag\_ao\_ceuab” pode ser entendida como uma variável que integre essas duas características destacadas pelos autores. Duflo et al. (2015) também defendem a provisão integrada de tais serviços como a forma mais eficiente de afetar os indicadores de saúde.

Os resultados dos efeitos marginais da estimação para a população de todas as idades para o segundo modelo, na versão SDMT com a matriz de contiguidade, encontram-se na Tabela 4.7.

**Tabela 4.7: Efeitos marginais do modelo SDMT com matriz W contígua – Modelo2**

INC DRSAI - pop todas idades	ED	EI	ET
esg_ao_ceuab	-0,0004	-0,055	-0,056
ag_ao_ceuab	-0,02683***	-0,139***	-0,166***
lixo_mor_inad	-0,07773	-1,204	-1,281
d_nbueiro	-0,00478	-0,076	-0,081
rendapc	0,00007	0,001	0,001
gini	3,15531	60,557***	63,712***
mu_resp_alfab	-0,04657***	-0,464***	-0,51***
tpopurb	0,05809***	0,565***	0,623***
dom_kmq	-0,00004	-0,001	-0,001
est_saude	0,00004	-0,002	-0,002
dummy_capital	-0,65814	-8,036	-8,694
dummy_rm	-1,3788***	-13,591***	-14,97***

\*\*\*Significante a 1%; \*\*Significante a 5%; \*Significante a 10%.

Fonte: Elaboração própria a partir de dados da pesquisa.

Tanto os resultados da Tabela 4.6 quanto os da Tabela 4.7 apontam que localidades que têm acesso à água por rede geral e não apresentam em suas condições de entorno esgoto a céu aberto têm menores taxas de incidência de internações. Novamente, a magnitude do coeficiente para o efeito total é maior para a população de 0 a 14 anos.

Neste caso também, as variáveis relativas ao acesso inadequado à drenagem de águas pluviais e ao lixo não foram significativas ao se controlar a censura dos dados. De fato, a maior parte dos estudos que relatam queda da taxa de incidência de diarreia, por exemplo, no Brasil e em outros países do mundo, apontam ser o esgotamento sanitário ou o processo de distribuição e tratamento de água como as ações de saneamento mais eficientes na promoção da saúde.

Por outro lado, a variável relativa à escolaridade das mulheres chefes de família (“mu\_resp\_alfab”) se mostrou relevante para diminuir a taxa de incidência de internações por DRSAI, como prevê a literatura. O coeficiente de Gini apresentou significância estatística na grande parte dos resultados apresentados nesse estudo. Ou seja, quanto maior a concentração de renda no distrito maior é a incidência de DRSAI no mesmo, indicando que variáveis relacionadas à distribuição de renda nos distritos pode ser uma política pública importante para a diminuição da incidência de DRSAI.

## 5. Conclusão

Nesse estudo a estimação do efeito do saneamento sobre a saúde foi realizada para o ano de 2010 e se valeu de análise espacial com o controle da censura dos dados que ocorre quando as informações da variável dependente não são observadas por causa da não ocorrência de casos de internações por DRSAI em alguns distritos. Ou seja, distritos saudáveis que não registraram nenhuma internação por DRSAI. A obtenção dos resultados relatados acima envolveu, inicialmente, a estimação de vários modelos espaciais, sem que fossem controlados os zeros da amostra, para nortear a escolha do melhor modelo e para verificar se os resultados iriam mudar ao se utilizar duas diferentes matrizes de ponderação. Nestas estimações, o modelo Spatial Durbin (SDM) com a matriz de contiguidade se mostrou o mais adequado, na maior parte dos casos. Além disso, as justificativas teóricas e hipóteses mencionadas também serviram de suporte para a escolha deste modelo. Para controlar os zeros da amostra, foi estimado um modelo Spatial Durbin Tobit Espacial (SDMT) por meio da abordagem Bayesiana.

A utilização de dados distritais, não disponível em domínio público e que foram construídos para melhor caracterizar a distribuição espacial das informações foi necessária para justificar de forma mais realista a hipótese acerca do efeito espacial de transbordamento dos agentes patogênicos das DRSAIs entre as regiões.

As variáveis de saneamento utilizadas representam o acesso inadequado aos serviços de água e esgoto e a variável de saúde é representada pela taxa de incidência de internações por DRSAI por mil habitantes. O modelo de Grosman (1972) foi utilizado como arcabouço teórico para a escolha das variáveis dos modelos que foram estimados para populações de duas faixas etárias diferentes: a de todas as idades e a de 0 a 14 anos.

Os resultados indicam, para a população de todas as idades, que o impacto total médio do aumento de 1% na parcela de domicílios servidos por água da chuva armazenada em cisterna ou caixa de cimento é de 0,217 casos por mil habitantes. Quando se considera a porcentagem dos domicílios que têm banheiro conectado a uma fossa rudimentar, esse impacto é de 0,035 casos por 1000 habitantes. Por outro lado, o impacto total médio de 1% na porcentagem de domicílios particulares permanentes conectados à rede de distribuição de água e que, além disso, na quadra onde estão localizados é inexistente condições de esgoto a céu aberto é de -0,166 casos por mil habitantes. Por fim, a variável de maior impacto total médio refere-se à porcentagem de domicílios que não têm banheiro e, além disso, possuem presença de esgoto a céu aberto em suas condições de entorno. O aumento em uma unidade percentual nessa variável gera aumento de 3,281 na variável de saúde, a taxa de incidência de internações por DRSAI.



O efeito de transbordamento espacial da taxa de incidência de internações por DRSAI foi positivo e significativo, com maior magnitude para o modelo SDMT (0,2) em comparação ao modelo SDM (0,14). Esse resultado era esperado pela própria definição de doenças infectocontagiosas das DRSAI. A transmissão dos agentes patogênicos causadores de tais doenças pode ocorrer tanto pelo deslocamento dos insetos de um distrito para o outro quanto pela água contaminada dos rios e mares que atravessam diversas localidades, gerando o efeito de *spillover* espacial, de tal forma que a incidência de DRSAI em um determinado distrito seja influenciada pela média dessa taxa nos distritos vizinhos. Além disso, as condições de infraestrutura de saneamento básico da vizinhança de um determinado distrito também são capazes de afetar essa variável de saúde, devido à importância de algumas variáveis explicativas defasadas espacialmente, conforme o modelo Durbin prevê. É o caso de distritos que lançam águas servidas sem tratamento, dejetos e outros resíduos coletados em suas regiões para fora de seus limites territoriais.

Em relação à sensibilidade dos efeitos para as populações das duas diferentes faixas etárias, pela magnitude das estimativas dos parâmetros dos efeitos marginais, a população de zero a 14 anos é a mais atingida pelo acesso a condições inadequadas dos serviços de saneamento básico, o que confirma uma das hipóteses estabelecidas nesse estudo, de que a população mais jovem é a mais frágil e a mais suscetível a contrair as DRSAIs.

## 6. Referências bibliográficas

- ANSELIN, L. Local indicators of spatial association – LISA. *Geographical Analysis*, v.27, n.2, p. 93-115, 1995.
- ALBERNAZ, A.; FERREIRA, F.H.G.; FRANCO, C. *Qualidade e equidade no ensino fundamental brasileiro. Pesquisa e Planejamento Econômico*, v.32, n.3, p.453-476, 2002.
- ALMEIDA, E. *Econometria Espacial Aplicada*, Campinas, SP: Editora Alínea, 2012.
- ALVES, D.; BELLUZZO, W. *Infant mortality and child health in Brazil. Econ Hum Biol*, 2(3): 391-410, 2004.
- ANSELIN, L. *Spatial econometrics: methods and models*. Boston: **Kluwer Academic**, 1988.
- ANSELIN, L.; BERA, A. *Spatial dependence in linear regression models with an introduction to spatial econometrics*. In: ULLAH A.; GILES D. E. (eds). **Handbook of applied economic statistics**, Nova York: Marcel Dekker, p. 237-289, 1998.
- ANDREAZZI, M. A. R.; BARCELLOS, C.; HACON, S. *Velhos indicadores para novos problemas: a relação entre saneamento e saúde. Revista Panam Salud Publica*, n. 22, 2007.
- BARRETO, M. L. et al. *Impact of a Citywide Sanitation Program in Northeast Brazil on Intestinal Parasites Infection in Young Children, Environ Health Perspect*, Nov; 118(11): 1637–1642, 2010.
- BERNDT, E.R.; Cutler, D.; Frank, R.G.; Griliches, Z.; Newhouse, J.; Triplett, J.E. *Medical care prices and output*. In: **Handbook of Health Economics**, vol. 1. Amsterdam: Elsevier, Cap. 03, p. 119-180, 2000.
- BLEAKLEY, H. *Disease and Development: Evidence from Hookworm Eradication in the American South, The Quarterly Journal of Economics*, 122 (1): 73-117, 2007.
- BLEAKLEY, H. *Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure, American Economic Journal: Applied Economics* v.2 (2), 1-45, 2010.
- CAIRNCROSS, S.; FEACHEM, R. G. *Environmental health engineering in the tropics: an introductory text*. John Wiley & Sons, Chichester, 1990.
- CAIRNCROSS, S.; R. FEACHEM. *Environmental Health Engineering in the Tropics*. 2. ed. Chichester, U.K.: John Wiley & Sons, 1993.
- CUTLER, D., & MILLER, G. *The Role of Public Health Improvements in Health Advances: The Twentieth-Century United States. Demography*, 42 (1), 1-22, 2005.
- DUFLO, E.; MICHAEL GREENSTONE, RAYMOND GUITERAS, AND THOMAS CLASEN. *Toilets Can Work: Short and Medium Run Health Impacts of Addressing Complementarities and Externalities in Water and Sanitation, NBER Working Paper*, No. 21521, 2015.

- ESREY, S.; POTASH, J., ROBERTS, L.; SHIFF, C. *Effects of Improved Water Supply and Sanitation on Ascariasis, Diarrhoea, Dracunculiasis, Hookworm Infection, Schistosomiasis, and Trachoma*. **Bulletin of the World Health Organization**, 69 (5), 1991
- FEWTRELL, L., KAUFMANN, R. *Water Sanitation and Hygiene Interventions to Reduce Diarrhoea in Developed Countries: A Systematic Review and Meta-Analysis*. **Lancet Infectious Diseases**, 5, 2005.
- FLORAX.; FOLMER, H.; REY,S. *Specification searches in spatial econometrics: the relevance of Hendry's methodology*. **Regional Science Urban Economics**, v.22, n.5, p. 557-579, 2003.
- GALLUP, J.; SACHS, J. *The Economic Burden of Malaria*, **American Journal of Tropical Medicine and Hygiene**, 64, 1, 85-96, 2001.
- GERUSO, M.; SPEARS, D. *Neighborhood Sanitation and Infant Mortality*, **National Bureau of Economic Research Working Paper Series**, n. 21184, 2015.
- GOLGHER, A. B. *Introdução à Econometria Espacial*, Jundiaí, Paco Editorial: 2015.
- GROSSMAN, M. *On the concept of health capital and the demand for health*. **Journal of Political Economy**, 80: 2, 223-255, 1972.
- HELLER L. *Saneamento e Saúde*. Brasília: **OPAS**; 1997.
- KASSOUF, A. *A demanda de saúde infantil no Brasil por região e setor*. **Pesquisa e Planejamento Econômico**, n. 24, 235-260, 1994.
- LESAGE, J.; PACE, R. K. *Introduction to spatial econometrics*. Boca Raton: Taylor & Francis Group, 2009.
- MARA, D. D.; FEACHEM, R. G. A. *Water and excreta related diseases: unitary environmental classification*. **Journal of Environmental Engineering**, n.125, 1999.
- MIGUEL, E; KREMER, M. *Worms: identifying impacts on education and health in the presence of treatment externalities*, **Econometrica**, 72 (1), 159-217, jan 2004.
- MORAES et al, L. R. S., CANCIO, J. A., CAIRNCROSS, S.; HUTTLY, S., *Impact of drainage and sewerage on diarrhoea in poor urban areas in Salvador, Brazil*, **Transactions of the Royal Society of Tropical Medicine and Hygiene**, v. 97, 153- 158, 2003.
- SOUSA, T. R. V.; LEITE FILHO, P. A. M. *Análise por Dados em painel do status de saúde no Nordeste brasileiro*, **Revista Saúde Pública**, v.45(2), 796-804, 2008.
- PHILIPPI JR; MALHEIROS, T. F. *Saneamento e Saúde Pública: integrando homem e ambiente*, p.3-31 in: PHILIPPI JR. A (ed). *Saneamento, Saúde e Ambiente: fundamentos para um Desenvolvimento Sustentável*, Barueri, SP: Manole, 2005.
- TOBIN, J. *Estimation of Relationships for Limited Dependent Variables*, **Econometrica**, 26 (1): 24–36, 1958.
- WATSON, T. *Public Health Investments and the Infant Mortality Gap: Evidence from Federal Sanitation Interventions on U.S. Indian Reservations*, **Journal of Public Economics**, 90: 1537-1560, 2006.

## ANEXOS

**Tabela A.1: Número de CEPs encontrados pelo CNEFE, Correios e agregações.**

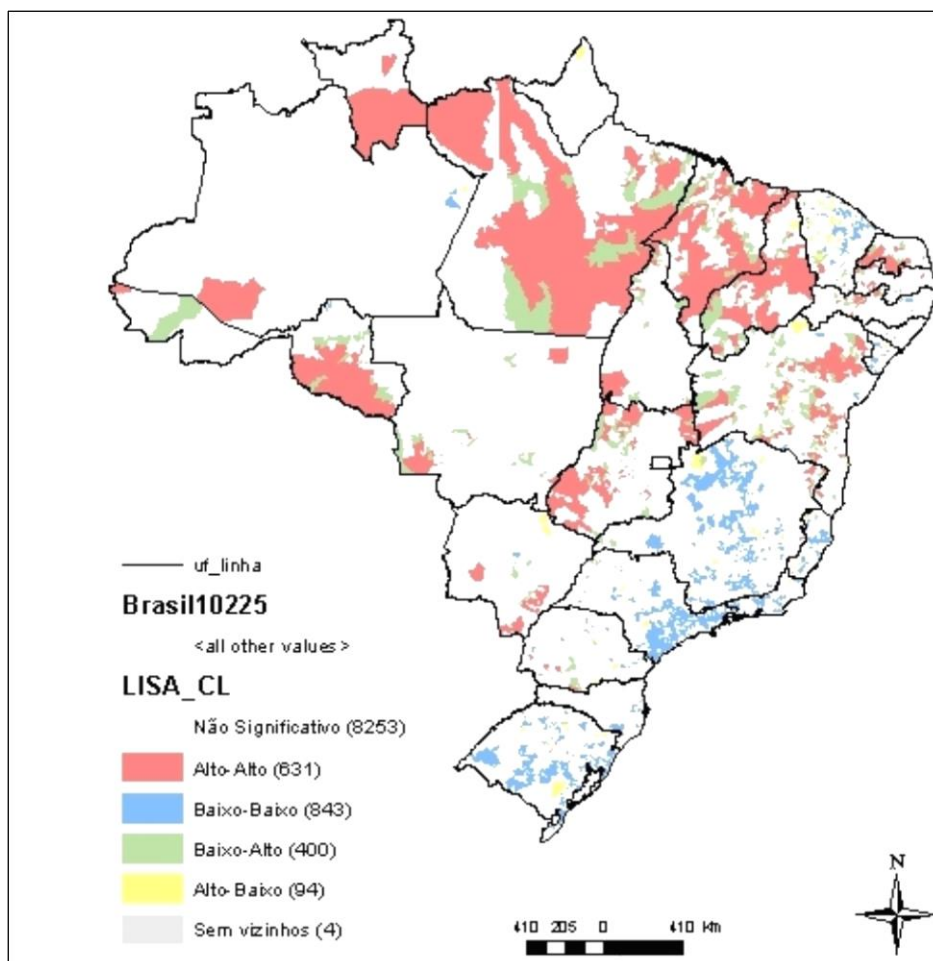
Cod_uf	Distritos agregados	Novos distritos	CNEFE	Correios	Distritos únicos	Distritos antes agregação	Distritos pós agregação
RO	31	20	9	29	35	104	84
AC	0	0	0	2	21	23	23
AM	0	0	11	35	48	94	94
RR	0	0	0	0	15	15	15
PA	27	17	15	118	90	250	233
AP	4	2	1	19	9	33	31
TO	9	5	2	14	128	153	148
<b>Norte</b>	<b>71</b>	<b>44</b>	<b>38</b>	<b>217</b>	<b>346</b>	<b>672</b>	<b>628</b>
MA	0	0	1	48	194	243	243
PI	0	0	6	0	223	229	229
CE	104	67	28	691	21	844	777
RN	2	1	4	28	152	186	185
PB	40	21	5	71	172	288	267
PE	36	22	40	250	78	404	382
AL	2	1	9	22	89	122	121
SE	0	0	0	13	70	83	83
BA	118	66	43	512	195	868	802
<b>Nordeste</b>	<b>302</b>	<b>178</b>	<b>136</b>	<b>1635</b>	<b>1194</b>	<b>3267</b>	<b>3089</b>
MG	273	162	149	872	466	1760	1598
ES	34	19	59	195	11	299	280
RJ	38	26	136	179	24	377	351
SP	34	16	208	339	454	1035	1019
<b>Sudeste</b>	<b>379</b>	<b>223</b>	<b>552</b>	<b>1585</b>	<b>955</b>	<b>3471</b>	<b>3248</b>
PR	44	24	92	470	211	817	793
SC	37	20	28	197	198	460	440
RS	177	117	142	790	238	1347	1230
<b>Sul</b>	<b>258</b>	<b>161</b>	<b>262</b>	<b>1457</b>	<b>647</b>	<b>2624</b>	<b>2463</b>
MS	9	5	11	118	35	173	168
MT	34	18	12	122	87	255	237
GO	12	6	73	93	201	379	373
DF	0	0	19	0	0	19	19
<b>Centro_Oeste</b>	<b>55</b>	<b>29</b>	<b>115</b>	<b>333</b>	<b>323</b>	<b>826</b>	<b>797</b>
<b>Total</b>	<b>1065</b>	<b>635</b>	<b>1103</b>	<b>5227</b>	<b>3465</b>	<b>10860</b>	<b>10225</b>

Fonte: Elaboração própria a partir de dados da pesquisa.

**Tabela A.2: Estatísticas descritivas das variáveis explicativas e explicadas**

Variáveis	Observações	Media	Desv. Padrão	Mín	Máx
Inc_todas_idades	10225	3,36	9,25	0	458,65
Inc_0a14	10225	5,86	19,36	0	1000
dcist	10225	1,41	5,03	0	85,97
sbceuab	10225	0,19	0,69	0	14,89
dfrud	10225	44,77	30,66	0	100
esg_ nao_ceuab	10225	30,25	31,44	0	100
ag_ nao_ceuab	10225	45,16	30,31	0	100
lixo_mor_inad	10225	0,06	0,41	0	13,71
d_nbueiro	10225	15,87	20,96	0	100
rendapc	10225	430,19	305,38	28,76	5373,99
gini	10225	0,60	0,06	0,30	0,92
mu_resp_alfab	10225	76,24	14,62	0	100
popurb	10225	56,89	29,98	0	100
dom_kmq	10225	153,46	681,49	0,02	16023,88
est_saude	10225	16,11	49,44	0	1259
dummy_cap	10225	0,04	0,19	0	1
dummy_reg_met	10225	0,05	0,22	0	1

Fonte: Elaboração própria a partir de dados da pesquisa.



**Figura A.1: Distribuição espacial das aglomerações das taxas de incidência de internações por DRSAI para a população de todas as idades (2010).**